

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Massey University, Auckland, New Zealand

# **Analysing Underpinning Patterns in Social Media Posts that Promote Fat Stigmatisation**

Vajisha Udayangi Wanniarachchi

2022

A thesis submitted in partial fulfilment of the requirements for the degree of  
Doctor of Philosophy (Ph.D) in Information Technology

## **Abstract**

Social media offers users an online platform to freely express themselves; however, when users post opinionated and offensive comments that target certain communities, this could instigate hatred towards them. With the global increase in obese/fat populations, social media discourses laced with fat hatred have become commonplace, leading to much fat stigmatising content being posted online. This research aims to investigate the patterns of fat stigma, and how female and male genders are positioned in fat stigmatising discourses that are being conducted over social media. To achieve this objective, a methodological framework is proposed for unearthing underlying stigmatising patterns prevalent in social media discussions, with specific focus on fat stigma. Methods incorporating natural language processing techniques such as sentiment analysis and topic modelling, along with discourse analysis have been described for classifying users' emotions and comprehending the stigma patterns embedded in social big data. The framework has been applied to weight-based textual data, extracted from Twitter and Reddit, to identify emergent gender-based themes, emotions and word frequency patterns that underpin the fat stigmatising content posted online. The experiential consequences of being considered fat across both genders have been analysed with objectification theory. The findings from this study have provided a holistic outlook on fat stigmatising content that is posted online which can further inform policymakers in planning suitable props to facilitate more inclusive social media spaces. This study showcases how lexical analytics can be conducted by combining a variety of data mining methods to draw out insightful subject-related themes that add to the existing knowledge base; therefore, has both practical and theoretical implications.

## **Acknowledgement**

3.5 years of a challenging yet fascinating life experience have been wrapped up with this thesis and many individuals have given their best guidance, support and encouragement to make it a success.

First and foremost, I would like to express my deepest appreciation to my tremendous supervisor Assoc. Prof. Anuradha Mathrani for her continuous support and guidance throughout this research. She has always encouraged me to look for positivity and motivated me to accept any challenge that comes on my way. Without her guidance, completing this thesis would have been hardly possible. I am also highly grateful to my co-supervisors, Dr Teo Susnjak and Prof. Chris Scogings, who provided invaluable insight, guidance and expertise.

I would like to extend my sincere thanks to the School of Mathematical and Computational Sciences for supporting my studies with a fee scholarship.

Special thanks to my parents, who have always believed in me, supporting me to push the cultural barriers and encouraging me to crave life adventures. You have no idea how much strength you have given me every day with a simple 'Hello'. Also, I would like to thank my brother for always being there for me.

Words cannot express my gratitude to my loving husband, Kalindu. You have been my biggest strength, and thank you for your love, emotional support, encouragement and constant patience when it was most needed. I love you!

## **List of Publications**

The presentations and publications that have been published from this research so far are:

1. Wanniarachchi, V. U., Scogings, C., Susnjak, T., & Mathrani, A. (2022). Hate Speech Patterns in Social Media: A Methodological Framework and Fat Stigma Investigation Incorporating Sentiment Analysis, Topic Modelling and Discourse Analysis. *Australasian Journal of Information Systems*. (Accepted)
2. Wanniarachchi, V. U., Mathrani, A., Susnjak, T., & Scogings, C. (2022). Methodological Aspects in Study of Fat Stigma in Social Media Contexts: A Systematic Literature Review. *Applied Sciences*, 12(10), 5045. doi:<https://doi.org/10.3390/app12105045>
3. Wanniarachchi, V. U., Scogings, C., Susnjak, T., & Mathrani, A. (2022). Fat stigma and body objectification: A text analysis approach using social media content. *Digital Health*, 8, 20552076221117404. doi:<https://doi.org/10.1177/20552076221117404>
4. Wanniarachchi, V. U., Mathrani, A., Susnjak, T., & Scogings, C. (2020). A systematic literature review: What is the current stance towards weight stigmatization in social media platforms? *International Journal of Human-Computer Studies*, 135, 102371. doi:<https://doi.org/10.1016/j.ijhcs.2019.102371>
5. Wanniarachchi, V. U., Mathrani, A., Susnjak, T., & Scogings, C. (2019). *Gendered objectification of weight stigma in social media: a mixed method analysis*. Paper presented at the Australasian Conference on Information Systems, Perth, Australia.

## Table of Contents

<b>ABSTRACT .....</b>	<b>3</b>
<b>ACKNOWLEDGEMENT .....</b>	<b>4</b>
<b>LIST OF PUBLICATIONS .....</b>	<b>5</b>
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>13</b>
1.1    INTRODUCTION.....	14
1.2    BACKGROUND OF THE RESEARCH.....	15
1.2.1    Social media and hate speech.....	15
1.2.2    Obesity and fat stigma.....	17
1.2.3    Fat stigma in social media .....	19
1.2.4    Gendered fat stigma observed in social media .....	20
1.3    OBJECTIFICATION THEORY.....	20
1.4    RESEARCH QUESTIONS .....	23
1.5    SCOPE .....	24
1.6    CONTRIBUTIONS.....	25
1.7    ROADMAP OF THE STUDY .....	27
1.8    OVERVIEW OF THE THESIS .....	29
<b>CHAPTER 2: LITERATURE REVIEW .....</b>	<b>31</b>
2.1    INTRODUCTION.....	32
2.2    METHOD .....	33
2.3    BIOLOGICAL, PSYCHOLOGICAL AND SOCIOLOGICAL CORRELATES.....	40

2.3.1	Biological correlates.....	40
2.3.2	Psychological correlates.....	42
2.3.3	Sociological correlates .....	44
2.3.4	Discussion of the findings.....	45
2.4	METHODOLOGICAL ASPECTS .....	50
2.4.1	Domain .....	50
2.4.2	Data characteristics.....	52
2.4.3	Analytical methods.....	52
2.4.4	Techniques.....	54
2.4.5	Tools/models.....	55
2.4.6	Features.....	56
2.4.7	Discussion of the findings.....	58
2.5	SUMMARY .....	63
<b>CHAPTER 3: PRELIMINARY ANALYSIS.....</b>		<b>65</b>
3.1	INTRODUCTION .....	66
3.2	ANALYSIS I: A TEXT ANALYSIS APPROACH TO IDENTIFY FAT STIGMA AND BODY OBJECTIFICATION IN SOCIAL MEDIA .....	66
3.2.1	Method.....	67
3.2.2	Analysis .....	69
3.2.3	Results.....	70
3.2.4	Discussion of the findings.....	79
3.3	ANALYSIS II: UNDERLYING THEMES OF FEMALE BODY OBJECTIFICATION IN SOCIAL MEDIA .....	84

3.3.1	Method.....	85
3.3.2	Computational Analysis.....	88
3.3.3	Qualitative Analysis.....	92
3.3.4	Discussion of the findings.....	99
3.4	SUMMARY .....	103
<b>CHAPTER 4: METHODOLOGICAL FRAMEWORK.....</b>		<b>106</b>
4.1	INTRODUCTION .....	107
4.2	METHODOLOGICAL FRAMEWORK.....	107
4.2.1	Data collection.....	109
4.2.2	Data pre-processing.....	111
4.2.3	Data analysis .....	113
4.3	Summary.....	118
<b>CHAPTER 5: APPLICATION OF PROPOSED FRAMEWORK ON FAT STIGMA DATA.....</b>		<b>119</b>
5.1	INTRODUCTION .....	120
5.2	DATA COLLECTION AND DATA PRE-PROCESSING .....	120
5.3	DATA ANALYSIS: OVERALL TEXT ANALYSIS OF FATNESS.....	121
5.3.1	Sentiment analysis.....	122
5.3.2	Emotion analysis .....	123
5.3.3	Sentiment labelling.....	125
5.4	PATTERN IDENTIFICATION OF FAT STIGMA DISCUSSIONS .....	125
5.4.1	Topic modelling .....	125
5.4.2	Topic Visualisation .....	128

5.4.3	Word Frequencies.....	130
5.4.4	Discourse Analysis.....	131
5.5	SUMMARY .....	133
<b>CHAPTER 6: DISCUSSION.....</b>		<b>135</b>
6.1	PATTERNS OF FAT STIGMA AND EMERGED GENDER DIFFERENCES .....	135
6.2	METHODOLOGICAL FRAMEWORK TO DETECT PATTERNS OF FAT STIGMA.....	138
<b>CHAPTER 7: CONCLUSION, LIMITATIONS AND FUTURE RESEARCH DIRECTIONS.....</b>		<b>142</b>
7.1	CONTRIBUTIONS.....	144
7.2	LIMITATIONS AND FUTURE RESEARCH DIRECTIONS.....	145
<b>APPENDICES .....</b>		<b>148</b>
<b>REFERENCES .....</b>		<b>155</b>

## List of Figures

Figure 1. Consequences of social media objectifying experiences .....	23
Figure 2. Roadmap of the study.....	28
Figure 3. Literature search process (adapted from (Page et al., 2021)).....	36
Figure 4. Biological, psychological and sociological correlates with fat stigma in social media.....	47
Figure 5. Number of articles by year .....	50
Figure 6. Domains used for data collection in selected literature: Social media domains (darkest blue), other online domains (blue) and online media not used (light blue). .....	51
Figure 7. Analytical methods used: Qualitative (darkest blue), quantitative (blue), experimental (light blue) and mixed methods (lightest blue).....	53
Figure 8. Features investigated by selected literature. ....	58
Figure 9. Histograms of sentiment values across the three corpora .....	73
Figure 10. Emotion values as percentage of emotion word count .....	73
Figure 11. Distribution of data excerpts among the identified themes .....	77
Figure 12. Social media objectifying experiences observed across males and females ...	82
Figure 13. Generated word co-occurrence maps for the data corpus .....	90
Figure 14. Sentiments associated with each identified theme .....	92
Figure 15. Methodological framework .....	108
Figure 16. Distribution of data.....	121

Figure 17. Sentiment value comparison in Twitter (left) and sentiment value comparison in Reddit (right) .....	122
Figure 18. Differences between males and females explained by each emotion. ....	124
Figure 19. Determining optimal number of topics.....	126
Figure 20. Word clouds of male and female datasets – Twitter .....	130
Figure 21. Word clouds of male and female datasets – Reddit .....	130

## List of Tables

Table 1. Search terms .....	35
Table 2. Systematic literature review .....	37
Table 3. Keyword distribution in Twitter, YouTube and Reddit.....	72
Table 4. Word co-occurrence maps .....	75
Table 5. Selected excerpts representing different sentiments.....	78
Table 6. Selected Reddit post .....	86
Table 7. Examples of analysing text content related to words.....	91
Table 8. Data sources and list of keywords.....	110
Table 9. Stemming and lemmatization .....	112
Table 10. Keyword combinations.....	120
Table 11. Plutchik's emotion wheels.....	123
Table 12. Sample data for sentiment labelling .....	125
Table 13. Model parameters and coherence scores .....	126
Table 14. Generated topics and topic labels for discourse in Twitter .....	127
Table 15. Generated topics and topic labels for discourse in Reddit.....	127
Table 16. Topic visualisation of each dataset.....	129

# **Chapter 1: Introduction**

## 1.1 Introduction

The proliferation of media platforms has been rapid over the past few decades; we have swiftly moved from one-way media platforms (i.e., newspapers and television) to rich two-way digital media platforms (i.e., online news media, social media). Digital platforms have provided us with an online presence and facilitated ease in publicly expressing our opinions, ideas and emotional feelings over matters that interest us. As such, social media platforms have been increasing in popularity since our opinions can reach a larger audience and thereby contribute to online discussions. Depending on the nature of the content being expressed, our posts can impact the ongoing discussions positively or negatively. Unfortunately, the emergence of hate speech content over social media platforms has gathered momentum. “Fat stigma” is one such form of hate speech that has surfaced online. The increased obesity rate around the globe has brought attention to weight-based online discussions, and we have witnessed many offensive comments targeting obese/fat individuals. Therefore, this study has conducted an investigation of the fat stigmatising content appearing in social media platforms to identify user attitudes and user beliefs towards obesity (fatness). The study has proposed a mixed-method framework which combines different text analytic methods to examine social media big data for fat stigma themes.

This chapter discusses the background of our research and the theoretical framework that has been used. It further highlights the main research questions that have been posed and explains the implications of this research study. Also, the roadmap for this study is presented to showcase how the study has been conducted to answer the research questions that have been posed.

## **1.2 Background of the research**

To comprehend the nature of fat stigma content found in social media platforms, it is important to first have knowledge of social media and its relation to hate speech propagation. Such awareness will provide the foundation for this research study, as it will build understanding on what is hate speech in a social media context, what is fat stigma, what aspects of fat ideologies are commonly expressed on social media platforms, and further which gendered aspects are being voiced in the context of obesity and fat stigma in social media. We have therefore carried out a background literature search on these areas to highlight their current stage of research. The following sub-sections provide a brief overview.

### **1.2.1 Social media and hate speech**

Social media has become an integral part of everyday lives enabling users to freely voice their opinions anytime over online platforms. The number of social media users are seen to be growing at a rate of 13% on year-on-year basis, with currently over 4.20 billion users, which represents 53% of the world's total population (Kemp, 2021). Online platforms have changed the way we interact with each other, and the way we acquire news (Gallacher, Heerdink, & Hewstone, 2021), enabling users to freely express themselves using emotive language when arguing over contested topics. Some scholars attribute the excessive use of social media to be a consequence of the “fear of missing out” (Tandon et al., 2021), as users often feel much pressure in maintaining online connectivity through regular social interactions. In fact, online social media discourses have been likened to the notion of ‘para-social interactions’ (Kumar & Benbasat, 2002) where technology takes the role of a non-human actant (Latour, 2005) that facilitates conversations for shaping out a shared social reality (Jin, Ryu, & Muqaddam, 2021;

Ledbetter & Meisner, 2021; Min et al., 2020). Online actions reflect our social identities, namely our cognitive (self-awareness), evaluative (self-worth) and affective (self-identification) (Cheung & Lee, 2010), when we engage in discussions on emergent topics (pertaining to health, education, politics or other societal issues). Online user conversations occur as and when users comment on each other's posts where they state their own stance on the topic under discussion which may express similar or diverging opinions.

Unfortunately, users may target certain community groups in these emotional exchanges which could promote negativity or discrimination towards them (Vigna et al., 2017). Communities can be aligned to multiple identities including that of race, religion, gender or sexual orientation, which can in turn lead to multiple levels of discrimination (Mossie & Wang, 2020). Behavioural researchers are therefore keen to make sense of how social media platforms are being used, such as, how public opinions are framed around certain topics or what prejudicial views are expressed in online conversations. Offensive behaviours are often motivated by the anonymity facilitated by social media platforms as people are known to be more comfortable in expressing provocative views when they know they are anonymous (Erjavec & Kovačič, 2012). While anonymity has a liberating impact on users, it also reduces their hesitation when expressing contempt towards those posts that they do not agree with (Tucey, 2010). Anonymity can lead to 'deindividuation and disinhibition' when users let go of their emotions without fear of being judged for their aggressive posts while discussing controversial topics (De Brún et al., 2014). Hate speech can be thus advanced and caused further divisions based on religion, colour, race and body type amongst others.

Users often experience more insults in the online space compared to physical spaces, as the online media is not designed to safeguard users from online abuse (Zhong, 2020). Therefore, social media platforms can promote stigmatisation if users feel safe in posting negative content online. It is indeed challenging to mitigate the impact of hateful commenting without having a proper understanding of the types and targets of hate speech, as comments contain overlapping targets (Salminen et al., 2018). Hateful comments may target various communities based on their sexual orientation, ethnic origin, religious faith, gender identity or body weight or any combination of these (Albadi, Kurdi, & Mishra, 2018; Matamoros-Fernández & Farkas, 2021; Ştefăniță & Buf, 2021; Wanniarachchi et al., 2019). The offensive content posted online is therefore very subjective with no clear definition of the target; hence, human interpretations of societal events must complement the data-centric computational approaches for analysing its hateful (Kocoń et al., 2021).

### **1.2.2 Obesity and fat stigma**

The past three decades have witnessed a startling increase in worldwide obesity rates (WHO, 2022). The steady global rise in obesity numbers have led to public discussions on obesity perceptions associated with 'ideal' body shapes and sizes too have evolved. Body image is described as cognitions, perceptions, and attitudes towards one's appearance, where an individual may measure their bodily features against some overall level of attractiveness (Izydorczyk et al., 2021). Such forms of 'ideal' body image perceptions reinforce thinness as a form of ideal body size, leading to more expectations about body shapes in society.

Recent works in behavioural science research have discussed widespread stigmatisation, specifically negative weight-based characterization, and how detrimental such content

can be to those who are struggling with weight issues (Ata & Thompson, 2010; Libbey et al., 2008; Puhl & Heuer, 2009). The social devaluation and denigration of individuals who carry excess weight is defined as weight stigma (Tomiyama, 2014). It can lead societies to form prejudices (e.g., negative perceptions from employers towards overweight employees), frame negative stereotyping (e.g., overweight individuals are slow and lazy) and encourage discrimination towards vulnerable people (e.g., ignoring overweight individuals or looking at them with contempt) (Puhl & Heuer, 2009). Weight stigmatising attitudes can be conveyed by close friends or family in typical environments such as in schools or healthcare settings (Puhl & Heuer, 2009), although disparaging overweight comments from strangers can often cause more emotional and physical harm (Vartanian, Pinkus, & Smyth, 2014).

Fat stigma is evident across multiple domains, be it employment, health care, education or from interpersonal relationships (Puhl & Heuer, 2009; Tomiyama, 2014). The most frequent source of fat stigma can be from the victim's family and close friends (Puhl & Brownell, 2006) who are quick to judge and comment on their body's shape and size. This is consistent with research indicating that individuals who are overweight report higher levels of relationship strain and lower levels of support from family members compared to thinner individuals (Carr & Friedman, 2006). Apart from the emotional strain of fat stigma, it limits opportunities, as studies have found that parents are less likely to financially support their overweight children, particularly daughters, while attending college (Crandall, 1995). These familial patterns of exclusion can push the individuals towards psychosis (Charles & O'loughlin, 2012). The stress associated with fat stigma causes numerous biological effects to obese people, including cortisol and oxidative stress (Tomiyama, 2014), systemic inflammation (Sutin et al., 2014), excess body fat

accumulation (Hunte & Williams, 2009), increased diastolic and systolic blood pressure (Lepore et al., 2006) and elevated ambulatory blood pressure (Puhl & Latner, 2007).

### **1.2.3 Fat stigma in social media**

One of the major reasons behind the spread of fat stigmatising attitudes, is the misconceptions and oversimplifications of 'obese' that is reinforced by media (Ata & Thompson, 2010). Opinions stating "obese could easily control their eating behaviours and lose weight, if they had more discipline' is a very prevalent misconception and oversimplification" (Zahn et al., 2014, p. 606). Further mass media preferences for slim body shapes (Brewis et al., 2011) are constantly being promoted which add to negative experiences for stigmatised individuals.

The rising popularity of social media has also been instrumental in the spread of negative attitudes towards obesity. These attitudes can be fuelled with specific social media movements such as "thinspiration" which mainly portrays skinny women endorsing specific lifestyle towards thinness (Alberga, Withnell, & von Ranson, 2018). Simultaneously, movements such as fat acceptance (FA) use blogs to spread messages among social media users that contest the dominant societal assumptions of beauty and body image and question the negative claims frequently associated with fatness (Afful & Ricciardelli, 2015; Donaghue & Clemitshaw, 2012). These types of social media movements further encourage users to engage in discussions on obesity and fat stigma. However, fat discourses in social media platforms have not been much studied so as to enable detection of covert linguistic expressions; interpretation of these discourses can aid policy makers in considering appropriate actions for mitigating the harmful impact of the hidden negative expressions.

#### **1.2.4 Gendered fat stigma observed in social media**

“Fat studies” have revealed higher rates of internalized stigma and risks of weight-based discrimination towards women in education and employment (Himmelstein, Puhl, & Quinn, 2017; Puhl, Andreyeva, & Brownell, 2008). Moreover, female children and adolescents considered obese, encounter more teasing and social marginalization (Almenara & Ježek, 2015; Tang-Péronard & Heitmann, 2008). The impact of weight bias and fat stigma on women is further studies by Fikkan and Rothblum (2012). Their review revealed fat stigma experiences, specifically received by women in employment, education and marital relationships. However, gender differences in obesity stigma have rarely been examined, and results are not consistent (Hebl & Turchin, 2005).

Fat men too face much social disapproval; they are perceived as less masculine and less desirable which undermines their status, pride and identity, resulting in “spoiling their social competence” (Lozano-Sufrategui et al., 2016). Therefore, irrespective of gender, those who do not fit within some specific ideal classifications (Cramblitt & Pritchard, 2013; Jackson, Jiang, & Chen, 2016; Rochelle & Hu, 2017) experience some form of weight-based discrimination that has a negative impact on their self-esteem.

### **1.3 Objectification theory**

The objectification theory proposed by Fredrickson and Roberts (1997) offers a framework for exploring how living in a culture that sexually objectifies females encourages them to self-objectify and evaluate their bodies. Objectification theory primarily has its roots on how “women’s bodies are looked at, evaluated, and always potentially objectified” (p. 175) that lead to self-surveillance behaviours where women internalize the observer’s gaze to judge their physical appearance and consequently may

experience anxiety, low self-esteem or body shame. Recently, the theory has been applied to explain internalization of body surveillance, self-objectification, muscular dissatisfaction among men (Heath et al., 2016). The twenty-first century female and male therefore, both face social pressure in relation to their appearance, that is, while females aspire for a slender and toned body, the males too may aspire for athletic figures, as they both engage in various body pursuits to enhance their body image (Grogan, 2021). The ongoing sociocultural pressures for having an ideal body shapes can result in equating one's self-worth with physical appearance (Tylka & Hill, 2004) to impact society's judgement wherein appearance dominates accomplishment (Ellemers, 2018) that in turn forms gender stereotyping attitudes.

Theorists assert that those women who fail to live up to their idealized body shapes may develop mental health issues and suffer from depression or eating disorders such as bulimia and anorexia. Exposure to thin body images over social media can lead to self-objectification, as the battle between 'fashion versus fitness' takes shape (Prichard et al., 2018). With different forms of inspirational fitness images ("fitspiration") surfacing on social media, Prichard et al. (2018) observed negative moods and body shame among Australian women when they were exposed to fitspiration images. Appearance-related online commentary and interactions including up/down (or like/dislike) votes can reinforce further body surveillance behaviours'. For example, women may develop selfie-related behaviour where they may retake many selfies or edit selfies before sharing them on social media (Chua & Chang, 2016). Over time, selfie-posting could lower one's self-worth, such as lead to facial dissatisfaction alongside body dissatisfaction (Wang et al., 2019). Similarly, conforming to a given masculine norm which is regularly endorsed by male models in mainstream media, can lead to internalization of athletic ideals among men (Grogan, 2021).

However, studies concerning men's body image have reported mixed results in connection to objectification theory. A study by Wiseman and Moradi (2010) reported self-objectification experiences of sexual minority men that indicated internalization of attractiveness levels, body surveillance behaviours, some sense of body shame, and symptoms related to eating disorders. Similar observations are presented by Heath et al. (2016), indicating internalization of a mesomorphic (or muscular) ideal arbitrated through self-objectification, although this study had recruited participants from gyms and bodybuilding groups. However, Davids, Watson, and Gere's (2019) investigation of body-related experiences of heterosexual men observed that they infrequently experienced self-objectification and body surveillance behaviour. A similar result has been reflected by Daniel and Bridges (2010) as their study reported no relationship of objectification theory with the drive for muscularity.

Figure 1 consolidates the objectification constructs namely, body shame, anxiety, low self-esteem and overall negativity (e.g., low motivation, body surveillance behaviours) that result as a consequence of internalization of the objectifying gaze and lead to depression, disordered eating and overall body dissatisfaction (Jones & Griffiths, 2015).

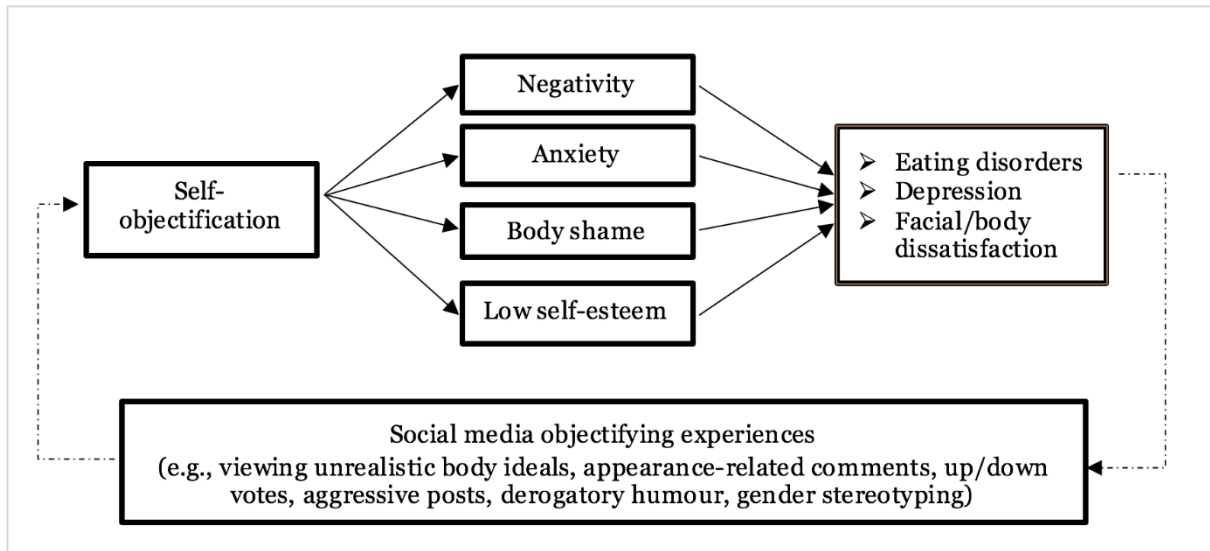


Figure 1. Consequences of social media objectifying experiences

Therefore, our study uses objectification theory as a foundation to monitor gender differences of fat stigma discourse in social media context.

#### 1.4 Research questions

The main purpose of this study is to investigate weight-based discussions in social media to identify fat stigmatising patterns. The background research has revealed the impact of fat stigma discussions and highlighted the importance of minimising the hate content being expressed. However, most research studies have tried to detect fat stigma content without analysing them further to gauge the underlying attitudes and perceptions being expressed. Recognition of the covert messages within this content can support policymakers in considering appropriate actions to mitigate the harmful impact of the hateful content.

Although gender plays a key role in verbal aggression towards fat people, the consideration of gender difference when investigating fat stigma content is limited.

Therefore, this research further used objectification theory as the underlying framework to explore male and female gender differences associated with fat stigma discussions.

In order to examine patterns of fat stigma discussions and how male and female genders are positioned in these discussions, this study considers both data mining and manual analysis techniques to gain deeper in text-based content. Prior studies assert that researchers must expand computational (or quantitative) capabilities with societal (or qualitative) interpretations when investigating fat stigma content (Brooker et al., 2018; Jeon et al., 2018; Lydecker et al., 2016). We have therefore proposed a novel mixed-methods framework that provides guidelines on integrating quantitative approaches to text analysis together with qualitative approaches through discourse analysis, in order to uncover fat stigma patterns in social media. The framework unravels different components of fat stigma discussions including sentiments, emotions and topics.

The following research questions clarifies the focus of this research.

- *RQ1: What are the patterns that emerged from fat stigma discussions in social media platforms?*
- *RQ2: How are male and female gender differences manifested in fat stigma content in social media?*
- *RQ3: How can we incorporate sentiment analysis, topic modelling and discourse analysis to detect fat stigma patterns in social media discussions?*

## **1.5 Scope**

The scope of this study was to investigate fat stigma content on two social media platforms Twitter and Reddit. Social media discussions relating to obesity, and female and male genders were captured in different time spans to perform preliminary analysis.

Data acquired in the period between June 2021 and June 2022 have been used to conduct the analysis to answer the research questions. Various keyword combinations have been used for this study, so as to encompass discussions associated with obesity, male and female genders.

## **1.6 Contributions**

The main contributions of this research are listed as follows.

- This study reveals different aspects of users' expressions including sentiments and emotions, and the type of topics underlying fat stigma discussions conducted over social media platforms. It further carries out a thorough analysis on the discourse to identify deeper conversational meanings. The findings of this research could aid social media platforms and policy creators to identify pathways that needs more attention when imposing social media policies, and filtering and modifying its content. It further contributes to fat research by highlighting the current societal stance towards fatness.
- The gender aspect of fat stigma in social media context has been weakly addressed in past research. With objectification theory as the theoretical framework, this study has investigated male and female gender biases in fat stigma discussions. The research creates knowledge of how males and females are positioned in fat stigma or obesity-related discussions. Therefore, by giving consideration to male and female genders when analysing fat stigma, this study can be expanded upon in future fat study research to explore other gendered aspects.
- To analyse fat stigma content, this study has proposed a novel methodological framework. The framework uses a mixed-method approach that combines quantitative and qualitative text analysis methods. It proposes combining

sentiment analysis, topic modelling and discourse analysis while accompanying these with emotion analysis and word frequencies to unravel patterns of fat stigma. The framework has been further validated by showing its application to larger datasets. Such methodological framework assists in bridging the gap between computational analysis and social sciences by establishing pathways for better comprehension of insights while conducting interdisciplinary studies.

- This thesis has contributed to the publication of the following research articles.

- [1] Wanniarachchi, V. U., Scogings, C., Susnjak, T., & Mathrani, A. (2022). Hate Speech Patterns in Social Media: A Methodological Framework and Fat Stigma Investigation Incorporating Sentiment Analysis, Topic Modelling and Discourse Analysis. *Australasian Journal of Information Systems*. (Accepted)
- [2] Wanniarachchi, V. U., Mathrani, A., Susnjak, T., & Scogings, C. (2022). Methodological Aspects in Study of Fat Stigma in Social Media Contexts: A Systematic Literature Review. *Applied Sciences*, 12(10), 5045. doi:<https://doi.org/10.3390/app12105045>
- [3] Wanniarachchi, V. U., Scogings, C., Susnjak, T., & Mathrani, A. (2022). Fat stigma and body objectification: A text analysis approach using social media content. *Digital Health*, 8, 20552076221117404. doi:<https://doi.org/10.1177/20552076221117404>
- [4] Wanniarachchi, V. U., Mathrani, A., Susnjak, T., & Scogings, C. (2020). A systematic literature review: What is the current stance towards weight stigmatization in social media platforms? *International Journal of Human-Computer Studies*, 135, 102371. doi:<https://doi.org/10.1016/j.ijhcs.2019.102371>

[5] Wanniarachchi, V. U., Mathrani, A., Susnjak, T., & Scogings, C. (2019). *Gendered objectification of weight stigma in social media: a mixed method analysis*. Paper presented at the Australasian Conference on Information Systems, Perth, Australia.

## **1.7 Roadmap of the study**

The following figure illustrates the roadmap of the study which was followed to answer the posed research questions and achieve study objectives. The steps of this roadmaps are further discussed in Chapter 3, 4 and 5.

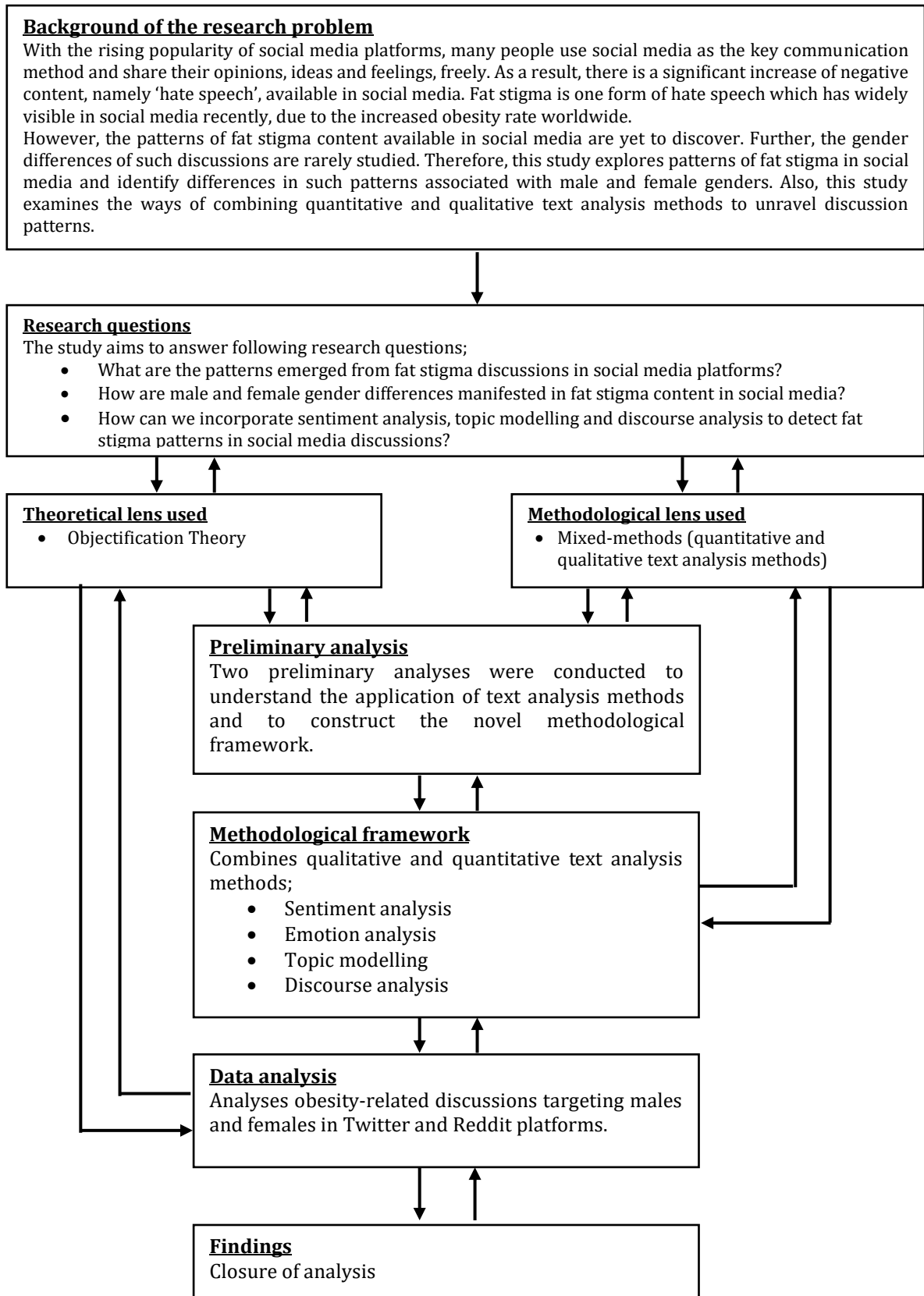


Figure 2. Roadmap of the study

## **1.8 Overview of the thesis**

The thesis is divided into seven chapters. This chapter discusses the background of the study, its objectives and research questions and the scope of the study. It further illustrates the contribution of the study and the roadmap that was taken to carry out this research.

