Factors Influencing Recipe Promotion on Facebook

Master Thesis submitted in fulfilment of the Degree

Master of Science

in Management

Submitted to Assoc. Prof. DI Dr. Christoph Trattner, BSc

Pamina Strobl

1311051

Vienna, 13th June 2018

AFFIDAVIT

I hereby affirm that this Master's Thesis represents my own written work and that I have used no sources and aids other than those indicated. All passages quoted from publications or paraphrased from these sources are properly cited and attributed.

The thesis was not submitted in the same or in a substantially similar version, not even partially, to another examination board and was not published elsewhere.

Date	Signature

ABSTRACT

Obtaining a healthy and nutritious diet has proven to prevent various diseases, yet many populations suffer from high obesity rates that continue to increase. One potential cause is that people take their cooking inspiration from online recipes which are rather unhealthy. Authors have previously examined food-related interactions online, and how users can be nudged into the direction of healthier food consumption. Additional literature has made use of social networks, such as Facebook, to infer health statistics. This thesis takes the study of nutrition in the online domain further by exploring the factors that influence interactions with recipe advertisements on Facebook, in order to eventually encourage people to make better health-related choices. The results show that some factors, for instance images and state healthiness, show a difference in clicks or impressions on advertisements. Other factors, including recipe healthiness and user interests, do not show differences in interactions. Users responding to recipe promotions tend to be of older age, and predominantly female. Results also reveal that advertising budget is important when promoting recipes. These findings can be useful to governmental bodies and other actors, as they reveal which factors influence recipe interactions. That knowledge can then be exploited to promote a healthier diet.

ACKNOWLEDGEMENTS

First, I would like to thank my family for supporting me and making my studies possible. I especially want to say thank you to my father, who has sponsored the advertisements in this thesis. Without him, an outcome of this extent would not have been possible. I also thank my sister for consultation and advice, and my mom for constant support.

I also want to thank my supervisor, Dr. Christoph Trattner, who was always of help when I needed him. Even after leaving the country, he was always available via Skype and listened to my countless questions and provided valuable answers.

I would like to acknowledge Dr. Ingmar Weber who, together with Dr. Christoph Trattner, had the idea to explore this topic. He also helped me a lot with precise references for related work on the topic.

Also, a big thank you to Dr. Lidija Lalicic, who has helped me a lot with statistical questions when analyzing my data. I am glad for the hours she took to help me with my thesis.

I additionally appreciate my friends taking the time to fill out my survey, and also providing me with support and advice when needed.

Lastly, I want to thank my boyfriend for being patient and encouraging me with this thesis. Also, for lending his ear to me when questions about Facebook advertising came up.

TABLE OF CONTENTS

Affidavit I

Abstract III

Acknowledgements.. V

List of Figures XII

1 Introduction 1
1.1 Motivation 1
1.2 Objectives 1
1.3 Contribution..2
1.4 Thesis Outline 3

2 Literature Review 4
2.1 Background on Health Decreasing due to Malnutrition..4
2.2 Studies on Nudging People into the Direction of a Healthier Diet 6
2.3 Studies on Online Interactions with Food 9
2.4 Exploiting Search Engines and Social Media to Monitor Epidemiological Patterns..........15
2.5 Summary, Differences to Previous Work & Contributions.......................................23

3 Methodology 25
3.1 Selection of Recipes 25
3.2 Selection of Images 30
3.3 Selection of Target Group..30
3.3.1 Unhealthy Interests..32
3.3.2 Healthy Interests..32
3.3.3 Interest Targeting Options 33
3.4 Selection of Study Site 34
3.5 Research instrument 37
3.5.1 Facebook Page..38
3.5.2 Campaign..39
3.5.3 Ad Sets 40
3.5.4 Ads 41
3.6 Statistical Analysis 43

4 Results 47
4.1 RQ1: What is the general response to advertisements promoting online recipes?.........47
4.2 RQ2: To what extent does recipe healthiness influence the interaction of users with advertisements?..49

4.3 RQ3: To what extent does the image used in a recipe advertisement influence the user's interaction? 50
4.4 RQ4: To what extent do user interests play a role in interactions with the recipe advertisement? 52
4.5 RQ5: To what extent does state healthiness play a role in the selection of recipe advertisements? 52
4.6 RQ6: How do reactions to advertisements differ among user characteristics? 56

5 Discussion & Limitations 62
5.1 Summary 62
5.2 Discussion 63
5.3 Limitations 68

6 Conclusion 70
6.1 Contribution to knowledge 70
6.2 Implications for relevant stakeholders 70
6.3 Future research 71

7 Bibliography 72

8 List of Abbreviations 81

Appendices 82
Appendix 1: Survey for Recipe Image Selection 82
Appendix 2: Demographic Data for high and low obesity prevalent states 83
Appendix 3: Advertisements as they appeared to targeted users 84
Appendix 4: Statistical Analyses of RQ1 85
Appendix 5: Statistical Analyses of RQ2 86
Appendix 6: Statistical Analyses of RQ3 87
Appendix 7: Statistical Analyses of RQ4 89
Appendix 8: Statistical Analyses of RQ5 90
Appendix 9: Statistical Analyses of RQ6 98

LIST OF TABLES

Table 3-1 Table of Nutritional Properties of Healthy & Unhealthy Pancakes with the matching health score 28

Table 3-2 Means and Standard Deviation for Ratings of images on “Healthy Pancakes” from the Survey 29

