FATORES ASSOCIADOS À MUDANÇA DE PESO EM REDES SOCIAIS: DE CORRELAÇÕES À CAUSALIDADE

TIAGO OLIVEIRA CUNHA

FATORES ASSOCIADOS À MUDANÇA DE PESO EM REDES SOCIAIS: DE CORRELAÇÕES À CAUSALIDADE

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do grau de Doutor em Ciência da Computação.

Orientador: Gisele Lobo Pappa

Belo Horizonte
Dezembro de 2016

TIAGO OLIVEIRA CUNHA

FACTORS ASSOCIATED WITH WEIGHT CHANGE IN ONLINE SOCIAL MEDIA: FROM CORRELATION TO CAUSAL ANALYSES

Thesis presented to the Graduate Program in Computer Science of the Federal University of Minas Gerais in partial fulfillment of the requirements for the degree of Doctor in Computer Science.

Advisor: Gisele Lobo Pappa

Belo Horizonte

December 2016

© 2016, Tiago Oliveira Cunha.
Todos os direitos reservados.

Cunha, Tiago Oliveira

C972f Factors Associated with Weight Change in Online Social Media: From Correlation to Causal Analyses / Tiago Oliveira Cunha. — Belo Horizonte, 2016
xxiv, 76 f. : il. ; 29cm

Tese (doutorado) — Federal University of Minas Gerais
Orientador: Gisele Lobo Pappa

1. Computação. 2. Redes sociais on-line.3. Comunidades online.4. Inferência causal. I. Orientadora. II. Título.

CDU 519.6*04(043)

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

FOLHA DE APROVAÇÃO

Factors associated with weight change in social media: from correlations to causal analyses

TIAGO OLIVEIRA CUNHA

Tese defendida e aprovada pela banca examinadora constituída pelos Senhores:

PROFA. GISELE LOBO PAPPA - Orientadora
Departamento de Ciência da Computação - UFMG

PROF. ALLINE MARIA REZENDE BELEIGOLI
Faculdade de Medicina - UFMG

PROF. ANA PAULA COUTO DA SILVA
Departamento de Ciência da Computação - UFMG

PROF. FABRÍCIO BENEVENUTO DE SOUZA
Departamento de Ciência da Computação - UFMG

PROFA. JONICE DE OLIVEIRA SAMPAIO
Departamento de Ciência da Computação - UFRJ

PROF. LUIS ENRIQUE ZÁRATE
Departamento de Ciência da Computação - PUC-MG

Belo Horizonte, 19 de dezembro de 2016.

To my parents. Antonio and Vilma.

Acknowledgments

First of all this thesis would not be possible without the support of my family. They were my safe port and provided a lot of encouragement in the difficult times and kept me motivated to conclude my work. Thus this thesis is dedicated to my mom Vilma, my dad Antonio, my sister Viviane and my girlfriend Veruska. Also for other close relatives that were cheering for my success.

My sincere thanks to Ingmar Weber that change the direction of my work and made a huge contribution to the success of this work and to my formation as a scientist. I will never forget the very productive meetings that we had in the 9 months I stayed in Qatar also the numerous Skype meetings after my return. A huge thanks for my advisor Gisele, like she said we had our ups and downs during the 4 years, but I always felt we were a team and she was there for me providing guidance and all sort of help. Thanks to Alline Beleigoli that showed a total different aspect of my work and made a huge contribution.

Of course I can not forget to thank all my DCC friends, they were very important to make UFMG a great place to be during my PhD. A special thanks to Rafael Colares, Rafael Santin, Mazza, Fernandinho, Walter, Zilton, Evelin, Mariane, Paulo Bicalho, Paulo Retore, Carlos, Leonardo, Pedro and a lot more that I possible forgot. Also to my friends from “Almeida”, Fillipe, Flávio and Giovanni that even from far were key persons to keep my sanity during those long 4 years. Thanks for my roommate Filipe for all the amazing “social” conversations that provided a different view about a lot of subjects.

Finally, a huge thanks to my QCRI friends, especially to Yelena, Aarti, Dat and Ankit that helped me to felt that I was at home despite being literally in the other side of the world.

"Ever tried. Ever failed. No matter. Try again. Fail again. Fail better."

(Samuel Beckett)

Resumo

A obesidade é reconhecida como um grande problema de saúde pública, aumentando o risco de doenças crônicas. Estudos têm associado o sucesso em perda de peso, bem como outros resultados positivos para a saúde a presença de suporte social. Com o surgimento da internet e comunidades *online* tornou-se mais fácil para os usuários encontrar grupos de apoio virtual para perda de peso. Embora haja uma série de estudos que analisam o efeito de receber suporte social sobre o engajamento em comunidades *online*, o efeito desse suporte sobre condições de saúde não foi estudado minuciosamente. É comum encontrar estudos que se baseiam apenas em correlações e assumem que essas correlações compartilham verdadeira relação causal. Esta tese preenche lacunas na literatura propondo um *framework* para estudar o efeito causal que receber suporte social na forma de comentários em *loseit*, uma comunidade de perda de peso no Reddit, tem sobre (i) a probabilidade do usuário ficar mais tempo no fórum e sobre (ii) a perda de peso relatada pelo usuário. Usando uma abordagem de *matching* para inferência causal, observamos uma diferença de 9 libras perdidas entre os usuários que recebem ou não comentários. Este efeito não é mediado por aumento na vida útil na comunidade ou por aumento no nível de actividade dos usuários. Estudamos as propriedades do *loseit* e mostramos seus mecanismos para fornecer suporte social. Além disso, construímos modelos estatísticos supervisionados utilizando características linguísticas do conteúdo compartilhado pelos usuários, bem como suas interações sociais para inferir a condição de obesidade. Descobrimos que a natureza do suporte social manifestado no Reddit, bem como o envolvimento do usuário e os tópicos discutidos são indicativos de mudança de peso com até 66 % de precisão. Nossos resultados mostram a importância que o suporte social tem ao usar fóruns de suporte *online* para alcançar melhorias na condição de saúde e o potencial das comunidades *online* para se tornar uma ferramenta fundamental para ajudar os profissionais de saúde.

Palavras-chave: Redes sociais online, Comunidades de suporte, Inferência causal, Análise de mediação.

Abstract

Obesity has long been recognized as major public health problem, increasing individual's risk for chronic diseases. Previous studies have tied successful weight loss as well as other positive health outcomes to the presence of social support. With the emergence of the internet and social media it has become easier for users to find virtual support groups for weight loss. Though there are a number of studies looking at the effect of receiving social support on engagement with an online community, the effect of such support on health outcomes has not been thoroughly studied. It is common to find studies that are only based on correlations and assume that these correlations share true causal relation. This thesis fills gaps in the literature by proposing a framework to study the causal effect that receiving social support in the form of comments in *loseit*, a Reddit weight loss community, has on (i) the probability of the user to stay longer in the forum, and, more importantly, on (ii) the weight loss reported by the user. Using a matching approach for causal inference we observed a difference of 9 lbs lost between users who do or do not receive comments. Surprisingly, this effect is mediated by neither an increase in lifetime in the community nor by an increased activity level of the user. We study the properties of *loseit* and show its mechanisms to provide social support. Further, we build supervised statistical models using linguistic features of the content shared by the users as well as their social interactions to infer the users obesity condition. We found that the nature of social support manifested in Reddit as well as the user engagement and the topics discussed are indicative of weight change with up to 66% of precision. Our results show the importance that social support has when using online support forums to achieve health outcomes and the potential of OSNs to become a fundamental tool to aid health professionals.

Palavras-chave: Online social media, Online health communities, Causal Inference, Mediation Analysis.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

Over the last decades, obesity rates have increased in many countries around the world, making the condition a major public health problem. Obesity is associated with significantly increased risk of more than 20 chronic diseases and health conditions [Thiese et al., 2015], and directly affects quality of life. In Brazil, the Risk and Protective Factors Surveillance System for Chronic Diseases by Telephone Interviews (VIGITEL) revealed that 52,5% of the population is overweight or obese[1] and that chronic diseases began to predominate in the death statistics, especially when compared with mortality from infectious diseases, which have had a significant decline. In the U.S. 68.8% of adults are overweight or obese[2] and one out of four adults has two or more chronic health conditions[3]. Financially, health issues related to obesity and lifestyle diseases impose an ever-increasing burden with medical costs linked to obesity estimated at USD 147 billion in 2008[4].

Though the percentage of U.S. adults self-reporting to be on a diet in any given week has fallen from 31% in 1991, 20% are still trying to lose weight through dieting at any point in time[5]. Previous studies have tied successful weight loss as well as other positive health outcomes to the presence of social support. In particular for long-term behavioral changes required for achieving and maintaining weight loss, user engagement is essential [Teixeira et al., 2015]. Previous studies showed that users who stay longer

[1] http://www.abeso.org.br/noticia/quase-60-dos-brasileiros-estao-acima-do-peso-revela-pesquisa-do-ibge

[2] http://www.niddk.nih.gov/health-information/health-statistics/Pages/overweight-obesity-statistics.aspx

[3] https://www.cdc.gov/chronicdisease/overview/

[4] http://www.cdc.gov/chronicdisease/overview/#sec3

[5] https://www.npd.com/wps/portal/npd/us/news/press-releases/the-npd-group-reports-dieting-is-at-an-all-time-low-dieting-season-has-begun-but-its-not-what-it-used-to-be/

in these programs have greater success in achieving their goals [Patrick et al., 2011].

Through the internet and social media it has become easier for users to find virtual support groups for anything from weight loss to drug addiction or depression. Though there are a number of studies looking at the effect of receiving social support on sustained engagement with an online community [Burke et al., 2009; Cunha et al., 2016], the effect of such support on health outcomes has not been thoroughly studied. In particular, it is not clear how large one would expect the effect of receiving online comments on observed health outcomes to be as, arguably, many other effects from one's social environment should dominate the benefits of receiving encouragement from potential strangers online.

To study the importance of online support, one would, ideally, set up a proper experiment with a randomized control and treatment group. The treatment group would then receive encouragement while the control group remains ignored. However, such studies both require access to an appropriate platform and they also come with certain ethical concerns [Song and Chung, 2010]. Pushed to the limit it would, for example, be unethical to withhold online social support from a suicidal person wanting to chat. Even not offering support to a person trying to lose weight could be questionable.

With the emergence of online social network platforms like Facebook and Twitter over the last decade, a huge amount of user-generated content from social interactions has been created with various topics related to specific diseases. These topics include people's experiences, recommendations and feedback about certain medications, medical procedures, diets or exercises, and support. Hence, this observational data offers the opportunity to study social processes when in a context where a randomized experiment is not feasible.

A central aphorism of statistics is that correlation is not necessarily causation, i.e., when we observe that two things are associated, it does not logically imply that changes in one of them cause changes in the other [Esarey, 2015]. However, it is common to find studies that are only based on co-occurrence and assume that co-occurring items share some true causal relation. Efforts to investigate the causality in observational studies on social media are rare, but recent studies tried to improve on correlational analyses by applying causal inference techniques that have come into extensive use in medicine, economics, and other sciences [Olteanu et al., 2016; De Choudhury et al., 2016; Tsapeli and Musolesi, 2015].

However, detecting causal effects in observational data is challenging. For example, users that received social support in an online community may systematically differ from users that did not. In order to eliminate any bias due to differences on the baseline characteristics of "exposed" and "unexposed" users, it is mandatory to control for several factors that could influence the result of the study.

In this thesis, we propose a framework for conducting causation studies on weight loss from social media data. We make use of a growing collection of methods for causal inference from observational data. Though not without their limitations, such methods allow to go beyond arguing about correlations, attempting to rule out as many confounding factors as possible. In particular, in this thesis we look at the effect of receiving social support in a popular Reddit[6] weight loss community, the /r/loseit[7], on users' self-reported weight loss. Reddit is a social media platform for internet communities where users called "redditors" share content and respond back and forth in a conversation-tree of comments. Though, our framework is general enough to be applied in other contexts. We look at whether users who receive support in the form of comments on their first post in the community are more likely than those who do not to (i) return for another activity in the community, and to (ii) later report a higher weight loss, as measured to the community's badge system.

We characterize the behavior of users in *loseit*, using a number of linguistic and social interaction measures. Then we analyze posts and comments content to identify types of social support and which subjects of topics drive the discussion around weight loss. We show that there is a positive correlation between social support and weight change aspects. Next, to correct for content differences in users' posts, which are linked to receiving more or less support, we apply a matching approach: a post receiving a number of comments bigger than a cutoff is paired with a post very similar in content that received a number of comments smaller than the cutoff. Similarity is defined in terms of posts exhibiting similar features derived via a statistical model using LDA topics [Blei et al., 2003], Linguistic Inquiry and Word Count (LIWC [8]) features, sentiment analysis [Hutto and Gilbert, 2014], question-centric words and posts length. To see if, for those users returning to the community, the difference in reported weight loss is mediated by a difference in (i) future lifetime in the community, or (ii) an increased engagement with the community we applied a so-called Sobel test [Sobel, 1986]. Based on the findings of our causal analyses and characterization, we build supervised learning models based on users real-world, interaction and linguistic characteristics to identify potential "at risk" users that could benefit from receiving social support.

Given the scarcity of studies on the effects of social feedback on continued engagement and weight loss in online health communities, we believe that these results contribute to a better understanding of the social dynamics underlying weight loss. We

[6] www.reddit.com
[7] www.reddit.com/r/loseit
[8] http://www.liwc.net

hope that the insights derived from our study lead to mechanisms further strengthening the social support online forums provide, especially to new users.

1.1 Thesis statement

The statement of this thesis is that online data generated by users can be effectively used to study causal effects and to infer users health condition. To show that, we characterize user behavior and social support in *loseit*. Then, we propose a causal framework to investigate the causal relationship between social support and (i) user retention, and (ii) weight loss. Finally, we build a supervised model to infer users potential obesity condition.

1.2 Contributions

Our main contributions can be summarized as follows.

- **Characterization and analysis of the properties of an online weight loss community**. We have characterized and analyzed properties of an online weight loss community. Among the results, we showed that users engage in helping each other and it is possible to identify social support in the content shared in the community. Then, we show that there is a positive correlation between social support manifested in the form of comments/voting score and weight change aspects.

- **A Framework for causal inference**. We proposed a general framework to study causal relationships in online communities. We applied our framework to study the effect of receiving social support in a weight loss community. In our findings, we observed an increase in return probability to the community for those users receiving feedback. Among returning users, those who had previously received social support report higher weight loss than the matched group who did not (46 lb vs. 37 lb). Somewhat surprisingly, only about 5% of the difference in reported weight loss appears to be mediated by an increased lifespan in the community, and this is not statistically significant even at $p = .1$. Instead, the *rate of weight loss* is the main difference between the two groups.

- **The use of social media data to infer users obesity condition**. From the results obtained in the previous stages, we identified interesting features that could be used to described the process of users weight loss, and built statistical models to infer the user's weight condition. These models presented a precision of up to 66% and can be used to identify "at risk" users and aid health professionals to better target those users.

- **Contributions to the medical area**. Our findings might aid health practitioners to design early warning systems or effective online health interventions strategies that can be incorporated to social media platforms and lead to more effective treatment. Further investigation is needed, but the success of these systems may provide great benefit to patients, for example, by integrating recommendation systems that can help users to make important decisions, such as choosing the right type of diet or exercise for their obesity condition.

1.3 Work Organization

The rest of this work is organized as follows. In Chapter 2, we discuss related studies focusing on the importance of social support for positive health outcomes, online support forums, studies of health communities in Reddit and causal inference with observational data. Chapter 3.2 presents the weight loss community investigated in this thesis and characterization of users behavior patterns in *loseit*. In Chapter 4, we present a characterization of mechanisms present in *loseit* to provide social support. Then, in Chapter 5 we describe the proposed framework for causal inference. Next, in Chapter 6, we illustrate how to build a supervised model to infer users obesity condition. Finally, in Chapter 7, we present conclusions, discuss limitations and future directions.

Chapter 2

Literature review

In their book “Social Media for Nurses: Educating Practitioners and Patients in a Networked World”, Nelson et al. [2012] predicted that social media would become a “large part of healthcare” due to the number of people with chronic illnesses who cannot access care anywhere other than in a virtual setting. Thus, population and health care providers need to be prepared to understand the benefits and challenges of social media and health.

In this chapter, we discuss the importance of social support for positive health outcomes and highlight the studies performed in online communities. Review the works regarding health communities in Reddit, our social media of interest. We also report a few studies that tried to improve on correlational analyses by applying causal inference techniques using observational data such as those available in *loseit*, and end the chapter with a discussion on how these studies relate to ours.

2.1 On the Importance of Social Support for Positive Health Outcomes

Many scientific studies have already shown there is a strong relationship between social interactions, social support and health outcomes. Conditions such as smoking, depression, coronary disease, among others, may be controlled if individuals receive enough social support [Umberson and Montez, 2010], or may deteriorate due to social influence [Christakis and Fowler, 2007]. These studies were initially performed into real social networks. However, given the popularity of online social media, one question remained [Laranjo et al., 2014]: does the relationship between social support and health outcomes also hold in online environments?

Social support, which can also be referred to as peer support or peer-to-peer support, can be defined, according to Kim et al. [2008], as supportive information from people with social ties, e.g., friends, within a given network. Typically it occurs when a given person talks about their needs for support with others, which leads to a response expressing that one is loved and cared for. For example, imagine a situation where a person shares a stressful personal situation on social media. Then his friends may give some responses as a way to help the one who is stressed [Medeiros and Bosse, 2016].

Social support can be measured in terms of functional and structural support [Cohen et al., 2000]. Functional support is a subjective measure of perception of support, and depends on the individuals characteristics and expectations. Structural support refers to social integration, and the availability of others to help. Structural social support, which is easier to quantify than functional support, can be subdivided into various types, which include information, esteem, emotional, tangible, appraisal and network support. Informational support can be provided in the form of suggestions or advice, by referring a person to resources, providing new facts or skills or information that is useful for self-evaluation. Esteem support is about making the person feel better by complimenting or relieving him from guilty. Emotional support includes encouragement, understanding, sympathy and close relationships. Network support happens when someone gives you access to new people, show willingness and availability to be with the person in need. Finally, tangible support refers to freeing the person from some responsibility, by performing an activity for him or lending something.

Prior research has extensively examined the role of social support in enhancing mental and physical health. It has been argued that receiving social support may reduce the rate at which individuals engage in risky behaviors, prevent negative appraisals, and increase treatment adherence [Fontana et al., 1989]. Research has shown that conditions such as smoking [Mermelstein et al., 1986], depression [Grav et al., 2012], and coronary disease [Lett et al., 2005] may be controlled with social support. Also, face-to-face support groups are positively correlated with desirable outcomes, such as lower blood pressure, and lower blood sugar levels, resulting indirectly from adaptive coping skills and responses [Sullivan, 2003].

In the context of obesity, improvements in healthy eating and physical activity, as well more successful outcomes in weight reduction programs, have been demonstrated in studies considering offline support groups [Wang et al., 2014]. Other studies also showed that there is an independent role of social support on health-related quality of life among obese individuals [Herzer et al., 2011], and support group attendance after bariatric surgery was associated with greater post-operative weight loss [Hildebrandt, 1998].

Turner-McGrievy and Tate [2013] examined the types of online social support utilized in a behavioral weight loss intervention and relationship of posting and weight loss. They performed a sub-analysis of the content and number of posts to Twitter among participants (n = 47) randomized to a mobile, social media arm as part of a 6-month trial among overweight adults, examining weight loss, use of Twitter, and type of social support (informational, tangible assistance, esteem, network, and emotional support) exchanged. As they had a rich dataset of user features, such as age, sex, marital status and education, they correlated these attributes with weight loss, and concluded that user age and engagement are good predictors of weight loss.