Chapter two combines two systematic literature reviews that were carried out to understand the theoretical and methodological aspects covered by fat stigma studies in the past decade. One systematic literature review identifies the biological, psychological and sociological correlates highlighted by fat stigma studies. The other systematic literature review observed the methodological aspects followed by fat stigma studies in the recent past. The chapter has combined these two systematic literature reviews (which have been published as separate papers).

Chapter three discusses the methods followed in the preliminary analysis and discusses the outcomes. The initial identification of theoretical and methodological gaps motivated the next stage where two preliminary analyses were conducted. The chapter combines these two analyses and illustrates the application of quantitative and qualitative text analysis methods on different social media platforms. The limitations of data sources and analysis techniques are identified.

Chapter four outlines a novel methodological framework proposed by this study design based on the findings of systematic literature reviews and preliminary analyses. It proposes a mixed-method approach which combines qualitative and quantitative text analysis methods.

Chapter five discusses the patterns that emerged from the application of the methodological framework on fat stigma discourse. It identifies the differences of male and female genders in fat stigma discussions and further contrast Twitter and Reddit platforms in relation to fat stigma discourse.

Chapter six conducts the discussion over the proposed methodological framework and the results of its application. The discussion provided in this section answers the research questions posed in the introduction chapter.

Chapter seven concludes the research, highlights the limitations of the study and provides future research directions.

## **Chapter 2: Literature Review**

## 2.1 Introduction

This chapter examines prior fat stigma research to lay out the first steps towards achieving the objectives of this research study and providing answers to the research questions that have been posed. When analysing fat stigma discourse in social media, prior research has been concentrated on biological, sociological and psychological correlates (Carels et al., 2009; Chung, Chiou, & Chen, 2015; Papadopoulos & Brennan, 2015). Biological health factors comprising body mass index (BMI) issues (Vartanian & Shaprow, 2008) and hypertension (Tomiyama, 2014) can result in physical disabilities. Moreover, low level of social support due to less education, poor employability, lower socio-economic status or lack of interpersonal relationships (Papadopoulos & Brennan, 2015) can make individuals easy targets of weight stigmatisation. The psychological impact from depression and anxiety (Puhl & Brownell, 2006; Vartanian & Shaprow, 2008), suicidal thoughts (Roberto et al., 2012), body dissatisfaction leading to bulimia (Jung et al., 2017) and poor self-esteem (Vartanian & Novak, 2011) further reinforces negative beliefs about oneself. The negativity may also occur due to poor psychological functioning (Durso et al., 2012) as a reaction to weight stigma. However, these consequences appear irrespective to BMI (Wott & Carels, 2010), with authors claiming that the stigmatising experience itself can result in psychological outcomes.

Further studies are needed to decipher the technical details of how these correlates can be discovered with social media big data. Various text mining and machine learning methods including sentiment analysis, topic modelling, emotion analysis and co-occurrence analysis have been implemented in hate speech detection and classification studies (e.g., Alshalan et al., 2020; Cao, Lee, & Hoang, 2020; Martins et al., 2018; Rodríguez, Argueta, & Chen, 2019, etc.). These technical methods promote the

investigation of large amounts of textual data which assist in unravelling the stigmatising information that is embedded in the social media discussions. As social media discourses on obesity evolve, examining the methodological approaches used for analysing obesity content in prior literature has enabled in gauging the current state of research in this field of study.

Therefore, two systematic literature reviews have been conducted next to identify the correlates analysed by prior research in the past decade and examine the methodological aspects that were used to identify these correlates. Both these reviews have been conducted separately during different stages of this study and the combined results have been presented in this chapter. This chapter discusses the methodology followed for conducting literature reviews, and details prior studies results and limitations that were observed.

## **2.2 Method**

Prior research studies covering fat stigma content in obesity-related discourses over social media have been identified in this work and reviewed by following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol (Page et al., 2021). PRISMA has provided researchers, methodologists, clinicians, medical editors and analysts with guidelines to enable them audit previous research studies and reach an un-biased and rational view of the subject at hand. The PRISMA protocol helps “improve the transparency and the scientific merit” (Swartz, 2011, p. 1) of research progression that is representative of a wider trend. This is further evidenced by the number of systematic reviews that have been conducted on various research topics including that of obesity (e.g., Lin et al., 2017; Papadopoulos & Brennan, 2015; Pawalia et al., 2016; Wanniarachchi et al., 2020). It facilitates integrating key elements of relevant

research articles by undertaking an iterative process that involves identifying, screening, checking eligibility and then critically appraising each article. An analysis of article metadata is performed to ensure that a proper publication-bias elimination process is followed during selection of the research articles. That is, a full electronic screening process is proposed for examining and scoping articles that need to be considered by specifying search limits (e.g., keywords, abstract, publication date, article library catalogues). Articles that meet the search criteria are confirmed as valid and relevant items for further review.

Our literature review used the EBSCO Discovery Service (EDS) (EBSCO, 2022). EDS is a web-enabled resource that integrates search capabilities to filter relevant articles from a wide range of electronic databases (i.e., library catalogues and local digital collections). Search results are presented in a relevancy ranked list that can be further refined by a variety of options (e.g., journal article, book, newspaper). These results are linked to content provider platforms that enable ease of access by further prompting users with links that could take them to full-text content (Foster & MacDonald, 2013).

Using the search words (shown in Table 1) that have been identified from literature, we obtained relevant research articles from EDS for this review. Further, we limited the article search to peer-reviewed articles that were published in the English language between 1 January 2012 and 1 January 2022. The search results comprised original quantitative and qualitative research studies that had examined fat stigma in any form and on any social media platform. Each search result's title and abstract were thoroughly examined to extract those articles that were relevant to the study's context, that is, they specifically dealt with fat stigma detection in social media settings. The articles that met the inclusion criteria have next been evaluated to seek out their research design and

methods, that is, we sought out the study’s purpose or concept, the data sources that were used, the analysis methods that were applied, which techniques were employed and how different tools/models aided their investigation. We also sought to understand the features that were probed during the study’s analysis along with the limitations identified by the study.

Table 1. Search terms

<b>Context</b>	<b>Search Terms</b>
Fat Stigma	“weight stigma”, “obesity stigma”, “weight bias”, “fat bias”, “fat shaming”, “body shaming”, “obesity”, “overweight”, “over weight”
Social Media/Online Media	“social media”, “social networks”, “twitter”, “facebook”, “youtube”, “reddit”, “social networking”, “online forums”, “online media”

Figure 3 describes the four steps - Identification, Screening, Eligibility and Inclusion - used in following the PRISMA protocol for this study. The initial EDS search resulted in 2070 articles. After eliminating 132 duplicates, a rough screening of the titles and the abstracts resulted in further removal of 1861 articles that were further assessed. Next, those wherein the research studies were not based on social media data (n= 55) were excluded from further review. Finally, 22 articles were found to meet our inclusion criteria.

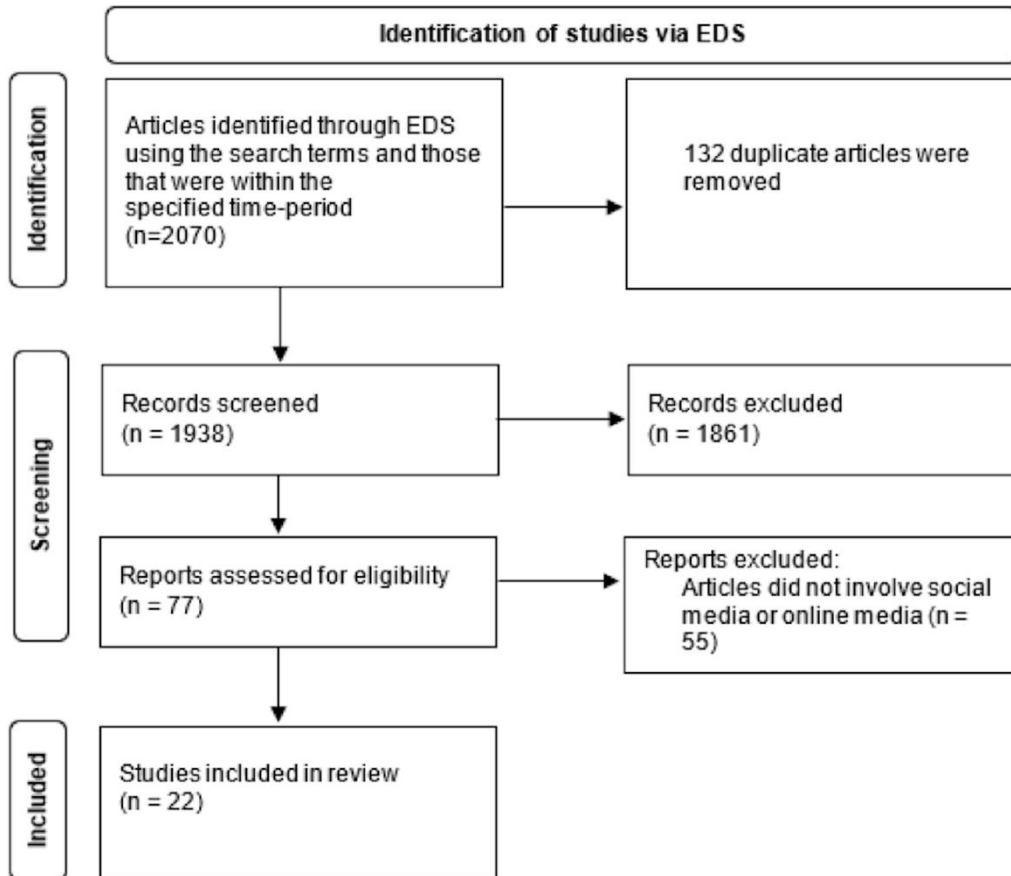


Figure 3. Literature search process (adapted from (Page et al., 2021))

The 22 selected articles have been assessed for the purposes of exploring the correlates, how experimental activities were conducted, what was the nature of the empirical datasets that were used, and which social media platforms were considered. Table 2 provides an overview of these assessments, which have been described in more detail in subsequent sections of this chapter.

Table 2. Systematic literature review

Citation	Domain	Dataset Characteristics	Correlates			Methodological Aspects			
			Psychological	Sociological	Biological	Analytical Methods	Techniques	Tools/ Models	Features
(Yoo & Kim, 2012)	YouTube	Videos; 417 videos	<ul style="list-style-type: none"> <li>- Humour</li> <li>- Stereotypical eating behaviours</li> </ul>			<ul style="list-style-type: none"> <li>Qualitative</li> <li>- Content analysis</li> </ul>	Manual coding	N/A	<ul style="list-style-type: none"> <li>- Themes</li> <li>- Frames</li> </ul>
(Lee et al., 2013)	Facebook	Structured text; 159 American and 137 Korean women participants	<ul style="list-style-type: none"> <li>- Body dissatisfaction</li> <li>- Psychological well-being</li> </ul>	<ul style="list-style-type: none"> <li>- Gender</li> <li>- Social comparison</li> </ul>		Experimental	Regression analysis (statistical analysis)	N/A	Effects of obesity discussions
(Chou, Prestin, & Kunath, 2014)	<ul style="list-style-type: none"> <li>- Twitter</li> <li>- Facebook</li> <li>- Blogposts</li> <li>- Forums</li> <li>- Website comments</li> </ul>	Unstructured text; 1.37 million posts	<ul style="list-style-type: none"> <li>- Self-derogation</li> <li>- Nexus of responsibility for obesity</li> <li>- Stereotypes</li> </ul>	Gender		<ul style="list-style-type: none"> <li>Mixed method</li> <li>- NLP</li> <li>- Qualitative linguistic analysis</li> </ul>	<ul style="list-style-type: none"> <li>- Discourse analysis</li> <li>- Sentiment analysis</li> <li>- Descriptive statistics</li> </ul>	N/A	<ul style="list-style-type: none"> <li>- Sentiments</li> <li>- Themes</li> </ul>
(De Brún et al., 2014)	YouTube	Unstructured text; 2872 comments	<ul style="list-style-type: none"> <li>- Self-indulgence</li> <li>- Stereotypes</li> <li>- Narrative resistance</li> </ul>			<ul style="list-style-type: none"> <li>Qualitative</li> <li>- Thematic analysis</li> </ul>	Manual coding	Nvivo	Themes
(Harris et al., 2014)	Twitter	Unstructured text; 1110 tweets	<ul style="list-style-type: none"> <li>- Internal and external attribution</li> <li>- Lifestyle choices</li> </ul>		Public health risks	<ul style="list-style-type: none"> <li>Qualitative</li> <li>- Content analysis</li> <li>- Network modelling</li> </ul>	<ul style="list-style-type: none"> <li>- Descriptive statistics</li> <li>- Visualization</li> <li>- ERGM</li> </ul>	<ul style="list-style-type: none"> <li>- IBM SPSS</li> <li>- Pajek 64</li> <li>- R statnet</li> </ul>	<ul style="list-style-type: none"> <li>- Themes</li> <li>- Network characteristics</li> </ul>
(Taniguchi & Lee, 2015)	Facebook	Structured text; 159 American and 102 Japanese women participants	<ul style="list-style-type: none"> <li>- Physical attractiveness</li> <li>- Self-esteem</li> <li>- Psychological well-being</li> </ul>			Experimental	Statistical analysis	N/A	Characteristics of obesity-related discussions
(Kent et al., 2016)	Twitter -Facebook	Unstructured text; 1382 posts	Life style choices		<ul style="list-style-type: none"> <li>- Perceptions towards cancer and obesity</li> <li>- Diet/Food factors</li> <li>- Public health risks</li> </ul>	<ul style="list-style-type: none"> <li>Mixed methods</li> <li>- Quantitative approach embedded descriptive qualitative analysis</li> </ul>	<ul style="list-style-type: none"> <li>- Manual coding</li> <li>- Sentiment analysis</li> <li>- Bivariate frequency analysis</li> </ul>	SAS	<ul style="list-style-type: none"> <li>- Themes</li> <li>- Sentiments</li> </ul>

(Lydecker et al., 2016)	Twitter	Unstructured text; 4596 tweets	<ul style="list-style-type: none"> <li>– Body dissatisfaction</li> <li>– Body-stereotypes</li> </ul>	Gender		Qualitative <ul style="list-style-type: none"> <li>– Content analysis</li> </ul>	Manual coding	N/A	<ul style="list-style-type: none"> <li>– Themes</li> <li>– Sentiments</li> </ul>
(So et al., 2016)	Twitter	Unstructured text; 120 tweets	Humour			Qualitative <ul style="list-style-type: none"> <li>– Content analysis</li> </ul>	Manual coding	N/A	<ul style="list-style-type: none"> <li>– Themes</li> <li>– Emotions</li> <li>– Causes</li> </ul>
(Webb et al., 2017)	Instagram	Images; 400 images	<ul style="list-style-type: none"> <li>– Body dissatisfaction</li> <li>– Body stereotypes</li> </ul>	<ul style="list-style-type: none"> <li>– Gender</li> <li>– Ethnicity</li> </ul>		Qualitative <ul style="list-style-type: none"> <li>– Content analysis</li> </ul>	Manual Coding	IBM SPSS	Themes
(Brooker et al., 2018)	The Guardian online	Unstructured text; 1452 comments	Lifestyle choices	Social responsibility	<ul style="list-style-type: none"> <li>– BMI</li> <li>– Illness</li> <li>– Disability</li> </ul>	Qualitative <ul style="list-style-type: none"> <li>– Frame analysis</li> </ul>	Co-occurrence analysis	Textometrica	Themes
(Holmberg et al., 2018)	N/A (Data not directly acquired from social media)	Structured text; 20 participants	Physical appearance	Gender		Qualitative <ul style="list-style-type: none"> <li>– Interviews with participants using multiple social media platforms</li> </ul>	Manual coding	N/A	Effects of fat stigma
(Jeon et al., 2018)	YouTube	Unstructured text; 316 comments from 2 videos		<ul style="list-style-type: none"> <li>– Gender</li> <li>– Heterosocial skills</li> </ul>		Qualitative <ul style="list-style-type: none"> <li>– Content analysis</li> </ul>	Manual coding	N/A	Characteristics of obesity-related discussions
(Karami et al., 2018)	Twitter	Unstructured text; 4.5 million tweets			<ul style="list-style-type: none"> <li>– Diet/food factors</li> <li>– Health risks</li> </ul>	Qualitative <ul style="list-style-type: none"> <li>– Content analysis</li> <li>– Topic modelling</li> </ul>	LDA Lexicon based approach	LIWC	Themes
(Lim & An, 2018)	N/A (Data not directly acquired from social media)	Structured text; 202 participants	<ul style="list-style-type: none"> <li>– Internal and external attribution</li> <li>– Body dissatisfaction</li> </ul>	<ul style="list-style-type: none"> <li>– Social media exposure</li> <li>– Gender</li> </ul>		Quantitative <ul style="list-style-type: none"> <li>– Survey questions given to adolescents who use social media</li> </ul>	Regression analysis (statistical analysis)	N/A	Effects of obesity discussions
(Yeruva, Junaid, & Lee, 2019)	Twitter PubMed	Unstructured text; 103609 Twitter and 6602 PubMed article abstracts			Diet/food factors	Qualitative <ul style="list-style-type: none"> <li>– Content analysis</li> <li>– NLP</li> <li>– Topic modelling</li> </ul>	<ul style="list-style-type: none"> <li>– TF-IDF</li> <li>– Word embeddings</li> <li>– Sentiment analysis</li> <li>– Co-occurrence analysis</li> <li>– LDA</li> <li>– Word2Vec</li> </ul>	<ul style="list-style-type: none"> <li>– Apache spark</li> <li>– Tensor Flow</li> <li>– CoreNLP</li> <li>– VADER</li> <li>– TextBlob</li> </ul>	<ul style="list-style-type: none"> <li>– Sentiments</li> <li>– Co-occurrences</li> <li>– Themes</li> </ul>

(Mitei & Ghanem, 2020)	Twitter	Unstructured text; 2500 tweets			Health risks	Quantitative – Cluster analysis – User analysis	Social media clustering	– Clauset-Newman-Moore clustering algorithm – NodeXL	Characteristics of obesity-related discussions
(Busam & Solomon-Moore, 2021)	Facebook	Unstructured text; 11 newspaper outlets, 30 news articles and 1104 responding comments	Lifestyle choices		– Health risks – Diet/food factors	Qualitative – Frame analysis	Manual coding	– R studio – Nvivo	– Frames – Themes
(Chansiri & Wongphothiphan, 2021)	Instagram	Structured text; 221 female participants	Self-esteem	Social comparison	BMI	Experimental	MMMA	N/A	Effects of idealized images
(Lazarus et al., 2021)	Twitter	Unstructured text; 18 274 NAFLD, 2621 NASH and 10 million tweets			Public health risks	Qualitative – Content analysis	– Sentiment analysis – Discourse analysis	– Self-developed NLP module – Daturks platforms	– Sentiments – Themes
(Lessard & Puhl, 2021)	N/A (Data not directly acquired from social media)	Structured text; 452 participants	Body dissatisfaction	– Gender – Social media exposure		Quantitative – Survey questions were given to adolescents who use social media	Statistical analysis	IBM SPSS	Characteristics of obesity-related discussions
(Bograd, Chen, & Kavuluru, 2022)	Twitter	Unstructured text; 2000 tweets		Social comparison		Qualitative – Content analysis	– Sentiment analysis – Manual coding	– AWD-LSTM – ULMFIT	– Sentiments – Supportiveness for Fat Acceptance Movement

## **2.3 Biological, psychological and sociological correlates**

The biological, psychological and sociological correlates stemming from the twenty-two articles have been analysed next. A summary of sampling methods, social media platforms used, research methods and research settings and key takeaways related to weight stigmatisation from each article are detailed in Table 2. The following subsections elaborate on the findings of biological, psychological and sociological correlates with weight stigmatisation.

### **2.3.1 Biological correlates**

The human physiology is conditioned by caloric consumption, energy demands, metabolism, physical activity and the individual's genetic makeup. Excessive weight gains or weight loss amongst populations translates to inequities in health care opportunities resulting in potential harmful stigmatisations for the concerned individuals. Two studies discuss how excess weight or obesity is a medicalized problem that may result from personal bodily irresponsibility, and which carries negative health effects that can strain our scarce public health funds (Brooker et al., 2018; Chansiri & Wongphothiphan, 2021). Brooker et al. (2018) revealed three underlying themes related to obesity: excess weight is a result of current lifestyles and can lead to disabilities, excess weight stems from irresponsible choices and can lead to health risks and that being overweight is a complex problem which cannot be treated with standard medical treatments. Chansiri and Wongphothiphan (2021) also took BMI into account to categorize study participants into two categories: healthy weight and overweight. From studying the interaction effects of #fitspiration and #thinspiration images in Instagram, their study revealed significant three-way interaction between the images, BMI and perceived weight of the participants.

The connection between obesity and healthy dieting, as discussed in social media, has been studied by Busam and Solomon-Moore (2021); Karami et al. (2018); (Yeruva, Junaid, & Lee, 2019). (Yeruva, Junaid, & Lee, 2019) have studied the sentiments towards healthy, unhealthy and compound food categories and indicated that unhealthy-negative food sentiment is highly visible in their dataset. Six studies have identified public health risks and health consequences related to obesity (Busam & Solomon-Moore, 2021; Harris et al., 2014; Karami et al., 2018; Kent et al., 2016; Lazarus et al., 2021; Mitei & Ghanem, 2020). The perceptions of social media users towards obesity and cancer and the resulting factors were discussed by Kent et al. (2016). According to the study results, many users believe that obesity increases cancer risk, which is consistent with results of a national survey discussing the connection between casual beliefs of obesity and lifestyle behaviour related to weight management (Wang & Coups, 2010). Karami et al. (2018) also studied public opinion on diabetes, diet, exercise, and obesity (DDEO). Their study highlighted that obesity is often linked with chronic diseases such as diabetes.

People who have Non-Alcoholic Fatty Liver Disease (NAFLD) and Non-Alcoholic Steatohepatitis (NASH) are partly stigmatised because of obesity as it is a common condition characterized by these people. Lazarus et al. (2021) studied social media content associated with NAFLD, NASH and obesity to investigate stigma and discrimination. The most commonly discussed obesity-related topics in Twitter were studied by Mitei and Ghanem (2020) and discovered that the topic of discussing obesity as healthcare complaint or lack of healthcare support for people with obesity draws greater social media attention compared to other common topics, like causes of obesity and childhood obesity. Busam and Solomon-Moore (2021) highlighted the medical, societal and behavioural frames of childhood obesity according to social media. The study

observed that social media users' views range from considering childhood obesity as a lifestyle choice such as healthy diet and exercise to appraising childhood obesity as far as the child is healthy.

### **2.3.2 Psychological correlates**

Scientific evidence has documented the psychological impact of weight stigmatisation on overweight and obese persons. These have been identified as body dissatisfaction, body stereotypes, self-derogation, eating disorders, poor self-esteem, derogatory humour and discrimination amongst many others.

Five articles identified the influence of weight stigmatisation on body dissatisfaction among overweight and obese individuals (Lee et al., 2013; Lessard & Puhl, 2021; Lim & An, 2018; Lydecker et al., 2016; Webb et al., 2017). Body dissatisfaction was examined on Twitter (Lydecker et al., 2016) when exploring the weight stigmatised contents, and on Instagram (Webb et al., 2017) for weight loss and fitness promoting hashtags. The fat shaming nature of these messages is pervasive and reinforces negative body images as gluttonous, not attractive, not desirable and sedentary to form a stigmatised identity. Lee et al. (2013) have identified the effect of weight stigma on body dissatisfaction from online media stories. The importance of thinness gets over emphasized as different people engage in social comparisons, and where the scope of thinness differs among different body types resulting in lower self-acceptance. Lessard and Puhl (2021) investigated the perceptions of adolescents in regards to their bodies and their weight stigma experience from peers and parents during the COVID-19 pandemic. The study revealed subjective body dissatisfaction as a result of stigmatising experience they receive from their parents and peers.

Some studies claim that often overweight and obese people are accused of being responsible for their weight gain over social media posts (Brooker et al., 2018; Chou, Prestin, & Kunath, 2014; De Brún et al., 2014; Kent et al., 2016). An analysis on online news media, discussion boards (De Brún et al., 2014), Facebook and Twitter data reveals perceptions of obese people to be lacking in self-control and being lazy. Two studies inspected the relationship between internal and external attributions among adolescents with excess weight stigma (Harris et al., 2014; Lim & An, 2018). Lim and An (2018) found evidence of interaction between external attribution and social media exposure, since perceptions of body images were influenced by advertisements associating beauty with thinness.

Two studies identified the effect of weight stigma on the psychological well-being of overweight and obese females across Asian (Korean and Japanese) and American cultures (Lee et al., 2013; Taniguchi & Lee, 2015). Online surveys revealed perceptions of low-self-esteem and poor body image to be associated with overweight profile pictures compared with thin profile pictures especially among Asian participants.

Online media also promotes unrealistic body-stereotyping imagery that makes heavier individuals feel marginalized and demoralized (Chou, Prestin, & Kunath, 2014; De Brún et al., 2014; Webb et al., 2017). Stereotypical portrayals of extreme obesity have been observed in Instagram by (Webb et al., 2017). Chou, Prestin, and Kunath (2014) witnessed negative sentiments on overweight individuals because of body-stereotypes on multiple social media platforms (such as Facebook and Twitter). One study identified how stereotypical eating behaviour has been portrayed in YouTube videos to convey weight stigmatisation to the viewers. "Eating unhealthy food and having a sedentary lifestyle have been portrayed as major causes of obesity in every single category of

YouTube videos” (Yoo & Kim, 2012, p. 94).

In their qualitative analysis on the comments of online message boards, De Brún et al. (2014) revealed how public opinion on “larger bodies threaten an individual’s cultural and symbolic capital” (p. 78). Body appearance norms in social media affects overweight and obese individuals who purposefully portray themselves as thin to avoid weight stigmatisation (Holmberg et al., 2018; Taniguchi & Lee, 2015).

Self-derogation too is a consequence of weight stigmatisation in social media (Chou, Prestin, & Kunath, 2014). In their qualitative analysis of stigmatising content from numerous social media platforms, the authors witnessed many derogatory posts regarding diet and perception of one’s own weight to be rather self-depreciating. Studies have also revealed derogatory humour to garner interest in social media (Chua & Chang, 2016; Yoo & Kim, 2012). One study identified “obese persons were the objects of humour in one out of every five YouTube videos” and videos of obese individuals with teasing tone have been mostly viewed, rated and commented by the users (Yoo & Kim, 2012, p. 95). A similar study too has noted the same effect on Twitter, where emotionally evocative obesity content are retweeted more compared to non-derogatory weight-related jokes (So et al., 2016).

### **2.3.3 Sociological correlates**

Social correlates of weight stigma include gender bias, age bias and ethnicity bias. It has an impact on their social relationships and interpersonal skills, where overweight individuals have reported more strained personal relationships and lower social support (Papadopoulos & Brennan, 2015).

Recent studies investigating the association of the target's gender with the stigmatising content note that women experience more weight stigmatisation (Chou, Prestin, & Kunath, 2014; Holmberg et al., 2018; Jeon et al., 2018; Lessard & Puhl, 2021; Lydecker et al., 2016). One study discovered that "Women were more likely than men to be referenced in negative Tweets, and were also more likely to be referenced in Tweets about others" (Lydecker et al., 2016, p. 233). One study detected gender factor in stigmatising content on YouTube but did not find much significant difference in the frequency of verbal attacks between both gender targets (Jeon et al., 2018). Yet another study identified sexism and misogyny in stigmatising content in different social media platforms and discovered that women are generally targeted more often (Chou, Prestin, & Kunath, 2014).

Another study investigated how overweight/obese men and women were attacked for their heterosocial skills in YouTube comments (Jeon et al., 2018). The study results "indicate that male targets are more likely to receive comments that attack their heterosocial or interpersonal skills" (Jeon et al., 2018, p. 6), while women are attacked for their capacities (e.g., laziness, maturity) with more use of swear words.

#### **2.3.4 Discussion of the findings**

The aim of this literature review is to identify to what extent the weight stigmatisation has been spread over social media platforms. Among the literature published in the past decade, the majority have identified the psychological correlates with weight stigma followed by sociological and biological correlates respectively. This finding aligns with the findings of the systematic literature review carried out by Papadopoulos and Brennan (2015, p. 1756) on "biopsychosocial correlates of weight stigma separately in the community and treatment-seeking samples of adults with overweight/obesity". The

included literature for the systematic literature review is categorised into two: the literature that directly identifies weight stigma in social media, and the literature which captures indirect weight stigma in social media mostly in weight-loss or clinical interventions (Kent et al., 2016; Webb et al., 2017). Though the number of such interventions captured in the literature review is minimal, the insights of indirect weight stigma should not be ignored. The proliferation of fat- versus fit-themed body images enhances risk for causing body image disturbance and encourages unrealistic expectations. Instead Webb et al. call for broader inclusion of diverse body sizes to promote fat acceptance and exemplify that “beauty and style are attainable and not privileges reserved only for thin women” (p. 60).

Figure 4 draws upon the themes extracted from this systematic literature review related to biological, sociological and psychological correlates with weight stigmatisation over social media platforms. Prior research has identified biological (BMI, disability, public health risks, etc.), psychological (body dis-satisfaction, psychological well-being, body-stereotypes, etc.) and sociological (social comparison, social media exposure, heterosocial skills, etc.) consequences of public discourses on overweight and obese individuals over different social media platforms.

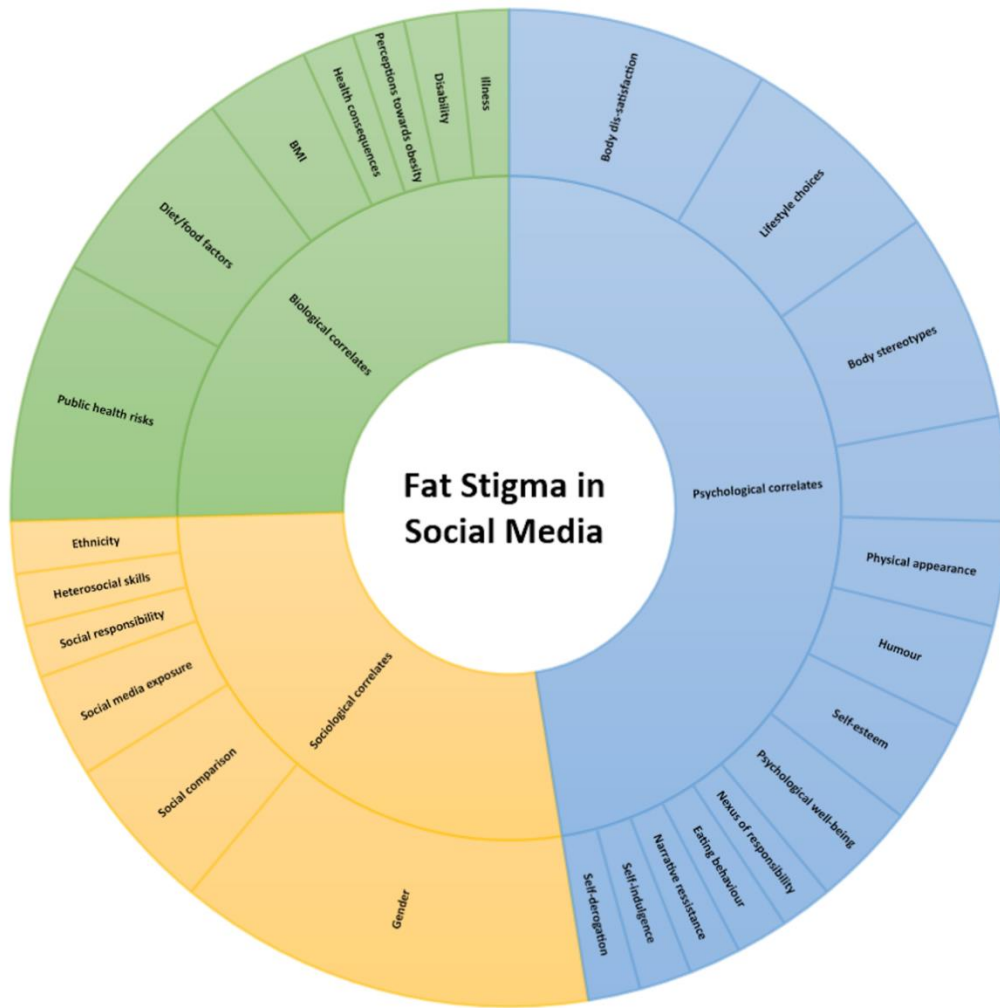


Figure 4. Biological, psychological and sociological correlates with fat stigma in social media

Majority of the selected articles have emphasized the anonymous nature of social media as one of the main reasons of weight stigma in social media platforms. Being anonymous, individuals can openly share their beliefs and attitudes with the rest of the world. With this unconditional freedom of anonymity in social media platforms, the spread of negative sentiments and hate speech is manifested by verbal aggression combined with stigmatising messages. The aim is to humiliate vulnerable persons (in this case the overweight individuals) by exacerbating stereotypical roles (like overweight people are

lazy, irresponsible and overindulgent) and presenting them as targets of ridicule to be laughed upon by general public (Vartanian, Pinkus, & Smyth, 2014).