Table 3-3 Means and Standard Deviation for Ratings of images on “Unhealthy Pancakes” from the Survey 29

Table 3-4 Unhealthy Interests on Facebook and their Esimated Audience Size 31

Table 3-5 Healthy Interests on Facebook and their Estimated Audience Size 33

Table 3-6 Target Audiences for Healthy vs. Unhealthy Interests 40

Table 3-7 Eight Facebook Advertisements and their properties 42

Table 4-1 Advertisement Outcome of the recipe promotions in terms of various indicators 47

Table 4-2 Mean Ranks of Differences in CTR between Ads....48

Table 4-3 Mean Ranks of Differences in Impressions between Images in Healthy Recipe....51

Table 4-4 Mean Ranks of Differences in Impressions between Images in Unhealthy Recipe....51

Table 4-5 Descriptives for Anova on CTR in Regions including Mean and Standard Deviation ..56

Table 4-6 Descriptives for Anova on Impressions in Regions including Mean and Standard Deviation 57

Table 4-7 Descriptives for ANOVA on CTR between Age Groups including Mean and Standard Deviation 59

Table 4-8 Kruskal-Wallis Ranks on Impressions between Age Groups 59

Table 4-9 Kruskal-Wallis Ranks on CTR between Gender groups 60

Table 4-10 Kruskal-Wallis Ranks on Impressions between Gender groups 61

Table A-1 Demographic Data for the 5 States with the highest and lowest obesity rates 83

Table A-2 four Advertisements as they appeared to the two targeted user groups 84

Table A-3 Kruskal-Wallis on differences between all advertisements in respect to CTR............85

Table A-4 Kruskal-Wallis on differences between advertisements in respect to Impressions ...85

Table A-5 Mann-Whitney U test on differences between the healthy and unhealthy recipe advertisements in terms of CTR...86

Table A-6 Mann-Whitney U test on differences between the healthy and unhealthy recipe advertisements in terms of Impressions...86

Table A-7 Independent Samples t-test on differences between the appealing and the unappealing image in terms of CTR 87

Table A-8 Mann-Whitney U test on differences between the appealing and the unappealing image in terms of Impressions 87

Table A-9 Independent Samples t-test on differences between the appealing and the unappealing image in terms of CTR on the healthy recipe 87

Table A-10 Mann-Whitney U test on differences between the appealing and the unappealing image in terms of Impressions on the healthy recipe...88

Table A-11 Independent Samples t-test on differences between the appealing and the unappealing image in terms of CTR on the unhealthy recipe 88

Table A-12 Mann-Whitney U test on differences between the appealing and the unappealing image in terms of Impressions on the unhealthy recipe...88

Table A-11 Mann-Whitney U test on differences between the healthy and the unhealthy interest group in terms of CTR 89

Table A-14 Mann-Whitney U test on differences between the healthy and the unhealthy interest group in terms of Impressions...89

Table A-15 Correlation Analysis on Diabetes Prevalence and CTR 91

Table A-16 Correlation Analysis on Diabetes Prevalence and Impressions 93

Table A-17 Correlation Analysis on Obesity Prevalence and Impressions 95

Table A-18 Correlation Analysis on Obesity Prevalence and CTR...97

Table A-19 ANOVA on differences in CTR and the user’s region in the advertisements 98

Table A-20 ANOVA on differences in impressions and the user’s region in the advertisements99

Table A-21 ANOVA on differences in CTR and the user’s age in the advertisement 100

Table A-22 Post hoc Scheffe test on differences in CTR and the user’s age in advertisements 100

Table A-23 Kruskal-Wallis test on differences in impressions and the user’s age in the advertisements 101

Table A-24 15 Post hoc Mann-Whitney U tests on differences in impressions and the user’s age in the advertisements 103

Table A-25 Kruskal-Wallis test on differences in CTR and the user’s gender in the advertisements 104

Table A-26 3 Post hoc Mann-Whitney U tests on differences in CTR and the user’s gender in the advertisements 104

Table A-27 Kruskal-Wallis test on differences in impressions and the user’s gender in the advertisements 105

Table A-28 3 Post hoc Mann-Whitney U tests on differences in impressions and the user’s gender in the advertisements...105

LIST OF FIGURES

Figure 1-1 Conceptual Model of the Research Design 3

Figure 2-1 Using tailored health communications along the behavioural pathways to colorectal cancer screening 5

Figure 2-2 Top-10 features in each of the the 3 studies 7

Figure 2-3 Distributions of Internet recipes in terms of WHO and FSA health scores.................9

Figure 2-4 FSA criteria of Recipe Categories 10

Figure 2-5 Popular Categories with the highest and lowest percentages of recipes provided by men 12

Figure 2-6 Example Instagram posts with their tags, matching canonical food names, and their derived nutritional profile 12

Figure 2-7 Caloric value of foods mentioned in tweets versus obesity rates............................14

Figure 2-8 Sample advertisement promoting smoking cessation aimed at users of the Bing search engine...16

Figure 2-9 Facebook marker interests for tracking tobacco use, obesity, and diabetes, along with placebo interests..21

Figure 3-1 Diagram of the Advertisement Structure for Recipe Promotions on Facebook 25

Figure 3-2 Homepage of the Allrecipes Website 26

Figure 3-3 Unhealthy Pancake Recipe with rating, reviews and photos...................................27