2.2 Online Support Forums

The main implication of the studies mentioned in the last section is that developing social support networks may help people manage their health conditions. Online health communities can be used to develop large social support networks, to understand and to promote health behavior. People have always tried to answer health related questions by themselves, now the Internet has become an important resource. Previous studies suggest that 30% of U.S. Internet users have participated in medical or health-related groups. Advantages of online communities include access to many peers with the same health concerns, and convenient communication spanning geographic distances. These communities present an interesting contrast to similar offline groups, as they provide an environment where people are more likely to discuss problems that they do not feel comfortable discussing face-to-face. In addition such online health communities are known to foster well-being, a sense of control, self-confidence and social interactions [Johnson and Ambrose, 2006].

Still, little is known about how the support provided in these communities can help enhance positive health outcomes, such as weight loss. The literature offers little information about how members of large online health communities experience social support for weight loss. Most works concentrate on showing that social support exists in online weight-management communities, and qualify the types of support present online [Turner-McGrievy and Tate, 2013; Ballantine and Stephenson, 2011]. For example, Hwang et al. [2010], analysed SparkPeople[1], a network for weight loss management. Their study was based on questionnaires and involved inquiries about the user assessment of support in the network. The great majority of users (out of 31) said people were always available for them when they needed help. Major social support themes were

[1] http://www.sparkpeople.com

encouragement and motivation, followed by information and shared experiences. Members valued convenience, anonymity, and the non-judgemental interactions as unique characteristics of internet-mediated support. Ballantine and Stephenson [2011] performed a simple analysis using the Facebook webpage of the Weight Watchers to recruit users to answer survey questionnaires, including demographic and support related questions, who also agreed the Facebook page is a source of support for users going through the process of weight loss. Black et al. [2010], in turn, worked with FatSecret[2], but analyzed how support is provided by the community. They showed that FatSecret offers many ways of communication and consequently, support, and performed a qualitative evaluation of support looking at style of posts and group oriented interactions.

Based on the fact that support exists, a few studies have tried to correlate online engagement and support with the effectiveness of weight loss [Turner-McGrievy and Tate, 2013], or to show that a network of engaged users is linked to persistent sharing of fitness related information [Park et al., 2016]. The latter is of great importance as self-monitoring is one of the factors already shown to be associated with increased weight loss [Hutchesson et al., 2016]. Poncela-Casasnovas et al. [2015] examined the relationship between individual and social media variables, and weight loss in a large, international online weight management (OWM) program. They studied the online activity and weight change of 22,419 members during a six-month period, focusing especially on 2,033 members with at least one friend within the community. They investigated if social contagion and social support were related to weight loss and if a linear model can predict weight loss. The data was modeled as a friendship graph and several network measures were computed to create a predictive model of weight loss. Their linear model achieved 27% of adjust. They found social embeddedness strongly correlated to weight loss. However, none of these studies were able to isolate the effects of online social support on weight loss, as there are many other underlying factors and real-world variables difficult to account for.

Apart from studies concerning weight management and support, a few other have investigated other issues. Ma et al. [2010], for example, looked at obesity propagation in FatSecret, a social media target to people interested in food and diet. Their study was inspired by Christakis and Fowler [2007], who analysed a real social network of people and found out that obesity "spreads through social ties". They showed that, similarly to real social networks, social ties in FatSecret did play a role on the success of weight loss, and that the probability of success in losing weight increases when your friends become successful.

[2]`http://www.fatsecret.com`

Chou et al. [2014], in contrast, collected posts based on a set of keywords from Twitter, Facebook, Flickr, Youtube and forums, and performed an analysis of discourse across these social media. They used very broad terms for data collection, such as obesity, fat and overweight, which can add significant noise to the analysis. They concluded that the way people talk about obesity differs across social media, and discussed ways to tackled weight stigmatization.

2.3 Correlation of online data with real-world statistics

Understanding the relationships among environment, behavior and health is a core concern of public health researchers. While a number of recent studies have investigated the use of social media to track infectious diseases such as influenza [Culotta, 2010; Ginsberg et al., 2009] or dengue fever [Gomide et al., 2011], little work has been done to determine if other health concerns, including obesity, can be inferred.

Eichstaedt et al. [2015] analyze social media language to identify community-level psychological characteristics associated with mortality from atherosclerotic heart disease (AHD). They used data from 1,347 U.S. counties for which AHD mortality rates, county-level socioeconomic, demographic and health variables, and at least 50,000 tweeted words were available. Data include 148 million county mapped tweets across the 1,347 counties, where more than 88% of the US population lives. Regularized linear regression (ridge regression) over topics and dictionary based variables (relative frequencies of psychologically related words) were used to create a predictive model. They had three major findings: first, language expressed on Twitter revealed several community-level psychological characteristics that were significantly associated with heart-disease mortality risk. Second, the use of negative emotions (anger, disengagement and negative-relationship) are associated with increased risk, whereas positive-emotion and engagement were protective. Third, the predictive results suggested that the information in Twitter language fully accounts for the relevant information in 10 representative assessed demographic socioeconomic and health variables.

Chunara et al. [2013] used the Facebook advertisement tool to create a data sample of users within a specific range of age, location and interests. They were particularly interested in detecting levels of activity versus sedentarism and interest on tv, and correlate these interests with obesity. One of the most interesting experiments performed was to correlate data collected using the Facebook tool and real obesity statistics obtained from different regions of the US. However, they concluded that their data was not rich enough to correlate to obesity data in the real world.

Mejova et al. [2015] used combined data from Foursquare and Instagram to analyze the relationship between food places and obesity. They found correlation between obesity statistics and fast food places in a county. Their study revealed the relationship between small businesses and local foods with better dietary health, with such restaurants getting more attention in areas of lower obesity. To check the explanatory power of the Instagram data, they built a linear model to predict obesity level of a county, but the model presented an extremely low coefficient of determination (R^2) illustrating that without the deeper demographic knowledge, the picture metadata was not sufficient to infer county-wide obesity.

Culotta [2014a] performed a linguistic analysis of Twitter data in the 100 most populous counties in the US considering 27 health-related statistics. They investigated the predictive power of the linguistic attributes in Twitter and found a significant correlation on held-out data for 6 of the 27 statistics, including obesity, health insurance coverage, access to healthy foods, and teen birth rates. They also showed models that augment demographic variables (race, age and gender) with linguistic variables (from Twitter) are more accurate than models using demographic variables alone for 20 of the 27 health statistics. Their results suggest that Twitter data and demographic variables are complementary.

Paul and Dredze [2011] proposed a more general approach that discovers many different ailments and learns symptom and treatment associations from tweets. To create structured information, they developed a new topic model that organizes health terms into ailments, including associated symptoms and treatments. The model uses explicit knowledge of symptoms and treatments to separate coherent ailment groups from more general topics. They used a collection 11.7 million tweets, and obtained training data for a supervised classifier using Mechanical Turk. They applied a SVM classifier to the corpus, which produced 1.63 million health related tweets. Based on LDA, they developed a structured model that uses lists of symptoms and treatments to uncover diseases (ailments), such as flu, allergies or cancer. They showed their model to learn to group symptoms and treatments into latent ailments, as well as grouping remaining words into health related topics. They applied it to several tasks, including tracking illnesses over times, measuring behavioral risk factors, localizing illnesses by geographic region, and analyzing symptoms and medication usage. Their results showed quantitative correlations with public health data, suggesting that Twitter has broad applicability for public health research.

2.4 Studies of Health Communities in Reddit

Reddit has been used to study different health conditions under different perspectives, including social support. For instance, Choudhury and De [2014] analyzed the discourse of Reddit posts and comments looking for indicatives of depression in several mental health subreddits. They considered in their analysis the role of social support, and emphasized the importance of anonymity in the case of depression in online forums. They found users in these communities to share, explicitly, information about mental health issues providing evidence of the use of this non-conventional tool as a medium that fulfills certain needs. Even though users are not compensated for their actions, they observed the feedback manifested in the comments to be of surprisingly high quality, and ranges from emotional and instrumental, to information and prescriptive advice. This is an important contrast to social media like Twitter, where sharing health information is most times a broadcast or an emotional outburst.

Nevertheless, the great majority of Reddit health community studies perform more exploratory analysis of users behavior or determine correlations between language and health outcomes, with the few studies looking at causality mentioned in the next subsection. In the first category, Eschler et al. [2015] perform a content analysis in the posts of patients in different cancer stages in the subreddit r/cancer. They found patient and survivor participants showing different types of information and emotional needs according to their illness phase, and suggested some community reorganization to make information access easier for people with different types of cancer in different stages. In the second category, Tamersoy et al. [2015] focused on the health challenge of addiction, specifically addiction to tobacco or alcohol. They collected data from two smoking and drinking abstinence communities on Reddit (r/StopSmoking and r/StopDrinking) and performed an analysis to identify key linguistic and interaction characteristics of short-term and long-term abstainers, focusing on tobacco or alcohol. Then, they built a supervised learning framework to distinguish long-term abstinence from short-term abstinence.

2.5 Causal Inference with Observational Data.

In this section, we focus on the works that try to investigate causality using matching, a common nonparametric methodology for causal inference from observational data borrowed from the domain of medicine/political sciences/economics [Imai et al., 2008], which will be introduced in detail on Section 5.1. One of the key benefits of randomized experiments for estimating causal effects is that the treated and control groups are guaranteed to be only randomly different from one another on all background covariates, both observed and unobserved. Work on matching methods has examined how to replicate this as much as possible for observed covariates with observational (nonrandomized) data [Stuart, 2010].

This methodology is used by De Choudhury et al. [2016] to identify shifts to suicidal ideation in Reddit mental health communities. They identify changes in linguistic structures, interpersonal awareness, social interaction and content between those who proceed to discuss suicidal ideation in the future, from those who do not. Tsapeli and Musolesi [2015] discuss the design, implementation and evaluation of a generic quasi-experimental framework for conducting causation studies on human behavior from smartphone data. They demonstrate the effectiveness of the approach by investigating the causal impact of several factors such as exercise, social interactions and work on stress level. Finally, Olteanu et al. [2016] present a more general framework for applying the methodology to distill outcomes of personal experiences from social media timelines. One downside of the aforementioned works is that all of them rely on the propensity score matching (PSM) to study the causal effect, and as shown by King and Nielsen [2015], this method is sub-optimal and can even increase the imbalance in the data. Also they do not perform a mediation analysis, in mediation, the relationship between the treatment and the outcome is hypothesized to be an indirect effect that exists due to the influence of a third variable (the mediator).

However, we did not find any studies looking at causality between online variables and weight loss in online communities.

2.6 Discussion

In the last several years, online health communities presented as an important tool to disseminate health messages and connect people with health conditions, specially people living with chronic conditions. Integrating social media into health communication campaigns and activities allows health communicators to leverage social dynamics and networks to encourage participation, conversation and community, all of which can help spread key messages and influence health decision making.

Similarly to previous works, we perform a *loseit* user characterization through interactions and analysis of the content and sentiment of messages. With this characterization, we were able to identify the mechanism that the community uses to provide support to their users. The analyses are general enough to be performed in other health-related Reddit communities, as long as a clear health outcome is available.

Our work differs from previous ones in several ways. To the best of our knowledge, we performed the first large scale study of causal effects of receiving social support on an online health community and (i) return for another activity in the community, (ii) later report a higher weight loss, (iii) perform a bigger number of activities, iv) stay longer in the community, and v) have a higher weight loss rate. We proposed a general framework for causal studies. Instead of using PSM, our matching occurs directly on the variables and not on the propensity scores which, as shown by King and Nielsen [2015], is preferable. Further, besides accounting for balance checking we also perform a mediation analysis where we can check if part of the causal effect of receiving treatment is being mediated by some variable.

Chapter 3

The *loseit* Community

Reddit[1] is a social news website and forum. Its content is organized in communities by areas of interest called subreddits. In 2015 it had 8.7 million users from 186 countries writing 73.2 million posts and 725.9 million comments in 88,700 active subreddits[2].

In Reddit users can submit content, such as textual posts or direct links to other sites, both collectively referred to as *posts*. The community can then vote posted submissions up (*upvotes*) or down (downvotes) to organize the posts and determine their position on the site's pages. Information on downvotes and upvotes is, however, not exposed via the API, instead they expose the aggregated number of votes, referred to as voting score (number of upvotes minus number of downvotes). Users can also reply to posts with *comments*. This thesis focus on the popular weight loss subreddit *losei*t[3].

In *loseit*, a participating user can add a "badge" (the icon which appears next to usernames, see Figure 3.1) to their profile that indicates self-reported information about their weight loss progress in pounds and kilograms. The badges can be updated by the users at any time, regardless posting/commenting, and 28,603 have reported at least one badge. Apart from the badges, another way to report weight loss in *loseit* is using community conventions, such as SW (start weight), CW (current weight) and GW (goal weight) or through posting things such as "I've lost 10lbs" in the posts body/title and comments. However, we acknowledge that collecting this type of information is challenging, as it requires a lot effort to avoid low precision when dealing with information extraction from free text. An example of that would be a user using one of those conventions, but actually describing someone else's weight or a hypothetical

[1] `https://www.reddit.com`

[2] "Active" is determined by having 5 or more posts and comments during at least one week in 2015.

[3] loseit - Lose the Fat, `https://www.reddit.com/r/loseit/`.

situation. In Figure 3.2 we present an example of post using the conventions.

Figure 3.1: Examples of the users' weight loss badges on the *loseit* subreddit. The weight loss value is displayed in pounds and kilograms.

Figure 3.2: Example of user using the community conventions to self-report his personal information. The information in the post (highlighted in the box) means that this user is a 29 years old male and 6'7" tall (M/29/6'7"). His start weight is 393 lbs, goal weight of 250 lbs and current weight of 247 lbs (SW:393. GW:250. CW:247.).

3.1 Data Collection

The data used in our analysis covers five years (August 2010 to October 2014) and was crawled from Reddit using PRAW (Python Reddit API Wrapper), a Python package that allows simple access to Reddit's official API in November 2014. The data collected include posts, comments and other metadata (i.e., timestamp, user name, voting score and badge). In total, we obtained 70,949 posts and 922,245 comments. These data were generated by 107,886 unique users, of which 38,981 (36.1%) wrote at least one post and 101,003 (93.6%) at least one comment. Table 3.1 shows the mean, median and standard deviation (SD) for basic statistics of the dataset, including the length of posts and comments and the number of daily messages.

3.2 Ethical considerations

For our study, we used only publicly available data that users chose to post online. All analysis is done in aggregate and we do not post results for any individual. However,

	Mean	Median	SD
Posts per day	45.5	45	22.7
Comments per day	586.6	599	264.3
Voting score per post	35.7	6	126.7
Voting score per comments	3.1	2	11.4
Words per posts	89.3	64	95.8
Words per comments	25.5	14	35.3

Table 3.1: Basic statistics of *loseit* dataset.

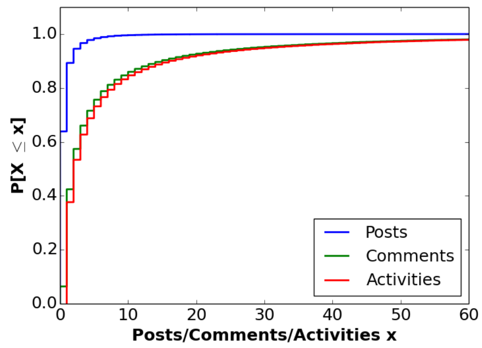
as is often the case with such data collection, users might not be aware of the fact that they are being studied by researchers. To at least partly alleviate such concerns, we reached out the moderators of *loseit* to inform them about our study. Their reaction was very positive ("Wow, I'm really looking forward to it") and they also pointed us to the community survey [denovosibi, 2016] that we had previously been unaware of. Once finalized, we will share our findings with the *loseit* community to encourage a positive atmosphere and, in particular, ensure a warm welcome of new members.

3.3 Users behavior in *loseit*

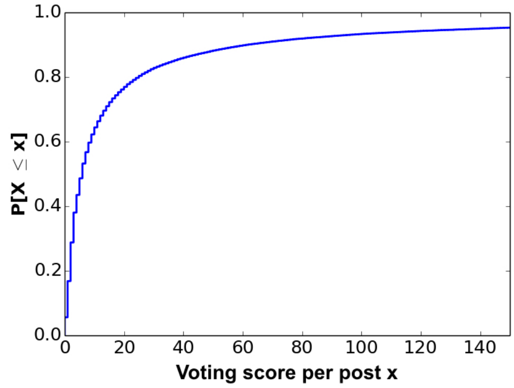
This section characterizes the behavior patterns of users in *loseit*. First we investigate users activities, and then build an interaction graph to investigate how users interact. Finally, we examine what are the main interest of *loseit* users expressed in the community. This characterization gives us insights that we use as input for our statistical model in Chapter 6.

3.3.1 Users activity

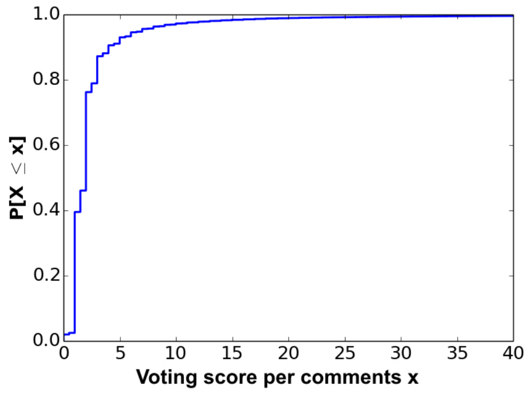
The dataset considered in this study encompasses $107,886$ active users (i.e. users with at least one post or comment in the community), while *loseit* has a total of 252,279 subscribed users, i.e., 57% (144,393) of users may be only readers (also known as lurkers [Tagarelli and Interdonato, 2013]). Figure 3.3a shows the cumulative distribution function (CDF) of the user distribution over posts, comments and activities (posts + comments) disregarding the tail of the function, as the number of posts of the remaining users (1%) grows up to 215 and the number of comments up to 3,353. Observe that around 36% of users posted at least once while 93% of users made at least one comment. In general, we observe that users comment more than create their own posts, indicating user engagement in helping others, and that the CDF of activities follows the same pattern of the CDF of comments. Approximately 7% of users made more



(a) Number of posts, comments and activities (post+comments).



(b) Voting score per post.



(c) Voting score per comment.

Figure 3.3: Cumulative Distribution Function (CDF) of *loseit* activities within a 5-year period.

than 20 comments, while the median number of comments is 2. For posts, 1.5% of users have made more than 5 contributions, suggesting that this group starts many possible discussions around weight loss problems. These users are the ones that remain active in the network for longest, with an average of 72 active weeks (and median of 62) against an average of 17 active weeks (and median of 1) for the remaining users. These users are also the ones with most comment, and 80% of them have at least 20 comments.

Figure 3.4 shows the distribution of the number of distinct active users in the community considering the 5-year period present in the dataset. Overall, the number of unique active users increases over time, with high values in January and July-August. High values in January may be due to the new year effect, as people usually set up new life goals in this period, and losing weight is commonly one of them. Note that the strong presence of users in the summer months may be a reflex of the season, where people are make more aware of their bodies.

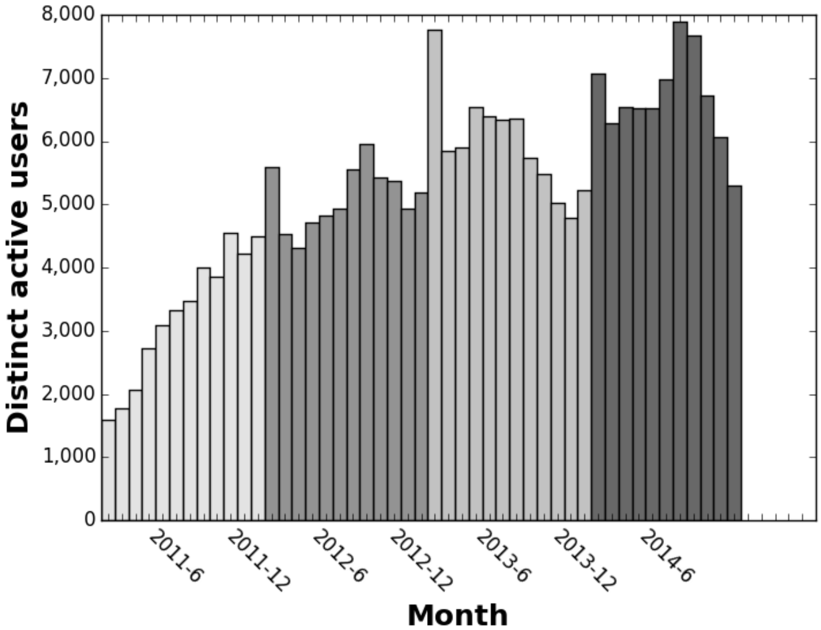


Figure 3.4: Number of distinct active users per month. Different colors correspond to different years.