There are different forms of weight stigma existing in social media as identified by the previously published literature. The main facility of any social media platform is to allow the user to discuss and share their personal opinions and know opinions of others' within a closed group of like-minded people or over an open public forum. As a result, most internet users use social media to share their personal experience with a larger audience. As identified by some of the existing literature, one form of weight stigma in social media is targeting these personal experiences, a personal video or an image of a specific person (Brooker et al., 2018; So et al., 2016). Sharing stigmatised attitudes based on a personal experience with others (i.e., seating next to an overweight individual) is also a form of weight stigma identified in the social media context. Another form of weight stigma, which can also explain why indirect weight stigma can be captured in clinical interventions, is the discussion of important health issues with use of stigmatised phrases (i.e., depression because of unattractive and overweight issues, eating disorders because of uncontrollable weight gain, etc.). Therefore, it is clear that the root of weight stigma in social media is spread across various forms and to identify the impact of weight stigma, all of these forms should be accounted for.

Most attacks have linked overweight individuals to lack of self-control or lack of self-responsibility (Brooker et al., 2018; Lydecker et al., 2016). It can be one of the reasons why overweight/obese individuals are stigmatised in public. The literature has discussed both positive and negative impacts and reasoning of having such an attitude towards overweight/obese individuals. One of the most commonly discussed positive reasons is motivation. By accusing an individual of being overweight because of lack of self-control,

some social media users believe that they are indirectly motivating that overweight individual to be more concerned and self-control. However, having such attitudes can cause harm to a person's self-respect, as they are being labelled with negative social desirable characteristics.

When considering the data corpus, the selected literature has used different methods to capture and analyse data associated with weight stigma in social media platforms. The most common keywords used for data extraction are "overweight", "obese", "obesity" and "fat". The authors have used several analysis methods such as natural language processing (NLP), content analysis, discourse analysis, statistical analysis, etc. to filter the most significant data for the discussion. Majority of the literature have centred their research on one selected social media platform or online media platform. There was one research which captured data from more than two social media platforms (Chou, Prestin, & Kunath, 2014). Some of the selected literature captured the comments made on a selected online article or on a selected topic to identify the stigmatising content. However, some articles have identified the need for analysis of social media discussions rather than single comments (Brooker et al., 2018). Two of the selected literature have employed machine learning techniques to analyse the data corpus (Chou, Prestin, & Kunath, 2014; So et al., 2016) and two research studies were conducted by using screenshots of mock Facebook profiles as stimuli (Lee et al., 2013; Taniguchi & Lee, 2015). The social media content is evolving exponentially with every passing minute and this has posed many technical and analytical challenges. Studies are ongoing to overcome social media governance challenges and to find ways of controlling the obesity epidemic by utilizing positive projections that do not undermine overweight members of society. However, this is still very much a work-in-progress.

## 2.4 Methodological aspects

This section consolidates findings from the exploration of the selected 22 research articles. Before the methodological and technological findings are described, all the articles are grouped based on their year of publication to highlight the growing interest in investigation of fat stigma content expressed over social media. Figure 5 illustrates an ongoing interest in this topic throughout the last decade with a noticeable rise of interest shown in 2018 and 2021. This trend is continuing with a publication on this topic in the first month of 2022 (which was the closing time of this review).

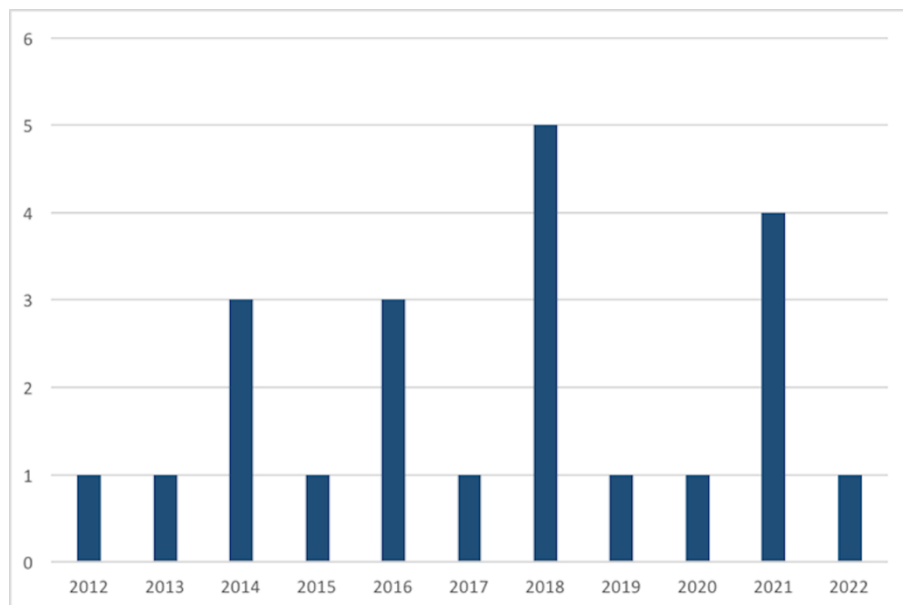


Figure 5. Number of articles by year

### 2.4.1 Domain

The domain refers to the social media platform that provided the researchers with the empirical dataset for their investigation. Ten of the review articles used Twitter as the data source (Bograd, Chen, & Kavuluru, 2022; Chou, Prestin, & Kunath, 2014; Harris et al., 2014; Karami et al., 2018; Kent et al., 2016; Lazarus et al., 2021; Lydecker et al., 2016; Mitei & Ghanem, 2020; So et al., 2016; Yeruva, Junaid, & Lee, 2019), while nine considered

other social media platforms, such as YouTube, Instagram, Facebook and public websites (e.g., newspaper comments) (Brooker et al., 2018; Busam & Solomon-Moore, 2021; Chansiri & Wongphothiphan, 2021; De Brún et al., 2014; Jeon et al., 2018; Lee et al., 2013; Taniguchi & Lee, 2015; Webb et al., 2017; Yoo & Kim, 2012). Three articles did not acquire data from social media platforms, rather, they surveyed and interviewed social media users (Holmberg et al., 2018; Lessard & Puhl, 2021; Lim & An, 2018). Almost all the articles that used empirical data extracts from social media domains considered only one social media platform for retrieving data, although two articles gathered data from two or more social media platforms (Chou, Prestin, & Kunath, 2014; Kent et al., 2016). We did not find evidence of data extracts from the Reddit platform, an upcoming social media platform, that has rich textual content (Proferes et al., 2021) which is organized into user-created communities referred as ‘subreddits’. Figure 6 illustrates a breakdown of the domains used in the selected literatures for extracting/collecting empirical data in their investigation of fat stigma content.

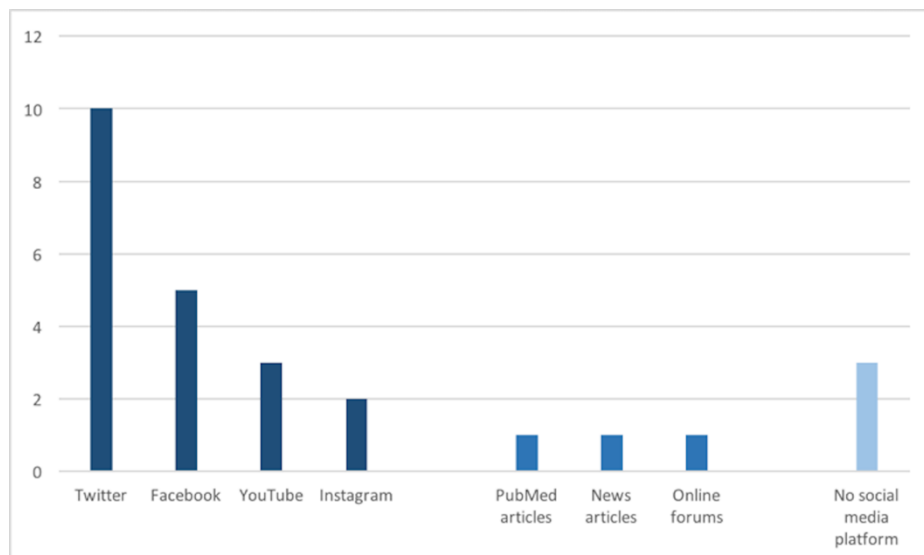


Figure 6. Domains used for data collection in selected literature: Social media domains (darkest blue), other online domains (blue) and online media not used (light blue).

### **2.4.2 Data characteristics**

As a variety of data sources were used, the characteristics of the datasets thus collected have exhibited considerable differences in both the size of the data samples and in their format (i.e., structured/unstructured text, images, videos). Only four articles indicated a sizeable data collection, that is, these studies have extracted more than 10,000 data records (i.e., posts, tweets, etc.) from social media platforms and among these, three articles had collected obesity-related data records in the order of millions (Chou, Prestin, & Kunath, 2014; Karami et al., 2018; Lazarus et al., 2021). Two articles closely analysed videos on YouTube to explore how fat people are portrayed in visual displays (Jeon et al., 2018; Yoo & Kim, 2012); further, one other article used images to analyse the effect of body ideals on fat individuals (Webb et al., 2017). Finally, several other research articles examined the effect of obesity-related discussions and experiences of being exposed to stigmatising content by interviewing and monitoring human participants (Chansiri & Wongphothiphan, 2021; Holmberg et al., 2018; Lee et al., 2013; Lessard & Puhl, 2021; Lim & An, 2018; Taniguchi & Lee, 2015). This variation in data characteristics across these articles being reviewed calls for different analytical methods, diverse techniques and a wide array of tools/models for conducting subsequent analysis.

### **2.4.3 Analytical methods**

The articles under consideration used different analytical methods for investigating obesity-related extracts from social media, with most of them having used qualitative methods. Only three articles employed quantitative methods (Lessard & Puhl, 2021; Lim & An, 2018; Mitei & Ghanem, 2020), while another three carried out experimental work (Chansiri & Wongphothiphan, 2021; Lee et al., 2013; Taniguchi & Lee, 2015) and two

other articles used mixed methods by incorporating both quantitative and qualitative research methods (Chou, Prestin, & Kunath, 2014; Kent et al., 2016).

Most of the qualitative research that was undertaken has indicated researchers' interest in employing content analysis that includes video analysis, textual analysis and visual content analysis. Apart from content analysis, two articles have considered frame analysis (Brooker et al., 2018; Busam & Solomon-Moore, 2021), two applied topic modelling methods (Karami et al., 2018; Yeruva, Junaid, & Lee, 2019) while another two employed NLP methods (Chou, Prestin, & Kunath, 2014; Yeruva, Junaid, & Lee, 2019). The distribution of analytical methods used are illustrated in Figure 7, where we grouped them as qualitative, quantitative, experimental and mixed methods.

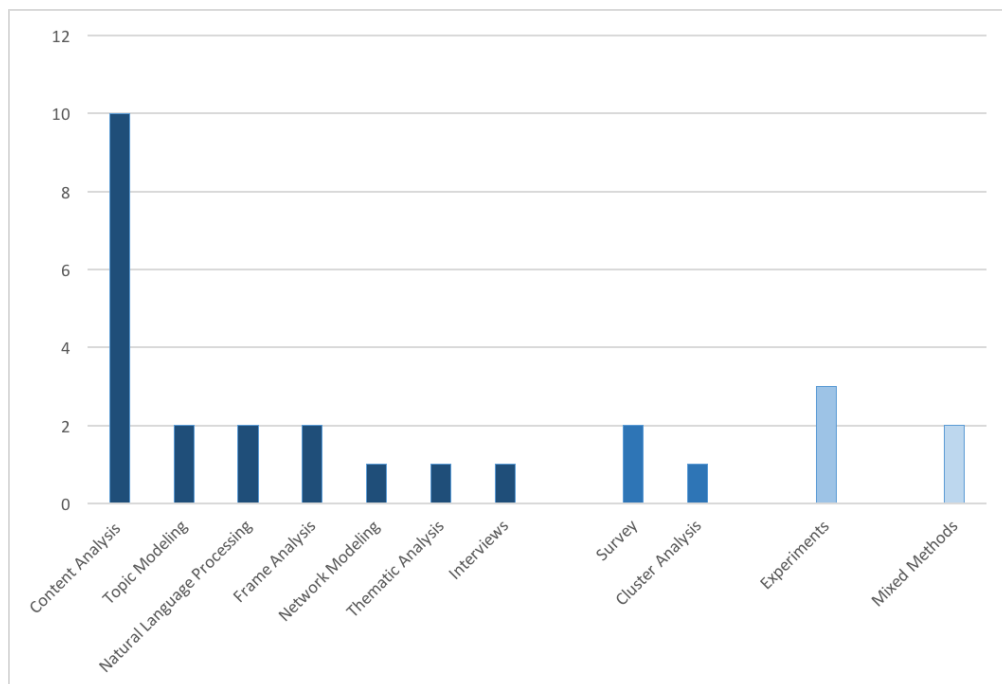


Figure 7. Analytical methods used: Qualitative (darkest blue), quantitative (blue), experimental (light blue) and mixed methods (lightest blue).

#### 2.4.4 Techniques

Among the techniques incorporated for the data analysis, it is highly noticeable that most of these articles were investigated by using manual coding techniques to carry out their content analysis. Sentiment analysis too has been employed by five research articles (Bograd, Chen, & Kavuluru, 2022; Chou, Prestin, & Kunath, 2014; Kent et al., 2016; Lazarus et al., 2021; Yeruva, Junaid, & Lee, 2019) where LDA and Mallet implementation of LDA were used to perform topic modelling by Yeruva, Junaid, and Lee (2019) and Karami et al. (2018) respectively. Statistical analysis has been conducted by four articles (Lee et al., 2013; Lessard & Puhl, 2021; Lim & An, 2018; Taniguchi & Lee, 2015) of which two considered regression analysis techniques. Although discourse analysis and co-occurrence analysis have been widely discussed techniques for hate speech detection, we find that only two studies have applied discourse analysis while another two applied word co-occurrence analysis amongst the selected literature.

Apart from these commonly used techniques, some studies employed specific techniques to analyse social media data. Harris et al. (2014) used ERGM (exponential random graph modelling) which is a widely used technique to analyse social network structures. ERGM assisted them in estimating the probability of a tie between any 2 Twitter users based on their characteristics and network structure. TF-IDF (term frequency inverse document frequency) and word embedding were among the few analytical techniques used by Yeruva, Junaid, and Lee (2019). Using TF-IDF, the study collected data from two social media domains that were then used in word embeddings to extract the context of words used in these collected datasets. Social media clustering has been used by Mitei and Ghanem (2020) to analyse relationships such as friends, followers, etc., to acquire details of obesity-related social network characteristics. Kent et al. (2016) have used bivariate

frequency analysis, along with sentiment analysis and manual coding techniques, to statistically analyse the number of occurrences of obesity and cancer in social media platforms. The application of MMMA (moderated moderated mediation analysis) to explore complex media effects by enabling multiple individual differences was studied by Chansiri and Wongphothiphan (2021). The authors applied MMMA with appearance comparisons as the mediator and further with BMI and perceived weight as the moderators to identify the effects of fitspiration and thinspiration on women's self-esteem.

#### **2.4.5 Tools/models**

While a majority of these articles have conducted manual coding to analyse the visual and textual contents, it is interesting to note that most of these studies have not used any type of software tool (or algorithm) to assist their analysis. Popular tools that have been prescribed for content analysis in hate speech detection to reduce datasets by classifying them into relevant categories (that are specific to the topic under investigation) is not evident in their usage in selected literature. Rather, we find much use of statistical analysis with IBM SPSS (IBM, 2022) in many research investigations (Harris et al., 2014; Lessard & Puhl, 2021; Lim & An, 2018; Webb et al., 2017) in comparison to R (Team, 2013) and SAS (SAS, 2022) which have recently emerged as strongholds of statistical measures. In having said this, we also witnessed a collective usage of several emergent modelling tools in five studies (Bograd, Chen, & Kavuluru, 2022; Busam & Solomon-Moore, 2021; Harris et al., 2014; Lazarus et al., 2021; Yeruva, Junaid, & Lee, 2019). For instance, Lazarus et al. (2021) developed a NLP model to analyse the sentiments of the data corpus and used DataTurks platform for manual data annotation. LIWC (Linguistic Inquiry of Word Count) is used by Karami et al. (2018) to identify health related topics as

it assisted in revealing thoughts, feelings, personality and motivations within a given corpus. Yeruva, Junaid, and Lee (2019) used Apache Spark and TensorFlow to implement their proposed framework. They further employed VADER (Valence Aware Dictionary and sEntiment Reasoner), CoreNLP and TextBlob for obtaining more accurate sentiment analysis in the corpus. The application of Textometrica on word co-occurrence analysis has been studied by (Brooker et al., 2018). NodeXL was used by Mitei and Ghanem (2020) to construct social network graph for the dataset and further examine the clusters formed as a result of retweets, replies and mentions using Clauset-Newman-Moore clustering algorithm. After experimenting with different machine learning algorithms, Bograd, Chen, and Kavuluru (2022) have used ULMFiT (Universal Language Model Fine-Tuning), a recurrent neural network-based model, to acquire better contextualized representations of words in a corpus. ULMFiT employs a language model (LSTM: long short-term memory) for processing sequential sequences using iteratively averaged weights (referred as weight-drop) to efficiently analyse large volumes of textual data (e.g., AWD-LSTM). Thus, NLP tools are gathering attention by researchers to enable them to delve into a variety of optimization strategies for language modelling and sequence learning tasks.

#### **2.4.6 Features**

Next, we examined the features of obesity-related discussions prevalent in the selected literature; that is, we inspected which aspects of fat stigma have been tackled by recent research studies. It was evident that most of these studies focused on identifying key themes (e.g., lifestyle, behaviour, attractiveness, diet, etc.) associated with fat stigma or obesity-related contents in social media. However, several studies investigated the underlying fat sentiments that are spread via obesity discussions (Bograd, Chen, &

Kavuluru, 2022; Chou, Prestin, & Kunath, 2014; Kent et al., 2016; Lazarus et al., 2021; Lydecker et al., 2016; Yeruva, Junaid, & Lee, 2019), with one study having tracked which emotions are mainly expressed within such discussions and the causes of fat stigma (So et al., 2016). Other characteristics of obesity-related discussions including perceptions, gender, user, network, etc., have been explored by five studies (Harris et al., 2014; Jeon et al., 2018; Lessard & Puhl, 2021; Mitei & Ghanem, 2020; Taniguchi & Lee, 2015). Determining the effects of fat stigma or obesity-related content seems popular among researchers as four studies have been conducted to unravel aspects of social media interactions and roles that impact vulnerable members of society who are on the receiving end (Chansiri & Wongphothiphan, 2021; Holmberg et al., 2018; Lee et al., 2013; Lim & An, 2018). Chansiri and Wongphothiphan (2021) studied the effects of idealized body images on fat people while the rest of the studies reviewed the effects of obesity-related discussions. Two studies framed the obesity content in social media into behavioural, societal, medical or healthy eating frames (Busam & Solomon-Moore, 2021; Yoo & Kim, 2012). One study focused on identifying the supportiveness received by fat people from social media (Bograd, Chen, & Kavuluru, 2022) while another revealed more details of word co-occurrences that exist in obesity discourses. Figure 8 illustrates prevalent features of obesity content as was discovered from these research studies.

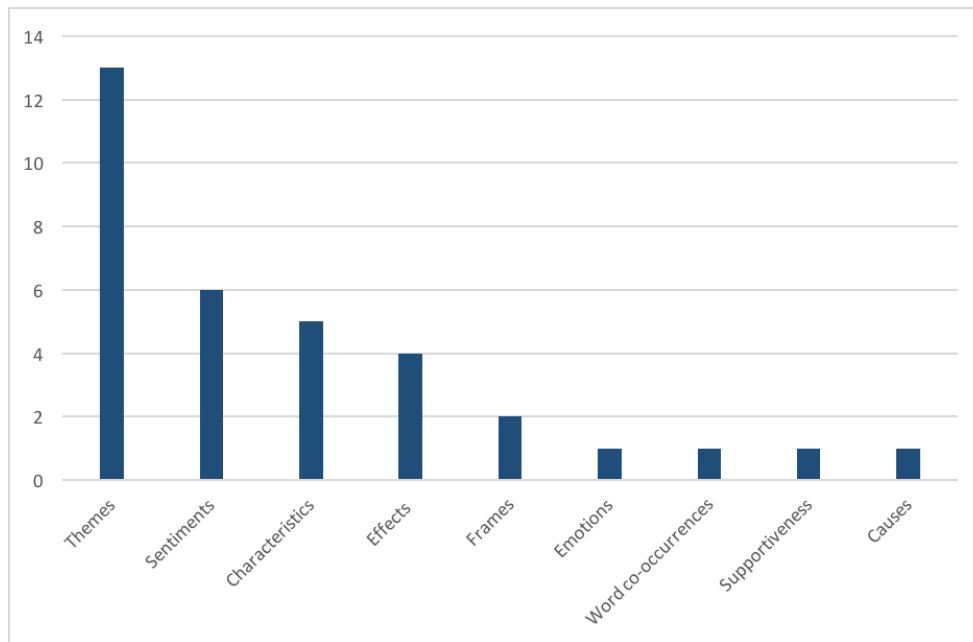


Figure 8. Features investigated by selected literature.

#### 2.4.7 Discussion of the findings

The findings of this review highlight how fat stigma studies have been conducted in recent years. It has identified elements of data acquisition, analytical methods, tools that have been incorporated and techniques utilized alongside the key features studied in each of the selected fat stigma research. These findings have aided in discovering trends showcasing common technical and methodological designs used and have thereby revealed possible gaps that need attention in future obesity research. Moreover, a snapshot of limitations (as expressed by the authors of these articles) is presented for interpreting and providing more substantive results. Therefore, this comprehensive analysis provides directions for the conduct of future research on the application of upcoming design methods in the study of data extracts from social media.

The review finds that research studies have mainly investigated how fat people are positioned in obesity content rather than analysing the effects of such content on fat

people. One of the reasons for such a tendency would be the ethical dilemma that researchers are faced with when closely studying the effect of negative content on human participants (Sales & Folkman, 2000). Closely monitoring fat person's feelings when exposed to negative obesity-related content in social media can indirectly affect these individuals who are participating in the study; therefore, a cautious approach towards maintaining ethical and social boundaries in the pursuit of this type of research is advocated. Such caution, on the contrary, is not an issue while studying publicly available discourses; rather, these discourses are enriched with obesity content and can be more easily examined to understand how fat people are being portrayed. Researchers are provided with more freedom to pursue in-depth analysis of sensitive topics to a larger extent.

The effect of close contacts such as peers and parents has been studied only once in the recent decade. However, more research on the effect of stigmatising behaviour of close contacts such as friends and relatives need to be studied, since the effect of such behaviours can greatly impact an individual in comparison to remarks made by any random stranger. Although social media is often considered as a platform where users are being exposed to a larger audience, it is also a representation of the society we live in since users are subjected to a breadth of ideas, opinions and sentiments from friends and relatives. As a result, there is a higher possibility of being stigmatised by a close relation than by a stranger. Having awareness of such online experiences can help us understand the effect of fat stigma on a fat individual by their friends or relatives.

Many studies have monitored the association between obesity and health issues based on the nature of the obesity content. This can be considered as a positive trend; with obesity having been recognized as an epidemic by WHO (2022), these online discussions

that revolve around health and obesity can further our understanding of public perceptions. It can help us better interpret the association between them and discern societal fears and apprehensions towards obesity to help remove uninformed myths that surround obesity-related health issues. We found that only two studies investigated the effects of obesity-related discussions on females. Unravelling fat stigma discussions can help us gain insights on female body objectifying content to assist in raising awareness of the impact of such objectification in society.

The domains used by selected research studies were observed. Twitter is visibly the most popular social media platform among researchers and the reason could be that Twitter facilitates easy data collection and showcases quick user responses to latest world events (Proferes et al., 2021). However, Reddit now facilitates the same criteria, besides, it offers more extensive textual content compared to the limited character of tweets. Reddit has become one of the most prominent social media platforms with 52 million active users (Reddit, 2020), therefore, it should also be considered as a useful fact-finding setting for conducting fat stigma research. Further, the rapid growth of social media popularity means that most individuals are likely to be active on more than one social media platform. In addition, the data types, data lengths, or data formats are different in each social media platform. Therefore, social media studies need to expand to multiple platforms; this will establish better comparisons and contrasts on obesity-related content.

Although millions of social media posts/comments are shared daily, only three studies have collected more than 10,000 comments/posts in selected literature. The main reason behind this could be the limitations of many social media APIs when retrieving data. For example, Twitter allows retrieval of past data for only up to 7 days only (Twitter, 2019).

This limits the number of data that can be retrieved at one time. Therefore, to obtain a larger dataset, the data needs to be collected daily for a considerable amount of time and this may pose time and resource limitations. Further, most of the datasets used in literature have consisted of textual datasets, with very few articles having analysed videos and images. This may be due to the rapid developments in textual data analysis methods compared to visual data analysis. However, as fat people can also be stigmatised visually (Heuer, McClure, & Puhl, 2011), more research on visual obesity-related content analysis is needed.

Majority of the articles indicated content analysis as their main analytical method; this can be attributed to them having collected textual data for their analysis. Analysis of textual social media data to detect hate speech has been a popular research topic in the recent past and many research studies have employed machine learning (Chatzakou et al., 2017; Nobata et al., 2016; Waseem & Hovy, 2016) and deep learning (Arango, Pérez, & Poblete, 2020; Founta et al., 2019; Gröndahl et al., 2018; Zhang, Wang, & Liu, 2018) methods to achieve hate speech detection. However, instead of using technical methods, many articles that studied fat stigma in social media have adopted manual coding methods, which is subjective to the perspectives of research teams. One reason for not employing machine learning or deep learning methods would be that these methods mostly consider single type textual features (Cao, Lee, & Hoang, 2020). Therefore, by using machine learning and deep learning methods, the researcher either has to neglect other rich textual information or combine different methods to attain improved fat stigma detection. Chen et al. (2018) reason that people who are trained in qualitative methods are generally not trained in machine learning techniques. Moreover, social scientists frame their research around deduction, causality or hypotheses, unlike

computational scientists who depend upon experiments wherein the data is frequently independent of any assumptions. While the precise reasons behind the use of manual coding in favour of technical approaches within our selected papers is unclear, the limitation in technology adoption in obesity research is evident. The growth of different machine learning tools and algorithms does not appear to have significantly influenced fat stigma research methodologies. As a result, many of the selected papers in this study did not use any software tool or algorithm for their analysis. Majority of studies have used statistical analysis tools such as IBM SPSS. Therefore, it can be concluded that a clear gap exists between obesity research methodologies and technology use; hence, more practical content analysis frameworks are required. We acknowledge that using a different academic database or using different search terms could yield another viewpoint. However, this review followed a systematic approach for selecting the 22 research articles to eliminate any publication bias in our selection process.

The limitations of methods applied in the selected articles provide grounds for future research directions and informs on what technical details require closer attention in future studies. Many studies considered limited data quantities in their evaluation of fat stigma themes. As most of these studies considered manual coding, the amount of data that could be manually coded or annotated is understandably limited. This may be the main reason behind the small data samples in the reviewed articles. Further, a biased or subjective coding process due to manual coding has been considered a limitation by few studies and some have even highlighted the limited number of annotators used in the process. These limitations are linked with the adaptation of manual coding techniques for processing and analysing the data. Therefore, the adaptation of existing text mining, machine learning and deep learning techniques could be further justified, since these

would minimize the limitations observed in this review. The temporal effect of trends on social media conversations is considered a limitation by few studies. Certain global events, such as pandemic situations, war, etc., could impact what people discuss on social media and also change the perception of certain topics such as obesity. Although one study has analysed the perception of obesity during the COVID-19 pandemic (Lessard & Puhl, 2021), more such studies could detect trends of how perceptions deviate based on different current events. Some studies have emphasized the inability to acquire different types of textual information such as geographical location, demographical information and BMI of users and different conversational methods such as replies, retweets, etc. These limitations mainly occur due to privacy constraints imposed by social media platforms to secure users' anonymity. However, the revelation of this information could benefit in developing hate speech detection strategies in this study.

## **2.5 Summary**

Although many studies have conducted to identify fat stigma or aspects of fat stigma using social media discourse, recognition of biological, psychological and sociological correlates of these studies and analysis of their methodological aspects have not been conveyed before. Therefore, as the first step in this research study, two separate systematic literature reviews assisted into analysing the biopsychosocial correlates that prior fat stigma studies have observed and the methodological aspects of their studies in relation to social media. The analysis of correlates further endorsed the use of objectification theory as the theoretical framework for this research as gender is the most common focus of studies among all the correlates. Further, the second systematic review identified the methodological aspects of fat stigma studies. It has revealed lack of technical adaptation by fat stigma studies in past decade. The findings of this review

provide motivation to combine multiple text analysis methods and focus our study on multiple social media platforms. It further aided in fine-tuning the methodological framework in the context of this study.

## **Chapter 3: Preliminary Analysis**

### **3.1 Introduction**

The findings of the literature review highlighted the key elements of fat stigma content in social media that have been tackled so far. However, before constructing the methodology, two preliminary analyses were conducted to assess the keyword structures, data sources and analysis methods used by previous studies and to examine how objectification theory provides a research lens to carry out fat stigma studies in social media. The first preliminary analysis was conducted to identify body objectification in fat stigma contents. Based on its findings, the second preliminary analysis was conducted specifying female body objectification.

### **3.2 Analysis I: A text analysis approach to identify fat stigma and body objectification in social media**

This analysis examines fat stigmatising discourses to see how men and women are positioned in obesity discussions over three popular social media outlets, namely Twitter, YouTube and Reddit. Objectification theory has provided the analytical lens for probing the experiential consequences of being overweight in a culture where individuals (mainly women) are treated as objects to be looked at and evaluated (Fredrickson & Roberts, 1997; Tylka & Hill, 2004). The analysis seeks to understand attitudes around gender in context of fatness. This includes examining underlying negative and positive sentiments associated with opinions often voiced around fat people on social media outlets. Perceptions of stigma in the context of physical appearance mediated by social interactions have been analysed using objectification theory constructs. We believe analysis of such social media encounters can provide insights on how male and female bodies are objectified in weight stigmatising content. Two research questions are posed to carry out this analysis:

- *RQ1: What user sentiments have been expressed in weight-based discussions and how are gender-differences manifested in fat stigmatising content?*
- *RQ2: What latent themes emerge within text from these social media exchanges which contain gender-based objectification around fatness?*

### **3.2.1 Method**

Social media posts capture a real-world perspective, as the user-generated data exposes user sentiments, beliefs and biases. Many weight stigmatising studies have been informed by social media (Chou, Prestin, & Kunath, 2014; So et al., 2016) by leveraging data analytic techniques, as rich user data corpus can be extracted from social media outlets without being intrusive. Further, the user data can be harvested frequently in changing contextual conditions such as coverage at a single point of time or over different times (Chang, Kauffman, & Kwon, 2014) enabling social scientists make meaningful interpretations of the real-world phenomena that is unfolding before them. This study too adopts a computational social science perspective by bridging social constructivist studies with data mining techniques on weight stigmatisation content to give a gendered objectification perspective. Online text mined from three popular outlets – Twitter, YouTube and Reddit – provided a first-hand view of users’ disposition towards people considered overweight.

Predetermined keywords (“obese”, “overweight” and “fat”) were used for text mining. We used appropriate crawling methods to browse user content on selected social media outlets, and publicly available APIs (Application Programming Interfaces) to scrape data (Google, 2019; Reddit, 2020; Twitter, 2019). Next, user sentiments on extracted data excerpts were computationally analysed. Computational modelling can help social scientists leverage research methodologies to achieve “control and precision in

measurement, while maintaining realism in application and generality in theory development” (Chang, Kauffman, & Kwon, 2014, p. 68) Ethical guidelines described by Townsend and Wallace (2016) informed in the selection of quotes and their reporting. Anonymity of online users has been preserved in our presentation of excerpts and phrases (e.g., users’ Twitter handles have been replaced with “@USERNAME”). Only those data excerpts not considered very derogatory have been quoted verbatim, and with expletives (if any) replaced by asterisks. Derogatory words have been blurred in word co-occurrence maps and in subsequent qualitative reporting, although all typos, misspellings, and slangs have been retained to illustrate authentic exchanges. The following subsections elaborate on data scraping methods and techniques that were applied for pre-processing data before conducting analysis across the three social media outlets.

#### *3.2.1.1 Data Scraping on Twitter*

The corpus extracted from Twitter included “girl” and “boy” keywords along with other predetermined keywords. Data were mined on 18th July 2019 by requesting the latest 5000 tweets containing different combinations of the keywords, for example, “fat+girl”. Crawling is an automated software technique, so the next step involved data cleaning which was conducted manually. In this step, irrelevant tweets, such as those texts which referenced topics other than human body weight were removed. (e.g., excluding the tweets that contained the requested keywords in the screen name).