Figure 3-4 Healthy Pancake Recipe with rating, reviews and photos.......................................27

Figure 3-5 Increasing Obesity Rates in the States of the USA from 1994 to 2014 35

Figure 3-6 Increasing Diabetes Rates in the States of the USA from 1994 to 2014...................35

Figure 3-7 Scatterplot of Diabetes vs. Obesity Prevalence rates in the States of the USA, 201436

Figure 3-8 Facebook Page “Best of Allrecipes” Layout that was used for recipe advertisements38

Figure 3-9 Facebook Ads Manager which was used to collect data from the eight placed advertisements 39

Figure 3-10 Creation Of an Advertisement – Choosing of the Ad Creative 41

Figure 3-11 Possible Options for Group comparison 44

Figure 4-1 Scatterplots of CTR and Obesity Prevalence and their corresponding correlation coefficient 53

Figure 4-2 Scatterplot of CTR and Diabetes Prevalence and the corresponding correlation coefficient 54

Figure 4-3 Scatterplots of Impressions and Diabetes Prevalence for Ad 5 & 6 and correlation coefficients 55

Figure 4-4 Scatterplots of Impressions and Diabetes Prevalence for Ad 7 & 8 and correlation coefficients 55

Figure 5-1 Average Click-Through Rate 63

Figure 5-2 Audience Overlap 65

Figure A-1 Pancakes Survey distributed via Facebook as it appeared to people 82

1 INTRODUCTION

1.1 Motivation

It has become apparent over the past years that health conditions of individuals in the United States of America are significantly getting worse. The World Health Organization (2003) reports that obesity prevalence, diabetes and heart diseases are rapidly increasing. In proportion to other diseases, the institution predicts the non-communicable disease burden to be increasing by 57% by the year 2020 (World Health Organization, 2003). Although various habits could be the underlying cause of this, studies have shown that living a healthy lifestyle and eating a nutritious diet can help to prevent this growing trend (Ornish et al., 1998). Cooking is a practice that populations have undergone for centuries, however what has changed in recent years is the source of inspiration, which no longer stems solely from traditional cookbooks. Individuals have expressed an increased interest in online recipes, as they provide a quick and easily achievable guide for at-home-cooking (Cunningham & Bainbridge, 2013). With millions of recipes circulating the Internet, research reveals that most of them are not healthy (Trattner & Elsweiler, 2017). Some users are not even able to judge which recipes fit in the category of healthy food. This puts individuals at risk of their own wellbeing, because a nutritious diet is considered of utmost importance for overall health (World Health Organization, 2003). Nudging people into the direction of healthy recipes and therefore a nourishing diet is one way to prevent the decreasing health status of the general population, with many studies like Elsweiler et al. (2017) and Yom Tov et al. (2016) already focusing on that goal. People looking for cooking inspiration have multiple platform possibilities including popular websites such as allrecipes.com with about 85 million users (Allrecipes.com, 2017). In addition to recipe websites, social media tools like Facebook, Instagram and Twitter connect users and are an outlet for sharing inspiration (Kamal et al., 2010). Those platforms can be used to stimulate people to consume wholesome foods, rather than unhealthy ones. With two billion users monthly, customer reach on Facebook, for instance, has a considerable size (Facebook, 2018e). While multiple businesses, pages and individuals share their favorite recipes, health advocates can potentially use it to promote a healthier way of eating. Pointing populations into a direction that will result in nutritious home cooked meals is a major step for governmental bodies, health industries and various other stakeholders.

1.2 Objectives

The aim of the thesis is to analyze the behavior of users who interact with recipes online. The main goal is to explore the factors that influence recipe selection in a social media setting, in order to later on exploit this information for promoting a healthy diet. Such factors include recipe healthiness, image, interests and health on a state level. In order to examine this, the thesis uses a social media advertising tool, the Facebook Advertising API. Through this tool it is within reach to see which kind of users click on which type of recipe promotions. This thesis was driven by the following research questions (RQ):

- **RQ1:** What is the general response to advertisements promoting online recipes?
- **RQ2:** To what extent does recipe healthiness influence the interaction of users with advertisements?
- **RQ3:** To what extent does the image used in a recipe advertisement influence the user's interaction?
- **RQ4:** To what extent do user interests play a role in interactions with the recipe advertisement?
- **RQ5:** To what extent does state healthiness play a role in the selection of recipe advertisements?
- **RQ6:** How do reactions to advertisements differ among user characteristics?

1.3 Contribution

Previous research by Trattner & Elsweiler (2017) has suggested that healthy recipes are less likely to be cooked than unhealthy ones, but this has yet to be explored in a social media setting. Many articles have used Facebook audience estimates before, for instance Fatehkia and colleagues (2018) and Araújo et al. (2017), but thus far Facebook advertising has not been used as a tool to examine cooking patterns. Clicks on advertisements are a way to measure actual user interaction, rather than just audience estimates by the platform. Using Facebook advertisements to reach out to a large consumer group will establish what kind of user interacts with certain recipes, and whether or not those users are affected by attributes of the advertisements as well. Knowing how to promote healthy eating among large population groups is important, as this will become a predominant issue in the future. A conceptual framework about the contribution of this thesis is depicted in Figure 1-1, where it can be seen that user's reactions to the eight advertisements differ depending on various factors. Differences in this behaviour are analyzed statistically.