Considering the date of the first and last activity (posts or comments) of a user during the five years of collected data, we calculated the user life time in the community in days. We know weight loss is a long and complicated process, and hence these numbers may be an indication of whether users are giving up during the process or keep going, although the process is much more complex than that. Figure 3.5a shows the CDF for user life time in days, where the life time of a user is defined as his initial activity date minus his final activity date plus one. Note that 51% of the users have a life time of one day, meaning they have a single activity in the network. Nevertheless, this does not mean they left the community, as they can be simply acting as lurkers. The complement of the CDF shows that 10% of users (which corresponds to 10,874 users) have a life time greater than one year (449 days), and 5% of almost two years (682 days). The user with the longest life time participated for 1,566 days (223 weeks), but notice that, in the way the user life time was defined, the user with the longest life time is not necessarily the oldest in the community, as older users may have been silent in the community for some time.

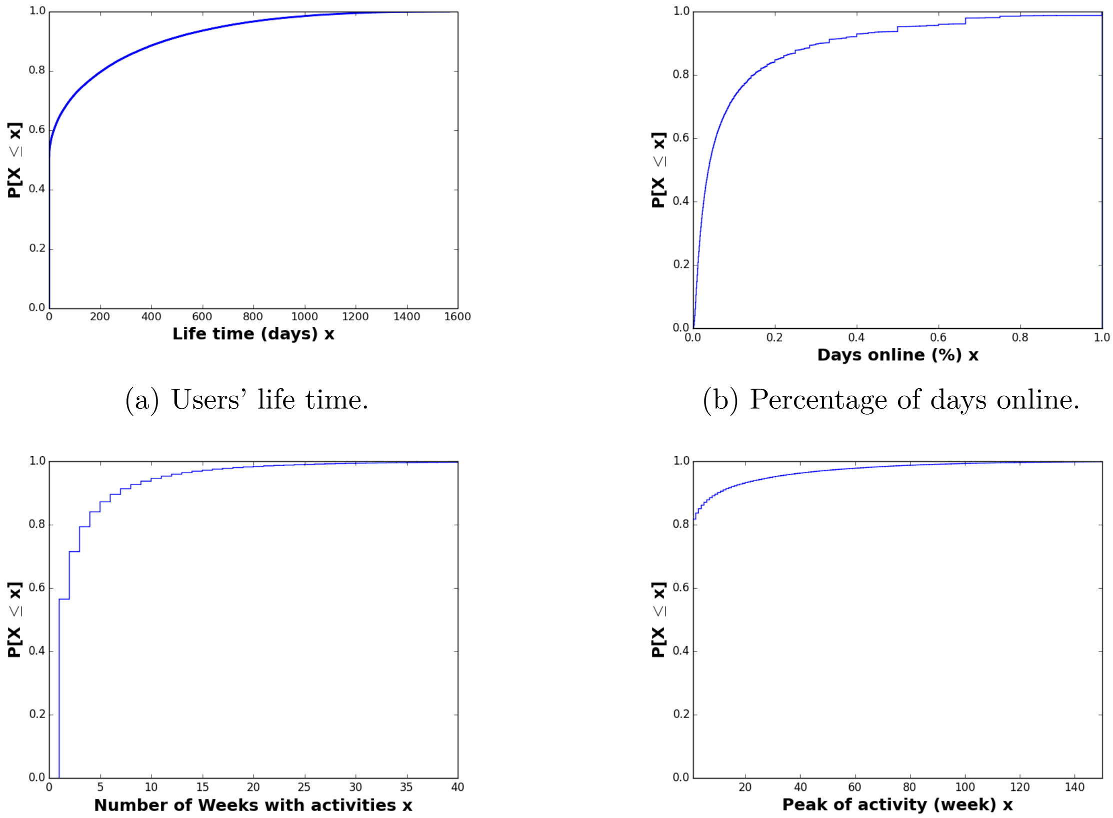
Figure 3.5b, in contrast, concentrates on the percentage of user active days during his life time, i.e., number of days he participated with a post or comment. In order to make this analysis fair, we removed from the dataset users with life times smaller than two days, which corresponded to 57,432 users (53% of the dataset). This was necessary because, the way we defined the life time of a user, these users would have an online time of 100%. Considering the remaining users, 5% were active in at least 50% of days. Among them, 59% have a life time shorter than four days. This is expected, as users with shorter life times can easily reach high levels of active time. Emphasis should be given to the 1% of users with active life time greater than 95 days. Together, these users are responsible for 2% of all community activities.

To investigate if there is a pattern in the users' activity in their lifetime, we created a weekly temporal series of uses' lifetime. Figure 3.5c presents the CDF of the number of weeks users were active, where a user is considered active if it has at least one participation in the corresponding week. Approximately 56% of users participated on a single week, while 5% of users in the end of the CDF participate in at least 10 weeks. Figure 3.5d shows the week number where the user had his peak of activity, considering week 1 as its first participation in the community. Not surprisingly, 82% of users have their highest number of activities in week 1. From these, 69% were active only this week. Among the users with highest activity level in week 1, 90% were active for at most 3 weeks. Also, 90% of users in the dataset had their highest activity level before week 10. In contrast, 4% of users had their most active week after week 40. These numbers can be related to shifts in the intention to lose weight by giving up and resuming the use of the social media [Raynor et al., 2008], but are more probable to be related to low user retention, which is a common problem even in successful online communities [Preece et al., 2004]

3.3.2 Users interactions

This section analyzes user interactions in *loseit*. For this analysis, we modeled the initial dataset as a weighted directed user interaction graph G, where each node represents a user and each directed edge represents a comment a user u made to a post from user v. Each edge is weighted according to the number of comments user u made on posts from user v. The interaction graph G has 106,276 nodes and 667,912 edges. Note that the number of nodes is different from the number of total users in the community, it happens because some users only made comments and those comments were made in posts of deleted users, thus they are not included in the graph.

Given a user u in G, the in-degree represents the number of users that interact with user u, i.e., the number of unique users who comment on posts from u. The out-degree, in turn, represents the number of interactions u had with the community, i.e., the number of posts u commented on. Figure 3.6 presents the CDF of the in- and out-



(a) Users' life time.

(b) Percentage of days online.

(c) Weeks of activity.

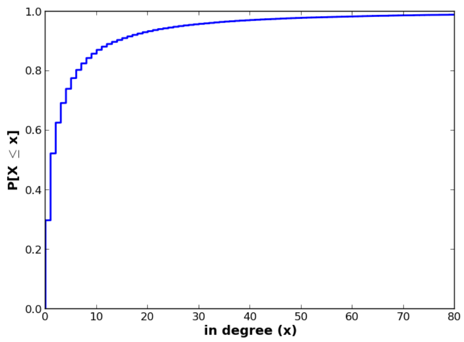
(d) Peaks of activity.

Figure 3.5: Cumulative Density Function (CDF) of the user life cycle in *loseit*.

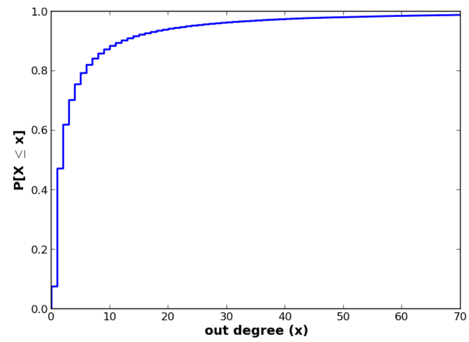
degrees of G. Regarding in-degree, around 70% of users received at least one comment in their posts, which indicates that, given the nature of this particular subreddit, data suffers less from a common effect, which is present in other popular subreddits [Gilbert, 2013]. The widespread underprovision effect, when too many people rely on others to contribute without doing so themselves, which leads posts to stay without any answer. Looking at the out-degree, we observe that the top 20% users in terms of number of comments commented more than 10 times. Moreover, 7.3% of users contribute with more than 20 comments.

Both in- and out-degrees are measures of frequency of interaction. However, we are also interested in their intensity. We analyzed the graph edge weights (Figure 3.6c), which represent the number of comments made/received by u. 87% of the users interacted only once with a user, while 13% made from two to 76 comments to the posts of a single user.

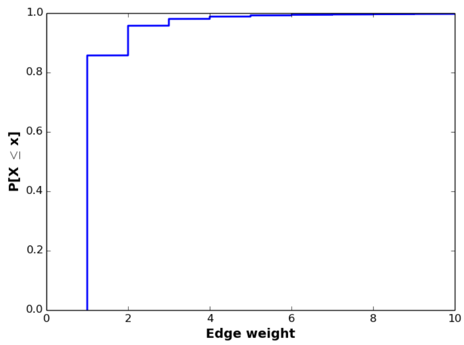
Table 3.2 shows a more complete set of metrics extracted from the interaction graph. The low transitivity and high diameter and node eccentricity corroborate the



(a) User in-degree.



(b) User out-degree.



(c) Edge weight.

Figure 3.6: Cumulative Density Function (CDF) of graph metrics.

fact that user interactions are based mostly on the content of the posts and comments, and do not depend on the users who generate them. This behavior is stimulated by Reddit, which does not have any content filter based on friendships and forbids “voting rings" (i.e., a group of people who vote for each others’ submissions in an organized manner) in its term of services. This may be also the reason for the low values of node closeness centrality.

Concerning the number of triangles, note that the high value of SD indicates it varies significantly from one node to another. This is a reflex of the high variation in the number of activities performed by different users. High active users tend to be present in a high number of triangles, as they have many interactions, while users with one or two activities are not likely to be included in any triangles.

In sum, the analysis of these metrics show that *loseit* interaction model is centered in the content of posts and comments, and not in users. This model helps new users to start their participation in the network, as it avoids closed groups. Moreover, all posts and comments have the same chance of having impact in the community, regardless of their author.

Table 3.2: Metrics from the user interaction graph

Graph metrics			
Metric	**Value**		
Diameter	13		
Transitivity	0.004		
Number of Cliques	486,252		
Metrics per node			
Metric	**Median**	**Mean**	**SD**
Eccentricity	9.000	8.822	0.54
Closeness Centrality	0.186	0.175	0.06
Clustering Coefficient	0.000	0.091	0.32
Number of Triangles	0.000	13.381	140.92

3.3.2.1 *Loseit* Subgroups

We just mentioned that Reddit centers its activities in content. But do subgroups of users emerge within the community after a first round of interactions according to common interests? These subgroups can be represented by connected components in the interaction graph, which were found using the Markov Clustering Algorithm (MCL) [van Dongen, 2008]. MCL is based on the idea that there will be as many links within a cluster as fewer links between clusters in a graph. This means that if you were to start at a node and then randomly travel to a connected node, you are more likely to stay within a cluster than travel between them. By doing random walks in the graph, it may be possible to discover where the flow tends to gather, and therefore, where clusters are. The random walks on a graph are calculated using Markov Chains. MCL found 7,732 strongly connected components, where 98% had less than 80 users and six had over 1,000 users. We focused our analysis on these six groups, and found an interesting pattern.

These six clusters have one user that acts as an attractor (or a network hub). When looking at how these communities first appeared, and if they had similar activity, interests or needs, we realized that all these groups emerged from an induced action, known as a *challenge*. A challenge has a predefined time spam and involves a set of users who set up goals for weight loss within its period. The winner of the challenge is the user who loses more absolute or relative weight. From these six groups, five were composed by people taking part of the same challenge, and the smallest number of comments received by the attractor was 660. The sixth group has similar characteristics, as it also poses a question and asks redditors to post their most interesting NSV (non-scale victory). An NSV is an achievement in the weight loss process not directly related to loosing weight, such as fitting in old pants or having to tight a belt. This particular attractor received 440 comments.

This finding is interesting because it shows that induced actions within the community are welcomed and successful, which means that health practitioners could used this pre-existent community to test their induced actions, and even recommend their patients to follow the group.

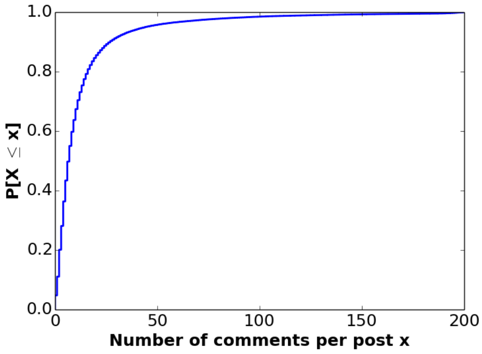
3.3.2.2 Community Discussion Patterns

Finally, we looked at users interaction from the point of view of discussions. User interactions can be modelled as discussion trees, where the root node corresponds to the post and child nodes to comments. We analysed the shape and size of these discussion trees. Figure 3.7a shows the CDF of the number of comments per post. Notice that the top 20% posts receive 16 or more comments, with highly commented posts reaching up to 200 comments. Highly commented posts are usually self-testimonies of people losing a lot of weight or weight loss challenge check-ins, which occur weekly during a challenge.

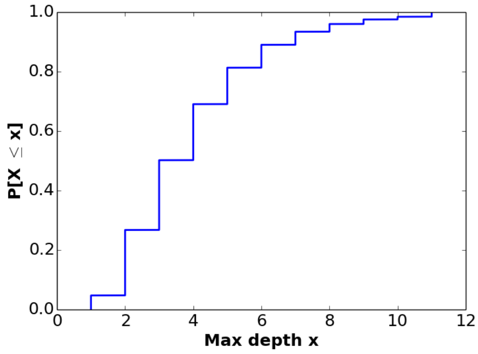
However, the total number of comments by itself only shows users interest in the post topic, and not necessarily represents interaction. A possible way to quantify the interaction level of a single post is to calculate the maximum number of comments written by a single user to that post. Figure 4.1a shows the maximum number of comments a single user posted to a thread. As observed, the top 20% received at least 10 messages from a single user, reaching a maximum number of 96 messages, indicating a high degree of interaction within discussion threads. Users do not only post/give an opinion, but also respond to other comments and the intensity of this behaviour pattern varies across different types of posts.

Another measure of interaction is the maximum depth of the discussions, as the intensity of the discussion might be concentrated in one of the branches due to, for example, a controversial comment that triggers even more reactions than the original post. Figure 3.7b shows the CDF of maximum depth of the discussions. Note that 50% of posts have at least 3 interactions, with threads reaching up to 11 levels. The deepest discussion threads are about emotional struggles that users face in real life, and their quest for support.

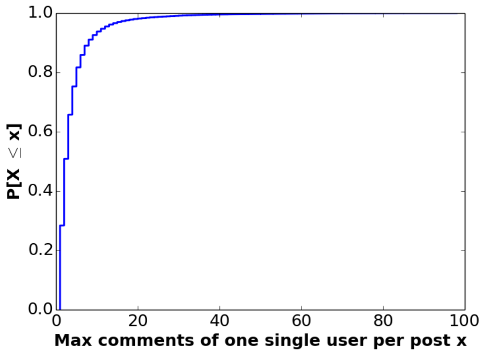
We also looked at the shapes of the discussion together with its content. We found that threads wide and shallow are normally about community challenges, the introduction of new users and NSV posts. Wide and deep threads, in contrast, are in



(a) Number of comments per post.



(b) Discussion threads depth.



(c) Maximum number of comments of a single user per post.

Figure 3.7: Cumulative Density Function (CDF) of discussion threads characteristics.

their majority about successful experiences losing large amounts of weight, and people telling how much their lives changed after their experience.

3.3.3 Users interests

The previous sections showed the *loseit* community users are active and engaged. This section analyses what are the most common user interests. We performed a semantic analysis of the messages with the Latent Dirichlet Allocation (LDA), an unsupervised statistical approach to document modeling that discovers latent semantic topics in large collections of text documents. LDA posits that words carry strong semantic information, and documents discussing similar topics will use a similar group of words. Latent topics are thus discovered by identifying groups of words in the corpus that frequently occur together within documents. In this way, LDA models documents as a random mixture probabilities over latent topics [Blei et al., 2003]. The required parameters for LDA – number of topics, number of iterations, α and β – were empirically defined as

Table 3.3: Most discussed topics into the loseit community.

	Posts		**Comments**	
Id	**Given Label**	**Top descriptive terms**	**Given Label**	**Top descriptive terms**
1	Healthy food	chicken,lunch,diner,breakfast,salad,fruit	Support	great,look,awesome,job,amazing,congrats
2	Clothing	size,fit,clothes,shirt,pant,new	Food	chicken,cheese,salad,vegies,cup,butter
3	Calories count	calorie,day,intake,eating,count,deficit	Clothing	size,fit,skin,clothes,loose,wear
4	Cardio workout	run,mile,minute,walk,bike,ran	Calories count	calorie,day,counting,eat,deficit,count
5	Weight loss apps	use,mfp,track,myfitnesspal,calorie,tracking	Weight loss apps	use,scale,track,using,mfp,weigh
6	Body changes	fat,body,muscle,stomach,skin,look	Weight control	week,pound,month,lost,two,last
7	Drinks	water,drink,soda,cut,day,diet	Self-esteem	see,look,picture,progress,difference,looking
8	Gym	cardio,workout,minute,training,lifting,muscle	Cardio workout	run,running,minute,mile,walking,walk
9	Support	loseit,thank,everyone,post,thanks,story	Support	feel,like,better,feeling,much,make
10	Unhealthy foods	food,eat,pizza,one,chocolate,sweet	Self-testimony	self,not,problem,control,issue,life

50, 2,000, 1 and 0.1. The set of posts and comments were analyzed separately, in order to understand how similar the topics in posts were to the topics in comments. Table 3.3 shows the top-10 topics describing these two sets of messages. For each topic, we manually associated a label according to the message given by the descriptive terms. For example, the most commented topic in posts refers to options of healthy food (topic 1) and types of unhealthy food (topic 10). The topics also include information about workout (topics 4 and 8) and the best ways to control calories intake (topics 3 and 5). The topics in the comments are not that different, although a strong presence of user support is frequently found in comments. However, most messages are a mixture of these topics all together. This may be challenging for LDA, as if most messages are a mixture of topics, telling them apart becomes a difficult task. For example, the comment below shows that clearly, as the user offers support while talking about exercise and diet:

"Good work dude, I am in the same boat you are .... I recently started strong lifts as well as the low carb high protein diet it's awesome and I rarely feel hungry. Keep it up dude and when you feel comfortable post some pics. When I hit my first milestone of 50lbs I will be posting some.... hopefully soon down to go"

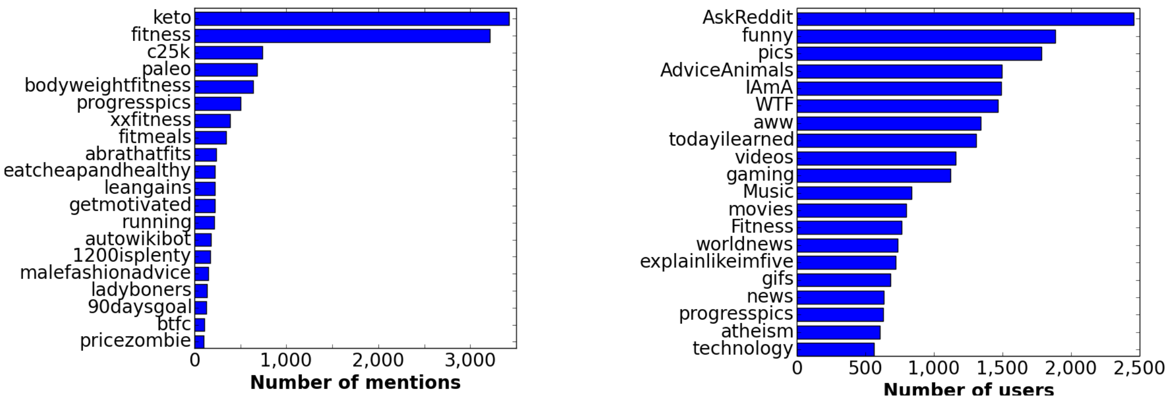
A lot of posts also pose questions to the community, which count on experience sharing. For example:

"Due to school placement and not currently having a car I eat subway for lunch 4-5 days a week. Anyways is this a healthy meal for everyday eating? If not what can I improve, have wheat toasted, Ham, American cheese, lettuce, tomatoes, chipotle sauce."