#### *3.2.1.2 Data Scraping on YouTube*

Two popular videos – “Fat Girl Tinder Date” and “Fat Guy Tinder Date” – were chosen from YouTube. These videos were a part of an online social experiment (as mentioned in each video’s title) which comprised male and female actors who deliberately made

themselves fat targets for online posts. Text-based comments that were posted for each video were scraped on 05<sup>th</sup> June 2019 and resulted in two datasets. The comment datasets thus produced from YouTube were different to that of Reddit and Twitter since these comments were contextualized to the corresponding video content. Moreover, both the YouTube videos were social experiments having the same gendered plot; therefore, they provided a similar context for conducting subsequent analyses on how obese male and female genders were being perceived by online communities. Next, predetermined keywords were run on each comment dataset to categorise user perceptions for this gendered plot setting.

### *3.2.1.3 Data Scraping on Reddit*

Discourses displaying fat stigma was extracted from Reddit using similar methods that were followed for Twitter. The data was mined on 16th January 2020 by scraping the 1000 most recently posted Reddit posts along with their comments using different keyword combinations. The data cleaning process was conducted manually by considering the topic discussed in the main posts.

## **3.2.2 Analysis**

A mixed method data analysis was conducted next to explore the social meanings in the stigmatising content. The existing large-scale analysis methods primarily cater to unravelling of the psychological characteristics of internet users (Guitton, 2019); therefore, several analysis methods were used to investigate fat discussions. The data was pre-processed to eliminate prepositions (of, to, for) and conjunctions (and, or, also). In the initial exploration, the relative distribution of keywords was examined visually. Subsequently, the R library “syuzhet” (Jockers, 2017) helped capture more in-depth sentiments associated with the linguistic corpus and quantify the intensity of emotions

expressed in the lexicons (Mohammad & Turney, 2010). The sentiment analysis tool helped quantify the mix of emotions underlying the human-produced texts. Human emotional states are informed by “a complex chain of loosely collected events that begins with a stimulus and includes feelings” to act in a certain way (Plutchik, 2001, p. 345). The Plutchik Wheel identifies eight emotions – anger, anticipation, disgust, fear, joy, sadness, surprise and trust – as an expression of one’s inner state. The tools assisted in detecting the intensity of the eight emotions across the large data corpus.

Textometrica, an automated tool that explores co-locations of words based on their frequency of occurrence in a body of text (Lindgren & Palm, 2011), allows filtering of the original dataset around selected words. This tool assisted in generating co-occurrence maps and provided an inductive approach for framing themes around gender-related stigmatising words contained within the large textual data. Word co-occurrence maps (for each social media outlet) illustrate the objectifying word occurrences alongside specified gender groups with different keyword combinations. The maps connected words that had some semantic correlation when users expressed their views on a particular topic, since these words tended to be near each other. Conversations around groups of connected keywords assisted in framing themes. Based on the common words observed in the generated word co-occurrence maps and their metadata, five underlying themes associated with fatness were drawn. Further, map visualisations assisted in deeper qualitative inquiry, that involved reading linked commenters’ opinions to understand word associations with male/female body objectification themes.

### **3.2.3 Results**

The investigation revealed weight/fat stigma sentiments that framed the raw rhetoric expressed over the selected social media outlets in a gendered context. Study findings are

presented next in the following order: relative distribution of the keywords, sentiment visualisations, word co-occurrence maps, underlying discussion themes and qualitative illustrations of selected posts.

### *3.2.3.1 Keyword Distribution*

Social media APIs were used to collect data having different parameters. The Twitter API request was set to capture at most 5000 tweets for each keyword combination against a sampling of recent tweets published in the past 7 days. A total number of 5000 tweets were attained for the keyword combinations; “fat+girl” and “fat+boy”, while other keyword combinations yielded fewer tweets. The Reddit API extracted the latest 1000 posts and the comments for each keyword combinations. It displayed a similar pattern of keyword distribution as on Twitter. The YouTube API extracted a total of 14613 comments from the video “Fat Girl Tinder Date (Social Experiment)” and 10312 from “Fat Guy Tinder Date (Social Experiment)”. These comments have been aligned with three keywords – fat, overweight and obese – and again show more associations with the word “fat”.

A detailed breakdown of comments with associated keywords is shown in Table 3. The most common keyword associated with human body weight that appeared in all social media outlets was “fat” compared to “overweight” and “obese”. Except in the case of Reddit, more posts for females (“girl”) compared to males (“boy”).

Table 3. Keyword distribution in Twitter, YouTube and Reddit

Twitter and Reddit			YouTube		
Keyword Combination	Number of Tweets Retrieved	Number of Reddit Posts and Comments Retrieved	Video Title	Keyword	Number of Comments
"fat" + "girl"	5000	11708	Fat Girl Tinder Date	fat	2944
"overweight" + "girl"	515	11341		overweight	122
"obese" + "girl"	129	8512		obese	156
"fat" + "boy"	5000	15429	Fat Boy Tinder Date	fat	1649
"overweight" + "boy"	321	11070		overweight	47
"obese" + "boy"	129	10595		obese	32

### 3.2.3.2 Sentiment visualisation

The extent of positive and negative sentiments expressed alongside eight emotion labels (i.e., anger, anticipation, disgust, fear, joy, sadness, surprise and trust) (Plutchik, 2001) are presented next. Sentiment analysis tools like the Plutchik Wheel utilise complex algorithms that draw from the fields of psychology, digital humanities, linguistics, annotations and visualizations to uncover the levels of the eight emotions expressed in a corpus (e.g., collection of tweets)(Semeraro, Vilella, & Ruffo, 2021).

Figure 9 first displays the distribution of overall sentiment (valence) scores of the three corpora associated with male and female genders for the three social media platforms, with values less than zero indicating negative sentiment. Statistical tests<sup>1</sup> indicated that the distributions of valance values between comments directed towards male and female genders were significant. For YouTube, 75% of the comments targeting females were negative as opposed to 60% for males. The reverse was true for Twitter, with 80% for

---

<sup>1</sup> The two Sample Kolmogorov-Smirnov test was used. All the results were significant at the 0.01 level with pvalue=2.22e-16 for YouTube, pvalue=1.34e-150 for Twitter and pvalue=6.43e-08 for Reddit.

males and 55% for females. For Reddit, both genders received a similar percent of negative comments of around 80%.

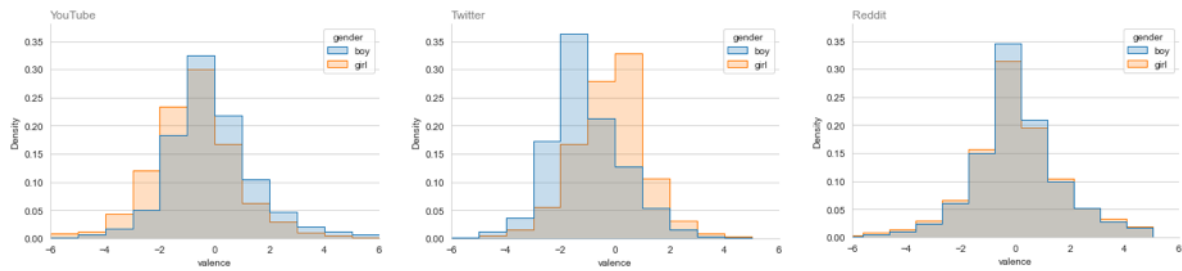


Figure 9. Histograms of sentiment values across the three corpora

Among the eight emotions, high values in anticipation, joy, surprise and trust indicate empathy towards persons at the receiving end of the social media post. On the other hand, high values for anger, disgust, fear and sadness exhibit a resentful attitude. Except for “disgust”, “sadness” and “trust”, all social media outlets indicated significantly lower emotion values for males. More intense and varied emotions were observed towards females compared to males. Twitter displayed drastically higher “disgust” emotions for males compared to other two outlets. Figure 10 gives an overview of the eight emotions (Plutchik, 2001) found to be associated with these discourses across the three corpora. The emotions are illustrated as percentages of emotion word count for further comparisons.

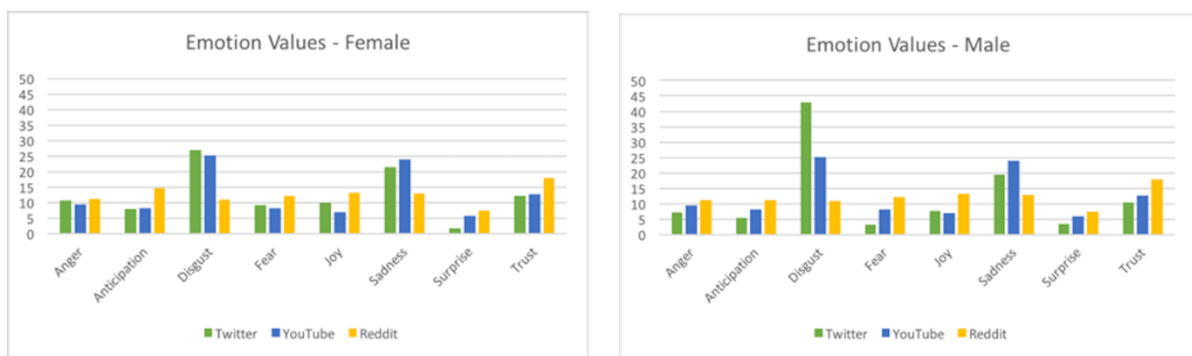


Figure 10. Emotion values as percentage of emotion word count

### *3.2.3.3 Word co-occurrence maps*

Word co-occurrence maps were generated next to identify words strongly associated with gender and human body weight. These maps were produced by selecting the strongest links by normalised co-occurrence frequency (i.e., the normalised value of the frequency occurrence of two-terms in a text corpus). Concepts were created in Textometrica by grouping words representing gender and body weight. Word co-occurrence maps for the keyword combinations “fat+girl” and “fat+boy” are displayed in Table 4. The maps reveal that tweets containing “fat+girl” have more derogatory content than “fat+boy”. Maps for YouTube and Twitter data demonstrate more objectifying words on female body than for Reddit. However, compared to Twitter, YouTube comments demonstrate less derogatory discourses towards males and females; although gender was more conspicuous in comments posted for “fat guy tinder date” video.



#### *3.2.3.4 Underlying discussion themes*

Since the initial data corpus used to identify user sentiments and investigate word co-occurrences is considerably large, an exemplar has been purposefully selected to discuss underlying discussion themes. Equal volumes of data excerpts were selected for each keyword combination from the data corpora extracted from Twitter, YouTube and Reddit. The data excerpts were analysed manually by examining word co-occurrence maps and their metadata to identify associated themes. The process involved investigator triangulation to mitigate researcher bias (Carter et al., 2014). Biological, psychological and sociological correlates of fat stigma in social media identified by Wanniarachchi et al. (2020) informed our analysis. Five themes, namely, attractiveness, physical appearance, lifestyle choices, health, and psychological well-being, were detected. Tweets, YouTube comments and Reddit posts that described fatness as an attraction barrier or discussed the link between fatness and attractiveness are categorised into the “attractiveness” theme. Under “physical appearance”, social media posts using “fat”, “overweight” and “obese” keywords to describe a person’s physical semblance are considered. Relatively higher number of excerpts discussed fatness as a lifestyle choice, where commenters emphasised changing diet or setting up exercise routines. Such content is categorised into “lifestyle choices”. Further health issues associated with fatness were also observed, while some users specifically expressed views on the association of psychological distress among fat people. These have been categorised into “health” and “psychological well-being” themes. A total of 570 data excerpts were categorised, and data distributions within each theme are presented in Figure 11.

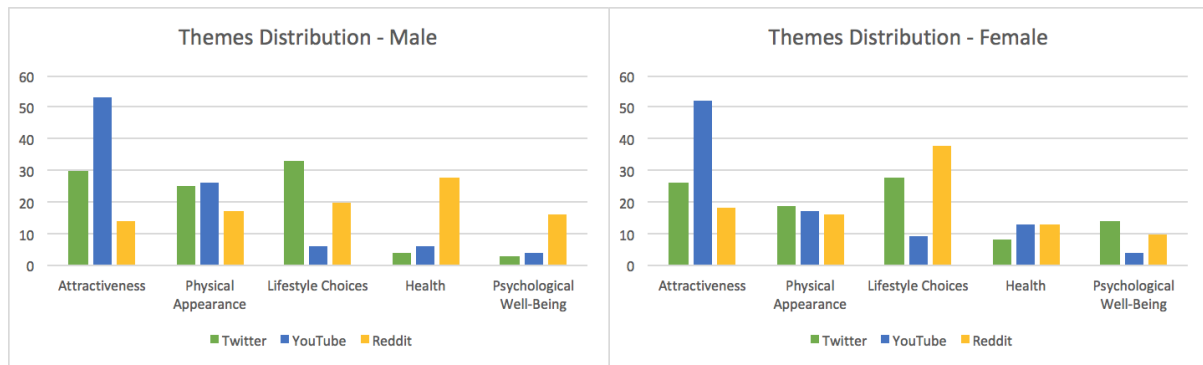


Figure 11. Distribution of data excerpts among the identified themes

Our data reveals that Twitter and YouTube users mainly discussed fatness with physical attraction for both genders, while Reddit users mostly discussed health in context of fat men and lifestyle choices for fat women. Compared to males, females overall had more varied discussions within the five identified themes. All the five themes exhibit levels of social, body and gender objectification that is based on an evaluation of the individual's weight.

### 3.2.3.5 Qualitative data analysis on sentiments

The data corpus was analysed for key sentiment expressions in weight discussions. Particular excerpts considered less derogatory have been selected from the discourse to report these sentiments (in Table 5). Data excerpts were selected through a consensus process by the study team.

Apart from some mean-spirited attacks, the selected tweets show how a part of society recognises overweight/obese individuals as out-grouped. However, female-centric tweets mostly targeted body-shaming for overweight girls, while male-centric tweets targeted physical abilities of the overweight boys.

The negative sentiments displayed different types of negativity. According to some users, losing weight is a self-responsibility (*"Omg I've been eating like a fat girl"*). Most

comments/posts showed negativity towards fatness by connecting it with being less attractive.

Table 5. Selected excerpts representing different sentiments

Sentiment	Gender	Comment
Negative	Girl	<i>Honestly can't believe a random boy just added me to tell me he isn't into me because he isn't into 'fat chicks' Omg I've been eating like a fat girl ... the self-esteem impact from the excess skin will be something she regrets forever...</i>
	Boy	<i>@USERNAME Nobody wants fat boy</i>
Anticipation	Girl	<i>@USERNAME Girl yes!!!! I've come up with a plan that allows me to still eat like a fat girl ! Just move and eat at the same time</i>
Anger/Disgust	Girl	<i>The girl must exercise. She looks expressionless, bloated, and overweight.</i>
	Boy	<i>@USERNAME Boy you stoopid lmaooo better work them 8 hrs a day and get a fat check</i>
Sadness/Fear	Girl	<i>also the boy i really liked rejected me because im still fat and im crying even more ... I looked in the mirror one day and was like I can't do this anymore....</i>
	Boy	<i>@USERNAME I don't look good. I'm so overweight and i feel i will never get a girl like you or anyone else</i>
Positive/Trust	Girl	<i>there is nothing more powerful then a fat girl who doesn't give a ****</i>
	Boy	<i>Fat men are sexy in my opinion (don't judge me ) ...I absolutely do agree with letting obese people live, be accepted, feel attractive, and have equitable access to healthcare without condemnation...</i>

Some comments also related fatness with self-esteem (*“the self-esteem impact from the excess skin will be something she regrets forever”*). Such negativity was observed in both female and male centric data excerpts.

Anticipation has been recognised mostly in tweets that targeted females. Though it was least presented within the captured data corpus, it is a sentiment that spreads positivism. By directing messages (as in the selected tweet) which though not motivational conveys that overweight/ obese females are not a separate group of the society (*“I've come up with a plan that allows ....”*).

Anger and disgust were highly observed emotions, although the levels did not seem to be gender specific. However, fat women commonly experience direct anger and disgust (*“...She looks expressionless, bloated, and overweight”*) compared to males. Sadness and

fear were also witnessed regardless of gender and mostly these emotions were associated with finding a partner from the opposite sex (*"...and i feel i will never get a girl like you"* or *"...the boy i really liked rejected me"*) and physical appearance (*"...I looked in the mirror one day and was like I can't do this anymore"*). Though negative sentiments surpassed the positives in the retrieved corpus, there were considerable posts spreading motivation and positivism.

### **3.2.4 Discussion of the findings**

The first research question seeks to understand how weight/fat stigmatisation content is formed around gender in weight-based discussions over social media. Text mining methods helped unravel the user sentiments to expose how gendered body objectification and weight stigmatisation is manifested. A mixed method analysis (comprising word co-occurrence and qualitative analysis) aided in investigating distributions of keywords and their correlations with male and female objectification. Findings reveal that "fat" was commonly used among the keyword distribution for excess body weight, while "overweight" and "obese" were mainly related to Body Mass Index. That is, content around words "overweight" and "obese" hinged on a clinical note while "fat" hinged as a social descriptor. Therefore, social media discussions about weight are mostly established around the word "fat".

Sentiment analysis further revealed how emotions related to gender-based fatness are scattered. Most common sentiments that emerged from the selected corpus were observed to be negative. Further, the analysis exposed emotions that exhibited higher intensity when discussing females as compared to males. However, it is unclear as to whose emotions were being represented since the posts could be either from the victims of weight stigma or from the stigmatisers. Overall, we found emotions voiced in YouTube

comments were found to be higher for females than for males, while Twitter posts demonstrated mixed results and indicated higher disgust for males compared to females. Next, the word co-occurrence maps revealed words associated with concepts 'female', 'male' and 'fat'. The maps visualisations indicated more derogatory words to be associated with these concepts in Twitter compared to YouTube and Reddit. The videos selected from YouTube were part of a social experiment comprising male and female actors who purposely portrayed themselves to be fat; therefore, the individuals involved in the experiment may not apparently be considered overweight or obese by commenters. This may have led commenters to discuss more on the overall video scenario rather than voice hateful comments targeting particular individuals. Also, YouTube and Reddit do not specify a text limit in their comment section, which allows users to be more expressive with their views. Although Twitter has doubled the character limit of a tweet from 140 characters to 280 characters, only 5% of tweets are longer than 190 characters (Perez, 2018). With a fewer number of words for commenting, the users may be using more terse words. This could also explain the use of more derogatory words on weight in Twitter discussions. While evidence regarding the moderation of user content (via tweets) in Twitter is not explicitly stated (Twitter, 2022), the video content and associated comments in YouTube can be filtered (i.e., removed or reported as spam) by the content creators thereby allowing users to decide whether any of the comments pose harm (Google, 2022). Conversely, Reddit's AutoModerator (Reddit, 2022) arbitrates the content posted on their subreddits to keep it free of hate speech and derogatory content. Our word co-occurrence maps too indicated fewer negative and more positive sentiments on Reddit compared to other social media outlets.

The metadata connected to each word and their links in the word co-occurrence maps were examined. In response to the second research question, five underlying themes were detected. These are, attractiveness, physical appearance, lifestyle choices, health, and psychological well-being. YouTube discussions were mostly scattered around the theme “attractiveness” while Twitter centred around the theme “lifestyle choices”, regardless of the gender. However, Reddit showed more focus on “health” when discussing fat men and on “lifestyle choices” when discussing fat women. The social experimental nature of YouTube videos for showcasing experiences of fat men and women on Tinder dates, could also explain its focus on attractiveness. The Twitter discussions mostly accused fat people for not taking control of their weight. Comments laced with sarcasm, stated fatness to be a consequence of lack of exercises or “unhealthy” diet (i.e.: *“You fight getting up for your doughnuts?”*). Reddit’s discussion towards fatness based on gender showed that users were more concerned about health of fat men (i.e.: *“This was all a result of my brother being an emotional eater, so he developed heart disease (a family issue), cholestrol, HBP & diabetes”*) and commented more on lifestyle choices of fat women (i.e.: *“Most of it is just bad eating culture and fast food”*). Although, these health-related comments on fat men were about someone else, the comments regarding lifestyle choices of fat women mostly came from commenters’ own experience (i.e.: *“I’m a bigger girl... I don’t smoke/do drugs and don’t eat a lot. I don’t exercise as much as I should, but I do on the days I am able to.”*).

More female body objectifying words were observed in Twitter discussions. Though most of these objectifying words were not directly co-occurring with the female concept, these words were discovered with the search terms “fat” and “girl”.

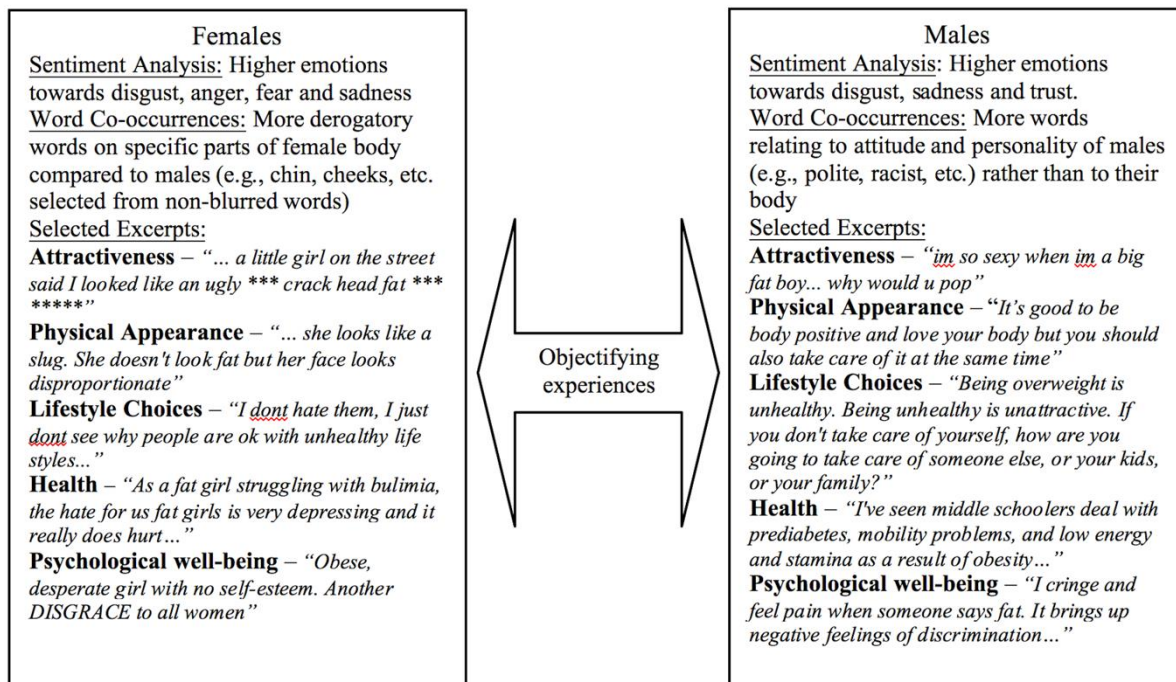


Figure 12. Social media objectifying experiences observed across males and females

This analysis has discovered several aspects of objectifying experiences across the three social media platforms. Figure 12 provides a gendered overview of sentiments expressed, word expositions used and related objectification themes that emerged from the fat stigmatising content. Such objectification can cause an overweight/obese individual psychological distress, which eventually leads to eating disorders (Chou, Prestin, & Kunath, 2014). These impacts were evident throughout the study, specifically when users claimed to have health issues such as bulimia. When any form of objectification occurs in social media, it can impact both the targeted user as well as readers. The existence of objectification, especially for females, should not be taken lightly as it directs women into internalising the objectifying perspectives on their own physical self, which could lead to mental health risks (Fredrickson & Roberts, 1997). The psychological distress of being targeted for fatness was also observed in Reddit. As Reddit is a discussion website, many conversations were on fatness and female body objectification. Some comments

emphasised different aspects of fat stigma and its relationship with female body objectification. Following examples illustrate this:

*Went on a date yesterday, we had a great time but towards end of the date he said I need to lose some weight. I know I m fat but it was hurtful. This video made me realize even though men say they want a girl with great personality they care more about the physique. Trying my best to lose weight...U know what I wish I found love but at same time i am ok being single for now until I find someone who loves me for how I look and who I am*

- YouTube Comment

*This this this. My mother has spent her whole life hating herself, putting herself down, and being hated and put down by others. She has tried to lose weight, and tried, and tried. At this point, her body might be too wrecked to really ever lose the weight. She is relatively healthy. I can't help but imagine how different her life could have been if she hadn't spent it under the terror of all that hatred.*

Reddit Comment

The sentiment analysis revealed many forms of objectifying themes. The availability of such content exposes overweight/obese women to society as unattractive and irresponsible individuals. Obesity in women is often associated with medical conditions including alterations in the reproductive cycle with a reduction in fertility, as well as an increased risk of the polycystic ovarian syndrome (PCOS), and maternal obesity (Templeton, 2014). Therefore, in such situations, the social pressure of objectification (that is often internalised) alongside some pre-existing medical conditions could make these women vulnerable to depression (Love et al., 2016). While medical reasons for

weight gain are outside the scope of this study, this aspect needs to be recognized and studied in future weight/fat stigmatisation research.

### **3.3 Analysis II: Underlying Themes of Female Body Objectification in Social Media**

The second analysis elaborates on how women are currently positioned in fat stigma content and how they are objectified in relation to their body shapes through an investigation of online content posted on Reddit. Tylka and Hill (2004) voice that body objectification cause women to adopt an outsider view to self-objectify themselves, which could add to their sense of body shame. This analysis examines the linguistic corpus that has been extracted from Reddit with use of sentiment analysis tools to find underlying objectification themes targeting women. It has examined the connection between fat stigma and female body objectification to answer the research question:

- *RQ: What are the common themes targeting women as observed in fat stigmatising content posted in social media platforms?*

According to the objectification theory, perceptions of fatness in social interactions often involve a woman's physical appearance. Social media interactions capture a variety of expressions and can provide researchers valuable insights on the pervasive nature of fat stigma and other biases that lead to objectifying fat body shapes (De Brún et al., 2014; Wanniarachchi et al., 2019). Accordingly, the data corpus extracted from Reddit platform, sentiment analysis and word co-occurrence analysis have been employed in this investigation.

### 3.3.1 Method

The usage of data mining techniques has opened possibilities for social scientists to make meaningful interpretations from large amounts of data. Therefore, online discussions on fat-based topics were extracted from the Reddit platform to analyse social views on fatness for constructing female body objectification themes. Brooker et al. (2018) approach for handling text data extracted from digital sources and analytic methods used for exploring its linguistic content for recognizing biases and ideologies, informed our analysis. The Reddit forums contain lengthy discussions on emergent subjects; hence discussion data were mined from Reddit using its publicly available API (Reddit, 2020) on 03rd March 2020. Moreover, with the anonymity facilitated by Reddit (i.e., it does not ask personal information such as gender, birthdate or email address), users can be more authentic and upfront (Gutman, 2018). This provides researchers much insight into emotions that are being expressed when online communities freely express themselves. The extraction process of the corpus included the keywords “fat”, “overweight” and “obese” along with the keyword “girl” to obtain relevant Reddit posts and comments, as the focus aligns with the female gender. As a result, 76 Reddit posts, along with their comments were obtained. Among the crawled posts, the study purposefully selected one Reddit post to analyse how discussions regarding fatness are structured around the female body (refer Table 6 below). Purposeful selection of the post was based upon Townsend and Wallace (2016) recommendation that researchers carefully consider the sensitivity of the content, and ask themselves whether it is suitable for re-use or re-publication.

Table 6. Selected Reddit post

<b>Title</b>	Why is there so much hate for fat people, especially girls?
<b>Post</b>	<p>As someone who always liked larger women it always startled and saddened me a lot to see how much hate, verbal and even physical violence they have to experience.</p> <p>I just don't get it, is it really that some people need some group of people that they can despise? Where does that anger stem from?</p> <p>Edit: Well since the most comments seem to take it for given that it's ok to hate fat people and thus focus on finding reasons to justify that with [health][URL included], laziness, it's unattractive (although I don't think so) and "They could easily fix it if they don't want to be hated" arguments, I don't buy into that. I mean I could ask you whether you also insult people as smokers or condemn procrastination, but that still misses the point.</p> <p>Is it really all about reinforcing some kind of social consensus? About disliking people who are different?</p>

The selected post questions the hate expressed towards fat people, specifically towards girls. The author mentions being witness to verbal and physical violence towards fat women; therefore, they question “why it happens” and “where does it stem from”. The post focuses on the nature of the negative comments towards “larger women” which is thematically related to the purpose of this study. While this data is publicly available over the Reddit platform and is not directly linked to individuals who can be identified, ethical considerations have been applied in our reporting (Moreno et al., 2013; Townsend & Wallace, 2016). Accordingly, usernames and subreddit names have been omitted to conceal the identity and maintain the anonymity of the Reddit users. Extracted posts containing sensitive content have been reported responsibly. That is, in our reporting of excerpts and phrases, all expletives have been replaced with asterisks while some belittling content has been paraphrased properly. We have retained certain typos, misspellings and slangs to give a more realistic illustration of emotions that were expressed.

User views have next been analysed computationally by using sentiment analysis and word co-occurrence analysis methods. Such data analysis methods can provide a deeper understanding of how social media discussions are structured. The word co-occurrence analysis assists in identifying words and concepts largely observed in a given data corpus and thereby can be linked to different themes, while sentiment analysis enables detection of negative and positive emotions associated with those themes. After extracting comments and replies posted to the selected Reddit post, the data was pre-processed to eliminate prepositions (of, to, for) and conjunctions (and, or, also). Subsequently, sentiment analysis and word co-occurrence analysis were separately conducted to investigate emergent patterns associated with fat stigma.

Word co-occurrence map has been generated using Textometrica (Lindgren & Palm, 2011) on the whole data corpus to identify the main words associated with the selected Reddit post and its comments. Connections between words were discovered from the generated word co-occurrence map and led to a close examination of metadata linked with those words. Four themes emerged namely, fatness as a lifestyle choice, fatness as an attraction barrier, general criticism of fatness, and fat people are easy targets. Comments and their corresponding replies have been individually categorized into the four themes. Further qualitative analysis helped gather more context for each theme. Sentiment analysis too was conducted on each comment and each reply using different python libraries (Jockers, 2017; NLTK, 2020) to highlight various positive and negative sentiments associated with each theme. Two sentiment analysis libraries were used to separate sentiments and also assisted in computational triangulation (by eliminating any possibility of bias from either library).

### **3.3.2 Computational Analysis**

The computational analysis provided a holistic and non-intrusive view of all emotions and sentiments underlying the user opinions that were expressed online. This section elaborates on results obtained from the two types of computational analysis, namely word co-occurrence and sentiment analysis.

#### *3.3.2.1 Word Co-occurrence Analysis*

Our data corpus comprises 481 comments and replies posted for the selected Reddit post. The corpus ran through Textometrica to build a model by including all the words between 1 to 25 characters and removing stop words such as “and” and “is”. In the next step, all 127 words that identified by Textometrica between 6 to 109 frequencies were selected to build the model. These selected words included words in different tenses and singular/plural forms which were then combined into one concept word-group. For example, when the words “eat”, “eating” and “ate” are included in the list, the concept “Eat” was created. By selecting the strongest links by normalized co-occurrences, 77 word pairs visuals were drawn (shown in the word co-occurrence map in Figure 13).

The strongest word connections displayed in the map have been closely analysed by cross checking the metadata associated to each link. The word connections and the content analysis of metadata resulted in identification of relevant themes associated with the online posts. Few examples of retrieving concepts, word frequencies and related text blocks that facilitate the in-depth analysis of the content has shown in Table 7. To perform this analysis investigator triangulation method that involved the thesis author and the supervisory team were used (Carter et al., 2014). The first and the most common theme has been constructed around those comments that reproach ‘fatness as a lifestyle choice’. This theme was identified with the support of words associated with a person’s lifestyle,

including exercising and eating habits. The second most discussed ideas on fatness were around the individual's appearance and their attractiveness.

The theme identified here is 'fatness as an attraction barrier'. Here the commenters engaged in discussions associated with concepts such as shape, size, beauty, etc. These concepts were derived from supportive words used in comments and replies that were based upon physical appearance. Some other comments focused on complex issues related to fatness such as health and psychological issues, and social acceptance of fat people.

Those comments have been categorized into a third theme, namely 'general critics of fatness'. Finally, a fourth theme was detected that related to discussions around 'fat people as easy targets' and so they can be freely criticized. The comments relating to the fourth theme were mostly observed as explanations to questions that were posed in the initial post.



Table 7. Examples of analysing text content related to words

Word	Frequency	Sample Text Block
Attractive	14	<b>Segments in 76527546:</b> [1]: ..."I think it's safe to say that <u>attractive</u> ness is, for the most part, socially constructed. Some cultures, I've heard, find fat people more attractive. Western culture, which is spreading thro.....
Fat	89	<b>Complete text of 76527520:</b> It the image people get of <u>fat</u> people as being lazy and hedonistic. Whether this is true or not, this is how they're viewed.
Obese	16	<b>Segments in 76527530:</b> [1]: ..."I find it ridiculous that so many here say they judge <u>obese</u> people because they are damaging their health. How many of those posters drink alcohol, speed while driving, ride motorcycles?.....

Prior literature too has framed fatness around themes related to personal (ir)responsibility, disability or illness, and based on individual variations between diet and metabolic rates (Brooker et al., 2018). Our exploration has reinforced some of these themes. Further, the affordances of the word co-occurrence maps yielded more empirically drawn themes on fatness. The study has identified themes around “fatness as a lifestyle choice”, “fatness as an attractive barrier” and “fat people as easy targets” besides grouping the “general critics of fatness” into another theme.