The thesis is relevant because knowing the significance that factors such as recipe healthiness, image attractiveness, user interests an state health statistics have in food promotions helps stakeholders know how to advertise more efficiently. Once it is established which factors influence recipe advertisement outcomes, health researchers and marketers can develop an innovative strategy on how to promote healthy eating to populations that are in need of it.

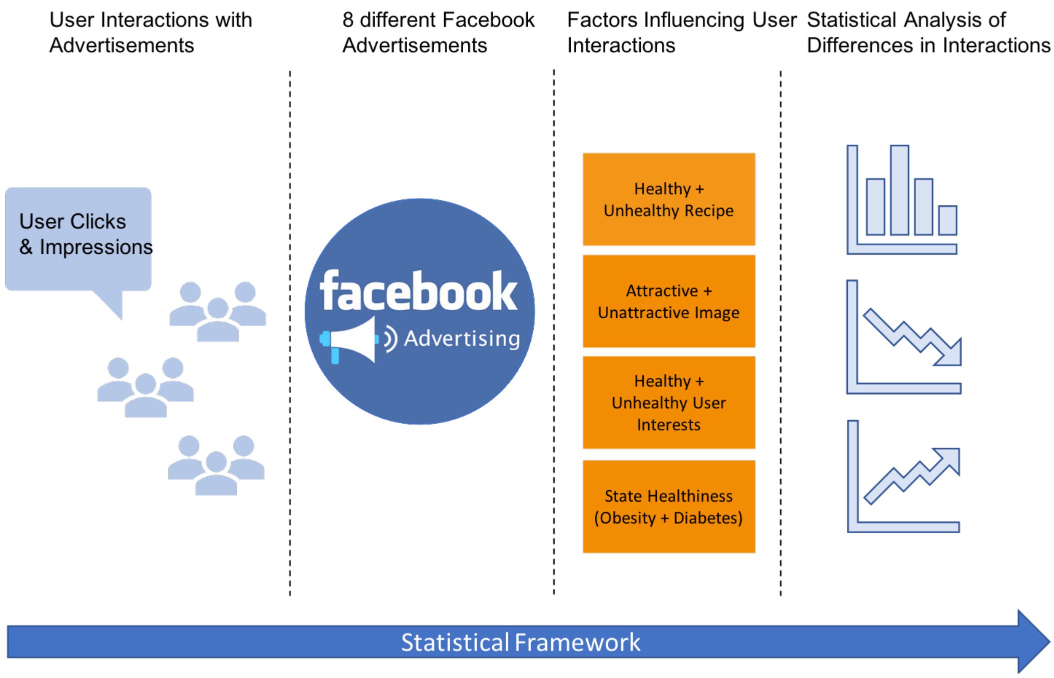


Figure 1-1 Conceptual Model of the Research Design

Picture Source: Murphy, 2015

1.4 Thesis Outline

Various chapters will explain the research undertaken. Chapter 2 focuses on work from authors who have previously done research in the field of the World Wide Web in connection to health-related activities or population studies. It reviews several important publications that have led to the choice of research in this thesis. Chapter 3, which explains the methods used throughout the thesis, shows why and how the advertisements for healthy eating promotion were created. The chapter includes the application of a fairly new tool, namely Facebook advertising. This can be used in order to gain insight into user's behavior in regard to online cooking and recipes. Statistical analyses of this data, like Pearson's correlation and t-test, as well as Mann-Whitney U's test and others then unveil which advertisements are preferred, and by which type of users. Chapter 4 reveals the results of the statistical analysis of the advertisements. The discussion in Chapter 5 covers a summary and insights into the results shown in the previous chapter. Limitations are also discussed. The final Chapter 6 covers concluding remarks that explain why the research is important to the study field of health and the Web. Lastly, further suggestions for research are implied.

2 LITERATURE REVIEW

Exploring how the previously mentioned factors influence the interaction with online recipes is a topic that has been investigated with other web-based tools before. Additionally, authors have already conducted similar research in the field of the Internet in combination with general health to explain and improve behavior.

The first section of this chapter covers literature that demonstrates why the declining health trend, mostly because of malnutrition, is a matter to be addressed. The latter sections explain how this issue can be solved by employing online tools like recommender systems. Authors have indicated that those can successfully be used to induce better cooking behavior. Other literature focuses on online interactions with food, and how those can be used to explain real-life health statistics. The last section shows how authors made use of search logs, the Web and even social media to monitor public health. Facebook, for instance, is increasingly being used for this purpose.

2.1 Background on Health Decreasing due to Malnutrition

The World Health Organization (WHO) has explicitly pointed out that with growing urbanization and globalization, obesity rates and chronic diseases are as high as never before. This is being caused mostly by malnutrition and lack of physical activity. In their report, it is stated that the key to effectively communicate a healthier lifestyle to people is to “create awareness, improve knowledge and induce long-term changes in individual and social behaviors --- in this case consumption of healthy diets and incorporating physical activity for health” (World Health Organization, 2003). “Well targeted communication” is also one of the main factors that is necessary to promote a healthier lifestyle, including a well-balanced diet (World Health Organization, 2003). The World Health Organization has deemed it appropriate to carry out more research regarding food consumption patterns, as well as whether consumers will then change their diets to a healthier version. According to them, there is a need to “change people’s behaviour towards adopting healthy diets and lifestyles, including research on the supply and demand side related to this changing consumer behaviour” (World Health Organization, 2003).