Finally, we investigated which other communities *loseit* users have an interest on. Two different analysis were performed: (i) looking at communities cited in users posts or comments (Figure 3.8a); (ii) looking at communities where active *loseit* users most frequently post (Figure 3.8b). The first analysis is within the context of weight loss, and ends up revealing communities related to *loseit*. The second analysis, in contrast, is much more general and shows other topics users losing weight pay attention to.

Knowing other user interests may be an important part of online clinical interventions, as it may be an indirect way of bringing people to the community.

Figure 3.8a shows the 20 most mentioned communities. Note that all of them have some relation to weight loss, including diet (e.g. keto, paleo), fitness (e.g. *c25k*, *running*) or motivation (e.g. *getmotivated*), except *pricezombie* and *autowikibot*. The former gathers people with sleeping problems looking for great sales and promotions online, while the latter is a bot that automatically answers users questions based on Wikipedia searches. The first may have some relation to the fact that obese people usually have sleeping problems. Among these communities, four appear as having a significant number of superposed users: fitness, progresspics, getmotivated and keto.



(a) Subreddits mentioned by users. (b) Subreddits with user participation.

Figure 3.8: List of common subreddits of interest of *loseit* users.

Considering other subreddits where *loseit* users are active, we found a miscellaneous of topics. The subreddit with most active users in common with *loseit* is *askReddit*, with almost 2,500 common users. This is the most general community in Reddit, where any question regarding any subject can be posted. Other communities include r/funny, pics, movies, music and games. One interesting idea is to use these interests to engage users into *loseit*, or even to start discussions on these subreddits about weight loss to recruit more users to *loseit*.

3.4 Discussion

Loseit is one of the most popular popular Reddit communities[4], thus is a natural choice while studying obesity in Reddit. The behaviour of users in loseit present the expected heavy tail trend observed in several social phenomena (see Figure 3.3a).

[4] http://redditlist.com

The way the community centers its attention on the content shared and the responsive nature to induced actions, may be of great value for the design of online interventions, which could greatly take advantage of the community reach, interactivity and cost-effectiveness.

These results illustrate the potential of loseit as a suitable environment to perform a large scale study of aspects related to weight change. Our characterization provides a better understanding of the properties of this particular community, showing its potential to drive various studies on obesity, as discussed in the next chapters.

Chapter 4

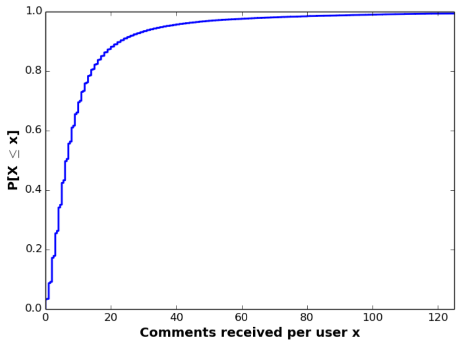
Understanding *Loseit* Mechanisms for Providing Social Support

In this chapter we show the mechanisms present in *loseit* to provide social support. In Section 4.1 we define quantitative measures of social support. We also characterize the various types of social social support present in *loseit* and investigate the content that drives social support in this community. Finally, in Section 4.2 we show the relation between social support and, (i) user retention, and (ii) weight loss.

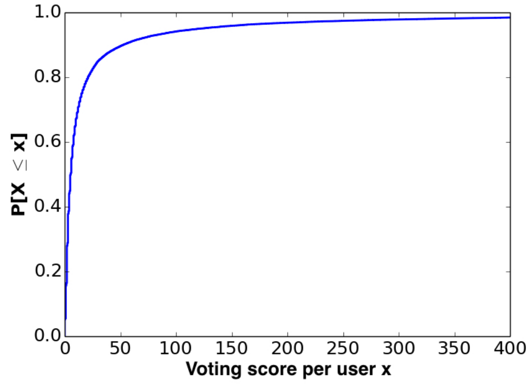
4.1 Measuring Social Support

In this section, we explore Reddit's content and features that indicate social support. Following previous literature [Choudhury and De, 2014] we also consider all types of social feedback as measures of social support. In Reddit there are two different ways in which a user can provide feedback on a post, i.e., voting and commenting. Thus, we define as our quantitative measures of social support: (i) number of comments received per post and (ii) voting score received per post. Figure 4.1 shows the CDF for these measures, which account for median values per user.

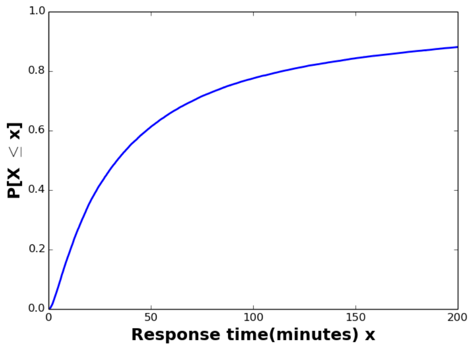
Let us start discussing the number of comments received. 40% of users with the smaller number of comments received up to 5 comments, while the next 40% received up to 14 comments. The top 20% user with most comments feedback received up to 200 comments. But most importantly 96% of posts received at least 1 answer, which indicate the responsive nature of the community. For median voting score received per user, we noticed a similar behaviour, 92% of users presented a voting score of at least one. In terms of posts, 95% presented voting score at least one. Also see Figure 4.1c for the time to receive an answer following post share. It illustrates the



(a) Comments received.



(b) Voting score per user.



(c) Average response time.

Figure 4.1: Cumulative Density Function (CDF) of median feedback received per user.

quick responsivity culture in the community we study, 60% of the posts receive an answer in at most 1 hour. Those characteristics are very important, as previous work has shown that frequent and faster community feedback reactions to a user's post contributed to user's perceived social support and ultimately alleviated loneliness [Seo et al., 2016].

Note that the CDFs presented do not necessarily represent all users, as for the ones that did not post, we were unable to observe voting score and comments received.

4.1.1 Characterization of social support

Next we performed a manual analysis to verify the different types of social support in the content shared in the community. Recall that, according to Cohen et al. [2000], social support is categorized into two types: structural and functional (Section 2.1).

Functional support is manifested in messages showing how much the users enjoy the community and feel better simply by reading their posts, as shown in Table 4.1.

Regarding structural support, four different types are likely be present in *loseit*.

Table 4.1: Functional Support Examples

1) Congratulations! Any time I feel low on motivation I'll just read some posts on here and I think it'll kick my ass back into the right direction :) 2) Pretty much I was looking for advice on how get started on the right path to being healthy. Boy am I glad I found this place. With all the support and the Loseit app, makes it seem like a really feasible goal.

Table 4.2: Structural Support Examples

Informational 1) You can try using greek yogurt instead - lots of protein. 2) If you want an easy way to lose weight and get started may I suggest the No Sugar No Flour diet. That's really all the explanation you need :) good luck!
Esteem–Emotional 1) Well done mate! just can't imagine the pride you must have felt. 2) I'm in the same boat, man. Unfortunately, family can make self-improvement very difficult. We're all here for you. You can definitely do it. You are a rock star.

Informational support is usually requested into posts that require advice. The easiest direct way to measure it, apart from reading the messages, is to count the number of URLs posted. Posts consisting exclusively of a URL (link post) are not frequent in *loseit*, and correspond to 3% of all posts. However, if we look at the URLs within posts and comments, we observe that 21% of posts (15,378) and 4% of comments (38,043) bring links.

Network support, which is related to user interactions and opportunities to meet new people, can be assessed by the measures described in Section 3.3.2. Low transitivity and high node eccentricity are signs of weak ties, which are very important to encourage sharing of information across different groups. Social systems that have more weak ties are more likely to be dynamic and innovative [Granovetter, 1983].

The two remaining types of structural support, esteem and emotional, are the most difficult to separate, as improving esteem usually includes encouragement and sympathy, and it is common to find comments with both types of support.

In order to better understand the type of support involved, three researchers familiar with the loseit content manually labeled 1,066 comments. Then we setup a system where human assessors (computer science students) were presented with instructions and definitions/examples of each type of social support then asked to classify the

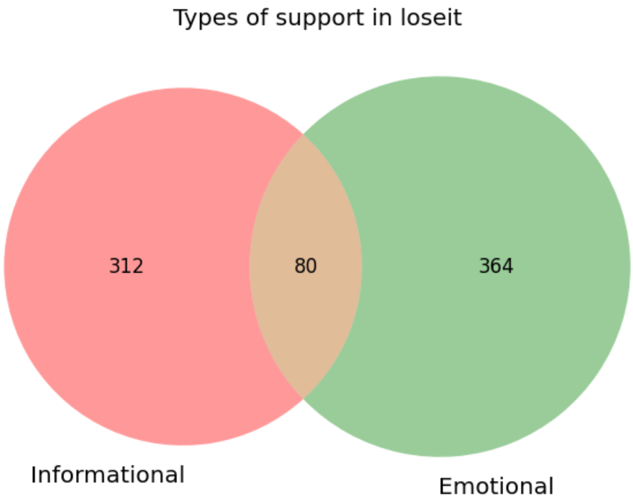


Figure 4.2: Categories of support in *loseit*

comments according to four classes: informational support, emotional/esteem support, both informational and emotional/esteem support or not sure. The sample size was chosen considering a random sample extracted from the population with a confidence interval of 3 and confidence level of 95%.

The results of the classification are shown in Figure 4.2. Considering all comments, 310 (29%) were classified as not sure, normally those are neutral messages discussing something about the original message. The rest of comments were classified as 29% (312) correspond to informational support, 34% correspond to emotional/esteem support, and 8% were classified as simultaneously providing both types of support. A few examples of these comments are presented in Table 4.2.

4.1.2 Self-diclosure in *loseit*

Prior studies suggested that self-disclosure in blogs or social media is beneficial to users in obtaining social support and establishing or maintaining friendship via positive interaction [Ko et al., 2013]. Hence, to understand which type of content users self-disclose in *loseit* posts and comments, we examine the common linguistic attributes manifested in their content. In Table 4.3, we present a list of the most frequent trigrams that appear in *loseit* posts titles and comments.

Analyzing the posts content, we have identified four main types of posts: (i) self-testimonies; (ii) advice request; (iii) victories and achievements; (iv) others, which in general are a mixture of the other types of posts. These categories are also reflected in the post titles. For example, *weight loss journey* posts are testimonies, while *trying lose weight*, *help losing weight* and *loseit need help* reflect advice requests. *First time since*

Table 4.3: Most popular trigrams in post titles and comments.

Post Titles	**Freq.**	**Comments**	**Freq.**
weight-loss-journey	257	keep-good-work	6,646
trying-lose-weight	207	trying-lose-weight	3,244
help-losing-weight	176	want-lose-weight	2,630
hey-r-loseit	153	calories-per-day	2,607
want-lose-weight	125	weight-loss-journey	1,898
need-help-losing	99	keep-great-work	1,878
need-lose-weight	96	congrats-weight-loss	1,655
need-help-getting	96	feel-free-add	1,385
loseit-need-help	96	body-fat-percentage	1,317
first-time-since	82	3-times-week	1,297

Table 4.4: Examples of posts from the categories previous defined.

Self-testimony Went to get a physical a couple weeks ago, and when I went back for a follow-up, they told me I was pre-diabetic and had high blood pressure. I'm only 25 (male)...I guess I need some encouragement. I've been lurking /r/loseit for more than a year now and have enjoyed hearing your success stories, but I guess now it's about time to start making my own. :/
Advice request I have been tracking calories and last night 1000 came from beer... how do you all manage to drink without taking in too many cals? I certainly do not drink like this often, but I do want suggestions for low calorie drinks. How does /r/loseit feel about drinking?
Victory/achievements Started at 263.1 Lbs, Currently at: 192.5 Lbs, Pounds Lost: 70.6, BMI Start: 34.7, BMI Now: 25.4. One year, 20 days to get this far. It's just eating better and moving more, that's it. Eating better can also mean eating the same stuff, but smaller portions.

posts are usually from category victories and achievements, and include, for example, "First time since Middle School I'm under 200 lbs!" or "Fit into an old pair of jeans for the first time in over four years!". Table 4.4 shows some examples of posts extracted from *loseit* and their respective categories.

Comments, in contrast, bring in their majority messages of encouragement - *keep good/great work* and *congrats weight loss* – or messages from users who are also taking measures to lose weight, reflecting emotional and esteem support. Note that the *feel free add* is commonly used to add that user into a different social media/app for them to keep talking and create closer ties.

4.2 Relation Between Social Support and Weight Loss

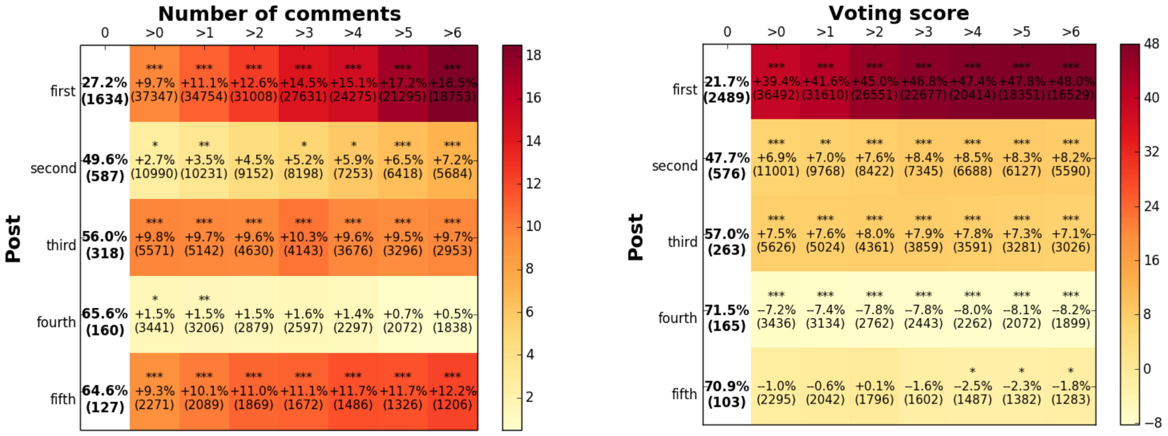
In the previous sections we showed that we can identify almost all types of social support in the *losejt* community. However, two questions arise from these results: if there is any relation between receiving social feedback in *loseit* and (i) staying in the community longer, which is associated to greater success in achieving their goals [Patrick et al., 2011] and (ii) the ultimate goal of losing weight.

To answer the first questions, we look for the probabilities of users to comeback to the community given the amount of social support ,i.e., comments, voting score, only positive comments and only negative comments they received. Specially, we look for effects of "diminishing returns", i.e., if receiving social feedback on a user's later posts gives less of a boost than receiving it earlier. We also look at the amount of social feedback received, as we hypothesize that any social feedback is better than none, but that receiving four comments is only marginally better than receiving three.

In Figures 4.3 and 4.4 we present the diminishing returns analysis for the amount of feedback received from the first until the fifth post. We show results for receiving comments, voting score, only positve comments and only negative comments as feedback. The first column shows the probability of a user's later return for activities in the community when receiving no feedback at all. The following columns present the relative difference (compared to receiving zero feedback) of receiving more than 0, more than 1, up to more than 6 feedback. For each of these columns, the return probability is also compared for its significance level against the baseline probability of receiving no feedback. The stars indicate the significance levels for a chi-square test for equality, with the number of asterisks corresponding to the p-values, *** for 0.1%, ** for 1% and * for 5%. For all the cells, the number of users considered appears between parentheses.

As is immediately evident from the color coding, receiving social feedback for the first post corresponds to by far the largest relative boost in return probability. Furthermore, after receiving feedback in the fifth post the further gains are largely irrelevant, in the case of voting score this gains stop even sooner, int third post. For some rows in the figure we also observe *negative* values indicating a *decrease* in return probability. We hypothesize that this is due to random noise induced by the smaller and smaller user sets for later and later posts. Surprisingly we observed an effect even when receiving negative comments.

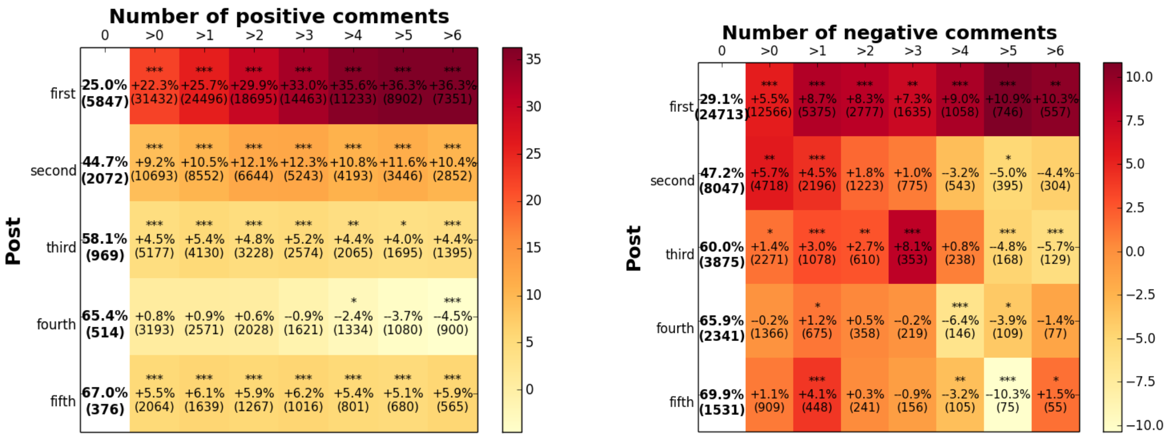
To verify whether there is a relation between receiving social support and weight loss, we computed the correlation between weight loss (extracted from the badge system) and (i) number of comments and (ii) the voting score. Both types of social



(a) Comments.

(b) Voting score.

Figure 4.3: Diminishing returns analysis for the number of comments and voting score received.



(a) Positive comments.

(b) Negative comments.

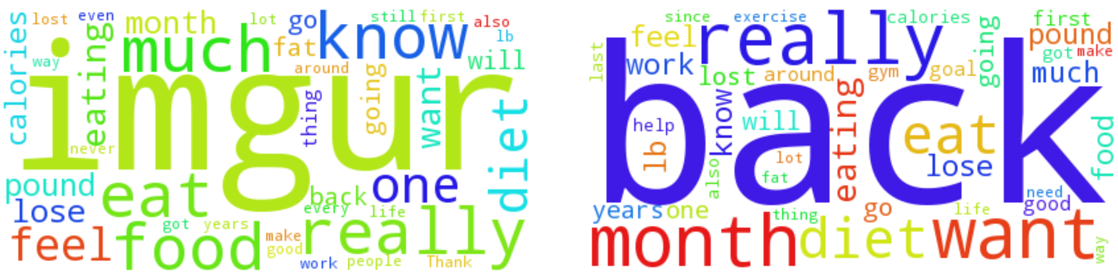
Figure 4.4: Diminishing returns analysis for the number of positive and negative comments.

feedback presented positive correlation with weight loss, 0.14 for number of comments received and 0.25 for voting score. We also very this relation for the feedback received only in the first posts. Again we observed a positive correlation for both types of feedback, 0.22 for number of comments and 0.28 for the voting score. All the correlations were statistically significant at $p = 0.001$. Although, we cannot claim causation, it seems that the more social support the users receive in the community, more weight they lose and this relation is stronger with the first post a user shares with the community.