### 3.3.2.2 Sentiment Analysis

For the selected linguistic corpus, positive and negative sentiments were identified. All themes showed a range of negative and positive sentiments, although findings indicate more negative sentiments than positive sentiments in each theme. Theme, ‘fat people as easy targets’ indicated more negative sentiments than the other themes while ‘fatness as an attraction barrier’ showcased more positive sentiments. Figure 14 illustrates the percentages of positive and negative sentiments scattered around each theme.

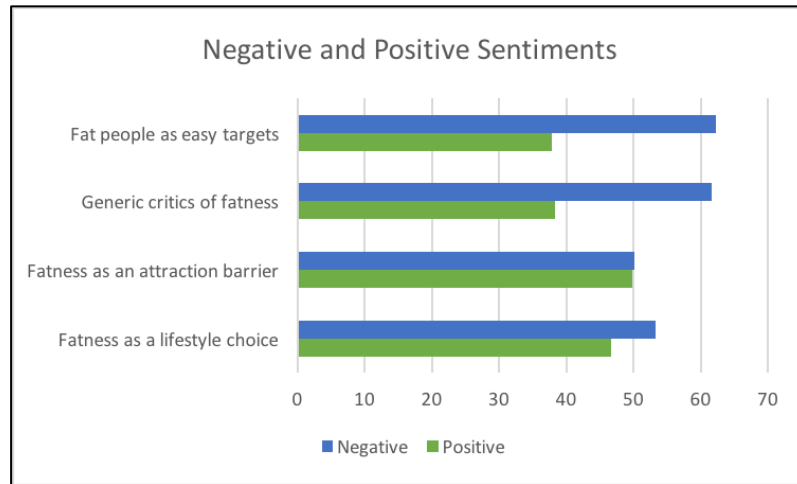


Figure 14. Sentiments associated with each identified theme

The sentiments analysis from the vast linguistic corpus provided understanding on prevalent biases and ideologies on how fat stigma is perceived by the online community. The breakup of positive and negative sentiments that underlined the four themes provided a visual analytic framework that set the next stage of our inquiry. Brooker, Barnett, and Cribbin (2016) contend that visual analytic should not be treated as an endpoint, instead researchers should ‘drill down’ further into the substantive social media content and engage in finer qualitative analysis.

### 3.3.3 Qualitative Analysis

Armed with the visual analytic results from two stages of computational analysis, a deeper qualitative analysis of the textual data was conducted to study the inner composition of each theme. The four themes have been elaborated next in the following subsections.

### 3.3.3.1 Fatness as a Lifestyle Choice

In opposition to the main post, many commenters shared either opinions or experiences regarding the connection between fatness and lifestyle choices. The connection was established in many ways where some users shared experiences on how changing their lifestyle (e.g., diet, exercise) had caused fluctuations in their weight, while some posted criticism where these users opined that fat people could change the way they look if they changed their lifestyle.

*But when someone sees someone obese and someone skinny, it's pretty clear who they think is lazier*

*I dont hate them, I just dont see why people are ok with unhealthy life styles. And when i think fat I'm thinking over obese fat. There's nothing wrong with being a little heavy, but when it effects your health then its a problem.*

Such comments directly criticized the lifestyle of fat people by accusing them of being lazy or following unhealthy lifestyles. Further, these comments portrayed the idea that being fat is a choice that could be changed with a healthy lifestyle.

*A good friend of mine always blamed her weight on thyroid problems, and yet when we'd go out for food she'd always order 'The Eighteen Wheeler' or something else with an obscene amount of food. At that point I have a hard time sympathizing with people's 'genetics'.*

*The least athletic, most immobile human could easily be non-obese by eating a diet that is "normal" to billions of people worldwide.*

Comments like these were mostly focused on diet and eating habits of either themselves or of people they knew. They implied that unhealthy eating habits are linked with fatness;

therefore, if people are responsible for maintaining their diet and eating habits, it would make a difference to their weight. Another segment of comments within this theme included:

*this claim of fat people being lazy is bull. The people I know that are obese, are so busy working 70+ hours a week they don't have time to care for themselves, cook healthy meals, or spend one hour at the gym three times a week. It's sad.*

*The issue I have with this is that my coworker literally could not work out more and she still has a little chub. She is training for an Iron man, teaches spin classes and regularly does triathlons. And no she doesn't just sit there stuffing her face all day. There are people who work out a lot and still have fat (mostly applies for women but applicable to some men too).*

These set of comments project a more positive spin, where the general criticism was not on fat people making irresponsible lifestyle choices. For instance, comments were on a person's inability to maintain a healthy lifestyle due to their busy work schedule. Another form of comments argued on the effectiveness of a healthy lifestyle, as some people did not show much change in weight, even though they kept a healthy routine. Some commenters also shared personal experience on effectiveness of lifestyle change.

*I apologize if i made it seem like i blamed anyone other than myself. My point was that the abundance of fast food/unhealthy food is popular/abundant and easy to get. With people's busy lifestyles its easy to justify taking the quick and easy route over something good for you, even if some healthy food is quickly made.*

*Besides, it was more of a hey, its not your fault, but if you make better decisions you'll be ok. Thus far I've had moderate success in weight loss.*

Therefore, discussions that were brought to the table for this theme included diet, eating habits, gym, exercise, busy work schedules and individual experiences on what can and what cannot be achieved by adopting a healthy lifestyle.

### 3.3.3.2 *Fatness as an Attraction Barrier*

The second most commonly discussed theme was based upon the attractiveness and appearance of fat people.

*The unfortunate answer is that many men don't see women for the beautiful lotus flowers that they are.. Women are objectified far too often. And what does a guy want? An iPhone or one of those Nokias with no color and blue lights on the side.. A fat Nokia.. Basically, you'd see less hate for chubbies if people changed perspective on women, overall.*

*This comes about because when you browse the internet or watch TV all you see is these fake, edited, and unbelievably tiny women/men. Our societies view on beauty is ridiculous. Yes having a good body makes you physically attractive, but that's all most people focus on. [URL included] That video is the perfect reason people hate on and tend to believe you have to be absolutely perfect to be beautiful when really all the models and actresses/actors you see have gone through HOURS of make-up and editing on a computer to produce the image they are. I guarantee half of them look nothing like the camera will make them look.*

Most comments openly discussed the perspectives that society bears on fat people, specifically women. Discussions mostly concentrated on female body objectification and the effect of media portrayals on unrealistic body ideals. Some comments meanwhile shared individual experiences and on being more accepting of their looks.

*I have always been the "chubby girl." Now, as an adult, I've embraced it. Men like the way I look. Granted, it isn't for everyone, but then again not everyone is attracted to white girls or gingers for that matter either. It comes down to this: some people are \*\*\*\*\*. I don't let that get me down, and I do not let myself be "victim" so to say. Am I healthy? Yes. Do I like how I look? Yes. To each their own =)*

These comments demonstrate how fat people avoid stigma as they embrace their bodies gracefully. Many comments were presented in a motivational manner. However, another set of commenters presented their experience of facing fat stigma from a different standpoint.

*It's really horrible as a girl who isn't really super fat but still overweight, that i can't get guys to like me because i'm not rail thin. People don't understand that you're supposed to have some meat on your bones for fertility and such.*

The negative effect of fat stigma was shared by commenters and presented the impact it has had on their lives. The theme largely focused on concepts such as 'beauty' and demonstrates negative sentiments with the use of derogatory words.

### *3.3.3.3 Generic critics of fatness*

The generic critics of fatness consists of comments on different perspectives of fatness, fat stigma and fat people. Such comments varied from negative to positive and from personal to societal.

*If your fat who cares as long as you don't whine and cry about it like most fat people do. "Oh I can't wear that cuz I'm fat :( " \*eats big ass cheese burger with triple cheese\**

The comments like these directly stigmatise the behaviour of fat people and their choices. However, another segment of comments has presented ideas that oppose these.

*What a lot of people don't realize is that all that hate actually makes it \*harder\* for overweight people to lose weight. It's not a motivator at all.*

However, some commenters shared depressing experience to showcase their difficulty in losing weight and how it affected their health as they tried to live up to others views on what is considered as acceptable weight.

*As a fat girl struggling with bulimia, the hate for us fat girls is very depressing and it really does hurt. I've had an ED since about 2006 and I was at my lightest weight when I met my husband. We would go on dates and my ED was so out of hand that I would eat two double cheeseburgers from In N Out, go home and barf it up..it went on like this behind his back for awhile. One time on a date a woman shouted from afar that I was huge...which really hurt because I was at my lightest weight. Anyway, I feel for all of us girls struggling to lose the weight. It's hard and it sucks to be considered less than human.*

*When I was in middle school I was a healthy weight/child. I can still remember one boy, the first person, who told me I was fat and needed to lose weight. I DIDN'T, but it was so traumatic I let it get to me. I started having strong self doubt and reservations about the development of my body. I found out about purging and got sucked into the cycle of bulemia. In \*\*\*\*\* middle school! I stayed thin for a while because of the constant purging, but I also started emotional eating. Now, at 26, I am 220 lbs, struggle with my weight and am completely self deprecating. You teach children to judge and they will in turn do it indiscriminately as a tool for subjecting social standards they don't understand. If you think I am lazy, disgusting, hedonistic,*

*fine! You made me this way! Go \*\*\*\* yourself and I will continue to strive for happiness and peace.*

The above comments highlight the health and psychological issues that are attached to fat stigma. Both selected data excerpts mentioned eating disorders such as bulimia faced by women. They further explain the psychological effect of fat stigma had made women constantly monitor their bodies and this in turn led to critical health issues.

#### *3.3.3.4 Fat People as Easy Targets*

According to the comments of the selected post, most commenters agreed that fat people are often stigmatised because most people believe that fat people are easy targets. Commenters were of the view that social media trolls looked down on fat people and therefore, they purposely bullied fat people who were more vulnerable than others. There were different ways of expressing this belief.

*People look for any reason to think they're superior.. it's called insecurity. Making fun of someone who is heavier than you is one of the easiest ways of convincing yourself you're not just another piece of \*\*\*\* trolling the internet.*

*Because bullies naturally pick on weaker prey, fat people are seen as weak prey because it is assumed that they have little willpower and generally women are more emotional than men*

*Most people are \*\*\*\*\* and need to put down someone to make themselves feel superior.*

The comments associated with this theme are mostly built around the words such as 'social' and 'beautiful'.

### 3.3.4 Discussion of the findings

This study has examined fat stigma discussions, specifically targeting women, by making use of a purposefully selected Reddit post in this context. Analysis of the selected data corpus was conducted using word co-occurrence and sentiment analysis. Moreover, a deeper content analysis of the metadata unravelled four themes, 'fatness as a lifestyle choice', 'fatness as an attraction barrier', 'generic critics on fatness' and 'fat people as easy targets'.

The generated co-occurrence maps illustrate ways around which societies construct opinions and express emotions around four themes. This informed on different notions associated with fat stigma that target women in particular. Although notions of being lazy and unattractive are commonly associated with fatness (Brooker et al., 2018; Puhl & Heuer, 2009), other notions related to health and psychology, and fat people being considered easy targets are new. Further, discussion themes targeting fat women unravelled concept words and their associations with other concept words. Subsequent research in weight-related social media discussions can use these concept words to observe if similar word co-occurrence patterns exist, which can further reveal existence of stigmatisation within each theme.

The comments on 'lifestyle choice' accused fat people, specifically women, for not exercising and having poor food choices. These posts (comprising comments and replies) were made mainly from women who self-identified themselves as fat and also from observations made by those Reddit users who self-categorized themselves as "non-fat". The nature of their reasoning of lifestyle choices varied. The common ideology of most of the commenters was that increased body-weight is solely an effect of non-healthy lifestyles such as laziness. Commenters provided advice on how people should eat and

exercise in order to reduce their weight. However, the discourse by itself provided enough evidence on fat stigma observed at places like gyms, where the victims are targeted for exercising, which can eventually lead the fat people to avoid such places. Following are two examples.

*...few weeks ago I decided that I was going to take up running. I went to \*\*\*\*\* to buy a new pair of running shoes and when I asked the young girl who was working there where the running shoes were located, she looks at me and says 'do you even \*run?\*' She had the worst grin on her face I've ever seen.*

*My husband and I were out for a bike ride last summer and I got a "fat \*\*\*\*\* on a bike" response. It really made me feel embarrassed, and it freaked me out about going out and trying to loose weight.*

While the effectiveness of a healthy lifestyle including healthy food and exercise cannot be ignored, some people experience excessive weight issues due to various reasons, including certain medical conditions. Specifically, obesity in women can be associated with alterations in the reproductive cycle, increased risk of the polycystic ovarian syndrome (PCOS), and maternal obesity (Templeton, 2014). Claims attributing people to be personally responsible for excess weight gain outcomes as a consequence of not conforming to responsible behaviour (also referred as Protestant work ethic) (Crandall, 1994; Mattingly, Stambush, & Hill, 2009) were noted. Such responses could be due to the pervasive belief that one's weight is controllable (Hegarty & Golden, 2008). Moreover, fat women who endorse Protestant work ethics and beliefs associated with it have been found to exhibit lower psychological well-being (Quinn & Crocker, 1999). However, sharing one's own experience of weight-loss due to lifestyle changes, like the below example, reveal aspects of positivity that came with a changed attitude.

*As a compulsive overeater AND alcoholic in recovery, I can confirm this. Took me two years of AA sobriety to even have the strength to look at my compulsive overeating. 100 lbs lost 5 years ago, haven't found those lbs back, yet.*

The second theme observed from the comments relate to 'attraction barrier'. This theme directly links objectification theory proposed by Fredrickson and Roberts (1997) and research findings of recently conducted studies (Fikkan & Rothblum, 2012; Smith, 2012). Many women shared personal experiences on how they have (or have not) been accepted by society as being attractive due to their excess weight. Prior studies have noted that women use their body attractiveness as a mating strategy to attract possible mating partners (Perilloux, Cloud, & Buss, 2013). However, direct fat stigma where women discussed their attractiveness was less evidenced in the selected discourse; therefore, self-stigma was not noticeably observed. This may be due to restraint exercised in sharing experiences on a platform where one can be subjected to onslaught of negativity by other Reddit users. Another reason that could be speculated is that people who are victims of fat stigma have less online space where they can share their ideas and experiences directly. Therefore, when they come across any such post, they tend to instantly share their experiences without considering much about the content which may indirectly indicate hints of self-stigma.

The theme 'generic critics' exhibits general ideas that were shared by commenters regarding fat women, including health issues, such as psychological distress and eating disorders. Women shared experiences of being victims of eating disorders such as bulimia due to societal pressures of not being slim-bodied. These commenters identified social media platforms to promote slim body ideals which eventually can become the social perception of beauty. Most women tend to compare their body with others', so

these promotions indirectly puts extra pressure if they do not belong to the promoted view of “beauty”. The comparisons eventually lead them on a dangerous path of weight-loss that involve unhealthy eating disorders (Tylka & Sabik, 2010). Therefore, mass media should take extra care when promoting slim body ideals. A considerable number of comments shared the idea that fat people are easy targets, although these comments were not specifically targeted towards women. However, most of these comments asked for a justification, such as “*why is there so much hate for fat people, especially girls?*”. This theme mostly targeted people who engaged in stigmatising fat people instead of the victims.

Early commenters to social media post can drive the direction of onward discussion. Brooker et al. (2018) observed that most replies were posted for comments that had been posted at earlier stages of discussion. The format of how discussions are presented in social media forums could be one reason behind such behaviour. Hence, a novel structuring method of highlighting the comments based on the comment score (or the number of upvotes) is essential in order to keep the freshness of the discussion ongoing. Misinterpretation of the word “fat” in the comment section of this post was also noticeable. The usage of the word “fat” is often declared as a verbal weapon that is dissociated from its definition. The word “fat” and its concept seemed to be used to stop fat people, specifically women, from embracing themselves publicly, or from being vocal about their true experience as fat individuals. The discussion on the importance of the knowledge and experience shared by fat people have risen with the concept fat epistemology (Cooper, 2016; Pausé, 2020). Fat epistemology recognizes fat people as the ones who know about fatness and suggest that they are the ones who are appropriate to

produce knowledge about fatness. Therefore, allowing fat people to speak up in public platforms such as social media is needed to understand the valid facts regarding fatness.

Sentiment analysis further brought out positive and negative emotions expressed within the identified themes. Though we did not witness a drastic change in emotions within each of the identified themes, women were found to face more negative criticism during discussions about fat people being easy targets. In having said this, women also faced more positive and reassuring comments on attractiveness. Further, discussions on attractiveness or appearance were mostly from self-stigma posts compared to other three themes. In such cases, indirect stigma may not be identified by the existing sentiment analysis tools. Therefore, critical discourse analysis could complement the analysis to identify direct and indirect stigma content. A combination of word co-occurrence, sentiment and qualitative analysis could inform on fat bias in social media platforms.

### **3.4 Summary**

The findings of both preliminary analyses displayed stigmatising contents related to fatness in social media. Both analyses assisted in refining the methodology of this research. The first preliminary analysis provided insights of keyword distribution, sentiments, word co-occurrences and qualitative aspects of fatness. Findings of this analysis revealed that “fat” was commonly used among the keyword distribution for excess body weight, while “overweight” and “obese” were mainly related to Body Mass Index. That is, content around words “overweight” and “obese” hinged on a clinical note while “fat” hinged as a social descriptor. Therefore, social media discussions about weight are mostly established around the word “fat”. As all three keywords, fat, overweight and obese, indicates positive direction of data retrieval, this research use these keywords to

retrieve data from social media. However, analysing YouTube comments on purposefully selected videos also can be biased toward the video content. The comments are directed towards visual content and indicate opinions and sentiments related to that content. Therefore, analysing the textual data on YouTube without analysing the visual content may not be effective. Hence, the research proceeds with only Twitter and Reddit data to identify discussion patterns of fat stigma.

Although analysing text data using word co-occurrences disclose important findings, using it to identify underlying themes can be subjective as the interpretation of the maps are subjective to the interpreter. Further, the generation of word co-occurrence maps using Textometrica can result in eliminating less frequent words that have a significant impact on the themes. In contrast, topic modelling is considered an effective topic discovery and analysis method by recent hate speech studies (Alshalan et al., 2020; Calderón, de la Vega, & Herrero, 2020; Cao, Lee, & Hoang, 2020; Mathew et al., 2020; Shaw Jr & Karami, 2017). Therefore, this research will replace word co-occurrence with topic modelling for the purpose of unravelling fat stigma discussion themes.

The second preliminary analysis uncovered four discussion themes related to the female fat stigma discussions in social media and their sentiments. Although sentiment analysis is largely used to identify hate speech in social media (Bograd, Chen, & Kavuluru, 2022; Lazarus et al., 2021; Pronoza et al., 2021; Shibly, Sharma, & Naleer, 2021), the incorporation of sentiment analysis and discourse analysis is poorly researched. When conducting the sentiment analysis, contradicting outcomes could be observed as a reply to a negative comment can be negative as it approves the negative comment but shows positive in the analysis due to its wording. This is a result of sentiment analysis being mostly dependent on words or pairs of words; therefore, this issue was hard to tackle.

However, by employing discourse analysis methods alongside sentiment analysis tools can aid in overcoming such issues. As a result, discourse analysis is further incorporated with machine learning techniques to examine fat stigma content in subsequent investigations.

## **Chapter 4: Methodological Framework**

## **4.1 Introduction**

The outcomes of the literature review and the preliminary analysis informed the methodological design for conducting the next phase of the research study. The two literature reviews conducted on the theoretical and practical aspects of previous research on fat stigma on social media revealed the value of in-depth analysis of the discourse (Brooker et al., 2018; Chou, Prestin, & Kunath, 2014; De Brún et al., 2014). Therefore, this study designed a methodological framework incorporating the latest hate speech detection techniques such as sentiment analysis, topic modelling and discourse analysis. Further, the limitations identified in preliminary analysis aided in finalizing the steps of the framework, such as identifying the data sources and techniques to perform analysis methods. This chapter provides an overview of the methodological framework and its key elements.

## **4.2 Methodological Framework**

The proposed methodological framework, outlined in Figure 15, provides a detailed representation of the different steps that are required, that is, data collection, data pre-processing, data analysis, and pattern identification. Tasks required within each step are stated. Each step is discussed in following sub sections.

This study aims to uncover underlying patterns within textual content extracted from social media to identify fat stigmatising patterns. Moreover, our proposed framework can be applied to other forms of hate speech as well. It follows NLP workflow for data collection and data pre-processing and subsequently combines sentiment analysis, emotional analysis, topic modelling, word frequencies and discourse analysis to assess fat stigma patterns in different social media platforms.

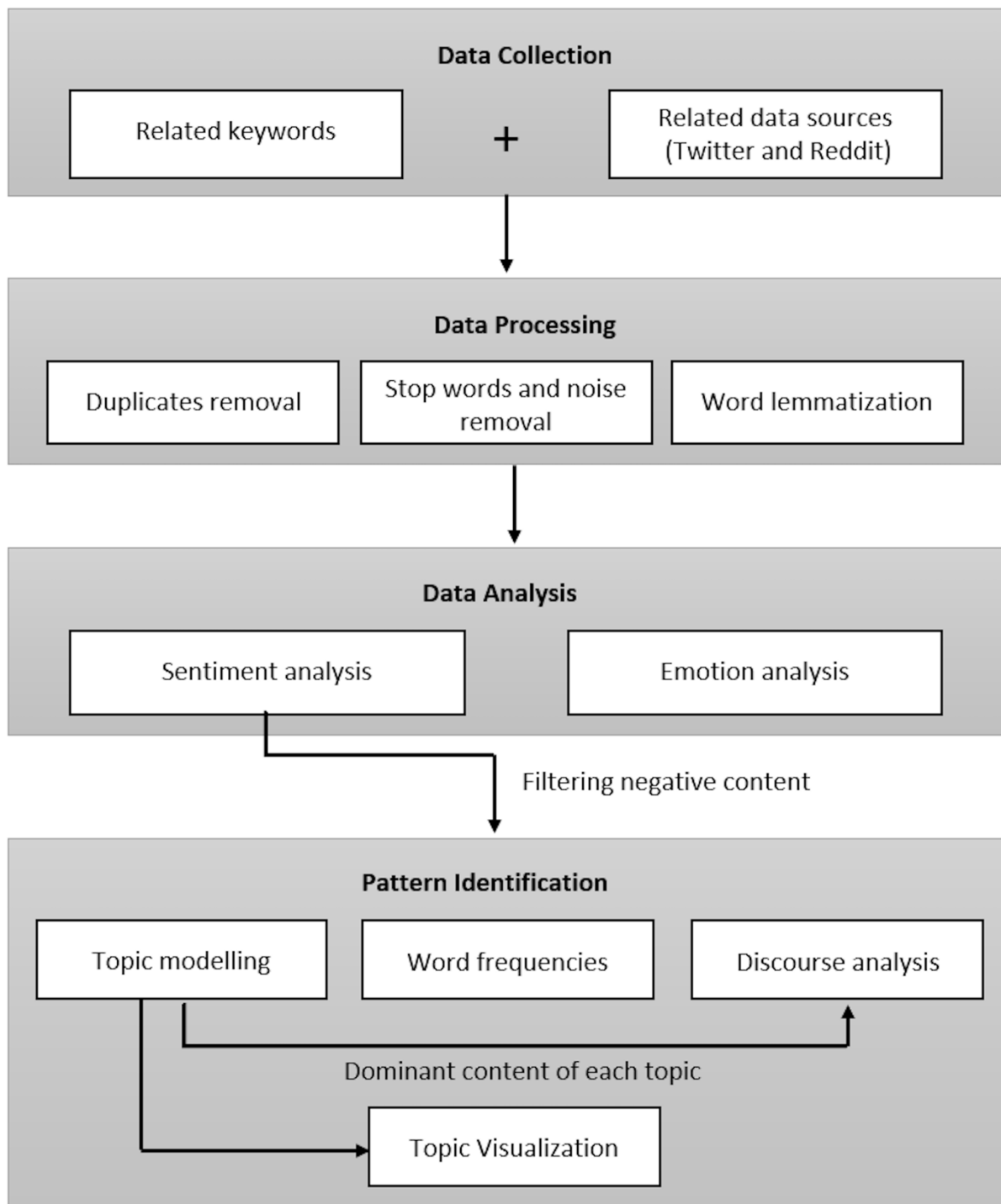


Figure 15. Methodological framework

The outputs of sentiment analysis can aid in filtering the discourse, for instance, only negative sentiments and the associated negative content can be used as inputs for conducting topic modelling to identify fat stigmatising themes. Apart from discovering

these topics/themes, topic modelling can highlight the most dominant posts/comments within each topic, which can then be used to carry out discourse analysis. This framework is based on the research directions implied by recent literature and is further validated by our experiences in analysing fat stigma content in social media. The following subsections describe the activities pertaining to each step.

#### **4.2.1 Data collection**

The main tasks of data collection are to identify suitable data sources and the relevant set of keywords to the study. With knowledge gathered from literature reviews and from the preliminary analyses, the first decision was to decide on the data sources and keywords that relate to fat stigma in social media.

##### *4.2.1.1 Data sources*

As different social media platforms provide various types of features and facilities, not all social media platforms may be relevant to a specific study. Accordingly, this study's context refers to textual data with the aim to identify fat stigma patterns from a range of social media platforms. Hence, this study has been scoped to obesity-related textual data from Twitter and Reddit, both of which have data scraping facility via publicly available APIs (Reddit, 2020; Twitter, 2019). Also, their textual content is mostly not dependent on non-textual content (i.e., image, video, link, etc.), making it easier for subsequently conducting sentiment analysis. Fat stigma studies have made significant use of Twitter in comparison to other social media platforms as explored in Chapter 2. Also, obesity-related discussions are gaining momentum on other platforms, such as in Reddit which appears "more as a discussion forum, and [its] comments have a wider range of document lengths than Twitter data" (Curiskis et al., 2020).

#### 4.2.1.2 Keywords

Two sets of keywords were combined to extract more specific fat stigma content from the selected social media platforms. Data collection started with prior understanding of the context; hence a selection of relevant related keywords was used for scraping and gathering data. As the context of this research focuses on uncovering patterns of fat stigma, those keywords that have been determined in previous literature on obesity-related text mining in social media (i.e., Chou, Prestin, & Kunath, 2014; So et al., 2016) were used. These keywords comprised 'fat', 'overweight' and 'obesity'. While these keywords are recognized as the primary keywords, the study further sought to investigate gender aspects in fat stigma discussions, hence keywords associated to gender too have been combined with the primary keywords (i.e., 'fat+girl', 'obese+man', etc.). The complete list of keywords that formed combinations and data sources are displayed in Table 8.

Table 8. Data sources and list of keywords

Data source	Keyword combinations
Twitter	[fat, obese, overweight] + [woman, women, girl, female, man, men,
Reddit	boy, male]

Data collection plays a crucial role and needs to be carefully executed. A major issue of collecting data from a social media platform would be acquiring irrelevant data. One example of irrelevant data could include the text content posted by a user whose username matches the keywords used for scraping the data. If the keyword combination is 'fat + girl', then posts or comments posted by a user named 'fatgirl123' would be scraped even though these may not be relevant to the topic being studied. Or, in another case, while the given keyword exists in the scraped post/comment, however, the context

of the keyword is different. That is, the keyword 'fat' could be used in a different scenario, such as, the comment has a line containing words like 'big fat liar'. Such cases can be partially solved using a combination of keywords relevant to the study's context instead of just one keyword; although, when the length of the text content is significant, this method may still not be workable. In this study's context, it was found that a combination of keywords helped to acquire more relevant text data from Twitter compared to Reddit, as Twitter enforces a word limit for each post.

#### **4.2.2 Data pre-processing**

Pre-processing steps are needed to present the data in a more structured format for performing effective and accurate analysis. Following pre-processing techniques, including duplicates removal, noise/stop words removal and word lemmatization, are described next.

##### *4.2.2.1 Duplicate removal*

When scraping data from social media, one of the key issues is receiving duplicates. In Twitter, most of duplicates occur when re-tweeting. Although re-tweets can be eliminated at data collection stage for Twitter, the duplicated posts in Reddit are to be explicitly removed. Duplicated posts occur on Reddit because users often post the same content on multiple subreddits. We suggest that the investigator removes duplicates by using 'remove duplicates' feature in Microsoft Excel (Microsoft Corporation, 2016).

##### *4.2.2.2 Stop-words and noise removal*

Removing stop-words is the first step of the pre-processing phase. The common words in a text including prepositions (of, to, for) and conjunctions (and, or, also), which are to be removed to enhance the accuracy of the analysis. Additional noise comprising

superfluous characters such as the symbols (@, ?, !, \*, etc.), emojis (:D, :(, etc.) and URLs in the text data can add more complexity and make the analysis process harder to execute. Automated analyses for creating insights from raw textual data is complicated; hence noisy data can add further distractions and produce results that are inconsistent in subsequent downstream tasks. Therefore, all noisy characters need to be removed.

#### 4. 2.2.3 Word lemmatization

Stemming and lemmatization involves converting a word to its basic form. Stemming usually refers to a crude heuristic process that chops off the ends of words in the hope of achieving the primitive meaning most of the time. This often includes the removal of derivational affixes. Lemmatization usually refers to doing things properly using a vocabulary and morphological analysis of words, normally aiming to remove inflectional endings only and return the base or dictionary form of a word, which is known as the lemma (Manning, Raghavan, & Schütze, 2008). However, using stemming on the dataset may sometimes occur because chopping off the end of some words could result in words with no meaning. Lemmatization does not result in meaningless words since token words are presented in their base form by using their origins, irrespective of whether the word is used in the text as a verb or as a noun. Therefore, we propose using lemmatization instead of stemming. Table 9 gives two examples to represent the difference between stemmed words and lemmatized words.

Table 9. Stemming and lemmatization

<b>Token word</b>	<b>Stemmed word</b>	<b>Lemmatized word</b>
troubled	troubl	trouble
saw	s	see

### **4.2.3 Data analysis**

The researcher needs to conduct meaningful analysis of the pre-processed data for establishing generalizations. Analysis, in this study's context, is deciphering of the fat stigma content expressed in social media discussions in a cumulative and holistic manner.

#### *4.2.3.1 Overall analysis of fatness*

Before filtering any of the fat stigma content, it is important to have a proper grounding of the overall themes around the discussions on fatness. The selected keywords extract discussions of fatness regardless of the type of emotion expressed. As a result, the data extracted from social media indicates positive, negative and neutral sentiments. However, as the main objective of this study is to detect patterns of fat stigma, the contents that display negative sentiments need to be filtered. Therefore, as the initial step, sentiment analysis and emotion analysis is conducted on the acquired social media data. NRC dictionary developed by Mohammad and Turney (2013) has been used to capture more in-depth sentiments and to identify emotions associated with a linguistic corpus.

##### *4.2.3.1.1 Sentiment Analysis*

Sentiment analysis is a subset of NLP that aims to extract qualitative characteristics from text data. Application of sentiment analysis on social media data facilitate to obtain societal standpoint on a specific topic as it is filled with subjective text (Zhang, Wang, & Liu, 2018). Sentiment analysis is most commonly used to detect the polarity in a given text as positive, negative or neutral (Zimbra et al., 2018). This study engaged sentiment analysis for two purposes: identification of overall sentiments regarding obesity, and

filtering text data that display negative sentiments of fatness (fat stigma). Each of these tasks were performed using NRC sentiment lexicon.

Two types of sentiment analysis have been performed for the initial dataset that represents discussions of fatness in social media. A document-level sentiment analysis has been conducted to identify the overall sentiments highlighted in discussions related to obesity. After acquiring the overall sentiment analysis of the entire data corpus, sentence-level sentiment analysis was also performed on each post and comment to determine sentiment scores and labels for each post/comment of the data corpus. Sentiment score or valence score is determined by assigning a sentiment value (ranging from -1 for negative to +1 for positive and 0 for neutral) to each word and then by aggregating all sentiment values of the specific post/comment. Therefore, posts/comments with more negative words indicate a higher negative valence score, and posts/comments with more positive words will indicate a higher positive valence. If there are equal amounts of negative and positive words in a text, the valence score would be '0' and be considered a neutral text. Based on the sentiment scores, each post/comment can be labelled to indicate whether it demonstrates a positive, negative or neutral sentiment. Next, by considering the labels, only those posts and comments that indicate negative sentiment is filtered. In this manner, the filtered linguistic corpus can be used as the input for the next stage, that is, to identify patterns of fat stigma.

#### 4.2.3.1.2 Emotion Analysis

While sentiment analysis reveal positive, negative and neutral sentiments of a given text data, emotion analysis can expose in-depth emotions associated to the text data. Emotion analysis facilitate to detect the psychological state and mood of the user and thereby adds more qualitative aspects of discussions of fatness retrieved from social media. The NRC

emotion lexicon (Mohammad & Turney, 2013) is used to perform the analysis which reveals eight primary emotions namely anger, anticipation, disgust, fear, joy, sadness, surprise and trust for the whole linguistic corpus. In this manner, emotions associated with obesity-related discussions in social media can be unravelled objectively.

The acquired scores of emotions are then interpreted using Pyplutchik (Semeraro, Vilella, & Ruffo, 2021) which illustrates Plutchik's wheel of emotions, to describe how these emotions are related. Plutchik's wheel of emotions organizes the eight identified emotions based on the psychological purpose of each and are grouped into polar opposites: joy and sadness, trust and disgust, fear and anger, and surprise and anticipation. It further can be used to compare the emotions displayed in a linguistic corpus and highlight which emotions are most commonly visible.