Coronary heart disease is an example of a disease which is brought about mostly from an unhealthy lifestyle. Ornish et al. (1998) have investigated this through a randomized controlled experiment, where one control group and one experimental group with an intensive change in lifestyle were put under trial. This lifestyle change included a change in nutrition. They came to the conclusion that in the group that had undergone an intensive change in their lifestyle there was more regression of the disease than in the control group. The control group had more cases of cardiac events and the coronary atherosclerosis that was investigated in this

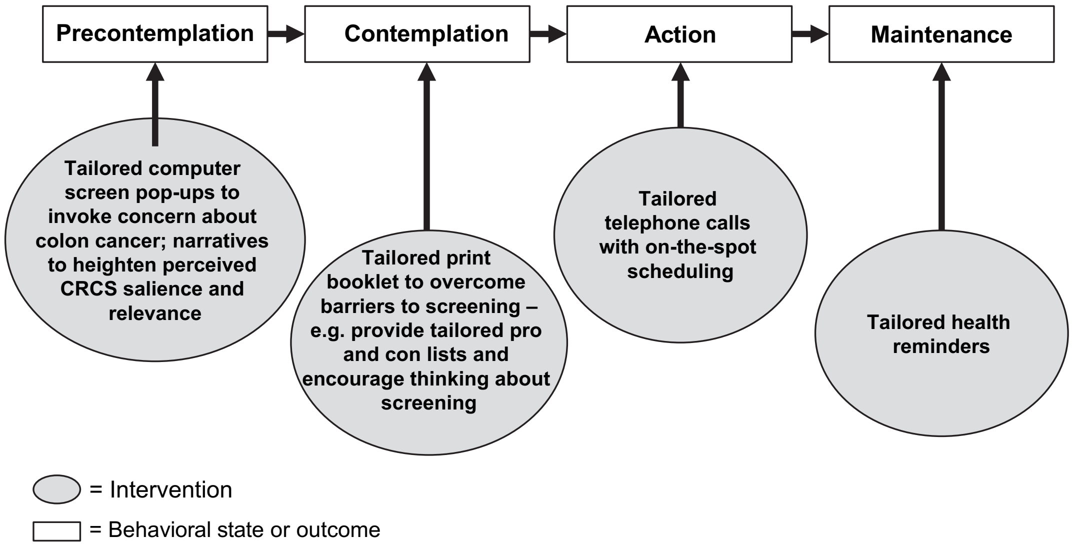


Figure 2-1 Using tailored health communications along the behavioural pathways to colorectal cancer screening

Source: Rimer & Kreuter (2006)

group progressed further than the experimental group. Their experiment proved that long term lifestyle changes can, in fact, reverse and surpass health issues (Ornish et al., 1998).

Food choices are options given to individuals multiple times daily. How people make those choices is up to them, but there are some factors that can influence them. According to the European Food Information Council (2006), the costs and time constraints of food are mostly the reason why people do not consume more nutritious meals. A lack of cooking skills and knowledge is often present, which is why lower income families do not know how to prepare food out of basic ingredients. According to the stages of change model, an intervention for dietary change is best made when the consumer is at a certain stage of making a food choice. This can be the “pre-contemplation, contemplation, preparation, action and maintenance” stage (EUFIC, 2006).

Knowing the stages of change model, Rimer and Kreuter (2006) studied tailored health communications (THC) and their advancement. THC is, ultimately, a method to reach particular demographic groups with effective messages relating health. Primarily, this was done through magazines sent per post or other types of traditional advertising. This tailoring, according to them, has positive effects on behavioral chance. Some effects would be that firstly, people get targeted only with information according to their interests and needs. What enhances the motivation of individuals to process this health information is the channel of delivery, as well as the type of content, which can be targeted to that which the individual prefers. This may be the biggest chance to induce the favored behavioral change. The main goal of THC is that “behavior change occurs through increasing motivation to process information” (Rimer & Kreuter, 2006). Figure 2-1 shows ways that tailored health communications can target individuals of each stage of change. The tailored intervention can be targeted according to “individual’s access to health care, their health care needs, or personal preferences” (Rimer & Kreuter, 2006). At the time of writing in 2006, the ways to communicate with individuals consisted of actions such as pop-up messages, phone calls or other outdated activities. In the year 2018, one major possibility would be Facebook advertising, as demonstrated in this thesis. Operational tasks to deliver those THC are, for instance, “choosing credible sources, developing a message strategy, defining the appropriate sources of data, and determining the settings and/or channels for optimal communication delivery” (Rimer & Kreuter, 2006).

The studies below serve as a background of how this problem of declining health can slowly be solved by making use of up-to-date research tools and practices.

2.2 Studies on Nudging People into the Direction of a Healthier Diet

The Web often uses recommender systems because they allow for personalised user experiences. Preferences of users and their likes or dislikes, in addition to feedback, are collected in order to make recommendations in various instances (Aggarwal, 2016).