4.2.1 Qualitative analysis of first posts

So far we have shown that (i) receiving social feedback seems to boost the return probability of a user, and that (ii) this boost shows "diminishing returns" and receiving social feedback on later posts is less impactful. One obvious question we have not yet answered though is: is there something special about the first post? Put differently, how should newcomers behave to boost their chances of receiving feedback?

We attempt to answer this question through qualitative analysis. We sort all first posts by their number of received comments. Then we compare the content for the top 10% posts (3,898 posts) to the content of the bottom 10% posts (3,898 posts). For better clarity we removed the most common words present in both groups (e.g. weight, year, day, time, now, week, calorie and started).



(a) Top 10% (b) Bottom 10%

Figure 4.5: Word clouds showing frequently used words in the most and least commented posts.

Examples of parts of highly-commented posts are "Super **fat** and **really** gross me `http://i.imgur.com/ANON.jpg`" and "**Feel** like sh*t so I **eat**, which makes me feel like sh*t, so I... and so on". Bottom posts include passages such as "I lost a lot of weight once but put it **back** on and much more. I'm finally ready to start again" and "I hit the **gym** and last year I hit 180. I've drifted **back** up to 190 since then".

Words present in the top posts are related to feelings, food and diet. Note that the most used word in this class is "imgur". This word is taken from an image sharing domain where people post pictures showing their body transformation.[1] This suggests that *loseit* users feel that the community is a safe enough environment for them to share images of themselves in swim wear. Interestingly, the corresponding loss of anonymity does not seem to play a huge role for these sharing users and they do not seem to be fearing discrimination. Indeed, previous work has showed that this type of online groups can be particularly appealing to people with disabilities or social stigma, as they report mistreatment and discrimination because of their conditionPuhl and Brownell

[1]Shared pictures are hosted on the website `http://imgur.com`.

[2001]. At the bottom end, there appear to be posts from users expressing their desire to get back on track and lose weight.

4.3 Discussion

Reddit has features that allow us to objectively measure social support in *loseit*, including all types of feedback the user gets from the community. We analyzed the content shared in the community and our results showed that almost all types of social feedback are present in *loseit* and we identified the main types of content that drives social support in the community. Based on these results we investigated if social support is related to (i) increase in returning probability and (ii) weight loss.

The boost in return probability is highest for social support received on the first post, rather than later posts. The relative, additional gains in probability level off after receiving feedback in the 4-5th post. Looking at the type of posts that are most likely to receive social feedback, we observe that posting images of oneself seems to be a good way to ensure a response from the community. We also observed a positive relation between weight loss and the number of social feedback received, which confirms previous research [Poncela-Casasnovas et al., 2015]. Given the scarcity of studies on the effects of social feedback on continued engagement and weight loss in online health communities, we believe that these results contribute to a better understanding of the social dynamics underlying weight loss. We also found a positive relation between feedback and weight loss reported via the community badge system.

Looking back at our analysis, we were surprised to see a positive effect of receiving negative comments. Though this might be partly explained by an inadequacy of using VADER for this automatic "supportive" vs. "discouraging" labeling, more likely the *loseit* Reddit community is "simply too nice". In our study we did not observe any outright bitter or hateful comments. An example negative comment is "You f***ing skinny son of a b*tch... good on you mate.", which despite the chosen terms is actually supportive.

Ultimately, we intend to go beyond correlations studying the causal effects of receiving social support on continued engagement in the weight loss community and actual weight loss. As a large number of community members regularly report their weight following community badge system, in Chapter 5 we propose a framework for causal studies to investigate the link between weigh loss and social support.

Chapter 5

Causal Analyses

In this chapter, we discuss our matching framework to investigate a potential causal effect of receiving social support in the Reddit *loseit* community over (i) probability of the user return to perform another activity, and (ii) users weight loss. We start by defining matching. Then, we discuss some limitations and design decisions. Later, we present our causal framework, which is composed of a matching approach and a mediation test.

5.1 Matching

Matching is a nonparametric method of controlling for the confounding influence of pretreatment control variables (also known as confounding or covariates) in observational data. The key goal of matching is to *prune observations* from the data so that the remaining data have better balance between the treated and control groups, meaning that the empirical distributions of the covariates in the groups are more similar and model dependency is reduced [King et al., 2014].

Without matching we may have imbalance, for example, a generally optimistic user might write a first post with a more positive tone than a more pessimistic counterpart. Let us imagine that, in response to their posts, the former user receives lots of support and the latter receives none. Now let us further imagine that the former user returns for more activity on the subreddit later, whereas the latter user is never to be seen again. The question then arises whether the support received "caused" the former user to return or, rather, whether that user was at a higher disposition to return anyway and the social support received was a mere correlate. Here the tone of the posts, an important covariate, is imbalanced and is generally more positive in the treated group (= those with social support) than in the control group (= those without social support). Matching approaches are applied in such scenarios to remove the relationship between the covariates and the supposed causal variable by reducing the imbalance.

In the simplest case, matching is applied to settings of a dependent outcome variable Y_i, a treatment variable $T_i(1 = treated, 0 = control)$ and a set of pretreatment covariates X_i [Rubin, 1976]. We want to observe the treatment effect for the treated observation i (TE_i), which is define as the value of i when i receives the treatment minus the value of i when it does not receive the treatment (control).

$$TE_i = Y_i(1) - Y_i(0) = observed - unobserved \tag{5.1}$$

Obviously if i is treated, we can not also observe i when it does not receive the treatment. Hence, matching estimates $Y_i(0)$ with a $Y_j(0)$, where j is similar to i. In the best case, each i is *matched* to a j with the exact same values for all the control variables. In practice, "similar enough" observations are being matched.

After applying a matching method, it is mandatory to check whether the treated and control groups are balanced. If balance has not been achieved, the matching method has to be revised. Finally, if any confounding bias has been sufficiently eliminated, the treatment effect can be estimated by comparing the effect variable Y of the matched treated and control units.

Matching can be viewed as trying to find hidden randomized experiments inside observational data. The most commonly used matching method is Propensity Score Matching (PSM) [King and Nielsen, 2015], which aims to approximate a complete randomized experiment. PSM first builds a model to predict the probability of a particular user to receive the treatment. Users are then matched according to their probability of receiving the treatment. However, recently King and Nielsen [2015] showed that this method is suboptimal and that PSM can, under certain circumstances, even increase the bias in the data.

5.2 Causal Inference Framework

In this section we introduce the proposed framework for causal inference. Figure 5.1 shows the diagram with the steps of the framework. We end this section discussing some limitations.

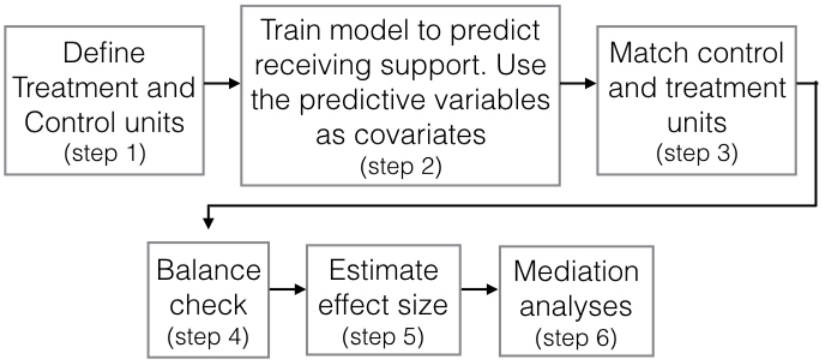


Figure 5.1: Causal framework diagram.

5.2.1 Treatment and control definition (Step 1)

The first step of our framework is choosing the appropriate definition of treatment. Here treatment is defined as receiving social support in the form of comments from the *loseit* community. Given the fact that 96% of the first posts received at least one comment, we experimented with different definitions of treatment to avoid that the number of remaining matched observations becomes too small to draw any statistically significant conclusions. After running the experiments, we defined the treatment group as those users who received at least 4 comments on their first post in *loseit*. The control group consists of all users who received 3 or less comments on their first post. With this definition, we can guarantee (i) statistical significance of our findings, and (ii) the balance (see Step 4) between the two groups after performing matching.

5.2.2 Statistical method for covariates selection (Step 2)

Choosing appropriate confounding variables is an important step in matching methods. Ideally, conditional on the observed covariates, there should be no observed differences between the treatment and control groups. To satisfy the assumption of ignorability, it is important to include in the matching procedure all variables known to be related to the treatment assignment. Generally poor performance is found by methods using a relatively small set of "predictors of convenience", such as gender only. Oppositely, including variables that are actually unassociated with the outcome can yield slight increases in variance. Commonly the confoundings' choice is based on previous research and scientific understanding, which can yield researcher discretion and consequently bias [Stuart, 2010].

Here instead, we propose to use a statistical model to select the most important covariates. We first examine whether attributes of the content of posts, are predictive of receiving treatment. We model a prediction task with the data being split into two categories, the ones that received treatment and the ones that did not. Then we use the variables that remained in the final model as the confoundings in the matching approach (see Step 3). Below, we detail the variables involved in this model.

The model's response variable represents if the user received or not the treatment.

Explanatory variables In our case, the definition of the predictive variables was motivated by the hypothesis that posts with similar content have a similar probability of receiving feedback. Since user attributes like demographics or profile images are not available in Reddit, and hence the sole focus on the post's content is natural. We used a topical representation of the first post's content (title + body) extracted by Latent Dirichlet Allocation (LDA) [Blei et al., 2003], counts of the various semantic categories provided by the psycholinguistic Linguistic Inquiry and Word Count (LIWC[1]), sentiment analysis computed with Valence Aware Dictionary for sEntiment Reasoning (VADER) [Hutto and Gilbert, 2014], counts of question-centric words (what, where, when, which, who, whose, why, how) and the length of a post (number of whitespace delimited words), a total of 98 variables. The required parameters for LDA – number of topics, number of iterations, α and β – were empirically defined as 20, 2,000, 0.4 and 0.1. LIWC establishes strong links between linguistic patterns and personality or psychological state. The rationale to get question words is to understand to what extent posts on weight loss seek explicit feedback or suggestions from the Reddit community.

Statistical model - We used a logistic regression with LASSO as our prediction method. Logistic regression is well-suited to handle binary dependent variables, while LASSO is a method that performs both variable selection and regularization in order to enhance the prediction accuracy and interpretability of the statistical model. To assess the quality of the model produced, we (i) computed the mean AUC (Area Under the Curve) over the 10-fold cross validation setting, and (ii) performed a qualitative analysis of similar posts, with the similarity computed using the features that remained in the model. As the goal of the matching is to pair posts that "look similar to a human reader", this qualitative analysis is important to understand whether features have succeeded to adequately identify similar texts.

The highest cross-validated AUC of 0.62 was obtained for 37 variables. However, based on our qualitative analysis and in order to further reduce the dimensionality of the space used for matching, we kept pushing the regularization and at the end, we chose features from a model with slightly lower AUC (0.61) but which used only 20

[1] www.liwc.net

Table 5.1: Final model to predict treatment

Feature	Coefficient
negative emotions	0.0128370430
anger	0.0058555923
sexual	0.0004610124
reward	-0.0002951086
work	-0.0048173004
sentiment	-0.4714861297
topic0	-1.2638822452
topic1	2.8146153201
topic2	-2.9367005663
topic4	-1.0781352275
topic6	1.7857991692
topic7	1.5670713048
topic8	1.1654609263
topic10	-0.1308221794
topic14	-0.8769023251
topic15	2.4472328942
topic16	-1.3819845665
topic17	0.3231376434
topic18	-0.6053494393
topic19	1.2732195551

variables. In terms of quantitative analysis, both models fared similarly without any noticeable difference.

The 20 variables that "survived" the shrinkage were: 5 LIWC categories (negative emotions, anger, sexual, reward and work), sentiment and 14 LDA topics (see Table 5.1). Additionally we use the coefficient values as covariates "weights" in the similarity computation in Step 3. This choice was motivated by the fact that the regression coefficients have two desirable properties. The first one is a scale normalization property, where something measured in, say, kilometers would have a larger coefficient than the same property measured in meters. This normalization is crucial for computing meaningful similarities in a metric space. Second, they reflect an importance of the predictive variable in relation to the response variable. This means that variables with more effect on receiving feedback will be given more importance on the post similarity.

5.2.3 Matching approach (Step 3)

Here, we apply a matching distance approach (MDA) [Rubin and Stuart, 2006], which aims to approximate a fully blocked experiment. In experimental research, a fully blocked experiment dominates the complete randomized experiment for: imbalance, model dependence, power, efficiency, bias, research costs and robustness [Imai et al., 2008]. As we mentioned before PSM tries to approximate complete randomized experiment, which means that it has lower standards than MDA.

In MDA, we apply cosine distance among the observations based on their covariates. Treated units are matched to their nearest control, assuming they pass a predefined similarity threshold, a.k.a. caliper. Ideally this similarity threshold should be as close as possible to 1, barring constraints related to data sparsity. To find an appropriate value, we gradually increase the value, starting from 0.9, until we are able to observe three conditions: (i) the matched posts are similar enough (based on a qualitative analysis), (ii) treatment and control groups are balanced (see Step 4), and (iii) results are statistically significant. We allow one-to-many matches, i.e., we match with replacement.

Pruning the unmatched observations makes the control variables matter less. In other words, it breaks the link between the confounding and the treatment variable, consequently reducing the imbalance, model dependence, research discretion and bias.

5.2.4 Balance check (Step 4)

One necessary condition for a successful application of a matching methodology is a balance of the covariates. If, say, one LDA topic was more strongly pronounced in the treatment group than in the control group then this imbalance, rather than any causal effect, could lead to an apparent treatment effect. To assess if the treatment and control groups are sufficiently balanced after the matching, we check the standardized mean difference [Austin, 2011] for each confounding variable c. For a continuous covariate, the standardized mean difference is defined as:

$$d_c = \frac{(\bar{x}_{treatment} - \bar{x}_{control})}{\sqrt{\frac{s^2_{treatment} + s^2_{control}}{2}}} \tag{5.2}$$

where $\bar{x}_{treatment}$ and $\bar{x}_{control}$ denote the mean of the covariate in the treatment and control groups, respectively. $s^2_{treatment}$ and $s^2_{control}$ denote the corresponding sample variances.

The standardized difference compares the difference in means in units of the pooled standard deviation. It is not influenced by sample size and allows for the comparison of the relative balance. The remaining bias from a confounding variable c is considered to be insignificant if d_c is smaller than 0.1 [Mimno et al., 2011].

5.2.5 Effect size estimation (Step 5)

After showing that any confounding bias has been sufficiently eliminated, we can estimate the effect of treatment on the matched treated and control units. Here for a given matching of treated and control units, we compute the estimated average treatment effect (EATE). Where N correspond to the number of units in treatment and control groups.

$$EATE = \frac{\sum_{i=1,j=1}^{N} \frac{(Y_i(1)-Y_j(0))*100}{Y_j(0)}}{N} \tag{5.3}$$

5.2.6 Mediation Analysis (Step 6)

Though matching methods can shed light on whether a change in the treatment condition T likely causes a change in the dependent variable Y, matching methods do not provide insights into whether (i) this causal relationship is "direct", or whether (ii) it is being mediated by another variable M. In our case, receiving social support might lead to an increased engagement with the community which, in turn, is responsible for an increase in weight loss success. Thus the weight loss success would be mediated by an increased engagement with the community.

Mediation analysis is the process of determining whether or not variables acting as an in-between step, called mediators, are present when looking at the relationship between an independent variable T (here the treatment condition) and a dependent outcome variable Y (here weight loss). As a result, when the mediator is included in an analysis model with the independent variable, the effect of the independent variable is reduced and the effect of the mediator remains significant [Boone, 2012].

To verify if any variable plays the role of a mediator and its significance in the relationship of social support and weight loss, we apply the Sobel test [Sobel, 1986]. The Sobel test assesses the statistical significance of the indirect effect (path $\alpha\beta$, in Figure 5.2). In other words it assesses if the significant part of the relationship of X and Y is explained by the mediator.

The procedure to perform the Sobel test is as follow: (i) show that the independent variable T is correlated with the outcome variable Y. Use Y as the response variable in a regression equation and T as a predictor (estimate and test path C). This step establishes that there is an effect that may be mediated, (ii) show that T is correlated with the mediator variable M. Use M as the response variable in the regression equation and T as a predictor (estimate and test path α). This step essentially involves treating the mediator as if it were an outcome variable, and, (iii) show that the

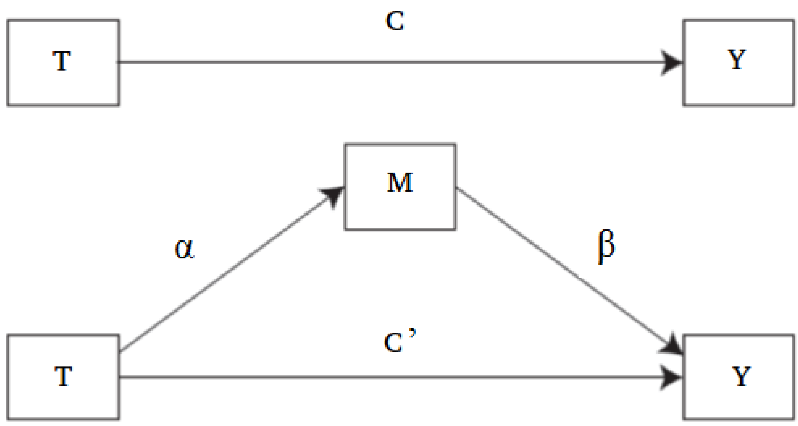


Figure 5.2: Basic mediation diagram, in the top the direct relationship between the treatment condition T and the outcome Y. In the bottom the the relation after the introduction of a mediator.

mediator affects the outcome variable. Use Y as the response variable in a regression equation and T and M as predictors (estimate and test path β). It is not sufficient just to correlate the mediator with the outcome; the mediator and the outcome may be correlated because they are both caused by the treatment. Thus, the treatment variable must be controlled in establishing the effect of the mediator on the outcome. If there is a mediation, as a result when the mediator is included in the regression analysis model with the independent variable, the effect of the response variable is reduced and the effect of the mediator remains significant.

Note that such analysis works naturally with matching as a preprocessing step: the matching reduces imbalance between the treated and control groups in terms of the covariates used for matching. Hence, the remaining unpruned observations are similar except for the treatment condition, and the treatment condition can be used as an independent variable in the Sobel test.

5.2.7 Limitations and Design decisions

Let us start discussing some limitations of matching for causal inference that impact in the design of our framework. The most obvious limitation is that matching has the assumption of ignorability, i.e., we can only match on those confounding factors that we (i) can observe (and think of), and (ii) have good measurements. Thus it is really important to know the context of study and have a good method to select the confounding variables. If you can identify relevant covariates so that ignorability is reasonable, you can assess causality by controlling for the covariates statistically. We addressed this issue in the Step 2 (Subsection 5.2.2) of our framework.

A second type of limitation is sample selection bias [Spirtes et al., 1995], which is a fundamental limitation of using social media for public health applications. For example, Twitter users are not a representative sample of the population, tending to skew towards young, urban, minority individuals [Culotta, 2014b]. Besides this general selection bias while working with social media data, here we have "our own" version of selection bias, as we only include in our weight loss analysis users that returned to the community for a second activity and disclosed their weight loss. In Chapter 6 we will discuss a model that could be used to adjust for this type of bias.