#### *4.2.3.2 Pattern identification*

The filtered social media content that indicates negative sentiments are used to identify prominent features (e.g., word frequencies, topics, themes) of fat stigma in social media. Topic modelling and topic visualization, word frequencies and discourse analysis aided to unravel these features.

##### *4.2.3.2.1 Topic modelling*

To discover the common topics of fat stigma discussions, topic modelling based on Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003) is used. LDA is an unsupervised, machine learning, clustering technique that is commonly used to reveal topics of a given document. In LDA, the topics are discovered through the probability distribution of topics over documents that are associated with word distribution. It categories the corpus into

a document and the words per topic by making two key assumptions: documents are a mixture of topics and topics are a mixture of words.

To construct topic models for the filtered fat stigma content, Gensim LDA implementation (Řehůřek & Sojka, 2011) in Python has been used. The initially generated topic models can be evaluated by first calculating the coherence score. Coherence score is the value of the relative distance between words within a topic. It is used to determine how interpretable the generated topics are to humans. There are two major coherence score types, C\_V (which typically range between 0 and 1) and uMass (that range between -14 and 14). This study used the C\_V scores as it provides the strongest correlation with human ratings and is considered a very reliable topic coherence evaluation measure (Röder, Both, & Hinneburg, 2015). In order to determine the optimal number of topics within the male and female corpus, we generated multiple models, each with a different number of topics ranging from 2 topics through to 10 topics. We then used the model with the highest coherence score as the most optimal for determining the number of topics for each corpus.

Once the optimal number of topics were identified, the models were further tuned using alpha and beta hyperparameters. Alpha represents document-topic density and beta represents topic-word density. The values of alpha and beta were recognized by generating multiple models for various alpha and beta value combinations with constant number of topics. Next, the coherence scores were compared. Those alpha and beta values which indicated the highest coherence score have been selected to tune the topic model and obtain topics.

Apart from identifying the discussion topics of the dataset, the most dominant topic of each post/comment and the most representative post/comment for each topic helped determine patterns from our analysis.

#### 4.2.3.2.3 Topic visualization

The determination of dominant topic of each post/comment assisted to visualize the distribution of topics within the document. The visualization was accomplished using UMAP (Uniform Manifold Approximation and Projection for Dimension Reduction) (McInnes et al., 2018). UMAP is a manifold learning technique for dimension reduction which construct a high dimensional graph representation of the data then optimizes a low-dimensional graph that is structurally similar. The visualization of topic distribution provides further insights of how the recognized topics are positioned within the corpus and how they are correlated.

#### 4.2.3.2.4 Word frequencies

The frequency of each word in the linguistic corpus are also determined to discover the most commonly used words in fat stigma discussions. Discovering the commonly used words gives insights of the language usage within fat stigma discussions by recognizing the words that are mostly associated with fatness. The Gensim library in Python facilitates to identify the word frequencies of a given corpus.

#### 4.2.3.2.5 Discourse analysis

The negative obesity-related social media content is further examined by carrying out discourse analysis, aimed to explore how social media users use language in relation to the fat stigma context. Discourse analysis is a research method that studies written or spoken language in relation to its subjective social context. The discourse analysis has

been carried out manually by closely examining the dominant sentence within each discovered topic. The need of an in-depth evaluation on fat stigma content is highlighted in the preliminary analysis. However, the existing text analysis methods are not facilitating automated discourse analysis. Hence, this study performed manual discourse analysis to obtain in-depth implications of fat stigma discussions.

### **4.3 Summary**

The main purpose of this study is to unravel patterns of fat stigma from social media discussions. To accomplish this, several latest machine learning techniques have been incorporated with manual discourse analysis to design a methodological framework. This chapter aims to provide the reader with details of the methodology (methodological framework) used in this study. This chapter describes the key elements of the data collection, including the data sources and the keywords used to collect data from the sources. It further explains data pre-processing steps and the software used to carry out the pre-processing. Additionally, the data analysis methods such as sentiment analysis, emotion analysis, topic modelling, word frequencies and discourse analysis have been explained in detail by briefly describing the theory behind each of these methods and the important steps.

On a final note, it should be added that while social media big data is not collected directly from participants, the data to be used for research must be completely anonymized and in no way should it identify any person's UID (user identity) or URL. This study obtained ethical approval from Massey University Ethical Approval Committee (MUHEC) which ensured that all ethical practices were followed. Since this study did not involve any direct interaction with participants, a low risk ethics approval was secured from MUHEC. A copy of the low risk ethics approval granted from MUHEC is attached in the Appendix.

# **Chapter 5: Application of Proposed Framework on Fat Stigma Data**

## 5.1 Introduction

The proposed methodological framework has been applied to fat stigma data acquired from Twitter and Reddit. This section presents and discusses the results generated at each step of the framework. It elaborates how NLP and qualitative methods have together provided multi-level insights from the textual data extracts. The patterns of fat stigma, and male and female gender differences of fat stigma content are studied simultaneously in this section. It further provides empirical evidence to showcase how our proposed methodological framework can be applied to practical real-world scenarios.

The sub-sections are aligned with each component of the methodological framework and presents with acquired results relating to that component.

## 5.2 Data collection and data pre-processing

The analysis has been carried out to identify fat stigma patterns in social media and how males and females are positioned in obesity-related discussions. Therefore, keyword combinations that represented fatness and gender were selected to scrape data (refer Table 10). As this analysis is focused on textual data only, the publicly available text-based content from Twitter and Reddit were scraped (via their APIs) for a year from 20<sup>th</sup> June 2021 to 20<sup>th</sup> June 2022. The data thus collected served as our data source.

Table 10. Keyword combinations

<b>Obesity-related keywords</b>	Fat, overweight, obese
<b>Gender-related keywords</b>	Female, girl, woman, women, male, boy, man, men
<b>Keyword combinations used</b>	fat + (female, girl, woman, women)
	overweight + (female, girl, woman, women)
	obese + (female, girl, woman, women)
	fat + (male, boy, man, men)
	overweight + (male, boy, man, men)
	obese + (male, boy, man, men)

The acquired data has categorised into four datasets as Twitter-Male, Twitter-Female, Reddit-Male and Reddit-Female. Data distribution among these four datasets has been displayed in Figure 16. The datasets were cleansed by removing stop words, symbols and emojis. Later, lemmatization was also applied to further cleanse the dataset for subsequent analysis. Next, the analysis has been carried out to all four datasets and the results were compared.

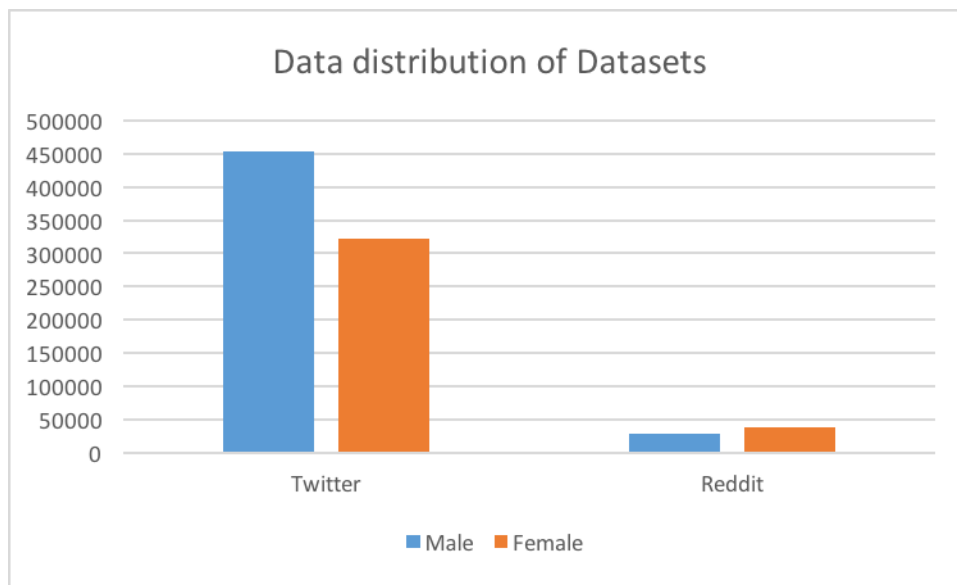


Figure 16. Distribution of data

Twitter contained larger amounts of obesity-related discussions compared to Reddit. Moreover, higher amounts of obesity-related discussions pertaining to the male gender were displayed in Twitter, while in Reddit, the obesity-related discussions were more for females.

### 5.3 Data analysis: overall text analysis of fatness

This section discusses the three data analysis approaches that were used to unearth sentiment values, conduct emotion analysis and generate sentiment labels for text-blocks

representing posts/comments. The following subsections expand on each of these approaches.

### 5.3.1 Sentiment analysis

The sentiment value for each selected comment/post has next been calculated using the NRC dictionary (Mohammad & Turney, 2013). The calculated sentiment values associated with females and males were then compared to identify patterns. The graphical representation of the distribution of overall sentiment (valence) scores associated with male and female genders in each social media platforms is presented in Figure 17.

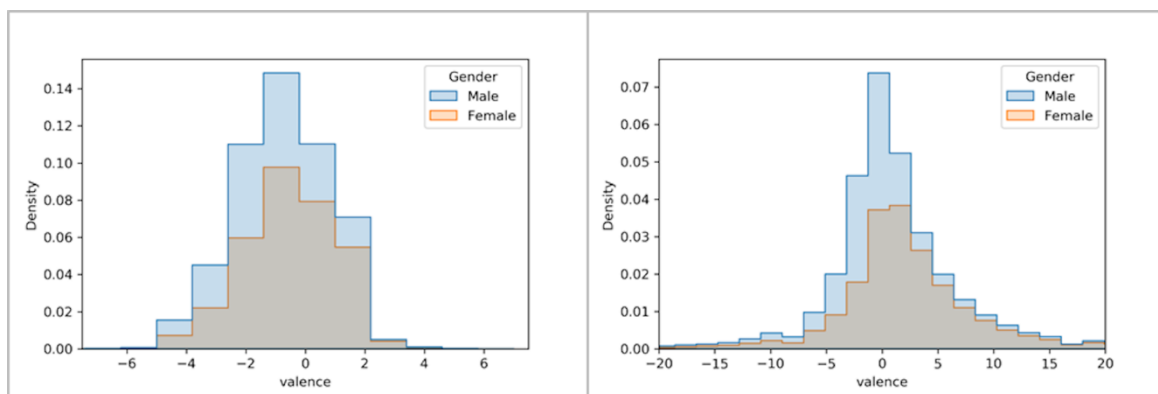


Figure 17. Sentiment value comparison in Twitter (left) and sentiment value comparison in Reddit (right)

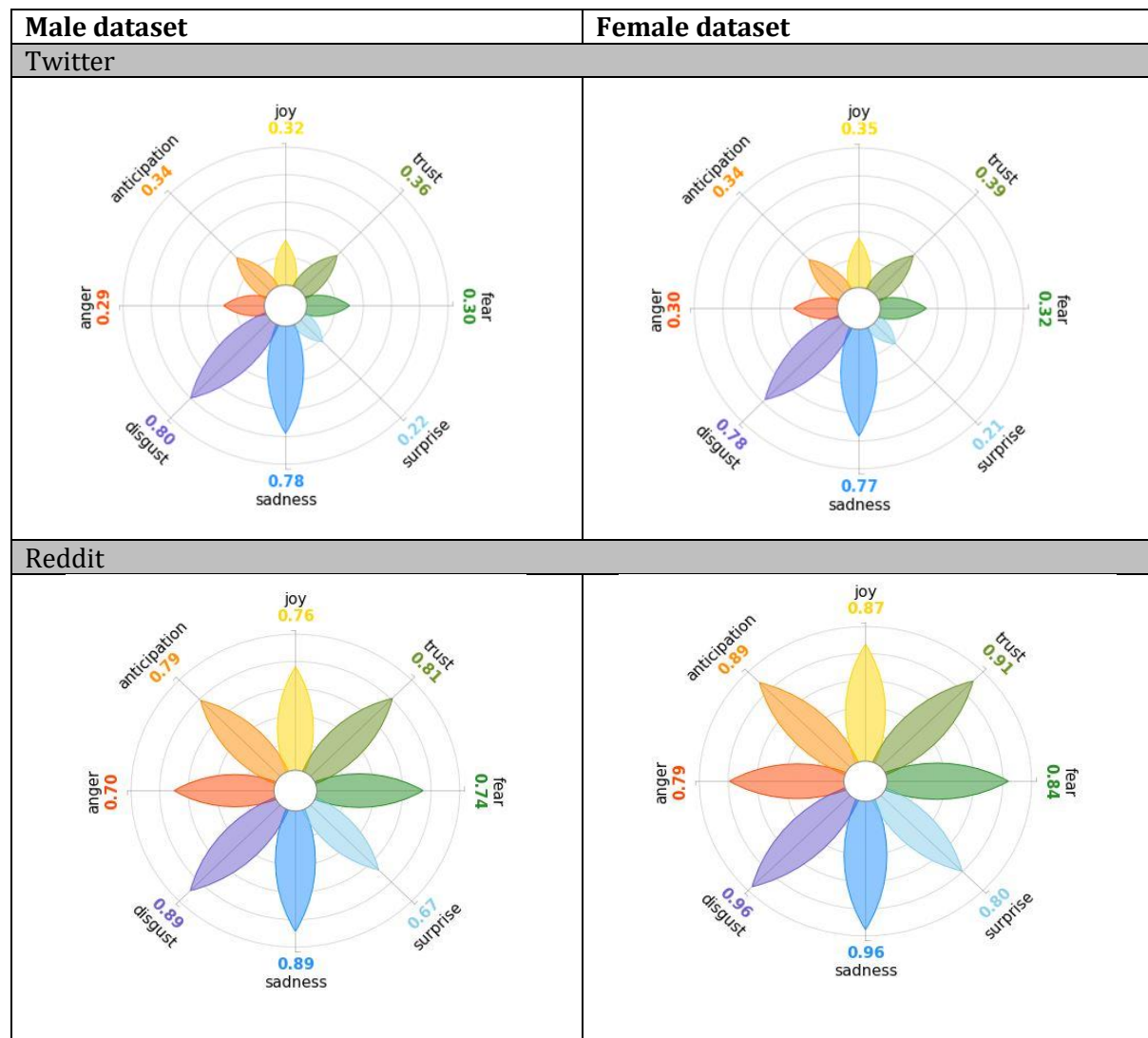
In both platforms, discussions relating to male gender display higher negative and positive sentiments compared to female. Twitter discussions on obesity relating to both male and female genders indicate more negative sentiments than positive sentiments, while in Reddit, a significant difference has not been observed. Also, the valence values in Reddit displayed wider spread compared to Twitter. Armed with this background information, the next step involved gaining further insight into the emotions that were

embedded within these sentiments. The following section expands on the emotion analysis results.

### 5.3.2 Emotion analysis

The eight primary emotions – joy, trust, anticipation, sadness, anger, fear, surprise and disgust – from Plutchik’s wheel were analysed for the whole linguistic corpus (Plutchik, 2001). The resulted Plutchik’s emotion wheels of all four datasets has been illustrated in Table 11.

Table 11. Plutchik's emotion wheels



Twitter indicates significantly higher disgust and sadness compared to other emotions for both male and female datasets. The high intensity values of the disgust category tended more towards loathing, while the sadness category tended towards grief. On the other hand, Reddit shows considerably higher values for all emotions. In both platforms, no significant differences were observed between male and female datasets.

The male and female emotions displayed in Plutchik’s wheels representing social media big data is spread across eight segments. While data amounts used during analysis can be huge, researchers are advised to scale their data to a coarser level with a smaller subset of data and consider such visuals that make their presentation more readable to their audience (e.g., use bar charts for ranking and comparisons) (Kosara, 2016). Therefore, in Figure 18, we have highlighted the differences in the expressed emotions between males and females. This figure depicts the overall differences between males and females that each emotion describes. Positive values describe emotions where the expression score was higher for females than males, and vice versa for negative values.

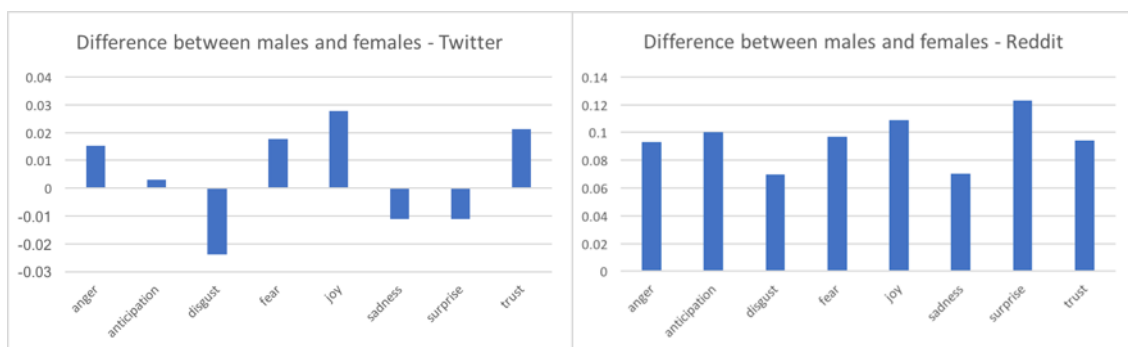


Figure 18. Differences between males and females explained by each emotion.

The graph highlights that female received greater expressions of emotions across all categories in Reddit. Disgust, sadness and surprise emotions are expressed higher for males than females in Twitter.

### 5.3.3 Sentiment labelling

For the purpose of extracting sentiment polarity, sentence-level sentiment analysis has been carried out next. Table 12 illustrates sentiment polarity labels (negative, neutral and positive) of three sentences which have been pre-processed (i.e., cleansed of stop words and other noisy characters). After obtaining the labels for each text-block, only the content which indicated negative sentiment labels were filtered and used as the input for pattern identification stage.

Table 12. Sample data for sentiment labelling

<b>Cleaned Text</b>	<b>Valence</b>	<b>Sentiment</b>
fat woman think theyre hot s**t please know die u	-1	negative
hit home skinny friend complain weight look like model	0	neutral
also im girl will will literally see beautiful plus size girl tv	1	positive

## 5.4 Pattern identification of fat stigma discussions

The underlying themes or patterns related to fat stigma have next been deduced with topic modelling approaches, word frequencies and discourse analysis to gain deeper insight on each topic. The following subsections elaborate on each of these methods.

### 5.4.1 Topic modelling

Topic models were generated by using Gensim LDA implementation for male and female datasets. The process outlined in Section 4.4 was used to determine the optimal number of topics for each dataset. Figure 19 illustrates the graphical representation of optimal number of topics acquired for each dataset.

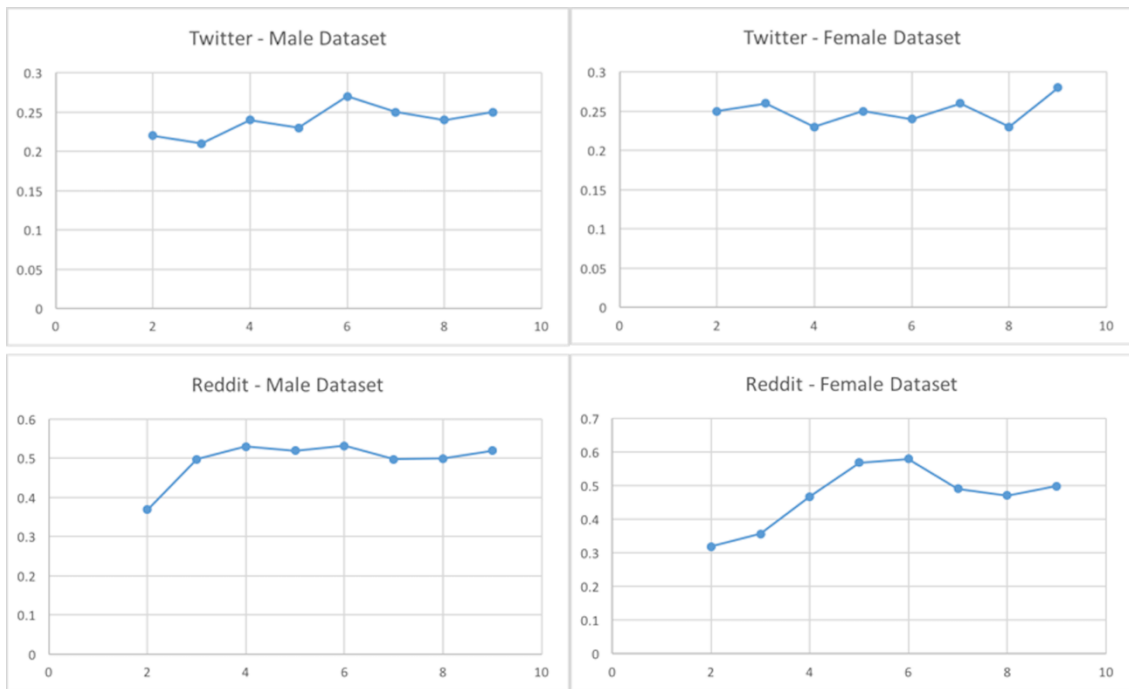


Figure 19. Determining optimal number of topics

The number of topics yielded for male and female datasets, their alpha and beta hyperparameters and coherence scores of each model are displayed in Table 13. It further highlights how model coherence scores are improved as a result of model tuning process.

Table 13. Model parameters and coherence scores

	Optimal Number of Topics	Coherence Score of Initial Model	Alpha	Beta	Coherence Score of Final Model
Twitter					
Male Dataset	6	0.26	symmetric	0.61	0.28
Female Dataset	9	0.28	0.31	0.91	0.29
Reddit					
Male Dataset	6	0.52	0.61	0.01	0.56
Female Dataset	6	0.58	0.61	0.01	0.62

After discovering the keywords generated in each topic, we reviewed these keywords to determine a suitable label for the topic. The set of keywords pertaining to a topic generated for each dataset along with topics' labels are displayed in Table 14 and Table

15 (representing Twitter and Reddit datasets separately). The derogatory and offensive words included in the generated topics are hidden by using the ‘\*’ character.

Table 14. Generated topics and topic labels for discourse in Twitter

Keywords of generated topic	Topic label
<b>Male dataset</b>	
look, male, man, know, age, guy, take, ugly, good, problem	Appearance
obese, man, morbidly, get, call, never, hate, try, life, f**k	Lifestyle
boy, see, bad, lose, want, get, weight, shit, play, body	Body
Still, hate, s**k, long, literally, joke, wrong, last, hear, crush	Hate/humour
Boy, fat, little, get, start, put, first, kid, check, run	Exercise/play
Obese, take, play, hit, show, life, funny, great, picture, post	Share
<b>Female dataset</b>	
woman, man, tell, body, even, many, shame, try, high, still	Body
ugly, big, girl, thing, much, stop, work, life, cause, hair	Appearance
Need, still, hot, wear, day, date, start, eat, dude, post	Dating/relationship
Lose, weight, show, leave, sure, different, pound, diet, late, month	Lifestyle
Girl, wear, never, date, need, take, white, shit, show, right	Dating/Appearance
Fat, girl, get, ass, say, look, big, see, s**t, f**k	Body
Feel, f**k, mad, boy, point, weird, death, long, hear, whole	Appearance
woman, obese, black, morbidly, skinny, think, fact, talk, comment, kill	Health
Girl, people, make, want, bad, look, get, literally, a*s, think	Appearance

Table 15. Generated topics and topic labels for discourse in Reddit

Keywords of generated topic	Topic label
<b>Male dataset</b>	
Weight, year, lose, eat, work, get, day, also, good, month	Lifestyle
Man, look, fat, old, guy, body, big, wear, a*s, see	Appearance
Window, train, police, gun, kill, search, government, back, death, speed	Crime
Man, mark, med, fin, active, dig, min, situation, folk	Activity
Slightly, several, monster, extra, deserve, diagnose, cliff, med, insecure, prefer	Disease
Already, problem, s**k, heart, obesity, case, breathe, complain, admit, totally	Health
<b>Female dataset</b>	
Feel, get, year, day, pain, time, overweight, female, also, weight	Time
Body, blood, large, high, hair, cause, long, foot, skin, breast	Body
F**k, s**t, a*s, c**k, big, p***y, c*m, female, d**k, t*t	Sexual
People, woman, get, want, female, man, make, good, think, say	Gender
Video, sex, porn, girl, black, woman, force, hot, naked, teen	Sexual
Grind, concern, city, power, fee, lab, vitamin, obesity, survive, gallbladder	Disease/me dication

The most topics resulted for Twitter indicate that fat stigma discussions are most commonly linked with appearance. However, some topics could not be labelled into any specific topic label. The low coherence scores that resulted for Twitter datasets indicate that the topic models generated for Twitter datasets cannot be considered as good models.

In Reddit, fat stigma discussions associated with female gender displaying topics that sexually objectify the female body were dominant patterns, while conversations associated with the male gender did not display such dominant pattern.

The frequencies of the most common words of fat stigma discussions have next been compared using word clouds.

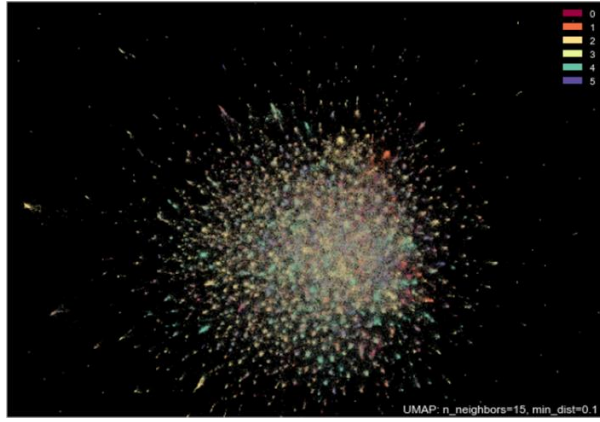
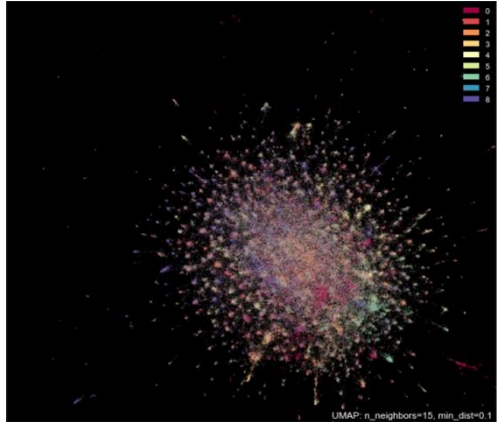
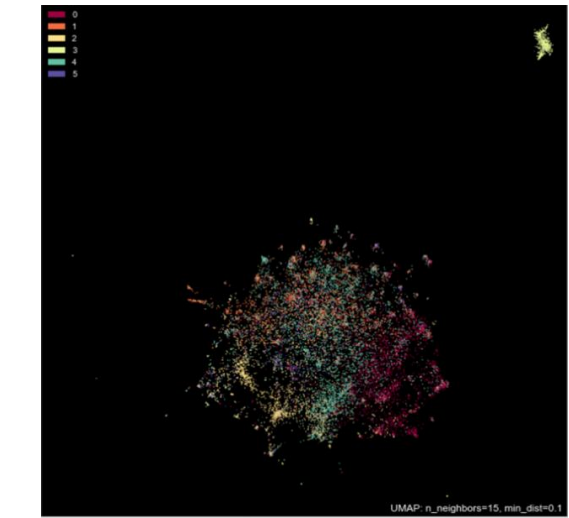
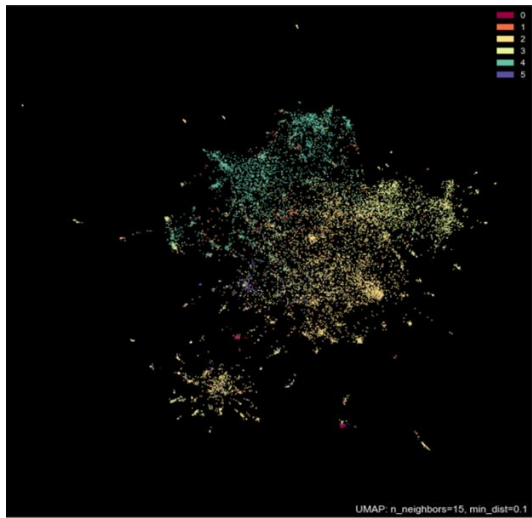
#### **5.4.2 Topic Visualisation**

As described in Chapter 4, the distribution of topics in each dataset has been visualised using Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP). The resulted plots highlight the topics that are mostly dominant in the datasets and how the topics are scattered around the fat stigma discourse. Table 16 illustrate the scattered plots generated using UMAP for each dataset.

Topic number 2, 3 and 4 indicate a vague domination around the male dataset in Twitter while the female dataset indicates stronger domination of topics 3, 5 and 6. The dominant topics of the male dataset represents the topic labels of body (2), hate/humour (3) and exercise/play (4). These are varied topic labels represent different aspects of fat stigma. The female dataset of twitter highlights the topic labels of body (0), dating/relationship (2) and appearance (6 & 8). These labels are commonly correlates with female bodies and appearances.

Male dataset of Reddit disclose the topic numbers 0, 2 and 4 as dominant topics which are represented with labels lifestyle (0), crime (2) and disease (4) respectively. The plot further highlights a distinct and separated mass of topic 3 represents the label "activity". Topic numbers 2, 3 and 4 are dominant in female dataset of Reddit which are belongs to the labels sexual (2 & 4) and gender (3). Therefore, this further demonstrate sexually objectifying content targeting females.

Table 16. Topic visualisation of each dataset

	Male Dataset	Female Dataset
Twitter	 <p>UMAP: n_neighbors=15, min_dist=0.1</p>	 <p>UMAP: n_neighbors=15, min_dist=0.1</p>
Reddit	 <p>UMAP: n_neighbors=15, min_dist=0.1</p>	 <p>UMAP: n_neighbors=15, min_dist=0.1</p>

### 5.4.3 Word Frequencies

The most prominent words of the corpus have been identified for both males and females in Figure 20 and Figure 21.

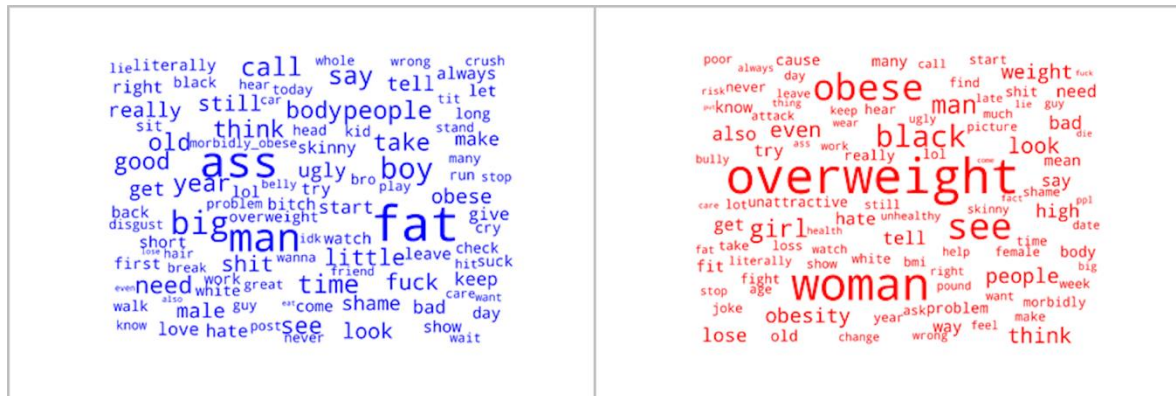


Figure 20. Word clouds of male and female datasets – Twitter

We noted that, in Twitter, fat discussions targeting male and female datasets indicate similar pattern of words. These words contain derogatory words that objectifying both male and female genders. However, the male dataset displays more words relating to body compared to female dataset.



Figure 21. Word clouds of male and female datasets – Reddit

We discovered more derogatory words were observed in the female dataset compared to male dataset in Reddit. The male dataset displayed words associated with health while the female dataset displayed words associated with body.

The keywords used to acquire data from social media are also largely observed in the presented word clouds. Although, removing these words after scraping data can be an option, it could affect when topic modelling is performed.

#### **5.4.4 Discourse Analysis**

Discourse analysis has been conducted by closely examining the most dominant comment/post for each of the identified topics. As a result, some more key features relating to fat stigma discussions were identified.

The fat stigma content targeting males indicates several features. For instance, some users shared scientific studies that they come across on social media; *“People with metabolically healthy obesity were at a substantially higher risk of diabetes, heart attack and stroke, heart failure, respiratory diseases, and all-cause mortality compared with people who were not obese and with a healthy metabolic profile”*. Such content cannot be directly linked with fat stigma as the main purpose of sharing such information would be to educate readers. However, there is a possibility that a fat person may feel victimized based on such claims. Some other content revealed that the social media users consider fatness as ‘unmasculine’; *“They are either smart, nerdy, short, fat, poor. Essentially unmasculine by society”*. Such statements further indicate that some users endorse this perspective as the society’s perspective. Fat individuals who therefore read such comments may consider these statements as a representation of themselves and believe that this is how society views them.

Obesity is found to be most commonly linked with diet and eating habits by social media users; *“ate fast food everyday (led to me becoming extremely obese for my age)”*. Some even provided tips on losing weight; *“I’ve been the obese 13 year old that wanted to lose weight and I asked everyone what to do about it. My parents didn’t know because they had bad*

*eating habits themselves.... The only one who told me the secret recipe was a friend. "Well I just don't eat more than once a day, I only eat boiled veggies and I don't drink much water because it bloats you". And that was my diet for the next three years (13-16) and I lost all the weight reaching 50kgs/110lbs".* Such statements connect body weight with lifestyle choices and mostly try to motivate people to adopt "healthy" lifestyles. Some users exhibited fear of rejection because of their fat bodies while at the same time indicating fat people are unattractive; *"To my horror I find myself being obese, bald, and unattractive. This shocked me, if I ever find my girlfriend, is this how I will look to her? How am I going to explain myself? Is she going to just dump me? And now, how am I going to reach her?"*. These statements are directly linked with body objectification as the connection of body weight with attractiveness is established here. It further highlights that body objectification occurs in males as well as in females.