In order to promote healthy nutrition, the following study was done for improving healthy recipe recommendation and leading people to employ a healthier lifestyle with different recipes focused on food choice biases. Using data from Allrecipes, Elsweiler et al. (2017) address the problem of recipes being rated the highest when in fact those are the unhealthiest ones. They aim to show the improvement of recommendations with different attributes. According to Gatti et al. (2014) “accurate wording is essential in persuasive verbal communication” which in this context means that the wording of recipes, like the use of positive adjectives, can indicate whether or not users will be likely to cook it (Gatti et al., 2014). Additionally, other factors play a role in how well a recipe recommendation is liked and accepted. During their research, Elsweiler et al., (2017) focus on finding recipes that are comparable in ingredients and similar in ratings, however do show a difference in nutritional properties. Those nutritional properties are again put in the categories of healthy or unhealthy according to the approach used in the paper above by Trattner and Elsweiler (2017) mentioned below. An example of recipes replaced would be “Ranch Crispy Chicken” and “Marinated Ranch Broiled Chicken”. In a survey, participants were given those recipes and asked to rate them and also give their opinions on their fat content. Most participants could not reliably differentiate between the fat recipes or leaner ones. Further testing showed that the image of a recipe, the ingredient list and the title has an influence on the bias of those specific recipes. By the means of those characteristics, participants could not distinguish which dishes had a higher fat content. The research question of Elsweiler et al. (2017) that will be important for this thesis is what kinds of characteristics make a user decide for an unhealthy recipe, rather than a healthy one. This was analysed by using machine learning techniques. The sets of characteristics were title, image, ingredients,

	Study 1		Study 2		Rand. Sample (rating)	
Rank	IG	Feature	IG	Feature	IG	Feature
1	.0933	IMG:contrast1	.0743	NUT:fat1	.1018	POP:sent2
2	.0829	IMG:brigthness1	.0634	IMG:contrast2	.1016	POP:sent1
3	.0719	IMG:entropy1	.0573	IMG:colorfullness1	.0679	IMG:colorfullness1
4	.0707	POP:rating2	.0568	NUT:cal1	.0609	NUT:fat2
5	.0703	IMG:entropy2	.0542	NUT:satfat1	.0605	NUT:cal1
6	.065	POP:sent2	.0512	NUT:fat2	.0562	POP:book1
7	.0612	POP:book2	.0484	NUT:salt2	.0549	POP:book2
8	.0568	NUT:cal2	.0454	IMG:entropy1	.0430	IMG:sharpness1
9	.0551	IMG:colorfullness2	.0417	ING:charCount2	.0361	POP:ratings2
10	.055	POP:ratings1	.0390	IMG:entropy2	.0344	NUT:satfat2

FIGURE 2-2 Top-10 features in each of the the 3 studies

Source: Elsweiler et al. (2017)

popularity and appreciation and finally nutrition. The features extracted from the title were length, text entropy and sentiment. The features for images were low level characteristics such as the sharpness, brightness, colourfulness, contrast, as well as the entropy. For the ingredient features, the numbers of ingredients and also the words to describe them were used. Average ratings of the recipes and the sentiment of the comments was used for the features in the popularity characteristics. Lastly, the nutritional features were calories, fat, saturated fat, sodium and sugar per 100g contained in the recipes, including also the overall FSA scores of a recipe. In three experiments, for instance, their findings indicate that the image features had a great influence on a user's recipe selection. The title features of recipes performed rather poorly compared to image and popularity features. Figure 2-2 shows which set of features performed the best according to information gain, where image features appear to be important in the first study. As seen on Figure 2-2, "nutritional features help most in the second study, whereas for the Allrecipes.com sample the most discriminative features are spread across the popularity, nutritional and image sets" (Elsweiler et al., 2017). In general, the findings of the classification experiments show that firstly, users tend to choose recipes that have a high fat content. Secondly, the choice often falls on recipes that are popular with other users. The visuals, including all the image features, play a big role in recipe choice. Their last research question answers whether it would be possible to nudge people into the direction of choosing a healthier recipe over an unhealthy one. They proved that this is possible, as the image features show that consumers are generally visual driven, and it is therefore possible to manipulate them into choosing a better recipe (Elsweiler et al., 2017). This thesis later on also investigates if image plays a role in recipe choice.

As a pre-requisite to recommending healthy meal plans, Harvey et al. (2012) have investigated what kind of foods correspond to users tastes. People are always surrounded with a great variety of cooking choices, so finding the one that is best suited for them on a health basis, but also appeals to them, is important. In their study, a recommender system approach was used where consumers could rate the internet-sourced recipes and subsequently give a reason to their rating. Those reasons could explain why a recipe did or did not fit their diet or taste. The various reasons were provided by the researchers, with the outcome that it is very complex to figure out the reasons of why meals were likely to be cooked by users. One important finding explains that a factor that influences the rating process is the ingredients that need to be used in a recipe. Not only single ingredients, but also combinations of ingredients can have an influence on the data collected. Another suggestion based on their findings was that "recipes could be assigned a healthiness score based on nutritional guidelines from health experts and learn which group a user belongs to based on the way they rate recipes with high or low health scores" (Harvey et al., 2012). This way, the recommendations would know more precisely which kind of recipes users prefer. However, their long-term goal is to build recommender systems that lead users to consume healthier recipes, rather than unhealthy ones (Harvey et al., 2012).