Last, matching is not a solution for endogeneity problems. When working with observational data, we do not have the opportunity to manipulate the explanatory variables; we just observe them. One consequence of this lack of control is endogeneity, i.e., the relationship between the independent variable x (e.g. treatment) and dependent variable y (e.g. outcome), where the influential relationship flows in both ways, putting in other words, x "causes" y and y "causes" x. With true experimental manipulation, the direction of causality is unambiguous. But for many areas of qualitative and quantitative research, endogeneity is a common and serious problem [King et al., 1994].

Our first idea was to investigate the effect of the social support received in all posts shared by a user in the community. However, we identified an endogeneity problem. For example, a user shares his first post with the community and receives a lot of support. In the occasion of a second post, he describes improvements in his health condition and again receives a lot of support in the form of comments from the community. It can occur the case that the support (x) received in the first post is "causing" an effect on the user's condition and consequently "affecting" the content of the second post (y). The content in the second post, in turn, might influence the subsequent support (x) received, which characterizes an endogenous situation. Hence, in order to filter out the degree y influences x and prevent our study of endogeneity problems, we decided to limit our analysis to investigate the causal effect of receiving social support only in the first post a user shares with the community. Another way of looking at our study is as an investigation of the effect of receiving a "warm welcome" in a weight loss community.

5.3 Results

In this section we present the results of the causal inference analysis we conducted to measure the effect of receiving social support for the first post a user shares with the community. In Figure 5.5 the stars indicate the significance levels for a permutation test, with the number of asterisks corresponding to the p-values, *** for 0.1%, ** for 1%, and * for 5%. In a permutation test, the labels (C)ontrol and (T)reatment are repeatedly randomly shuffled and for each (fake) control-treatment assignment the effect size is measured. The significance level indicates the fraction of permutations that lead to an effect size bigger than the one actually observed. One advantage of permutation test is that it exists for any test statistic, regardless of whether or not its distribution is known [Hesterberg et al., 2005].

To perform our causal analysis we created two sets of users.

Group 1 (G1). We extracted the list of unique users whose first recorded activity in the community was a textual post (self post), rather than a comment or a post consisting exclusively of a URL (link post). This gave us a set of 25,647 users who had no public activity in the community prior to their post. We use this user set to study the effect of receiving comments on this post on the probability to return later.

Group 2 (G2). From users in Group 1, we extracted the list of unique users that both (i) returned to the community later to comment or post and (ii) also had badge information indicating weight loss. This left us with a set of 6,143 users. Figure 5.3 shows the weight loss distribution displayed in the badges. We use this set of returning users to study the effect of receiving comments on their initial post on the weight loss they achieve.

User engagement

We start by analyzing the effect of receiving social support on the probability of a user to return to the community to comment or post again. For this analysis we used the 25,647 users present in **Group 1**. Among those users, 18,000 received the treatment (at least 4 comments) and 7,647 did not (control).

We applied our one-to-many matching approach with a similarity threshold of 0.965 to ensure that for **Group 1** our method was matching similar enough posts and balancing the groups (see Figure 5.4). The matching produced 14,570 similar pairs (14,570 unique treatment and 5,279 unique control users). Our results indicate that receiving social support increases the relative probability of a user returning to the community by roughly 66% (see Figure 5.5, red bar). This analysis is statistically

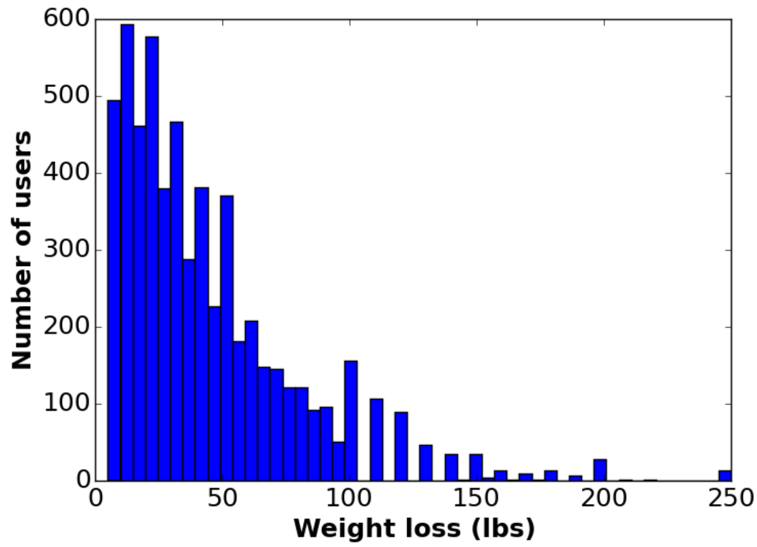


Figure 5.3: Users' weight loss distribution displayed in the badges.

significant at 0.1%. These results provide evidence for how important social support is to increase user engagement, which is associated with better chances of obtain success in weight loss programs [Patrick et al., 2011]. Note that this effect sizes is bigger than the one reported in a similar study Cunha et al. [2016]. There the authors used a different definition of control and treatment group based on ranking the posts by the number of comments received than getting "top 40%" vs. "bottom 40%" comments, rather than our "at least 4" vs. "at most 3".

To show that the matching approach indeed matches similar posts, we present in Table 5.2 parts of a pair of posts matched according to our approach, this pair had a cosine similarity of 0.97.

Weight loss

Next we investigated the effect of receiving social support on weight loss. For this analysis we used the set of 6,143 users in **Group 2**, among those users 4,657 received at least 4 comments (treatment group) and 1,486 did not (control group). Here, to ensure similar enough posts and balanced groups in the matching for **Group 2**, we used a similarity threshold of 0.955. Figure 5.5 (green bar) shows that receiving social support in the first post leads to a relative increase in the achieved weight loss of 26%, or an absolute mean difference of 9 lbs. The observed effect size is statistically significant and after applying the balance check (see Figure 5.6) we confirmed that all covariates were balanced between the control and treatment groups.

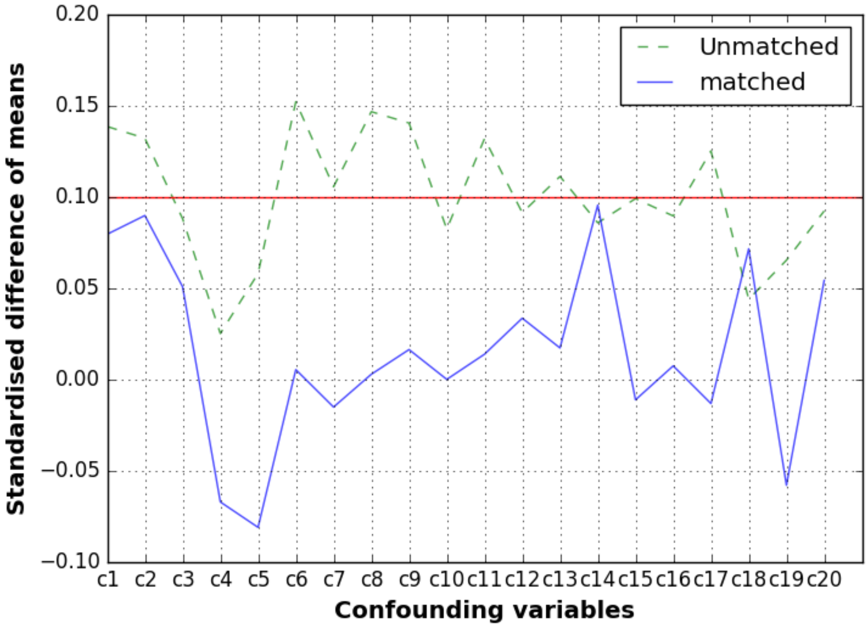


Figure 5.4: Standardized difference of means for each confounding variable (Group 1). Note that after the matching all the values are below 0.1, thus the groups are balanced.

Treatment: Hi guys, new here, I'm on a low carb and dairy (pale) diet, but recently i just went on a vacation, I ate..- Wrap with grilled chicken, lettuce, and a small amount of buffalo sauce, on whole grain wrap,Banana, Apple, Orange, and plain oatmeal - Tossed salad with grilled chicken. Is this healthy eating on the pale diet? I also did not exercise, but we did some walking around...
Control: Hey everyone, I'm new to loseit. I'm starting my first workout/diet routine ever. My doctor says I need to get my cholesterol under control. So far, I've been doing 30 minutes of cardio and taking care of my diet. I've cut out soda, beer, and red meat. I've switched to skim milk, olive oil, whole grain, and brown rice...

Table 5.2: Parts of similar posts matched.

However, note that the frequency with which people update their badges may interfere in this analysis. Maybe users who do not get comments do not update their badges as often, even if they lose as much weight as others. In other words, receiving social feedback might simply lead to more active "profile management" than to more weight loss. To test this alternative explanation, we computed the number of badge updates (every change in the badge information) for users in the treatment and control groups. Afterwards, we ran a permutation test to check if the two groups' badge

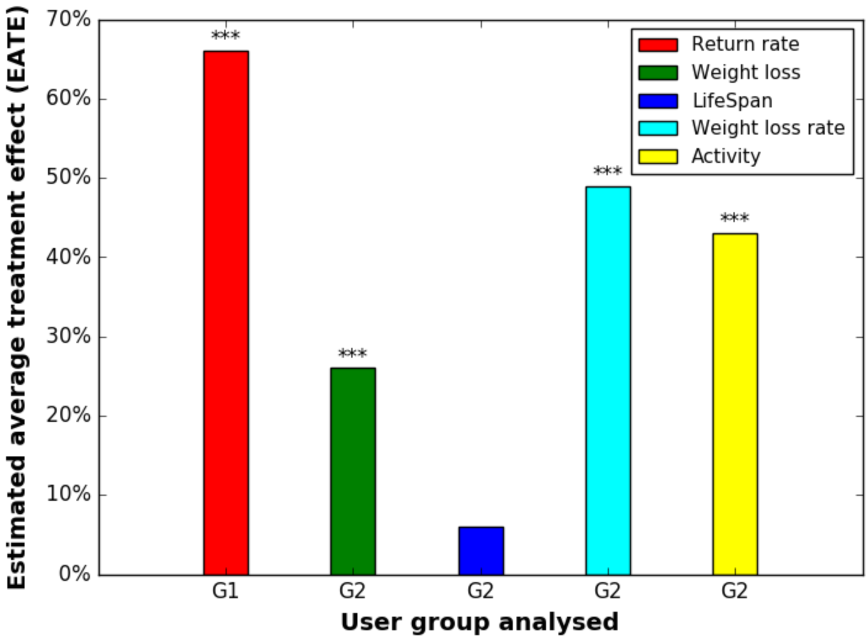


Figure 5.5: The effect size for the factors analyzed.

updating behaviors were similar. The groups presented a mean number of updates of 1.75±3.93 (treatment) and 1.60±2.70 (control), but this difference was not statistically significant, (p=0.16). As the two groups' badge updating behaviors are similar and it does not seem to affect our analysis.

We also experimented with different definitions for the treatment cutoff to see if there is an effect of "diminishing returns": receiving at least one comment (vs. none) could have a bigger impact than receiving 10 (vs. 9 or less). Figure 5.7 presents the effect size for different treatment definitions, although we can not guarantee the balance property for all the cutoffs, these results show that as expected when we increase the cutoff the effect size drops.

Mediation test

After estimating the causal effect of receiving social support on weight loss, we focus on checking if certain variables that were not included in the set of covariates could act as a mediator, explaining part of the observed effect of social support on weight loss. Conceptually and based on prior work, receiving social feedback could cause the effected user to (i) show a higher activity level in the community, and (ii) remain longer in the community. Therefore we first check if, indeed, receiving a comment on a user's first post has an effect on these variables and, if yes, if this effect mediates the observed effect on the reported weight loss.

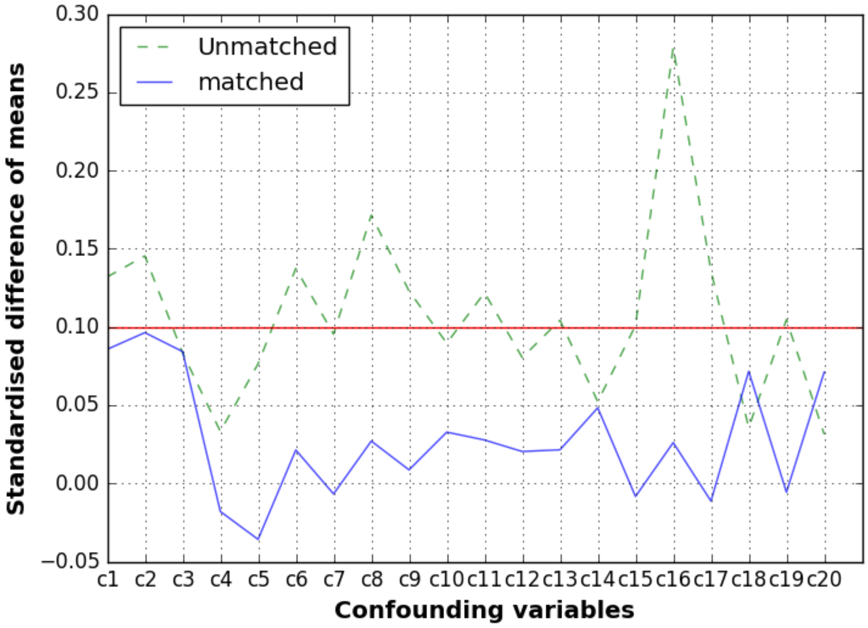


Figure 5.6: Standardized difference of means for each confounding variable (Group 2). Note that after the matching all the values are below 0.1, thus the groups are balanced.

Figure 5.5 shows that receiving social support also has an effect over the users' lifespan (i.e., the difference in days between the date of their last and first activity observed in the community) and the of number activities (i.e., the sum of the number of comments and posts). The effect size for lifespan is 6% (see Figure 5.5; blue bar), but this effect was not statistically significant, for the number of activities the effect size was 43% (see Figure 5.5; yellow bar).

Since the observed effect on the lifespan is small, we estimated if the social support has an effect on the users' weight loss rate (i.e., the weight loss in lb divided by the lifespan). As the rate is an unstable estimate for users who are only active for one or two days in the community, we chose to look at the median rather than the average effect. Concretely, we computed the median of the individual paired ratios of (weight loss rate treated individual / weight loss rate control individual). We then use the median of these medians as an estimate of the effect size. As expected there is an effect on the weight loss rate, where users that received at least 4 comments (treatment) lose weight roughly 35% faster than the ones that did not received (0.48 lb/day vs. 0.35 lb/day).

Finally, we applied a Sobel test to verify if lifespan and number of activities act as mediators in the relationship of the social support and weight loss, i.e., if they explain a significant part of the causal effect of social support on weight loss. The results of the Sobel test showed that the proportion of the effect of social support over weight loss due to lifespan and the number of activity is small – 5.6% and 3.4% respectively.

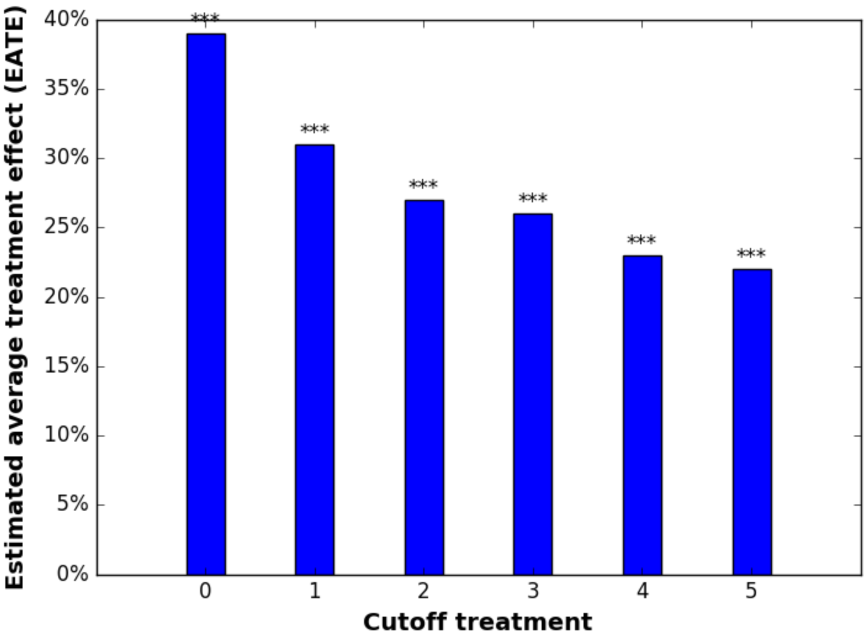


Figure 5.7: Effect size for different treatment definitions.

However, these results were not statistically significant even at $p = 0.1$. Surprisingly, among the users that return to post again, the difference in achieved weight loss does not seem to be linked to either lifespan or engagement in the community. Rather, the *rate of weight loss* seems to be affected for those users returning to the community.

5.4 Discussion

Our results show the importance of the social support exchange in online health communities to help people to improve their obesity condition. We performed a mediation analysis, which is a very important step in causal inference, but often overlooked. In this analysis we found very interesting to see that effect of social support over weight loss is not mediate by neither time in the community nor activity levels. Although we checked for these important variables, which previous literature showed that are associated with weight loss, we acknowledge that there are more variables that could mediate this effect, but we are unable to measure them.

Does weight loss equal success? Our analysis crucially assumes that members of the *loseit* community *want* to lose weight. If that was not the case then talking about "weight loss success" would be meaningless. However, results from a recent survey of the *loseit* community [denovosibi, 2016] indicate that 91% of the respondents were currently trying to lose weight, with another 7% trying to maintain their weight.

Therefore, it seems adequate to consider a higher level of weight loss as a desirable outcome.

Qualitative evidence. Though our analysis is deliberately using quantitative methods, there is also qualitative evidence to further support the claim that social support received in the community effects weight loss. Figure 5.8 shows a tag cloud from the aforementioned community survey [denovosibi, 2016], which includes the question "What do you like about /r/loseit?". The topics most emphasized by the survey participants were related to terms such as "community", "people", "supportive" and "support". Similarly, one can easily find posts explicitly acknowledging the perceived importance of the social support such as: "I have visited this page almost daily over the past 15 months, and it was really helpful in keeping my motivation. I hope this may provide similar motivation to those just starting! " .

Figure 5.8: Tag cloud from *loseit* community survey for the questions "What do you like about /r/loseit?".

What type of social support matters? One could also try and extract the topics or tone from the comments on a given post to see if particular types of comments have a larger effect on the reported weight loss. This, however, comes with endogeneity problems as the type of comments received is likely correlated with the subject matter of the post. Given large enough data sets one could hope to correct for this using our matching framework where the treatment is no longer binary – receiving a comment or not – but is multi-variate. We chose not to explore this route due to sparsity concerns.

We did, however, experiment with using another type of social feedback based on votes (see Figure 5.9): Reddit has a voting system with up- and down-votes and an aggregate "voting score" combining these two types of votes, positive - negative, can be obtained via the API. In our data set, this voting score was never negative. Using the same matching setup (see Chapter 5), this gave an effect size of 16% to users to return to the community and 45% to weight loss ($p < 0.01$) . The balance condition also held for this experiment. Though most of the limitations discussed below also apply to this setup, the fact that a "similar" effect is observed for a different definition of social feedback indicates that our results might hold more general.