When targeting females on fat stigma, few social media users linked fatness with appearance; *"But I can confidently say she is not the sperm whale you make her sound like. From my general search, she is what some folks would call 'thick,' which is not technically considered fat, as she does not have a belly sticking out"* and sometimes with health and well-being; *"your girlfriend is overweight, but she is not obese. Anytime you are overweight, there is some possible significant health risk, and it can be great that you care about her well-being"*. These comments further highlight the connection of body weight with lifestyle choices. The dataset also revealed that users share scientific and statistical studies relating to obesity to educate the readers; *"5.8 percent of healthy-weight females with no sexual assault history attempted suicide. The percentage rose to 27.1 percent for healthy-weight girls with a history of sexual assault. Weight influenced the suicide rate among women: 8.2 percent of overweight girls with no sexual assault history attempt suicide"*. However, compared to males, fat stigma content targeting females exposed

higher amount of sexual content. In most of these comments/posts, the words 'fat' and 'obese' were used to describe a person or a female body part, sexually. These contents were mainly posted either to share users' sexual preferences or to describe a visual content.

## **5.5 Summary**

The proposed methodological framework has been applied to discourses that were scraped from Twitter and Reddit. To analyse the male and female gender differences in fat stigma discussions, separate datasets representing male and female gender have been created for both social media platforms. The emerged patterns of fat stigma discussions are then compared between male and female genders, and also for Twitter and Reddit platforms simultaneously. The data distribution amongst four datasets has indicated a considerably higher number of Twitter data compared to Reddit. The overall sentiment analysis further illustrates higher negative and positive sentiments associated with male datasets in both platforms, while Twitter indicated higher negative sentiments of both male and female genders than positive sentiments. However, the emotion analysis, which was conducted to take one step deeper into the sentiments, resulted in more varied emotions for females compared to males. Patterns of emotions have not indicated a considerable difference between male and female but illustrated a noticeable difference between the two platforms.

Next, the topic models were generated for the four datasets separately and the results compared. The models produced for Twitter datasets revealed lower coherence values and thereby resulted in significantly weaker models compared to Reddit. Further, no notable contrast between topics of male and female datasets could be witnessed on Twitter. However, the topic models of Reddit differ between male and female datasets,

reflecting body objectifying content in the female dataset. These results are further clarified with the word clouds constructed using the frequencies of most common words. Finally, discourse analysis has been conducted manually by closely analysing the dominant discourse of each topic that emerged from topic modelling. The analysis unravelled hidden patterns pertaining to the identified topics in the social media discussions. The next chapter discussed these unfolded results in-depth by aligning them with current research.

## Chapter 6: Discussion

The methodological framework of this study and its application to fat stigma discourse have been discussed in this chapter. The discussion has been carried out by aligning the findings with the research questions posed in the introduction chapter. As *RQ1: What are the patterns emerged from fat stigma discussions in social media platforms?* and *RQ2: How are male and female gender differences manifested in fat stigma content in social media?* are associated with the patterns that emerged from the analysis, discussions carried out to answer RQ1 and RQ2 are given in section 6.1. Next, the discussion of the proposed methodological framework has been conducted to answer *RQ3: How can we incorporate sentiment analysis, topic modelling and discourse analysis to detect fat stigma patterns in social media discussions?* in section 6.2.

### 6.1 Patterns of fat stigma and emerged gender differences

To understand the emerging fat stigma patterns of social media platforms, this study compared fat stigma discourse collected from two social media platforms: Twitter and Reddit. Within one year of data collection period, we could acquire a considerably larger dataset for Twitter compared to Reddit. It indicates that there are more discussions of fatness on Twitter rather than on Reddit. Further, Twitter has displayed more negative sentiments than positive sentiments, while Reddit has illustrated more positive sentiments. One of the reasons behind this difference, as discussed in our 1st preliminary analysis, could be that Twitter allows a fewer number of words for commenting (Twitter, 2022). As a result, users may be using terser words to communicate. Meanwhile, Reddit uses auto moderation (Reddit, 2022) of its content to avoid hate speech and derogatory words. This reason could have also explained the considerable difference of emotions in

Twitter and Reddit. While Twitter illustrated higher emotions only for disgust and sadness, Reddit data displayed more varied emotions.

One of the key purposes of this study is to analyse the perceptions towards male and female who are considered to carry excessive weight in the online weight-based discussions in social media platforms. To gain insight on gender related stigmatised content, first, keywords relating to male and female genders were used to scrape online data. Subsequent analysis has disclosed noteworthy themes regarding the positioning of males and females from a fat stigma context.

The overall sentiment analysis of male and female datasets clearly indicates that males receive a higher range of sentiments compared to females in fat stigma discussions. However, the emotion analysis revealed overall more varied emotions towards females compared to males. The emotion analysis results tuned the sentiment analysis outputs and highlighted that although males received a slightly higher number of positive sentiments, females tend to get more varied emotions compared to males. Such findings further justify the decision to use multiple text analysis methods to analyse the fat hatred content.

Fat-based themes targeting females consisted of appearance, body and health-related information, with most of these comments/posts having some degree of sexual content. Fat-related words were often used to describe female body parts in highly sexed manners. Often, these words were paired with other derogatory words that spread more negativity toward females. Further, the dominant topics of female datasets are more related to the topics labelled as sexual, body, appearance and lifestyle while the dominant topics of male datasets are widely scattered around different aspects of fat stigma discussions such as exercise, lifestyle, crime or disease. These findings align with the

objectification theory established by Fredrickson and Roberts (1997) which takes account of how female bodies are sexually evaluated based on “their bodies or body parts” (p. 176). Such body objectification can cause women to adopt an outsider view, where they equate their self-worth with their physical appearance and start self-objectifying and self-monitoring their body shapes. Objectification theory has been used as a theoretical framework for studying the experiential consequences of being a female in a culture which sexually objectifies the female body (Tylka & Hill, 2004). The topics that emerged from this study provide insights on how weight prejudices are reflected in fat-based discussions. We found common topics from both male and female datasets. These include general categories including body, appearance and time. Closer examination of these categories further indicates the extent of sexual objectification. While themes relating to male datasets indicate common words that are associated with general aspects of fatness such as weight loss, exercises, time, medication, etc. However, with the female datasets more objectified keywords appeared in topics such as words relating to appearance, body and sex. These fat stigma patterns illustrate the extent of female body objectification taking place and further demonstrate the application of our methodological framework for advancing of the knowledge base and its alignment with current social theories.

While the findings mostly align with female body objectification, the results also indicates male body objectification aspects. High frequency words found in the male dataset revealed those words that bodily objectify males (e.g., c\*\*k, t\*t, etc.). Also, some dominant topics and topic keywords represent body objectifying content relating to males. However, there was not enough evidence to indicate whether these objectifying words were clearly demonstrating sexual objectification. Studies of recent past have also

explored the male body objectification in relation to the objectification theory and displayed mixed results (Daniel & Bridges, 2010; Davids, Watson, & Gere, 2019; Heath et al., 2016; Wiseman & Moradi, 2010). Therefore, there is not much research on the positioning of males in fat stigma discussions in prior research. This study has filled this gap as it provides a clear comparison of male and female genders illustrated in fat stigma discourse in social media and thereby touches upon aspects of male body objectification, an area that requires more focused research in future.

## **6.2 Methodological framework to detect patterns of fat stigma**

The study has proposed a methodological framework to examine hate speech patterns underpinning discussions in social media. Although sentiment analysis and topic modelling methods have been integrated into prior hate speech (or fat stigma) research (Cao, Lee, & Hoang, 2020; Pronoza et al., 2021; Shibly, Sharma, & Naleer, 2021), the application of such integration to identify patterns of fat stigma is poorly discussed as discovered by the systematic literature review. The review further noted that most studies of fat stigma analysis for observing fat hate patterns in social media were conducted with manual techniques. As a result, most prior studies were limited to small datasets that consist of less than thousand number of social media posts, comments or tweets. Social media spread millions of unstructured data across platforms in a second (Abdul-Rahman et al., 2021; Tighe et al., 2015) targeting various topics. Therefore, analysing limited number of data may not clearly expose the societal stance on a specific topic. Therefore, this study has expanded on using multiple text analysis methods (quantitative and qualitative) to analyse larger amount of fat stigma data to unravel stigmatising patterns.

Fat stigma studies commonly focused on statistical data analysis methods to detect stigmatised content. However, recent novel quantitative analysis methods such as NLP and machine learning techniques can also reveal different dimensions of stigmatised data. Yet, very few studies have incorporated NLP and machine learning techniques with quantitative methods such as discourse analysis for uncovering societal aspects that are associated with fat stigma. The proposed methodological framework overcomes these limitations to show how a combination of sentiment analysis, topic modelling and discourse analysis can be applied to a given textual dataset. to unravel fat stigma patterns in social media. Moreover, this framework can be applied to any textual hate speech content to discover hidden patterns. Study findings have exposed that the use of novel text mining and machine learning methods along with in-depth text analysis techniques such as discourse analysis can aid in the discovery of hate speech patterns. However, we find a gap in the provision of guidelines or a frame of reference that encompasses multiple methods for facilitating detection and enabling researchers to comprehend the stigma patterns embedded in social big data. Therefore, this methodological framework shows conceptual relevancy of how different steps can be undertaken to detect hate speech patterns in social media. The framework details a four-step process that reveals specific activities within each step, explains the conduct of each activity and provides insights on overcoming shortcomings with existing techniques, to report on data-driven findings.

The first step of the proposed methodological framework is to identify the most relevant keywords from published literature that align with the study's context. It may be noted however, that the usage of jargon language has evolved; hence, the nature of online commenting specifically in social media too has changed (Eisenstein et al., 2014). Consequently, the linguistic awareness around words expressed in everyday

conversations over social media platforms needs proper examination for currency and relevancy. Moreover, in collecting relevant social media data by conducting keyword searches, we are very likely to acquire the same comment/post multiple times. While Twitter re-tweets can be easily eliminated, the duplicate comments/posts from Reddit need to be explicitly removed in the data pre-processing step. We advise that to avoid duplicates from Reddit, data acquisition process consider data from a single subreddit, since duplicates mainly occur due to posting the same content on multiple subreddits. Next, we have highlighted several methods for noise removal. By removing stop words, symbols, emojis and URLs, much of the noise in the text can be eliminated. However, in modern-day texting and social media postings, emojis or emoticons play a considerable role in conveying emotions. For example, “I had a rough day LLL” has more emotions than “I had a rough day”. Emojis or emoticons can enhance digital communication by adding non-verbal cues (Alshenqeeti, 2016). Another prominent characteristic of digital text communication is amplifications or abbreviated words, commonly known as SMS language, textspeak, or texting language (Stieglitz & Dang-Xuan, 2013). For example, the canonical form of the words ‘goood’ and ‘gud’ is ‘good’. Normalizing social media text is important to achieve deeper text analysis, specifically on social media data. Satapathy et al. (2017) observed that the accuracy of sentiment classification has increased by ~4% when using such normalization. Therefore, it is clearly important to enhance available text mining techniques to cope with emojis and textspeak for more accurate text analysis. This framework further highlights the use of discourse analysis as an extension to computer-mediated tools and techniques to gain deeper insight on the hate content. Discourse analysis can be combined with topic modelling by connecting filtered comments/posts with strength/size of the topics to thematise topic categories (Shirazi,

2013; Törnberg & Törnberg, 2016). The knowledge categories thus discovered can provide more societal background information of the specified subject to enable better interpretation (e.g., hate content relating to an incident that has occurred). When seeking hate speech patterns, researchers need to focus on who the targets of offensive expressions are, or which content or words cause the most impact, or are intense discussions occurring because of some event. The manual exploration in carrying out discourse analysis cannot be overlooked; hence having relevant data extracts are important. Therefore, technological solutions, such as sentiment analysis can aid in filtering out relevant posts/comments for attaining more accurate discourse analysis that can be used for larger datasets. The dataset acquired from keyword search needs to be further contextualized in the context of the study objective. Hence, to observe hate speech patterns, we must first remove the content that exhibits positive sentiments. The positive sentiments will otherwise act as noise in the detection of hate speech patterns. The content that indicates negative sentiments is carried forward for understanding the hate content. Finally, discourse analysis helps to provide a richer understanding of the subjective negative content, rather than labelling all content as offensive content.

# Chapter 7: Conclusion, Limitations and Future Research Directions

Current social media platforms have fuelled discussions on various emergent issues that are spread in different directions. With features that allow users to share their opinions, ideas and emotions on these issues, the social media has unfortunately also become a place that can foster hate speech. Fat stigma is one such form of hate speech that targets fat people and seeks to shame them for their weight. Some studies have tried to detect fat stigma discussions in social media forums and capture the motivation of users in voicing such content. However, the identification of patterns underlying fat stigma using larger datasets (i.e., social media big data) has been rarely studied over the past decade. This research has therefore provided rich insight on how we can identify fat stigma patterns in a social media context. After a thorough observation of background information relating to fat stigma, three research questions have been posed. The main objectives were set based on the proposed research questions to identify patterns of fat stigma and investigate what gender differences were observed, that is, what patterns underlie males and females who are considered to be fat. This study has further developed a novel methodological framework to guide future research studies in investigating hate speech. The study was initiated by conducting two separate systematic literature reviews. One systematic literature review has been conducted to identify the biological, psychological and sociological correlates associated with fat stigma studies in the social media context. Next, the methodological aspects of these studies were analysed by conducting a second systematic literature review. The findings of these reviews aided this research study in identifying research gaps in existing fat stigma research. With the knowledge acquired

from the background studies and systematic literature reviews, this study has taken one step forward by conducting two preliminary studies. The acquired information was practically tested with these studies and the outcomes thus achieved further facilitated in the design of a novel methodological framework.

The proposed step-by-step framework provides methodological grounding on conducting NLP with social media big data extracts. The context of the problem domain needs to be articulated explicitly by the researcher so that data-driven insights can be revealed. A key part of using automated analytics is that relevant data is collected and is cleansed properly, following which the analysis is context-based and enables the creation of real-world insights. Creating insights is not straight-forward; hence, this study has laid out guidelines on the use of mixed approaches that include different data-driven as well as human-centred qualitative analysis methods. The framework takes hate speech detection a step forward by combining existing NLP techniques to investigate discussion patterns of fat stigma (or hate speech) in a social media context. It bridges the gap between technology and social sciences to enable a deeper understanding on the severity of the socio-psychological impact that fat stigma can cause on fat people. Though this study targets fat stigma, the suggested framework can be applied to various categories of hate speech for identifying relevant themes and patterns in that context. The outputs thus identified (i.e., themes, patterns) can assist in developing machine learning techniques for mitigating hate speech content in social media platforms or other public forums.

The framework has been applied on Twitter and Reddit data associated with fat stigma discussions. The datasets have been further categorised into male and female genders to identify the differences in patterns targeting male and female. The results indicate clear differences of how males and females positioned in fat stigma discussions and aligns with

objectification theory. The results of the study identified higher body objectification targeting females than males. However, the results further provided hints of male body objectification. The application of the framework on fat stigma discussions in social media platforms highlights the implications for conducting more forms of interdisciplinary research. This study bridges data mining and social construction studies with embedded analytics to extend our understanding on male and female objectification. Our study calls on policy makers to plan props for reconfiguring the online space as a place where societies “can function without fear of being threatened, abused, sexually exploited or silenced” (Ging & Siapera, 2018, p. 523).

## **7.1 Contributions**

The outcomes of this study contribute to the practical and theoretical research on hate speech and fat stigma. These contributions align with the research questions that have been answered throughout the study. First, the study generates new knowledge of fat stigma in the social media context. It identifies the sentiments scattered around fat stigma discussions on Twitter and Reddit. It further reveals levels of the eight emotions: anger, trust, anticipation, surprise, disgust, joy, sadness and fear, thereby creating new knowledge on how these emotions are expressed in fat stigma discussions. Future studies could acquire this knowledge when designing fat stigma conventions on how fat people feel when being stigmatised on social media platforms. Also, by unravelling the topics and the frequent words associated with fat stigma discussions, this study informs the most common aspects of fat stigma in the social media context. This information could be used by existing social media platforms when updating their moderating and filtering algorithms of hate speech. It also facilitates social media policymakers to carefully craft social media policies relating to hate speech towards fat people.

Secondly, this study found male and female gender differences in fat stigma discussions. Although the results imply that female bodies are more objectified compared to males, it further highlights varied sentiments associated with males more than females. Such findings hint at the existence of male body objectification as well. Therefore, by extracting male and female gender differences, the study indicates the connection between fat stigma discussions and body objectification. Further, it demonstrates the possibilities of male body objectification in social media, which contributes to the existing objectification theory.

Next, to analyse the fat stigma discussions, this study employed a novel methodological framework. This framework adopts a mixed-method approach and combines sentiment analysis, topic modelling and discourse analysis to recognise patterns of fat stigma. This methodological framework can be applied to any form of hate speech to detect its patterns. Therefore, the proposed methodological framework contributes to hate speech research in conducting analysis on larger datasets and identifying different aspects of hate speech. Also, by combining qualitative and quantitative text analysis methods, it bridges the gap between technology and social studies.

## **7.2 Limitations and future research directions**

Although the research outputs resulting from this study shows promising results and provides answers the posed research questions, there are some limitations that needs consideration in future fat stigma studies.

One of the limitations of this research is that the proposed tools relating to each method of the framework are solely based on prior literature and researchers' experiences. Future studies could compare multiple tools relating to each method and obtain the tools that indicate the most effective results. The proposed framework is focused on text-based

hate speech analysis, while hate speech can be observed in many forms, including visual and audio. In future, the research could expand this framework to incorporate methods of visual and audio analysis to overcome this limitation. Also, the application of the framework has been tested on Twitter and Reddit. However, various types of text-based online media such as forums and weblogs are currently available. Therefore, in future, studies could be conducted to identify patterns of hate speech (fat stigma) on these online media platforms.

Further, the topic models generated for Twitter are not strong models compared to Reddit. We have used LDA to generate topic models for all dataset. However, LDA does not work well with short and sparse text such as Tweets (Hong & Davison, 2010; Zhao et al., 2011). One of the reasons for such output could be that short and sparse text only generate limited word co-occurrence information compared to long documents (Wang & McCallum, 2006). Therefore, future research could employ more relevant topic modelling techniques to obtain strong topic models for Twitter datasets. Further, this research focus on identifying male and female gender differences in fat stigma discourse. However, the discussion of fat stigma in social media could be further expanded by identifying how other gender identities positioned in fat stigma discussions in social media.

Finally, new libraries are emerging to enhance the techniques used for natural language and speech processing fields. These fields are a cross-over of learning that encompass social linguists, psychologists and computational scientists. With new jargon language (comprising emojis, abbreviations, informal prose, etc.) gaining momentum over online discourses, this learning involves ongoing build-up of libraries and methods for refining existing NLP toolkits. Therefore, content analysis in the context of social media big data is a moving target with researchers continually applying new methods and techniques.

This study too is one such attempt to enable better interpretation of the textual content and highlight the underpinning themes by using multiple computational processes along with human-centred inference processes.

## **Appendices**

## STATEMENT OF CONTRIBUTION DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Vajisha Udayangi Wanniarachchi
Name/title of Primary Supervisor:	Assoc. Prof. Anuradha Mathrani
In which chapter is the manuscript /published work:	Chapter 2
Please select one of the following three options:	
<input checked="" type="radio"/> The manuscript/published work is published or in press <ul style="list-style-type: none"> <li>• Please provide the full reference of the Research Output: Wanniarachchi, V. U., Mathrani, A., Susnjak, T., &amp; Scogings, C. (2020). A systematic literature review: What is the current stance towards weight stigmatization in social media platforms? <i>International Journal of Human-Computer Studies</i>, 135, 102371. doi:https://doi.org/10.1016/j.ijhcs.2019.102371</li> </ul>	
<input type="radio"/> The manuscript is currently under review for publication – please indicate: <ul style="list-style-type: none"> <li>• The name of the journal:</li> <li>• The percentage of the manuscript/published work that was contributed by the candidate:</li> <li>• Describe the contribution that the candidate has made to the manuscript/published work:</li> </ul>	
<input type="radio"/> It is intended that the manuscript will be published, but it has not yet been submitted to a journal	
Candidate's Signature:	Vajisha Udayangi Wanniarachchi <small>Digitally signed by Vajisha Udayangi Wanniarachchi Date: 2022.07.13 02:29:46 +12'00'</small>
Date:	13-Jul-2022
Primary Supervisor's Signature:	Anuradha Mathrani <small>Digitally signed by Anuradha Mathrani DN: cn=Anuradha Mathrani, c=NZ, o=Massey University, ou=School of Mathematical &amp; Computational Sciences, email=a.s.mathrani@massey.ac.nz Reason: I am approving this document Date: 2022.07.13 11:25:07 +12'00'</small>
Date:	13-Jul-2022

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

## STATEMENT OF CONTRIBUTION DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Vajisha Udayangi Wanniarachchi
Name/title of Primary Supervisor:	Assoc. Prof. Anuradha Mathrani
In which chapter is the manuscript /published work:	Chapter 2
Please select one of the following three options:	
<input checked="" type="radio"/> The manuscript/published work is published or in press <ul style="list-style-type: none"> <li>• Please provide the full reference of the Research Output: Wanniarachchi, V. U., Mathrani, A., Susnjak, T., &amp; Scogings, C. (2022). Methodological Aspects in Study of Fat Stigma in Social Media Contexts: A Systematic Literature Review. Applied Sciences, 12 (10), 5045. doi:https://doi.org/10.3390/app12105045</li> </ul>	
<input type="radio"/> The manuscript is currently under review for publication – please indicate: <ul style="list-style-type: none"> <li>• The name of the journal:</li> <li>• The percentage of the manuscript/published work that was contributed by the candidate:</li> <li>• Describe the contribution that the candidate has made to the manuscript/published work:</li> </ul>	
<input type="radio"/> It is intended that the manuscript will be published, but it has not yet been submitted to a journal	
Candidate's Signature:	Vajisha Udayangi Wanniarachchi <small>Digitally signed by Vajisha Udayangi Wanniarachchi Date: 2022.07.13 02:32:33 +12'00'</small>
Date:	13-Jul-2022
Primary Supervisor's Signature:	Anuradha Mathrani <small>Digitally signed by Anuradha Mathrani DN: cn=Anuradha Mathrani, c=NZ, o=Massey University, ou=School of Mathematical &amp; Computational Sciences, email=a.s.mathrani@massey.ac.nz Reason: I am approving this document Date: 2022.07.13 11:21:17 +12'00'</small>
Date:	13-Jul-2022

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

## STATEMENT OF CONTRIBUTION DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Vajisha Udayangi Wanniarachchi
Name/title of Primary Supervisor:	Assoc. Prof. Anuradha Mathrani
In which chapter is the manuscript /published work:	Chapter 3
Please select one of the following three options:	
<input checked="" type="radio"/> The manuscript/published work is published or in press <ul style="list-style-type: none"> <li>• Please provide the full reference of the Research Output: Wanniarachchi, V. U., Mathrani, A., Susnjak, T., &amp; Scogings, C. (2019). Gendered objectification of weight stigma in social media: a mixed method analysis. Paper presented at the Australasian Conference on Information Systems, Perth, Australia.</li> </ul>	
<input type="radio"/> The manuscript is currently under review for publication – please indicate: <ul style="list-style-type: none"> <li>• The name of the journal:</li> <li>• The percentage of the manuscript/published work that was contributed by the candidate:</li> <li>• Describe the contribution that the candidate has made to the manuscript/published work:</li> </ul>	
<input type="radio"/> It is intended that the manuscript will be published, but it has not yet been submitted to a journal	
Candidate's Signature:	Vajisha Udayangi Wanniarachchi <small>Digitally signed by Vajisha Udayangi Wanniarachchi Date: 2022.07.13 02:34:03 +12'00'</small>
Date:	13-Jul-2022
Primary Supervisor's Signature:	Anuradha Mathrani <small>Digitally signed by Anuradha Mathrani DN: cn=Anuradha Mathrani, c=NZ, o=Massey University, ou=School of Mathematical &amp; Computational Sciences, email=a.s.mathrani@massey.ac.nz Reason: I am approving this document Date: 2022.07.13 11:16:45 +12'00'</small>
Date:	13-Jul-2022

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

## STATEMENT OF CONTRIBUTION DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Vajisha Udayangi Wanniarachchi
Name/title of Primary Supervisor:	Assoc. Prof. Anuradha Mathrani
In which chapter is the manuscript /published work:	Chapter 3
Please select one of the following three options:	
<input checked="" type="radio"/> The manuscript/published work is published or in press <ul style="list-style-type: none"> <li>• Please provide the full reference of the Research Output: Wanniarachchi, V. U., Scogings, C., Susnjak, T., &amp; Mathrani, A. (2022). Fat stigma and body objectification: A text analysis approach using social media content. <i>Digital Health</i>, 8, 20552076221117404.</li> </ul>	
<input type="radio"/> The manuscript is currently under review for publication – please indicate: <ul style="list-style-type: none"> <li>• The name of the journal:</li> <li>• The percentage of the manuscript/published work that was contributed by the candidate:</li> <li>• Describe the contribution that the candidate has made to the manuscript/published work:</li> </ul>	
<input type="radio"/> It is intended that the manuscript will be published, but it has not yet been submitted to a journal	
Candidate's Signature:	Vajisha Udayangi Wanniarachchi <small>Digitally signed by Vajisha Udayangi Wanniarachchi Date: 2022.12.26 09:16:36 +13'00'</small>
Date:	26-Dec-2022
Primary Supervisor's Signature:	Anuradh Mathrani <small>Digitally signed by Anuradh Mathrani Date: 2023.01.02 16:25:18 +13'00'</small>
Date:	2-Jan-2023

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

## STATEMENT OF CONTRIBUTION DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Vajisha Udayangi Wanniarachchi
Name/title of Primary Supervisor:	Assoc. Prof. Anuradha Mathrani
In which chapter is the manuscript /published work:	Chapter 4 & Chapter 5
Please select one of the following three options:	
<input checked="" type="radio"/> The manuscript/published work is published or in press <ul style="list-style-type: none"> <li>• Please provide the full reference of the Research Output: Wanniarachchi, V. U., Scogings, C., Susnjak, T., &amp; Mathrani, A. (2022). Hate Speech Patterns in Social Media: A Methodological Framework and Fat Stigma Investigation Incorporating Sentiment Analysis, Topic Modelling and Discourse Analysis. Australasian Journal of Information Systems. (Accepted)</li> </ul>	
<input type="radio"/> The manuscript is currently under review for publication – please indicate: <ul style="list-style-type: none"> <li>• The name of the journal:</li> <li>• The percentage of the manuscript/published work that was contributed by the candidate:</li> <li>• Describe the contribution that the candidate has made to the manuscript/published work:</li> </ul>	
<input type="radio"/> It is intended that the manuscript will be published, but it has not yet been submitted to a journal	
Candidate's Signature:	Vajisha Udayangi Wanniarachchi <small>Digitally signed by Vajisha Udayangi Wanniarachchi Date: 2022.12.26 09:24:54 +13'00'</small>
Date:	26-Dec-2022
Primary Supervisor's Signature:	Anuradh Mathrani <small>Digitally signed by Anuradh Mathrani Date: 2023.01.02 16:24:35 +13'00'</small>
Date:	2-Jan-2023

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

14/02/2020

Dear: Vajisha Wanniarachchi

**Re: Low Risk Notification - 400022178 - Analysing underpinning patterns in social media posts that promote weight stigmatisation**

Thank you for your notification which you have assessed as Low Risk.

Your project has been recorded in our database for inclusion in the Annual Report of the Massey University Human Ethics Committee.

The low risk notification for this project is valid for a maximum of three years.

If situations subsequently occur which cause you to reconsider your ethical analysis, please contact a Research Ethics Administrator.

Please note that travel undertaken by students must be approved by the supervisor and the relevant Pro Vice-Chancellor and be in accordance with the Policy and Procedures for Course-Related Student Travel Overseas. In addition, the supervisor must advise the University's Insurance Officer.

**A reminder to include the following statement on all public documents:**

*"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named in this document are responsible for the ethical conduct of this research.*

*If you have any concerns about the conduct of this research that you want to raise with someone other than the researcher(s), please contact Professor Craig Johnson, Director - Ethics, telephone 06 3569099 ext 85271, email [humanethics@massey.ac.nz](mailto:humanethics@massey.ac.nz)."*

Please note, if a sponsoring organisation, funding authority or a journal in which you wish to publish requires evidence of committee approval (with an approval number), you will have to complete the application form again, answering "yes" to the publication question to provide more information for one of the University's Human Ethics Committees. You should also note that such an approval can only be provided prior to the commencement of the research.

Yours sincerely



Professor Craig Johnson  
Chair, Human Ethics Chairs' Committee and Director (Research Ethics)

## **References**

- Abdul-Rahman, M., Chan, E. H., Wong, M. S., Irekponor, V. E., & Abdul-Rahman, M. O. (2021). A framework to simplify pre-processing location-based social media big data for sustainable urban planning and management. *Cities, 109*, 102986.
- Afful, A. A., & Ricciardelli, R. (2015). Shaping the online fat acceptance movement: Talking about body image and beauty standards. *Journal of Gender Studies, 24*(4), 453-472. doi:<https://doi.org/10.1080/09589236.2015.1028523>
- Albadi, N., Kurdi, M., & Mishra, S. (2018). *Are they our brothers? analysis and detection of religious hate speech in the arabic twittersphere*. Paper presented at the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).
- Alberga, A. S., Withnell, S. J., & von Ranson, K. M. (2018). Fitspiration and thinspiration: a comparison across three social networking sites. *Journal of Eating Disorders, 6*(1), 1-10. doi:<https://doi.org/10.1186/s40337-018-0227-x>
- Almenara, C. A., & Ježek, S. (2015). The source and impact of appearance teasing: an examination by sex and weight status among early adolescents from the Czech Republic. *Journal of school health, 85*(3), 163-170. doi:<https://doi.org/10.1111/josh.12236>
- Alshalan, R., Al-Khalifa, H., Alsaeed, D., Al-Baity, H., & Alshalan, S. (2020). Detection of Hate Speech in COVID-19–Related Tweets in the Arab Region: Deep Learning and Topic Modeling Approach. *Journal of Medical Internet Research, 22*(12), e22609. doi:<https://doi.org/10.2196/preprints.22609>
- Alshenqeeti, H. (2016). Are emojis creating a new or old visual language for new generations? A socio-semiotic study. *Advances in Language and Literary Studies, 7*(6).

- Arango, A., Pérez, J., & Poblete, B. (2020). Hate speech detection is not as easy as you may think: A closer look at model validation (extended version). *Information Systems*, 101584.
- Ata, R. N., & Thompson, J. K. (2010). Weight bias in the media: A review of recent research. *Obesity Facts*, 3(1), 41-46. doi:<https://doi.org/10.1159/000276547>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *the Journal of machine Learning research*, 3, 993-1022.
- Bograd, S., Chen, B., & Kavuluru, R. (2022). Tracking sentiments toward fat acceptance over a decade on Twitter. *Health Informatics Journal*, 28(1), 14604582211065702.
- Brewis, A. A., Wutich, A., Falletta-Cowden, A., & Rodriguez-Soto, I. (2011). Body norms and fat stigma in global perspective. *Current anthropology*, 52(2), 269-276. doi:<https://doi.org/10.1086/659309>
- Brooker, P., Barnett, J., & Cribbin, T. (2016). Doing social media analytics. *Big Data & Society*, 3(2), 2053951716658060.
- Brooker, P., Barnett, J., Vines, J., Lawson, S., Feltwell, T., & Long, K. (2018). Doing stigma: Online commenting around weight-related news media. *New Media & Society*, 20(9), 3201-3222. doi:<https://doi.org/10.1177/1461444817744790>
- Busam, B., & Solomon-Moore, E. (2021). Public Understanding of Childhood Obesity: Qualitative Analysis of News Articles and Comments on Facebook. *Health communication*, 1-14.
- Calderón, C. A., de la Vega, G., & Herrero, D. B. (2020). Topic Modeling and Characterization of Hate Speech against Immigrants on Twitter around the Emergence of a Far-Right Party in Spain. *Social Sciences*, 9(11), 188. doi:<https://doi.org/10.3390/socsci9110188>

- Cao, R., Lee, R. K.-W., & Hoang, T.-A. (2020). *DeepHate: Hate speech detection via multi-faceted text representations*. Paper presented at the 12th ACM Conference on Web Science.
- Carels, R. A., Young, K., Wott, C., Harper, J., Gumble, A., Hobbs, M. W., & Clayton, A. (2009). Internalized weight stigma and its ideological correlates among weight loss treatment seeking adults. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity, 14*(2), e92-e97. doi:<http://dx.doi.org/10.1007/BF03327805>
- Carr, D., & Friedman, M. A. (2006). Body weight and the quality of interpersonal relationships. *Social Psychology Quarterly, 69*(2), 127-149. doi:<https://doi.org/10.1177/019027250606900202>
- Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The use of triangulation in qualitative research. *Oncology nursing forum, 41*(5), 545. doi:10.1188/14.ONF.545-547
- Chang, R. M., Kauffman, R. J., & Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision support systems, 63*, 67-80. doi:<https://doi.org/10.1016/j.dss.2013.08.008>
- Chansiri, K., & Wongphothiphan, T. (2021). The indirect effects of Instagram images on women's self-esteem: The moderating roles of BMI and perceived weight. *New Media & Society, 14614448211029975*.
- Charles, M., & O'loughlin, M. (2012). The complex subject of psychosis. *Psychoanalysis, Culture & Society, 17*(4), 410-421. doi:<https://doi.org/10.1057/pcs.2012.30>
- Chatzakou, D., Kourtellis, N., Blackburn, J., De Cristofaro, E., Stringhini, G., & Vakali, A. (2017). *Mean birds: Detecting aggression and bullying on twitter*. Paper presented at the Proceedings of the 2017 ACM on web science conference.