It is important for the physical condition of people worldwide that those recommender systems do not only recommend what people like and dislike, but also focus on recommending medical solutions or other healthcare information. Schäfer et al. (2017) have investigated the use of recommender systems and why the traditional system is not the same as a health recommender system (HRS), as "ratings given by users do not necessarily reflect the actual intent of the users" (Schäfer et al., 2017). They explain that a user that might like one kind of food, for example ice cream, may actually be in need for recommendations on diabetes friendly options. Medical utility functions could be included in HRS, such as "treatment duration" or "pain relief". A few concepts have already made use of recommendations based on health awareness. Some of those concepts include personalisation based on health records, empowerment and persuasion of users and also the medical evaluation of patients and interventions on their lifestyle. Patients would, for HRS to work well, need a better profile with more information on their behaviour, health statistics, lifestyle changes and others. HRS could also aim to improve health and comfort of patients throughout cycles of a disease with "disease progression modelling". Through health aware recommendations, many challenges come up for the systems, the users and also evaluation challenges. Once those are solved, society can move towards "digital health assistants or medical advisors" (Schäfer et al., 2017).

Because not all recipe platforms are modern and do contain nutritional properties, it is important for the wellbeing of users to find a mechanism to estimate them. In order to estimate nutritional properties of recipes and further be able to use recommender systems accordingly, Müller et al. (2012) have done research by focusing on the detailed values of ingredients used in the recipe. They then matched the ingredients to 91% of all the recipes from their chosen recipe platform "Kochbar" and therefore were able to predict its nutritional values (Müller et

	Total (Percentage)		Total (Percentage)
WHO score	Recipes n =58,263	FSA score	Recipes n =58,263
0	3319 (.06)	4	2309 (.04)
1	22,009 (.38)	5	4305 (.07)
2	17,403 (.30)	6	8012 (.14)
3	8977 (.15)	7	6834 (.12)
4	4211 (.07)	8	8613 (.15)
5	1767 (.03)	9	11,068 (.19)
6	498 (.01)	10	10,950 (.19)
7	79 (0)	11	5359 (.09)
		12	813 (.01)

FIGURE 2-3 Distributions of Internet recipes in terms of WHO and FSA health scores

Source: Trattner & Elsweiler (2017)

al., 2012). Predicting these nutritional properties may help people to understand better what they consume when cooking with online recipes.

One further randomised-controlled trial inspects food choices and the willingness to exercise by taking a look at online advertisements once more. Yom-Tov et al. (n.d.) direct their research to the growing research field of trying to prevent diseases before they happen. Making use of Bing ads, they targeted users that looked for search terms that indicated low levels of sports or people with poor diet habits, living in the United States. Search terms like “High cholesterol” and “Plus size”, as well as “Exercise” were targeted. The maximum budget was set to US$1 for a click. The outcome of the experiment shows that the ads were clicked 1024 times, with an average age between 35 and 64 and more often female. A key result was that people who were exposed to their advertisements were more likely to perform a key word search based on health promotion than the ones from the control group. The paper introduces the possibility that those behavioural changes can be measured online. It shows that it is possible to target individuals based on their characteristics like shopping behaviour, previous search terms or even e-mail content. Based on this, stakeholders like health departments can effectively advertise health matters and reach people that are susceptible to change (Yom-Tov et al., n.d.).

2.3 Studies on Online Interactions with Food

While the World Health Organization has found that a healthier diet can prevent diseases, it is also important to know how this information can be translated to online recipes and applied by consumers to their diet. Trattner and Elsweiler (2017) have investigated the healthiness of online recipes by standards set from the Food Standard Agency and also the World Health