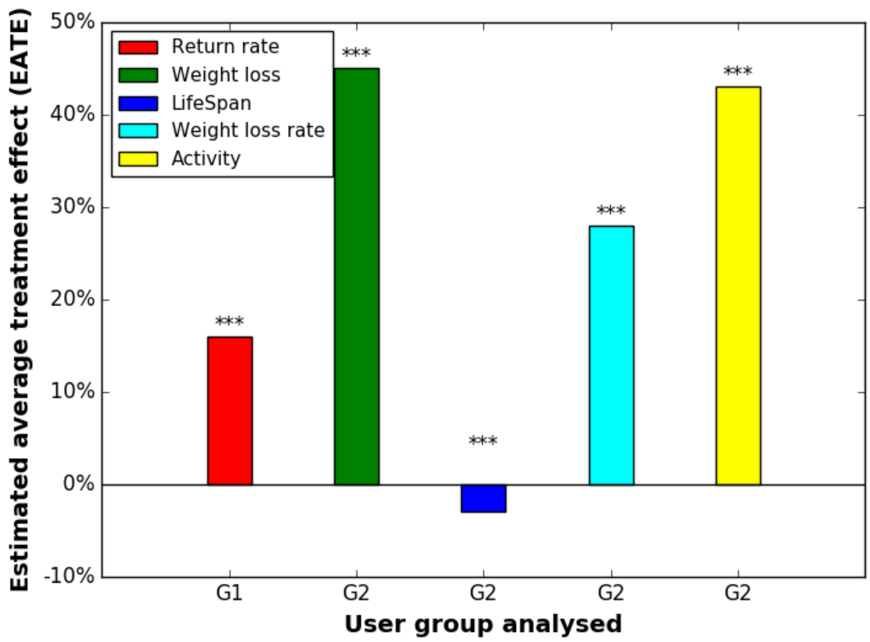


Figure 5.9: The effect size for the factors analyzed when considering voting score as the feedback.

Impact of pruning on effect size. We started our analysis with the assumption that, as we had previously observed for the effect on return-to-post probability [Cunha et al., 2016], pruning via matching would lead to a *lower* estimate of the effect of social feedback on weight loss when compared to the unmatched analysis. Intuitively, matching, and hence pruning, should reduce the effect of confounding user variables such as "positive outlook on life" which might affect writing style and have a positive effect on both return-to-post behavior as well as on weight loss success. However, we observed the *opposite*: the raw effect size for the unmatched data was 22% (treatment cutoff of 4 comments) whereas it was 26% for the matched analysis (treatment cutoff of 4 comments and similarity threshold of 0.955). In particular, users who were treated and who, eventually, lost less than the median weight loss were pruned more often (27%) than their treated counter-parts who lost a lot of weight (23%). At the same time for the untreated users the differences in pruning rates for less-than-median (28%) and more-than-median (28%) weight loss were small. Though this is surprising, it actually helps to make the overall claim, i.e., that social feedback supports weight loss in an online community, more robust.

But one question still unanswered, can we identify the users that are in more need of social support? To answer this question, in Chapter 6 we build a statistical model to infer users' weight condition, our goal is to provide a tool that might aid health professional to find a potential group of users that could benefit more from this type of "treatment".

Chapter 6

Inferering Users Weight Change

We showed in the last chapter that social support has an causal effect over (i) users retention in the community, and (ii) users weight loss. In this chapter, we investigate whether it is possible to infer users weight condition, which is an important task that can aid health professionals to target potential users that could benefit from the support received in the community. We look at whether the content shared in the community along with users interactions and real-world information are indicative of weight change.

However, knowing users weight change alone is not as informative as knowing their obesity condition. With the information of the obesity condition, health professionals can help users to reach and maintain a healthy weight, which is very important for overall health and can help to prevent and control many diseases and conditions. Hence, in this phase of the study instead of looking for users that used the community badge system to declared their weight change, we changed our focus to users that self-declared their personal information (i.e. age, gender, height and weights) in the body of posts and comments. This way, we were able to compute their body mass index (BMI), which is proportional to mass and inversely proportional to the square of the height (i.e., $BMI = weight(kgs)/(height(cm))^2$) and know their obesity condition. According to World Health Organization criteria the obesity classes are: normal (BMI from 18.5 to 25), overweight (BMI from 25 to 30), class I obesity (BMI from 30 to 35), class II obesity (BMI from 35 to 40) and class III obesity (BMI over 40). As we mentioned in Chapter 3.2, extracting this type of information from free text is a very difficult task, but we took advantage of the community conventions (e.g. SW/CW/GW) and created regular expressions to extract height, age, gender and weights. Besides applying the regular expressions, we also performed a manual analysis to ensure the precision of our strategy, and only considered users that had a minimum of three weights reported and at least 30 days between first and last report (minimum time necessary for the user to

Table 6.1: Before and after users obesity categories

	Final category					
Initial category	Normal	Overweight	Obese1	Obese2	Obese3	Total
Normal	25	2				27
Overweight	70	58	1			129
Obese1	37	105	46	1		189
Obese2	10	52	76	32	1	171
Obese3	4	36	61	63	74	238
	Females					
Normal	23	1				24
Overweight	60	39	1			100
Obese1	21	65	23	1		110
Obese2	8	23	35	11	1	78
Obese3	2	14	30	27	41	114
	Males					
Normal	2	1				3
Overweight	10	19				29
Obese1	16	40	23			79
Obese2	2	29	41	21		93
Obese3	2	22	31	36	33	124

actually start a relevant weight loss process). This preprocessing step generated a set of 754 users.

Table 6.1 shows the final user distribution according to obesity categories and gender, considering their initial category (computed with the first weight reported) and final category (computed with the last weight reported). Note that females are more present than males in this sample (55%-45%) and the median age is 25. Those characteristics are in accordance with an online study[1], which found that women are more likely than men to go online to figure out a possible diagnosis. Other group that have a high likelihood of doing so are younger people.

In Figure 6.1, we show the CDF of initial and final BMI for the 754 users. The median BMI value dropped from 38.35 to 28.94.

6.1 A methodology for assessing weight change

Here we introduce the methodology we followed to investigate determinants of weight change percentage in *loseit* considering users with at least 30 days between first and last weight reported. Below, we detail our models.

[1] http://www.pewinternet.org/2013/01/15/health-online-2013/

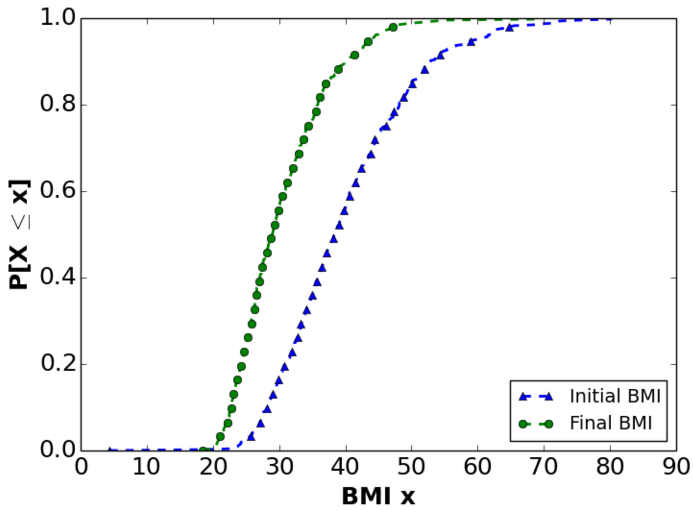


Figure 6.1: CDF of Users' BMI changes during their life time in *loseit*.

The model's response variable represents the weight loss percentage the users achieved while in the community.

The explanatory variables are divided in three categories: **Users real-world:** Our first set of explanatory variables focuses on extracting real world (i.e., age, gender and BMI) characteristics from a users posts and comments received in *loseit*. We used regular expressions to extract age, gender, height and weights shared in posts and comments. Then, with the start weight and height we computed the BMI of users. **Users online behavior:** Our second set of explanatory variables focuses on the various aspects of users activities in *loseit*. We introduce variables for the number of activities, comments received, voting score received, number of weight reported and users lifetime (defined as date of last activity minus date of first activity in *loseit* plus one). Also, we include variables that represent the number of online days (number of days with at least 1 activity), percentage of online days (i.e., number of days with at least one activity), days offline (i.e., number of days without any activity) and number of weeks active (weeks with at least one activity). **Users linguistic characteristics:** Our last set of explanatory variables focuses on extracting linguistic attributes from users posts and comments received in *loseit*. We have a set of 50 LDA topics to posts and a 50 LDA topics for the comments received (see Section 3.3.3). In order to summarize the most discussed topics per user, the probabilities returned by the LDA for each post or comment received by each user were summed up for each topic and then scaled to the 0-1 range. These probabilities per topic per user were than given as input to the regression model.

Statistical model. Having defined the set of variables, we applied a linear regression model to infer the users percentage of weight loss. To prevent collinearity, which reduces the accuracy of the estimates of the regression coefficients. We calculated the variance inflation factor (VIF) [James et al., 2014], it provides an index that measures how much the variance of an estimated regression coefficient is increased because of collinearity. To compute the VIF, if all your feature variables are X_1, X_2, ..., X_n, we build a linear model for each X_i as predicted by all the others Xs, and get the coefficient of determination (R^2) of the model. The VIF for X_i is defined as 1/(1 - R^2), if the VIF is bigger than 5, the convention is to drop the variable. Then, to avoid model complexity (overfitting), we performed a variable selection, which is the task of determining the predictors that are associated with the response. Among the various techniques that can be used to perform the variable selection, we choose the mixed selection. This is a combination of forward and backward selection. It starts with the null model (zero variables) and search through models lying in the range between the null and full model using the forward selection algorithm. For model selection, it uses the Akaike Information Criterion (AIC), which offers a relative estimate of the information lost when a given model is used to represent the process that generates the data [James et al., 2014]. We also tried LASSO regression, but its variable selection and shrinkage step generated a model with weaker explanatory power.

6.1.1 Regression models

After checking for collinearity and performing the mixed selection, we ended up with the following variables in the model: 3 real world characteristics (age, gender, BMI), 6 activity behaviors (number of activities, comments received, voting score received, number of weight reported, number of days offline and number of weeks active), 12 post topics and 15 comment received topics, in a total of 36 variables.

To show the importance of our dependent variables we consider three statistical models: (i) Real-world variables (W), (ii) Real-world + Online behavior variables (WA) and (iii) Real-world + Online behavior + Language variables (WAT). The three models are compared using their coefficient of determination (R^2) [Nagelkerke, 1991]. The first two models were motivated by previous work [Poncela-Casasnovas et al., 2015] that also attempted to infer users weight loss, and through the third, we examine the additional explanatory power of linguistic variables to infer weight loss. Table 6.2 shows the three linear models and their R^2. Note that the values of R^2 vary from 0.18 to W to 0.39 to WAT. The R^2 obtained by WAT is considered large for a social system, because there are a large number of other factors we do not have information about, such as diet, exercise level, education level, motivation and so on. For example, the R^2 reported in [Poncela-Casasnovas et al., 2015] for the same task was 0.27.

Table 5 shows the regression coefficients found for the models with different variables along with their standard deviation of errors and significance level. A positive coefficient indicates the variable is correlated to an increase in weight loss, while a negative coefficient indicates the variable is related to a decrease in weight loss. The p-value for each term tests the null hypothesis that the coefficient is equal to zero (no effect). A low p-value (< 0.05) indicates that you can reject the null hypothesis, i.e., the predictor is likely to be a meaningful addition to the model as changes in its value are related to changes in the response variable.

The estimated coefficients confirmed that BMI and number of weights reported, as expected, have significant explanatory power, being correlated with increasing weight loss. The amount of time spent in the community also has important role in our model, given the fact that offline days is associated with less weight loss and weeks active is associated with more weight loss. Age was not found significant for models WA and WAT. Perhaps the range of age in our data, concentrated around 25, might be too restrictive to identify the influences of age in weight loss. Instead, support from the community, in the form of comments and voting score received, showed a significant explanatory power and was strongly correlated with weight loss. Finally, the topics of comments and posts are important to explain the weight loss percentage, some (i.e. feelings, start again) are associated with less weight loss, others presented the function of protective factors (i.e. motivation, exercises) being associated to increasing weight loss.

Residual evaluations are good ways of looking at models that only need to fit to the current population. However, in this study we are dealing with a sample of online users, which can grow over time. In this case, residual evaluations do not give an indication of how well the model will perform when asked to make new predictions for data it has not already seen. Aiming to evaluate the generalization power of WAT model, we executed a 10-fold cross validation. After performing the cross-validation, we found a mean R^2 of 0.39 and standard deviation of 0.01. The low standard deviation indicates that indeed the model generalizes.

Besides checking the goodness of fit of our models, we also checked for precision in our inference. We modeled a classification task of the obesity condition of users, using the weight change found by our models to compute the "inferred" obesity category, then we compared it with the actual final obesity category of the users. We applied a 10 fold cross-validation procedure and to ensure that the we are not dealing with a very simple task, we created two naive baselines for comparison: (i) a model that assumes the users

Table 6.2: The three linear models considering different factors associated with weight loss percentage. Each cell shows the regression coefficient found by the model for the variable in that row followed by its standard error. The stars are shorthand for significance levels, with the number of asterisks displayed according to P (*** for p<0.001** for p<0.01, * for p<0.05)

Variables/Criteria	W R^2 0.18	WA R^2 0.26	WAT R^2 0.39
Real world			
BMI	0.51717 ± 0.09 ***	0.4151969 ± 0.09 ***	3.712e-01 ± 0.08***
Gender	-1.35705 ± 0.14 .	-1.9796563 ± 1.57 *	-1.471e+00 ± 1.56 .
Age	0.23238 ± 1.63 **	0.1263763 ± 0.14 .	1.180e-01 ± 0.13 .
Online behaviour			
Activity		-0.008 ± 0.005**	-6.652e-03 ± 0.005*
Comments received		0.12 ± 0.007	5.218e-03 ± 0.006
Voting score received		0.002 ± 0.001***	1.924e-03 ± 0.001 **
Number of weights		0.27 ± 0.12 ***	2.876e-01 ± 0.11 ***
Weeks active		0.008 ± 0.07 **	8.974e-02 ± 0.06 **
Days off		-0.004 ± 0.002***	-3.648e-03 ± 0.002 **
Posts topics			
Weight control apps			-5.726e+01 ± 53.36 *
Self-steem (1)			1.185e+02 ± 48.98 ***
Self-steem(2)			4.803e+01 ± 54.94 .
Workout			3.703e+01 ± 40.10 .
Friendship			8.372e+01 ± 56.60 **
Health information			4.484e+01 ± 62.98
Weight goals			7.790e+01 ± 53.96 **
Feelings			-1.250e+02 ± 75.14 ***
Asking for advice			6.897e+01 ± 89.40
Motivation			6.851e+01 ± 52.50 **
Weight check-ins			4.202e+01 ± 51.68
Family			-4.716e+01 ± 61.88
Comments topics			
Body transformations			1.126e+02 ± 119.36 .
Asking for help			2.357e+02 ± 192.76 *
Body measures			-2.909e+02 ± 216.40 **
Changing lifestyle			-2.192e+02 ± 282.80
Counting calories			1.874e+02 ± 143.10 **
Asking for support			-1.342e+02 ± 127.00 *
Friendship			2.234e+02 ± 210.20 *
Workout			2.953e+02 ± 144.86 ***
Motivation			3.048e+02 ± 299.20 *
Weight goals			2.430e+02 ± 133.00 ***
Motivation			2.518e+02 ± 200.80 *
Healthy foods			1.737e+02 ± 182.78 .
Gratitude			1.602e+02 ± 144.14 *
Starting again			-2.314e+02 ± 212.00 *
Lifestyle change			2.210e+02 ± 199.68 *

Table 6.3: Average precision of the models considering obesity categories

Model	Precision
WAT	0.66 ± 0.03
WA	0.52 ± 0.05
W	0.47 ± 0.06
Baseline 2	0.44 ± 0.01
Baseline 1	0.29 ± 0.03

will not change their weight and consequently not change their obesity category, and (ii) a model that assumes the users will drop one obesity category and if they are already in the normal category they will not change categories.

Table 6.3 shows the classification results in terms of mean precision and standard deviation for the models and baselines. In general, we observe that the best performing model WAT, which achieves the mean accuracy of 66% with standard deviation of 0.03. This model is followed by WA, W and then the baselines in terms of performance.

6.2 Discussion

Knowing about characteristics, experiences and interests of users of online applications might be of great value for improving weight loss strategies. Our analyses of real-world, behavioral, and linguistic characteristics of users of *loseit* that reported weight change within a 30-day minimum interval revealed that women were predominant, as well as overweight and obese subjects. The addition of a semantic analysis of the terms used in posts and comments led to a great improvement in the prediction of weight change among *loseit* users.

Regarding users characteristics that were shown to predict weight change, gender and initial BMI were significantly independent ones. In real-world interventions, gender differences in regard to success rates following short-term weight loss programs are controversial [Forster and Jeffery, 1986], despite males' higher metabolic rates [Wyatt et al., 1999]. BMI was positively associated with weight loss. Since weight and BMI were self-reported, we might imply that the higher the baseline BMI, the greater the perception of excessive weight and/or dissatisfaction with weight status. Weight perceptions accuracy and dissatisfaction with weight status have been consistently associated with trying to lose weight and better weight control among obese and overweight adults in population-based studies [Millstein et al., 2008].

Various measures of online activity were independent predictors of weight loss. Higher participation levels in the online social media might unveil higher levels of self-motivation, which have been associated with better weight loss outcomes following real-world interventions as well [Elfhag and Rossner, 2005]. Therefore, developing strategies that maintain users active in the online community should be a cornerstone of web-based weight loss communities.

Social support to health has been shown to be associated with maintenance of health behaviour change and better weight loss outcomes in real-world interventions [Verheijden et al., 2005; Flores-Gomez et al., 2012]. Our findings also convey this idea, the voting score received for users, which might lead to a sense of social belonging, was an independent determinant of weight loss among *loseit* users. Further, the fact that a great relative improvement (R^2 0.26 to 0.39) at explaining weight change occurred after adding topics discussed in posts and comments to the model underpins this idea.

Finally, statistical models that identify users weight condition are very important, it can be used to determine the potential obesity class of users for whom badge or other self-reported information on weight loss is not available. It might be particularly valuable for health professionals to provide help and support for individuals who intend to lose weight and use a social media platform.

Chapter 7

Conclusions and future directions

This thesis performed a large scale study of factor associated to weight change in a Reddit weight loss community and provided a better comprehension of the properties of this type of community, with an in depth characterization of users activities and the correlations to weight loss.

We presented an analysis of the effect of receiving social feedback in the form of comments on the reported weight loss in the /r/loseit community on Reddit. Correcting for confounding factors through a matching methodology, users who receive at least 4 comments on their first post in the community were (i) 66% more likely to return for a future activity in the community, and (ii) conditional on the user returning, those who had previously received feedback end up reporting on average 9 lb more in weight loss. For these returning users, this effect is not mediated by neither (i) an increased level of activity in the community, nor by (ii) a longer lifespan in the community. Though observational studies have inherent limitations on causal inference, our work helps to illustrate the importance of receiving feedback in online support forums, in particular for users new to the community.

In this direction, we built a supervised statistical model to identify potential users that could benefit from this social support received in the community. As major strengths, we must highlight the use of a quantitative methodology, which automatically looks at our measures of interest and allows for an analysis of a big sample, the long-term follow-up of the social network and the investigation of the influence of real-world, online behaviors and semantics of online discussions on weight change reported by users. This technique illustrates how to create tools that might aid health professionals to target groups of "at risk" users in this type of communities.

Considering the increasing access to computer, applications and internet worldwide, as well as the substantial amount of healthcare resources demanded by the high prevalence of obesity, particularly at young ages, online communities might be important public health strategies to obesity treatment and prevention. In order to potentialize the benefit of these communities, specific features that increase online activity should be investigated and incorporated to the online social network. Furthermore, activities that stimulate social support among users might be key points in the planning of online weight loss interventions.