- Chen, N.-C., Drouhard, M., Kocielnik, R., Suh, J., & Aragon, C. R. (2018). Using machine learning to support qualitative coding in social science: Shifting the focus to ambiguity. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 8(2), 1-20. doi:<https://doi.org/10.1145/3185515>
- Cheung, C. M., & Lee, M. K. (2010). A theoretical model of intentional social action in online social networks. *Decision support systems*, 49(1), 24-30. doi:<https://doi.org/10.1016/j.dss.2009.12.006>
- Chou, W.-y. S., Prestin, A., & Kunath, S. (2014). Obesity in social media: a mixed methods analysis. *Translational behavioral medicine*, 4(3), 314-323.
- Chua, T. H. H., & Chang, L. (2016). Follow me and like my beautiful selfies: Singapore teenage girls' engagement in self-presentation and peer comparison on social media. *Computers in Human Behavior*, 55, 190-197. doi:<https://doi.org/10.1016/j.chb.2015.09.011>
- Chung, K.-H., Chiou, H.-Y., & Chen, Y.-H. (2015). Psychological and physiological correlates of childhood obesity in Taiwan. *Scientific reports*, 5(1), 1-9. doi:<https://doi.org/10.1038/srep17439>
- Cooper, C. (2016). *Fat activism: A radical social movement*: Intellect Books.
- Cramblitt, B., & Pritchard, M. (2013). Media's influence on the drive for muscularity in undergraduates. *Eating behaviors*, 14(4), 441-446. doi:<https://doi.org/10.1016/j.eatbeh.2013.08.003>
- Crandall, C. S. (1994). Prejudice against fat people: ideology and self-interest. *Journal of personality and social psychology*, 66(5), 882.
- Crandall, C. S. (1995). Do parents discriminate against their heavyweight daughters? *Personality and Social Psychology Bulletin*, 21(7), 724-735. doi:<https://doi.org/10.1177/0146167295217007>

- Curiskis, S. A., Drake, B., Osborn, T. R., & Kennedy, P. J. (2020). An evaluation of document clustering and topic modelling in two online social networks: Twitter and Reddit. *Information Processing & Management*, 57(2), 102034. doi:<https://doi.org/10.1016/j.ipm.2019.04.002>
- Daniel, S., & Bridges, S. K. (2010). The drive for muscularity in men: Media influences and objectification theory. *Body image*, 7(1), 32-38. doi:<https://doi.org/10.1016/j.bodyim.2009.08.003>
- Davids, C. M., Watson, L. B., & Gere, M. P. (2019). Objectification, masculinity, and muscularity: A test of objectification theory with heterosexual men. *Sex Roles*, 80(7), 443-457. doi:<https://doi.org/10.1007/s11199-018-0940-6>
- De Brún, A., McCarthy, M., McKenzie, K., & McGloin, A. (2014). Weight stigma and narrative resistance evident in online discussions of obesity. *Appetite*, 72, 73-81. doi:<https://doi.org/10.1016/j.appet.2013.09.022>
- Donaghue, N., & Clemitshaw, A. (2012). *'I'm totally smart and a feminist... and yet I want to be a waif': Exploring ambivalence towards the thin ideal within the fat acceptance movement*. Paper presented at the Women's Studies International Forum.
- Durso, L. E., Latner, J. D., White, M. A., Masheb, R. M., Blomquist, K. K., Morgan, P. T., & Grilo, C. M. (2012). Internalized weight bias in obese patients with binge eating disorder: Associations with eating disturbances and psychological functioning. *International Journal of Eating Disorders*, 45(3), 423-427. doi:<https://doi.org/10.1002/eat.20933>
- EBSCO. (2022). EBSCO Discovery Service. Retrieved 2022 from <https://www.ebscohost.com/discovery>
- Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E. P. (2014). Diffusion of lexical change in social media. *Plos one*, 9(11), e113114.

- Ellemers, N. (2018). Gender Stereotypes. *Annual Review of Psychology*, 69(1), 275-298.  
doi:<https://doi.org/10.1146/annurev-psych-122216-011719>
- Erjavec, K., & Kovačič, M. P. (2012). "You Don't Understand, This is a New War!" Analysis of Hate Speech in News Web Sites' Comments. *Mass Communication and Society*, 15(6), 899-920. doi:<https://doi.org/10.1080/15205436.2011.619679>
- Fikkan, J. L., & Rothblum, E. D. (2012). Is fat a feminist issue? Exploring the gendered nature of weight bias. *Sex Roles*, 66(9-10), 575-592. doi:<https://doi.org/10.1007/s11199-011-0022-5>
- Foster, A. K., & MacDonald, J. B. (2013). A tale of two discoveries: Comparing the usability of Summon and EBSCO Discovery Service. *Journal of Web Librarianship*, 7(1), 1-19.
- Founta, A. M., Chatzakou, D., Kourtellis, N., Blackburn, J., Vakali, A., & Leontiadis, I. (2019). *A unified deep learning architecture for abuse detection*. Paper presented at the Proceedings of the 10th ACM conference on web science.
- Fredrickson, B. L., & Roberts, T. A. (1997). Objectification theory: Toward understanding women's lived experiences and mental health risks. *Psychology of women quarterly*, 21(2), 173-206. doi:<https://doi.org/10.1111/j.1471-6402.1997.tb00108.x>
- Gallacher, J. D., Heerdink, M. W., & Hewstone, M. (2021). Online engagement between opposing political protest groups via social media is linked to physical violence of offline encounters. *Social Media+ Society*, 7(1), 2056305120984445.  
doi:<https://doi.org/10.1177/2056305120984445>
- Ging, D., & Siapera, E. (2018). Special issue on online misogyny. *Feminist Media Studies*, 18(4), 515-524. doi:<https://doi.org/10.1080/14680777.2018.1447345>
- Google. (2019). YouTube Data API. Retrieved 2019 from <https://developers.google.com/youtube/v3/>

- Google. (2022). Review and reply to comments. Retrieved from <https://support.google.com/youtube/answer/9482367?hl=en>
- Grogan, S. (2021). *Body image: Understanding body dissatisfaction in men, women, and children*: Routledge.
- Gröndahl, T., Pajola, L., Juuti, M., Conti, M., & Asokan, N. (2018). *All you need is" love" evading hate speech detection*. Paper presented at the Proceedings of the 11th ACM workshop on artificial intelligence and security.
- Guillon, M. J. (2019). Manipulation through online sexual behavior: Exemplifying the importance of human factor in intelligence and counterintelligence in the Big Data era. *The International Journal of Intelligence, Security, and Public Affairs*, 21(2), 117-142. doi:<https://doi.org/10.1080/23800992.2019.1649122>
- Gutman, R. (2018). Reddit's Case for Anonymity on the Internet. *Technology*. Retrieved 2020 from <https://www.theatlantic.com/technology/archive/2018/06/reddit-anonymity-privacy-authenticity/564071/>
- Harris, J. K., Moreland-Russell, S., Tabak, R. G., Ruhr, L. R., & Maier, R. C. (2014). Communication about childhood obesity on Twitter. *American journal of public health*, 104(7), e62-e69.
- Heath, B., Tod, D. A., Kannis-Dymand, L., & Lovell, G. P. (2016). The relationship between objectification theory and muscle dysmorphia characteristics in men. *Psychology of Men & Masculinity*, 17(3), 297. doi:<https://doi.org/10.1037/men0000022>
- Hebl, M. R., & Turchin, J. M. (2005). The stigma of obesity: What about men? *Basic and applied social psychology*, 27(3), 267-275. doi:[https://doi.org/10.1207/s15324834basp2703\\_8](https://doi.org/10.1207/s15324834basp2703_8)

- Hegarty, P., & Golden, A. M. (2008). Attributional beliefs about the controllability of stigmatized traits: antecedents or justifications of prejudice? 1. *Journal of Applied Social Psychology, 38*(4), 1023-1044.
- Heuer, C. A., McClure, K. J., & Puhl, R. M. (2011). Obesity stigma in online news: a visual content analysis. *Journal of health communication, 16*(9), 976-987.
- Himmelstein, M. S., Puhl, R. M., & Quinn, D. M. (2017). Intersectionality: an understudied framework for addressing weight stigma. *American journal of preventive medicine, 53*(4), 421-431. doi:<https://doi.org/10.1016/j.amepre.2017.04.003>
- Holmberg, C., Berg, C., Hillman, T., Lissner, L., & Chaplin, J. E. (2018). Self-presentation in digital media among adolescent patients with obesity: Striving for integrity, risk-reduction, and social recognition. *Digital Health, 4*, 2055207618807603.
- Hong, L., & Davison, B. D. (2010). *Empirical study of topic modeling in twitter*. Paper presented at the Proceedings of the first workshop on social media analytics.
- Hunte, H. E., & Williams, D. R. (2009). The association between perceived discrimination and obesity in a population-based multiracial and multiethnic adult sample. *American journal of public health, 99*(7), 1285-1292. doi:<https://doi.org/10.2105/ajph.2007.128090>
- IBM. (2022). IBM SPSS Software.
- Izydorzyc, B., Walenista, W., Kamionka, A., Lizińczyk, S., & Ptak, M. (2021). Connections Between Perceived Social Support and the Body Image in the Group of Women With Diastasis Recti Abdominis. *Frontiers in Psychology, 3182*. doi:<https://doi.org/10.3389/fpsyg.2021.707775>
- Jackson, T., Jiang, C., & Chen, H. (2016). Associations between Chinese/Asian versus Western mass media influences and body image disturbances of young Chinese women. *Body image, 17*, 175-183. doi:<https://doi.org/10.1016/j.bodyim.2016.03.007>

- Jeon, Y. A., Hale, B., Knackmuhs, E., & Mackert, M. (2018). Weight stigma goes viral on the Internet: Systematic assessment of YouTube comments attacking overweight men and women. *Interactive journal of medical research*, 7(1), e6. doi:<https://doi.org/10.2196/ijmr.9182>
- Jin, S. V., Ryu, E., & Muqaddam, A. (2021). I trust what she's# endorsing on Instagram: moderating effects of parasocial interaction and social presence in fashion influencer marketing. *Journal of Fashion Marketing and Management: An International Journal*. doi:<https://doi.org/10.1108/jfmm-04-2020-0059>
- Jockers, M. (2017). Extracts Sentiment and Sentiment-Derived Plot Arcs from Text. Retrieved from <https://cran.r-project.org/web/packages/syuzhet/index.html>
- Jones, B. A., & Griffiths, K. M. (2015). Self-objectification and depression: An integrative systematic review. *Journal of affective disorders*, 171, 22-32. doi:<https://doi.org/10.1016/j.jad.2014.09.011>
- Jung, F., Spahlholz, J., Hilbert, A., Riedel-Heller, S. G., & Luck-Sikorski, C. (2017). Impact of Weight-Related Discrimination, Body Dissatisfaction and Self-Stigma on the Desire to Weigh Less. *Obesity Facts*, 10(2), 139-151. doi:<https://doi.org/10.1159/000468154>
- Karami, A., Dahl, A. A., Turner-McGrievy, G., Kharrazi, H., & Shaw Jr, G. (2018). Characterizing diabetes, diet, exercise, and obesity comments on Twitter. *International Journal of Information Management*, 38(1), 1-6.
- Kemp, S. (2021). *Digital 2021: Global Overview Report*. Retrieved from <https://datareportal.com/reports/digital-2021-global-overview-report>
- Kent, E. E., Prestin, A., Gaysynsky, A., Galica, K., Rinker, R., Graff, K., & Chou, W.-Y. S. (2016). "Obesity is the new major cause of cancer": connections between obesity and cancer on Facebook and Twitter. *Journal of Cancer Education*, 31(3), 453-459.

- Kocoń, J., Figas, A., Gruza, M., Puchalska, D., Kajdanowicz, T., & Kazienko, P. (2021). Offensive, aggressive, and hate speech analysis: From data-centric to human-centered approach. *Information Processing & Management*, 58(5), 102643. doi:<https://doi.org/10.1016/j.ipm.2021.102643>
- Kosara, R. (2016). Presentation-oriented visualization techniques. *IEEE computer graphics and applications*, 36(1), 80-85. doi:<https://doi.org/10.1109/mcg.2016.2>
- Kumar, N., & Benbasat, I. (2002). Para-social presence and communication capabilities of a web site: a theoretical perspective. *E-Service*, 1(3), 5-24. doi:<https://doi.org/10.2979/esj.2002.1.3.5>
- Latour, B. (2005). An introduction to actor-network-theory. *Reassembling the Social*.
- Lazarus, J. V., Kakalou, C., Palayew, A., Karamanidou, C., Maramis, C., Natsiavas, P., . . . Carrieri, P. (2021). A Twitter discourse analysis of negative feelings and stigma related to NAFLD, NASH and obesity. *Liver International*, 41(10), 2295-2307.
- Ledbetter, A. M., & Meisner, C. (2021). Extending the personal branding affordances typology to parasocial interaction with public figures on social media: Social presence and media multiplexity as mediators. *Computers in Human Behavior*, 115, 106610. doi:<https://doi.org/10.1016/j.chb.2020.106610>
- Lee, H. E., Taniguchi, E., Modica, A., & Park, H. (2013). Effects of witnessing fat talk on body satisfaction and psychological well-being: A cross-cultural comparison of Korea and the United States. *Social Behavior and Personality: an international journal*, 41(8), 1279-1295.
- Lepore, S. J., Revenson, T. A., Weinberger, S. L., Weston, P., Frisina, P. G., Robertson, R., . . . Cross, W. (2006). Effects of social stressors on cardiovascular reactivity in Black and White women. *Annals of Behavioral Medicine*, 31(2), 120-127. doi:[https://doi.org/10.1207/s15324796abm3102\\_3](https://doi.org/10.1207/s15324796abm3102_3)

- Lessard, L. M., & Puhl, R. M. (2021). Adolescents' exposure to and experiences of weight stigma during the COVID-19 pandemic. *Journal of Pediatric Psychology, 46*(8), 950-959.
- Libbey, H. P., Story, M. T., Neumark - Sztainer, D. R., & Boutelle, K. N. (2008). Teasing, disordered eating behaviors, and psychological morbidities among overweight adolescents. *Obesity, 16*(S2), S24-S29. doi:<https://doi.org/10.1038/oby.2008.455>
- Lim, Y., & An, S. (2018). Effects of attributions and social media exposure on obesity stigma among Korean adolescents. *Social Behavior and Personality: an international journal, 46*(12), 2049-2061.
- Lin, H., Zhang, L., Zheng, R., & Zheng, Y. (2017). The prevalence, metabolic risk and effects of lifestyle intervention for metabolically healthy obesity: a systematic review and meta-analysis: a PRISMA-compliant article. *Medicine, 96*(47). doi:<https://doi.org/10.1097/md.00000000000008838>
- Lindgren, S., & Palm, F. (2011). Textometrica Service Package. Retrieved from <http://textometrica.humlab.umu.se>
- Love, J. G., McKenzie, J. S., Nikokavoura, E. A., Broom, J., Rolland, C., & Johnston, K. L. (2016). The experiences of women with polycystic ovary syndrome on a very low-calorie diet. *International journal of women's health, 8*, 299. doi:<https://doi.org/10.2147/ijwh.s100385>
- Lozano-Sufrategui, L., Carless, D., Pringle, A., Sparkes, A., & McKenna, J. (2016). "Sorry Mate, You're Probably a Bit Too Fat to Be Able to Do Any of These": Men's Experiences of Weight Stigma. *International Journal of Men's Health, 15*(1), 4. doi:<https://doi.org/10.3149/jmh.1501.4>
- Lydecker, J. A., Cotter, E. W., Palmberg, A. A., Simpson, C., Kwitowski, M., White, K., & Mazzeo, S. E. (2016). Does this Tweet make me look fat? A content analysis of weight

- stigma on Twitter. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, 21(2), 229-235. doi:<https://doi.org/10.1007/s40519-016-0272-x>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval* (Vol. 1): Cambridge University Press, Cambridge.
- Martins, R., Gomes, M., Almeida, J. J., Novais, P., & Henriques, P. (2018). *Hate speech classification in social media using emotional analysis*. Paper presented at the 2018 7th Brazilian Conference on Intelligent Systems (BRACIS).
- Matamoros-Fernández, A., & Farkas, J. (2021). Racism, hate speech, and social media: A systematic review and critique. *Television & New Media*, 22(2), 205-224. doi:<https://doi.org/10.1177/1527476420982230>
- Mathew, B., Illendula, A., Saha, P., Sarkar, S., Goyal, P., & Mukherjee, A. (2020). Hate begets hate: A temporal study of hate speech. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2), 1-24.
- Mattingly, B. A., Stambush, M. A., & Hill, A. E. (2009). Shedding the pounds but not the stigma: negative attributions as a function of a target's method of weight loss. *Journal of Applied Biobehavioral Research*, 14(3), 128-144.
- McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, 3(29), 861. doi:<https://doi.org/10.21105/joss.00861>
- Microsoft Corporation. (2016). Microsoft Excel. Retrieved from <https://office.microsoft.com/excel>
- Min, J., Yoo, Y., Hah, H., & Lee, H. (2020). Social network technology (SNT) as a tool and a social actor: from self-verification to SNT use. *Internet Research*. doi:<https://doi.org/10.1108/INTR-07-2019-0285>

- Mitei, E., & Ghanem, T. (2020). *Leveraging Social Network Analysis to Explore Obesity Talks on Twitter*. Paper presented at the 2020 IEEE International Conference on Big Data (Big Data).
- Mohammad, S. M., & Turney, P. D. (2010). *Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon*. Paper presented at the Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text.
- Mohammad, S. M., & Turney, P. D. (2013). Nrc emotion lexicon. *National Research Council, Canada, 2*.
- Moreno, M. A., Goniou, N., Moreno, P. S., & Diekema, D. (2013). Ethics of social media research: common concerns and practical considerations. *Cyberpsychology, behavior, and social networking, 16*(9), 708-713.
- Mossie, Z., & Wang, J.-H. (2020). Vulnerable community identification using hate speech detection on social media. *Information Processing & Management, 57*(3), 102087. doi:<https://doi.org/10.1016/j.ipm.2019.102087>
- NLTK. (2020). NLTK Documentation. Retrieved from <https://www.nltk.org>
- Nobata, C., Tetreault, J., Thomas, A., Mehdad, Y., & Chang, Y. (2016). *Abusive language detection in online user content*. Paper presented at the Proceedings of the 25th international conference on world wide web.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . Brennan, S. E. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic reviews, 10*(1), 1-11. doi:<https://doi.org/10.31222/osf.io/v7gm2>

- Papadopoulos, S., & Brennan, L. (2015). Correlates of weight stigma in adults with overweight and obesity: A systematic literature review. *Obesity, 23*(9), 1743-1760. doi:<https://doi.org/10.1002/oby.21187>
- Pausé, C. (2020). Ray of Light: Standpoint Theory, Fat Studies, and a New Fat Ethics. *Fat Studies, 9*(2), 175-187.
- Pawalia, A., Savant, S., Kulandaivelan, S., & Yadav, V. S. (2016). Effect of pre-pregnancy BMI (Obesity) on pregnancy related complications with specific emphasis on Indian Studies: Systematic review based on PRISMA Guideline. *Indian Journal of Obstetrics and Gynecology Research, 3*(3), 239-252. doi:<https://doi.org/10.5958/2394-2754.2016.00054.0>
- Perez, S. (2018). Twitter's doubling of character count from 140 to 280 had little impact on length of tweets. Retrieved from <https://techcrunch.com/2018/10/30/twitters-doubling-of-character-count-from-140-to-280-had-little-impact-on-length-of-tweets/>
- Perilloux, C., Cloud, J. M., & Buss, D. M. (2013). Women's physical attractiveness and short-term mating strategies. *Personality and Individual Differences, 54*(4), 490-495.
- Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist, 89*(4), 344-350.
- Prichard, I., McLachlan, A. C., Lavis, T., & Tiggemann, M. (2018). The impact of different forms of# fitspiration imagery on body image, mood, and self-objectification among young women. *Sex Roles, 78*(11-12), 789-798. doi:<https://doi.org/10.1007/s11199-017-0830-3>
- Proferes, N., Jones, N., Gilbert, S., Fiesler, C., & Zimmer, M. (2021). Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media+ Society, 7*(2), 20563051211019004.

- Pronoza, E., Panicheva, P., Koltsova, O., & Rosso, P. (2021). Detecting ethnicity-targeted hate speech in Russian social media texts. *Information Processing & Management*, 58(6), 102674. doi:<https://doi.org/10.1016/j.ipm.2021.102674>
- Puhl, R. M., Andreyeva, T., & Brownell, K. D. (2008). Perceptions of weight discrimination: prevalence and comparison to race and gender discrimination in America. *International journal of obesity*, 32(6), 992. doi:<https://doi.org/10.1038/ijo.2008.22>
- Puhl, R. M., & Brownell, K. D. (2006). Confronting and coping with weight stigma: An investigation of overweight and obese adults. *Obesity*, 14(10), 1802-1815. doi:<https://doi.org/10.1038/oby.2006.208>
- Puhl, R. M., & Heuer, C. A. (2009). The stigma of obesity: a review and update. *Obesity*, 17(5), 941-964. doi:<https://doi.org/10.1038/oby.2008.636>
- Puhl, R. M., & Latner, J. D. (2007). Stigma, obesity, and the health of the nation's children. *Psychological bulletin*, 133(4), 557. doi:<https://doi.org/10.1037/0033-2909.133.4.557>
- Quinn, D. M., & Crocker, J. (1999). When ideology hurts: effects of belief in the protestant ethic and feeling overweight on the psychological well-being of women. *Journal of personality and social psychology*, 77(2), 402.
- Reddit. (2020). Reddit API Documentation. Retrieved 2020 from <https://www.reddit.com/dev/api/>
- Reddit. (2022). Full AutoModerator documentation. Retrieved from <https://www.reddit.com/wiki/automoderator/full-documentation>
- Řehůřek, R., & Sojka, P. (2011). Gensim—statistical semantics in python. Retrieved from *gensim.org*.
- Roberto, C. A., Sysko, R., Bush, J., Pearl, R., Puhl, R. M., Schvey, N. A., & Dovidio, J. F. (2012). Clinical correlates of the Weight Bias Internalization Scale in a sample of obese

adolescents seeking bariatric surgery. *Obesity*, 20(3), 533-539.  
doi:10.1038/oby.2011.123

Rochelle, T. L., & Hu, W. (2017). Media influence on drive for thinness, body satisfaction, and eating attitudes among young women in Hong Kong and China. *Psychology, health & medicine*, 22(3), 310-318. doi:<https://doi.org/10.1080/13548506.2016.1226507>

Röder, M., Both, A., & Hinneburg, A. (2015). *Exploring the space of topic coherence measures*. Paper presented at the Proceedings of the eighth ACM international conference on Web search and data mining.

Rodríguez, A., Argueta, C., & Chen, Y.-L. (2019). *Automatic detection of hate speech on facebook using sentiment and emotion analysis*. Paper presented at the 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC).

Sales, B. D., & Folkman, S. E. (2000). *Ethics in research with human participants*: American Psychological Association.

Salminen, J., Almerikhi, H., Milenković, M., Jung, S.-g., An, J., Kwak, H., & Jansen, B. J. (2018). *Anatomy of online hate: developing a taxonomy and machine learning models for identifying and classifying hate in online news media*. Paper presented at the Twelfth International AAAI Conference on Web and Social Media.

SAS. (2022). SAS Analytics Software & Solutions. Retrieved from [https://www.sas.com/en\\_nz/home.html](https://www.sas.com/en_nz/home.html)

Satapathy, R., Guerreiro, C., Chaturvedi, I., & Cambria, E. (2017). *Phonetic-based microtext normalization for twitter sentiment analysis*. Paper presented at the 2017 IEEE international conference on data mining workshops (ICDMW).

- Semeraro, A., Vilella, S., & Ruffo, G. (2021). PyPlutchik: Visualising and comparing emotion-annotated corpora. *Plos one*, *16*(9), e0256503. doi:<https://doi.org/10.1371/journal.pone.0256503>
- Shaw Jr, G., & Karami, A. (2017). Computational content analysis of negative tweets for obesity, diet, diabetes, and exercise. *Proceedings of the Association for Information Science and Technology*, *54*(1), 357-365. doi:<https://doi.org/10.1002/pr2.2017.14505401039>
- Shibly, F., Sharma, U., & Naleer, H. (2021). *Classifying and measuring hate speech in Twitter using topic classifier of sentiment analysis*. Paper presented at the International Conference on Innovative Computing and Communications.
- Shirazi, F. (2013). Social media and the social movements in the Middle East and North Africa: A critical discourse analysis. *Information Technology & People*.
- Smith, C. A. (2012). The confounding of fat, control, and physical attractiveness for women. *Sex Roles*, *66*(9-10), 628-631.
- So, J., Prestin, A., Lee, L., Wang, Y., Yen, J., & Chou, W.-Y. S. (2016). What do people like to “share” about obesity? A content analysis of frequent retweets about obesity on Twitter. *Health communication*, *31*(2), 193-206.
- Ștefăniță, O., & Buf, D.-M. (2021). Hate speech in social media and its effects on the LGBT community: A review of the current research. *Romanian Journal of Communication and Public Relations*, *23*(1), 47-55. doi:<https://doi.org/10.21018/rjcp.2021.1.322>
- Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of management information systems*, *29*(4), 217-248.

- Sutin, A. R., Stephan, Y., Luchetti, M., & Terracciano, A. (2014). Perceived weight discrimination and C - reactive protein. *Obesity*, 22(9), 1959-1961. doi:<https://doi.org/10.1002/oby.20789>
- Swartz, M. K. (2011). The PRISMA Statement: A Guideline for Systematic Reviews and Meta-Analyses. *Journal of Pediatric Health Care*, 25(1), 1-2. doi:<https://doi.org/10.1016/j.pedhc.2010.09.006>
- Tandon, A., Dhir, A., Almugren, I., AlNemer, G. N., & Mäntymäki, M. (2021). Fear of missing out (FoMO) among social media users: a systematic literature review, synthesis and framework for future research. *Internet Research*. doi:<https://doi.org/10.1108/intr-11-2019-0455>
- Tang - Péronard, J., & Heitmann, B. (2008). Stigmatization of obese children and adolescents, the importance of gender. *Obesity Reviews*, 9(6), 522-534. doi:<https://doi.org/10.1111/j.1467-789x.2008.00509.x>
- Taniguchi, E., & Lee, H. E. (2015). Individuals' perception of others' self-esteem, psychological well-being and attractiveness: Role of body size and peers' comments among Japanese and Americans. *The Social Science Journal*, 52(2), 217-228.
- Team, R. C. (2013). R: A language and environment for statistical computing.
- Templeton, A. (2014). Obesity and Women's Health. *Facts, views & vision in ObGyn*, 6(4), 175-176.
- Tighe, P. J., Goldsmith, R. C., Gravenstein, M., Bernard, H. R., & Fillingim, R. B. (2015). The painful tweet: text, sentiment, and community structure analyses of tweets pertaining to pain. *Journal of Medical Internet Research*, 17(4), e3769.
- Tomiya, A. J. (2014). Weight stigma is stressful. A review of evidence for the Cyclic Obesity/Weight-Based Stigma model. *Appetite*, 82, 8-15. doi:<https://doi.org/10.1016/j.appet.2014.06.108>

- Törnberg, A., & Törnberg, P. (2016). Muslims in social media discourse: Combining topic modeling and critical discourse analysis. *Discourse, Context & Media, 13*, 132-142.
- Townsend, L., & Wallace, C. (2016). Social media research: A guide to ethics. *University of Aberdeen*, 1-16.
- Tucey, C. B. (2010). *Online vs. face-to-face deliberation on the global warming and stem cell issues*. Paper presented at the Western Political Science Association 2010 Annual Meeting Paper.
- Twitter. (2019). Twitter API. Retrieved 2019 from <https://developer.twitter.com/>
- Twitter. (2022). About offensive content. Retrieved from <https://help.twitter.com/en/safety-and-security/offensive-tweets-and-content>
- Tylka, T. L., & Hill, M. S. (2004). Objectification theory as it relates to disordered eating among college women. *Sex Roles, 51*(11-12), 719-730. doi:<https://doi.org/10.1007/s11199-004-0721-2>
- Tylka, T. L., & Sabik, N. J. (2010). Integrating social comparison theory and self-esteem within objectification theory to predict women's disordered eating. *Sex Roles, 63*(1-2), 18-31.
- Vartanian, L. R., & Novak, S. A. (2011). Internalized societal attitudes moderate the impact of weight stigma on avoidance of exercise. *Obesity, 19*(4), 757-762. doi:<https://doi.org/10.1038/oby.2010.234>
- Vartanian, L. R., Pinkus, R. T., & Smyth, J. M. (2014). The phenomenology of weight stigma in everyday life. *Journal of Contextual Behavioral Science, 3*(3), 196-202. doi:<https://doi.org/10.1016/j.jcbs.2014.01.003>
- Vartanian, L. R., & Shaprow, J. G. (2008). Effects of weight stigma on exercise motivation and behavior: A preliminary investigation among college-aged females. *Journal of Health Psychology, 13*(1), 131-138. doi:10.1177/1359105307084318

- Vigna, F. D., Cimino, A., Dell'Orletta, F., Petrocchi, M., & Tesconi, M. (2017). *Hate me, hate me not: Hate speech detection on facebook*. Paper presented at the Proceedings of the First Italian Conference on Cybersecurity (ITASEC17).
- Wang, C., & Coups, E. J. (2010). Causal beliefs about obesity and associated health behaviors: results from a population-based survey. *International Journal of Behavioral Nutrition and Physical Activity*, 7(1), 19.
- Wang, X., & McCallum, A. (2006). *Topics over time: a non-markov continuous-time model of topical trends*. Paper presented at the Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining.
- Wang, Y., Xie, X., Fardouly, J., Vartanian, L. R., & Lei, L. (2019). The longitudinal and reciprocal relationships between selfie-related behaviors and self-objectification and appearance concerns among adolescents. *New Media & Society*, 1461444819894346. doi:<https://doi.org/10.1177/1461444819894346>
- Wanniarachchi, V. U., Mathrani, A., Susnjak, T., & Scogings, C. (2019). *Gendered objectification of weight stigma in social media: a mixed method analysis*. Paper presented at the Australasian Conference on Information Systems, Perth, Australia.
- Wanniarachchi, V. U., Mathrani, A., Susnjak, T., & Scogings, C. (2020). A systematic literature review: What is the current stance towards weight stigmatization in social media platforms? *International Journal of Human-Computer Studies*, 135, 102371. doi:<https://doi.org/10.1016/j.ijhcs.2019.102371>
- Waseem, Z., & Hovy, D. (2016). *Hateful symbols or hateful people? predictive features for hate speech detection on twitter*. Paper presented at the Proceedings of the NAACL student research workshop.

- Webb, J. B., Vinoski, E. R., Bonar, A. S., Davies, A. E., & Etzel, L. (2017). Fat is fashionable and fit: A comparative content analysis of Fatspiration and Health at Every Size® Instagram images. *Body image*, 22, 53-64.
- WHO. (2022). Data and statistics. Retrieved 2022 from <https://www.euro.who.int/en/health-topics/noncommunicable-diseases/obesity/data-and-statistics>
- Wiseman, M. C., & Moradi, B. (2010). Body image and eating disorder symptoms in sexual minority men: A test and extension of objectification theory. *Journal of Counseling Psychology*, 57(2), 154. doi:<https://doi.org/10.1037/a0018937>
- Wott, C. B., & Carels, R. A. (2010). Overt Weight Stigma, Psychological Distress and Weight Loss Treatment Outcomes. *Journal of Health Psychology*(4), 608. doi:<https://doi.org/10.1177/1359105309355339>
- Yeruva, V. K., Junaid, S., & Lee, Y. (2019). Contextual Word Embeddings and Topic Modeling in Healthy Dieting and Obesity. *Journal of Healthcare Informatics Research*, 3(2), 159-183.
- Yoo, J. H., & Kim, J. (2012). Obesity in the new media: a content analysis of obesity videos on YouTube. *Health communication*, 27(1), 86-97.
- Zahn, C., Schaeffeler, N., Giel, K. E., Wessel, D., Thiel, A., Zipfel, S., & Hesse, F. W. (2014). Video clips for YouTube: Collaborative video creation as an educational concept for knowledge acquisition and attitude change related to obesity stigmatization. *Education and Information Technologies*, 19(3), 603-621. doi:<https://doi.org/10.1007/s10639-013-9277-5>
- Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4), e1253.

- Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011). *Comparing twitter and traditional media using topic models*. Paper presented at the European conference on information retrieval.
- Zhong, B. (2020). Social consequences of internet civilization. *Computers in Human Behavior*, *107*, 106308. doi:<https://doi.org/10.1016/j.chb.2020.106308>
- Zimbra, D., Abbasi, A., Zeng, D., & Chen, H. (2018). The state-of-the-art in Twitter sentiment analysis: A review and benchmark evaluation. *ACM Transactions on Management Information Systems (TMIS)*, *9*(2), 1-29.