7.1 Limitations

Limitations of using badges to track weight loss. To infer a user's weight loss success or failure we are currently relying on the badges used in the *loseit* community. These badges only capture *self-reported* weight loss progress. The first issue imposes an important limitation as, one could imagine, receiving social feedback leads to a heightened sense of self-awareness and a feeling of "being watched" in the community. Though this could lead to positive peer pressure, it could also increase the probability of over-reporting weight loss progress. Badges are also always visibly displayed next to a user's screen name, increasing the likelihood of social signaling effecting their use. One solution to this issue could be to perform a similar analysis using auto-generated weight updates from smart scales as used by Wang et al. [2016]. Such data sources are less likely to be prone to misreporting errors.

Limitations to determining the start date of weight loss journey. Our analysis, especially that related to the rate of weight loss (Section 5.3), assumes that a user's weight loss journey starts the day they first announce themselves to the community in the form of a post. However, in practice, users might well first observe the community before deciding to post. This means that the actual rate of weight loss is likely to be lower, as the time period over which the weight loss is achieved is longer. A more subtle issue related to this passive use of the support community is that it could affect the writing style. Put simply, users who have been following the community for a while might (i) have a "head start" as far as weight loss is concerned, and (ii) they might write in a style more in tune with the community which, in turn, could lead to more social feedback. If these stylistic differences are not represented in the extracted covariates, this could lead to an overestimate of the effect size. Though we cannot completely rule out this possibility, we believe that the effect sizes are large enough to make it unlikely that they are fully explained by this hidden adaption to the community.

Limitations of our matching approach. When applying a matching approach, there are a number of choice one needs to make such as (i) the selection of covariates, (ii) how to normalize and potentially weight different covariates, (iii) which distance metric to use and whether to use "blocking", and (iv) which similarity threshold to choose. Of these, any choices for (ii), (iii) and (iv) should asymptotically converge to the same result as the matched pairs become more and more similar, being identical on all the covariates in the limiting case. We therefore did not include experiments with higher similarity thresholds because of issues of data sparsity.

Concerning the covariates used, we believe our statistical method is a reasonable choice. Introducing too many additional covariates can lead to problems of high dimensionality when attempting to match similar posts along, say, hundreds of dimensions. In such settings it becomes difficult to balance all the covariates considered. Furthermore, many other potential covariates should be balanced at least "on average", if they are correlated with the covariates in the final model. Still, the exclusion of unknown but potentially crucial covariates is always a concern when applying a matching methodology.

Limitations of observability of returning users. Our main analysis relies on users returning to the *loseit* community to post or comment again after their initial post. Assuming the user also has badge information, then this second data point provides an estimate of both the absolute weight loss (in the badge) and the weight loss rate (in the difference of time stamps). This means that we cannot make any statements concerning weight loss for users who do return to the community for a second, public activity. Though our main results are conditional on users returning, we also observe social feedback leading to an increase in return probability (see Figure 5.5). However, the issue of predicting who will or will not return to an online community has been studied before [Cunha et al., 2016; Choudhury and De, 2014] and so we did not analyze this further.

7.2 Future research directions

Who benefits most from social support? All of our results are aggregates indicating that, on average, users seem to benefit from receiving social support in the form of comments. For future work it would be interesting to look into what type of users are most or least likely to benefit from such support, for example looking for gender-specific effects. Though gender is not an attribute of a user's profile, it can sometimes be inferred from their posts ("I'm a mother of two ...") or from shared progress pictures. In some cases a user's chosen screen name such as "john123" or "mary456" also provide hints. Similarly, one could look for cases of users indicating their starting weight, rather than just the weight loss, to study whether the effect of social support is tied to a user's initial weight.

Designs implications. Our findings suggest that whether or not a user receives feedback on their initial post affects (i) their probability to return to the community, and (ii) given that they return, the amount of weight loss they report. Therefore mechanisms that increase the likelihood to receive a "warm welcome" are expected to lead to more engagement with the community and to better health outcomes. Fortunately, the vast majority of initial posts (96%) already receive at least one comment. It could be worth considering a mechanism where posts that do not are brought to the moderators' attention so that they can provide an adequate reply. It could also be worthwhile to construct a "positivity bot" which provides non-generic positive feedback on posts overlooked by the community [van der Zwaan et al., 2012]. Other researchers are exploring the creation of a framework to allow formal testing of theories of different moderation styles.[1] Our research contributes by providing a theory to test.

Relevance of our results. The results presented are based on a single Reddit community. Then, one concern that arises is whether /r/loseit may be specialized and have strong community norms that encourage a particular style of first-post-reply. Although, in our qualitative analyses we did not see any sign of community norm to give a "warm welcome" to newcomers. One way to combat this is to look at a different community where participants track progress and may have different norms. We have identified two potential communities, /r/keto and /r/xxketo, both present similar characteristics of *loseit* and are related to the ketogenic diet [Dashti et al., 2006].

Applying the framework to real-world data. Another important future direction is to apply our framework to real-world data from surveys conducted by governments and verify if the effect of social support also holds. One way to perform this study is using data from the Behavioral Risk Factor Surveillance System (BRFSS), an annual survey that measures the health-related behaviors of adult Americans. In this survey, the concept of social support is captured in the question, "How often do you receive the social and emotional support your need?". Respondents' weight loss is measured in two ways: with a binary variable that captures whether or not they reached the ten percent benchmark and a continuous variable that represents the

[1] See `https://civic.mit.edu/blog/natematias/reddit-moderators-lets-test-theories-of-moderation-together` for thoughts by Nathan Matias' and `https://np.reddit.com/r/TheoryOfReddit/comments/456503/want_to_test_your_theories_of_moderation_lets/` for a subreddit on the topic.

actual percentage of weight lost.

Bibliography

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*, 46(3):399--424.

Ballantine, P. W. and Stephenson, R. J. (2011). Help me, i'm fat! social support in online weight loss networks. *Journal of Consumer Behaviour*, 10(6):332--337. ISSN 1479-1838.

Black, L. W., Bute, J. J., and Russell, L. D. (2010). *Cases on Online Discussion and Interaction: Experiences and Outcomes*, chapter "The Secret Is Out!": : Supporting Weight Loss through Online Interaction. IGI Global.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993--1022. ISSN 1532-4435.

Boone, B. (2012). Mediation analysis: Expanding from one mediator to multiple mediators.

Burke, M., Marlow, C., and Lento, T. (2009). Feed me: Motivating newcomer contribution in social network sites. In *CHI*.

Chou, W.-y., Prestin, A., and Kunath, S. (2014). Obesity in social media: a mixed methods analysis. *Translational Behavioral Medicine*, 4(3):314--323. ISSN 1869-6716.

Choudhury, M. D. and De, S. (2014). Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Proc. of the 8th Int. Conf. on Weblogs and Social Media, ICWSM*.

Christakis, N. A. and Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine*, 357(4):370--379.

Chunara, R., Bouton, L., Ayers, J. W., and Brownstein, J. S. (2013). Assessing the online social environment for surveillance of obesity prevalence. *PloS one*, 8(4):e61373.

Cohen, S., Underwood, L., and Gottlieb, B., editors (2000). *Social support measurement and interventions: A guide for health and social scientists*. New York: Oxford.

Culotta, A. (2010). Towards detecting influenza epidemics by analyzing twitter messages. In *Proceedings of the First Workshop on Social Media Analytics*, SOMA '10, pages 115--122, New York, NY, USA. ACM.

Culotta, A. (2014a). Estimating county health statistics with twitter. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 1335--1344, New York, NY, USA. ACM.

Culotta, A. (2014b). Reducing sampling bias in social media data for county health inference. In *Joint Statistical Meetings Proceedings*.

Cunha, T. O., Weber, I., Haddadi, H., and Pappa, G. L. (2016). The effect of social feedback in a reddit weight loss community. In *DigitalHealth*, pages 99--103.

Dashti, H. M., Al-Zaid, N. S., Mathew, T. C., Al-Mousawi, M., Talib, H., Asfar, S. K., and Behbahani, A. I. (2006). Long term effects of ketogenic diet in obese subjects with high cholesterol level. *Molecular and cellular biochemistry*, 286(1-2):1--9.

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., and Kumar, M. (2016). Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2098--2110. ACM.

denovosibi (2016). [results] winter/spring 2016 user survey. `https://www.reddit.com/r/loseit/comments/4acxxn/results_winterspring_2016_user_survey/`.

Eichstaedt, J. C., Schwartz, H. A., Kern, M. L., Park, G., Labarthe, D. R., Merchant, R. M., Jha, S., Agrawal, M., Dziurzynski, L. A., Sap, M., Weeg, C., Larson, E. E., Ungar, L. H., and Seligman, M. E. (2015). Psychological language on Twitter predicts county-level heart disease mortality. *Psychological Science*, 26:159–169.

Elfhag, K. and Rossner, S. (2005). Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. *Obes Rev*, 6(1):67--85.

Esarey, J. (2015). Causal inference with observational data.

Eschler, J., Dehlawi, Z., and Pratt, W. (2015). Self-characterized illness phase and information needs of participants in an online cancer forum. In *Proc. of the Eighth International Conference on Weblogs and Social Media*.

Flores-Gomez, I., Bacardi-Gascon, M., Armendariz-Anguiano, A. L., Perez-Morales, M. E., and Jimenez-Cruz, A. (2012). [Evidence of social support as therapy for weight loss: a systematic review]. *Nutr Hosp*, 27(5):1422--1428.

Fontana, A. F., Kerns, R. D., and *et al.* (1989). Support, stress, and recovery from coronary heart disease: a longitudinal causal model. *Health Psychol*, 8(2):175--193.

Forster, J. L. and Jeffery, R. W. (1986). Gender differences related to weight history, eating patterns, efficacy expectations, self-esteem, and weight loss among participants in a weight reduction program. *Addict Behav*, 11(2):141--147.

Gilbert, E. (2013). Widespread underprovision on reddit. In *Proc. of the Conf. on Computer Supported Cooperative Work*, CSCW '13, pages 803--808.

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012--1014.

Gomide, J., Veloso, A., Meira, Jr., W., Almeida, V., Benevenuto, F., Ferraz, F., and Teixeira, M. (2011). Dengue surveillance based on a computational model of spatio-temporal locality of twitter. In *Proceedings of the 3rd International Web Science Conference*, WebSci '11, pages 3:1--3:8, New York, NY, USA. ACM.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological theory*, 1(1):201--233.

Grav, S., Hellzèn, O., and *et al.* (2012). Association between social support and depression in the general population: the hunt study, a cross-sectional survey. *J. of clinical nursing*.

Herzer, M., Zeller, M. H., Rausch, J. R., and Modi, A. C. (2011). Perceived social support and its association with obesity-specific health-related quality of life. *Journal of developmental and behavioral pediatrics: JDBP*, 32(3):188.

Hesterberg, T., Moore, D. S., Monaghan, S., Clipson, A., and Epstein, R. (2005). Bootstrap methods and permutation tests. *Introduction to the Practice of Statistics*, 5:1--70.

Hildebrandt, S. E. (1998). Effects of participation in bariatric support group after roux-en-y gastric bypass. *Obesity surgery*, 8(5):535--542.

Hutchesson, J. M., Tan, Y. C., and *et al.* (2016). Enhancement of self-monitoring in a web-based weight loss program by extra individualized feedback and reminders: Randomized trial. *J Med Internet Res*, 18(4):e82.

Hutto, C. J. and Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *ICWSM*.

Hwang, K. O., Ottenbacher, A. J., Green, A. P., Cannon-Diehl, M. R., Richardson, O., Bernstam, E. V., and Thomas, E. J. (2010). Social support in an internet weight loss community. *Int. Journal of Medical Informatics*, 79(1):5--13.

Imai, K., King, G., and Stuart, E. A. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the royal statistical society: series A (statistics in society)*, 171(2):481--502.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2014). *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated. ISBN 1461471370, 9781461471370.

Johnson, G. J. and Ambrose, P. J. (2006). Neo-tribes: The power and potential of online communities in health care. *Communications of the ACM*, 49(1):107--113.

Kim, H. S., Sherman, D. K., and Taylor, S. E. (2008). Culture and social support. *American Psychologist*, 63(6):518.

King, G., Keohane, R. O., and Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton university press.

King, G., Lucas, C., Nielsen, R., King, G., Pan, J., Roberts, M., King, G., King, G., Grimmer, J., King, G., et al. (2014). The balance-sample size frontier in matching methods for causal inference. *PS: Political Science and Politics*, 42:S11--S22.

King, G. and Nielsen, R. (2015). Why propensity scores should not be used for matching. *http://j.mp/1sexgVw*.

Ko, H.-C., Wang, L.-L., and Xu, Y.-T. (2013). Understanding the different types of social support offered by audience to a-list diary-like and informative bloggers. *Cyberpsychology, Behavior, and Social Networking*, 16(3):194--199.

Laranjo, L., Arguel, A., Neves, A. L., Gallagher, A. M., Kaplan, R., Mortimer, N., Mendes, G., and Lau, A. (2014). The influence of social networking sites on health behavior change: a systematic review and meta-analysis. *J Am Med Inform Assoc*, Online First(0).

Lett, H., Blumenthal, J., and *et al.* (2005). Social support and coronary heart disease: epidemiologic evidence and implications for treatment. *Psychosomatic medicine*.

Ma, X., Chen, G., and Xiao, J. (2010). Analysis of an online health social network. In *Proc. of the 1st ACM International Health Informatics Symposium*, pages 297--306.

Medeiros, L. and Bosse, T. (2016). Empirical analysis of social support provided via social media. In *International Conference on Social Informatics*, pages 439--453. Springer.

Mejova, Y., Haddadi, H., Noulas, A., and Weber, I. (2015). #foodporn: Obesity patterns in culinary interactions. In *Proceedings of the 5th International Conference on Digital Health 2015*, DH '15, pages 51--58, New York, NY, USA. ACM.

Mermelstein, R., Cohen, S., and *et al.* (1986). Social support and smoking cessation and maintenance. *Journal of consulting and clinical psychology*, 54(4):447.

Millstein, R. A., Carlson, S. A., Fulton, J. E., Galuska, D. A., Zhang, J., Blanck, H. M., and Ainsworth, B. E. (2008). Relationships between body size satisfaction and weight control practices among US adults. *Medscape J Med*, 10(5):119.

Mimno, D., Wallach, H. M., Talley, E., Leenders, M., and McCallum, A. (2011). Optimizing semantic coherence in topic models. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 262--272. Association for Computational Linguistics.

Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3):691--692.

Nelson, R., Wolf, D., and Joos, I., editors (2012). *Social Media for Nurses: Educating Practitioners and Patients in a Networked World*. Springer.

Olteanu, A., Varol, O., and Kıcıman, E. (2016). Towards an open-domain framework for distilling the outcomes of personal experiences from social media timelines. In *Tenth International AAAI Conference on Web and Social Media*.

Park, K., Weber, I., and *et al.* (2016). Persistent sharing of fitness app status on twitter. In *CSCW*, pages 184--194.

Patrick, K., Calfas, K., Norman, G., and *et al.* (2011). Outcomes of a 12-month web-based intervention for overweight and obese men. *Ann Behav Med*.

Paul, M. J. and Dredze, M. (2011). You are what you tweet: Analyzing twitter for public health. *ICWSM*, 20:265--272.

Poncela-Casasnovas, J., Spring, B., McClary, D., Moller, A. C., Mukogo, R., Pellegrini, C. A., Coons, M. J., Davidson, M., Mukherjee, S., and Nunes Amaral, L. A. (2015). Social embeddedness in an online weight management programme is linked to greater weight loss. *Journal of The Royal Society Interface*, 12(104). ISSN 1742-5689.

Preece, J., Nonnecke, B., and Andrews, D. (2004). The top five reasons for lurking: Improving community experience for everyone. *Computers in Human Behavior*, 20(2):201–223.

Puhl, R. M. and Brownell, K. D. (2001). Bias, discrimination, and obesity. *Obes Res*, pages 788--805.

Raynor, H. A., Jeffery, R. W., Ruggiero, A. M., Clark, J. M., and Delahanty, L. M. (2008). Weight loss strategies associated with BMI in overweight adults with type 2 diabetes at entry into the Look AHEAD (Action for Health in Diabetes) trial. *Diabetes Care*, 31(7):1299--1304.

Rubin, D. B. (1976). Multivariate matching methods that are equal percent bias reducing, i: Some examples. *Biometrics*, pages 109--120.

Rubin, D. B. and Stuart, E. A. (2006). Affinely invariant matching methods with discriminant mixtures of proportional ellipsoidally symmetric distributions. *The Annals of Statistics*, pages 1814--1826.

Seo, M., Kim, J., and Yang, H. (2016). Frequent interaction and fast feedback predict perceived social support: Using crawled and self-reported data of facebook users. *Journal of Computer-Mediated Communication*.

Sobel, M. E. (1986). Some new results on indirect effects and their standard errors in covariance structure models. *Sociological methodology*, 16:159--186.

Song, J. W. and Chung, K. C. (2010). Observational studies: cohort and case-control studies. *Plastic and reconstructive surgery*, 126(6):2234.

Spirtes, P., Meek, C., and Richardson, T. (1995). Causal inference in the presence of latent variables and selection bias. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 499--506. Morgan Kaufmann Publishers Inc.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 25(1):1.

Sullivan, C. F. (2003). Gendered cybersupport: A thematic analysis of two online cancer support groups. *Journal of health psychology*, 8(1):83--104.

Tagarelli, A. and Interdonato, R. (2013). Who's out there?: identifying and ranking lurkers in social networks. In *Proc. of the ASONAM'13*, pages 215--222. ACM.

Tamersoy, A., De Choudhury, M., and Chau, D. H. (2015). Characterizing smoking and drinking abstinence from social media. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, HT '15, pages 139--148, New York, NY, USA. ACM.

Teixeira, P. J., Carraça, E. V., Marques, M. M., Rutter, H., Oppert, J.-M., De Bourdeaudhuij, I., Lakerveld, J., and Brug, J. (2015). Successful behavior change in obesity interventions in adults: a systematic review of self-regulation mediators. *BMC medicine*, 13(1):1.

Thiese, M. S., Moffitt, G., Hanowski, R. J., Kales, S. N., Porter, R. J., and Hegmann, K. T. (2015). Commercial Driver Medical Examinations: Prevalence of Obesity, Comorbidities, and Certification Outcomes. *J. Occup. Environ. Med.*, 57(6):659--665.

Tsapeli, F. and Musolesi, M. (2015). Investigating causality in human behavior from smartphone sensor data: a quasi-experimental approach. *EPJ Data Science*, 4(1):1--15.

Turner-McGrievy, G. M. and Tate, D. F. (2013). Weight loss social support in 140 characters or less: use of an online social network in a remotely delivered weight loss intervention. *Translational Behavioral Medicine*, 3(3):287--294.

Umberson, D. and Montez, J. K. (2010). Social relationships and health a flashpoint for health policy. *Journal of health and social behavior*, 51(1 suppl):S54--S66.

van der Zwaan, J., Dignum, V., and *et al.* (2012). A conversation model enabling intelligent agents to give emotional support. In *Modern Advances in Intelligent Systems and Tools*.

van Dongen, S. (2008). Graph clustering via a discrete uncoupling process. *SIAM J. Matrix Analysis Applications*, 30(1):121–141.

Verheijden, M., Bakx, J., Van Weel, C., Koelen, M., and Van Staveren, W. (2005). Role of social support in lifestyle-focused weight management interventions. *European Journal of Clinical Nutrition*, 59:S179--S186.

Wang, M. L., Pbert, L., and Lemon, S. C. (2014). Influence of family, friend and coworker social support and social undermining on weight gain prevention among adults. *Obesity*, 22(9):1973--1980.

Wang, Y., Weber, I., and Mitra, P. (2016). Quantified self meets social media: Sharing of weight updates on twitter. In *Proceedings of the 6th International Conference on Digital Health Conference*, pages 93--97. ACM.

Wyatt, H. R., Grunwald, G. K., Seagle, H. M., Klem, M. L., McGuire, M. T., Wing, R. R., and Hill, J. O. (1999). Resting energy expenditure in reduced-obese subjects in the National Weight Control Registry. *Am. J. Clin. Nutr.*, 69(6):1189--1193.