		Mean											
			FSA front of package label				User Interactions					Health scores	
Category	n	Energy (kCal)	Fat (grams)	Sat. Fat (grams)	Sugar (grams)	Sodium (grams)	Comment Sentiment	Num Bookmarks	Rating	Num Ratings	User Health Perception†	WHO score	FSA score‡
Desserts	11,317↑	331.48↑	16.27 ↑	7.27 ↑	27.92 ↑	0.21 ↓	1.67	298.59↓	4.27	19.35	$2.06^{(0)}$	1.61	$9.64^{(1)}$
Ingredients	2039	265.06↑	14.13 ↑	5.84 ↑	16.44 ↑	0.36 ↑	1.92↑	1913.21↑	4.57↑	133.66↑	$4.28^{(-15)}$	1.59	$9.06^{(2)}$
Dinner	1033↓	166.61	9.07	3.44	2.59 ↓	0.35	1.94↑	2553.92↑	4.53↑	163.28↑	$4.31^{(-15)}$	1.41	$8.43^{(3)}$
Holidays and events	11,185	218.42↑	11.33 ↑	4.52 ↑	12.62 ↑	0.28	1.76	526.6↑	4.39	31.81	$2.66^{(+1)}$	1.87	$8.38^{(4)}$
Trusted brands	1744	200.45	10.06	4.08 ↑	8.73	0.32	1.77	111.02↓	4.37	6.57↓	$3.13^{(7)}$	1.83	$8.2^{(5)}$
Bread	2972	261.86↑	9.95	3.53	12.72 ↑	0.35 ↑	1.7	438.66	4.29	32.37↑	$3.63^{(-4)}$	2.42	$8.18^{(6)}$
Meat and poultry	12,672↑	151.97	8.46	3.09	2.62	0.33	1.74	465.88	4.3	26.79	$3.47^{(-2)}$	1.62	$8.17^{(7)}$
Breakfast and brunch	2167	188.8	9.26	3.56	7.82	0.28	1.69	377.25	4.31	22.86	$4.16^{(-6)}$	2.11	$8.09^{(8)}$
Main dish	13,188↑	159.51	8.36	3.08	2.48 ↓	0.31	1.73	438.92	4.27	25.59	$4.22^{(-7)}$	1.77	$8.09^{(9)}$
Appetizers and snacks	4162	226.67↑	15.73 ↑	5.79 ↑	4.8	0.44 ↑	1.74	428.86	4.35	25.4	$3.03^{(+4)}$	1.82	$8.08^{(10)}$
US recipes	3556	185.89	9.76	3.52	8.3	0.36 ↑	1.65↓	313.67	4.32	16.1↓	$2.19^{(+9)}$	1.92	$8.08^{(11)}$
Grilling	1682↓	156.72	8.74	2.77	4.83	0.54 ↑	1.83↑	481.01	4.41↑	22.68	$2.84^{(+8)}$	1.64	$8^{(12)}$
Allrecipes magazine	842↓	190.79	10.08 ↑	3.84	9.27	0.33	1.86↑	1952.1↑	4.54↑	142.78↑	$4.22^{(-2)}$	2	$7.94^{(13)}$
Everyday cooking	22,657↑	187	9.69	3.71	8.66	0.28	1.73	506.92	4.32	31.74	$4.47^{(-5)}$	2	$7.97^{(14)}$
Quick and easy	1955	167.82	8.65	3.23	2.39 ↓	0.32	1.7	404.72	4.25↓	23.55	$3.25^{(+7)}$	1.83	$7.86^{(15)}$
Pasta and noodles	2692	186.21	8.62	3.28	2.79	0.27	1.68	388.21	4.21↓	22.53	$3.84^{(+5)}$	2.31	$7.82^{(16)}$
Fruits and vegetables	19,574↑	171.44	8.7	3.25	9.06	0.24 ↓	1.73	373.59	4.32	21.85	$6.34^{(-9)}$	2.15	$7.76^{(17)}$
World cuisine	7444	178.05	9.05	3.26	7.46	0.29	1.68	361.72	4.28	19.53	$4.59^{(-3)}$	2.16	$7.68^{(18)}$
Lunch	693↓	158.36	9.1	2.78	3.11	0.32	1.94↑	515.8	4.6↑	26.54	$3.94^{(+6)}$	2.07	$7.63^{(19)}$
Slow cooker	1283↓	121.26↓	5.66 ↓	2.17 ↓	3.67	0.3	1.6↓	709.98↑	4.18↓	37.16↑	$5.19^{(-2)}$	1.89	$7.6^{(20)}$
Seafood	3237	157.6	8.94	3.05	1.79 ↓	0.32	1.75	298.29↓	4.31	16.95↓	$5.50^{(-2)}$	1.9	$7.46^{(21)}$
Salad	3031	146.84	9	1.93 ↓	4.48	0.24	1.78	247.46↓	4.36	13.17↓	$6.00^{(-3)}$	2.33	$7.22^{(22)}$
Vegetarian	4889	159.09	8.47	3.01	5.95	0.26	1.66↓	417.68	4.22↓	23.87	$5.50^{(-1)}$	2.58	$7.15^{(23)}$
Side dish	4006	128.99↓	6.64 ↓	2.69	3.71	0.24	1.71	324.4	4.3	19.1	$3.84^{(-12)}$	2.58	$6.97^{(24)}$
Soups stews and chili	3605	82.93↓	3.89 ↓	1.59 ↓	1.65 ↓	0.22 ↓	1.69	323.19	4.32	20.12	$4.56^{(+5)}$	2.29	$6.87^{(25)}$
Drinks	1801	86.37↓	1.5 ↓	0.82 ↓	10.22 ↑	0.03 ↓	1.57↓	126.26↓	4.36	6.51↓	$2.88^{(+21)}$	2.51	$6.01^{(26)}$
Healthy	3175	107.83↓	2.34 ↓	0.56 ↓	6.77	0.2 ↓	1.65↓	340.03	4.21↓	17.97	$6.53^{(0)}$	3.43	$5.6^{(27)}$
All recipes	58,263	204.87	10.58	4.10	10.55	.31	1.70	295.05	4.29	17.72	4.10	1.94	8.13

Note: Top-5 values in respect to macro nutr. content (i.e. Fiber, Sodium, Fat,...) and user interactions marked with ↑, bottom-5 in the corresponding column highlighted with ↓.
† Superscripts denote differences in ranking when compared to the FSA ranking of the actual category. ‡ Superscripts denote category ranking in respect to the FSA score.

FIGURE 2-4 FSA criteria of Recipe Categories

Source: Trattner & Elsweiler (2017)

Organization. In their study, the recipes investigated came from the platform Allrecipes, where 60,983 of them were used and the standards of the WHO, as well as the Food standard agency (FSA) "traffic light system" was applied. In the WHO standards, the seven macro-nutrients that were the most important are used to determine whether a recipe is considered healthy through a scale of zero to seven, with "0 meaning none of the WHO ranges are fulfilled and 7 meaning all ranges are met "(Trattner & Elsweiler, 2017). In the traffic light system with four chosen macro-nutrients, green means healthy, amber a middle score and red means un-healthy. For a single metric to measure healthiness, the method used by Sacks et al. (2009) is applied. It is to "assign an integer value to each color (green=1, amber=2 and red=3) then sum the scores for each macro-nutrient resulting in a final range from 4 (very healthy recipe) to 12 (very unhealthy recipe)". As seen in Figure 2-3, there were only few recipes that could be rated with the highest health score. Even more unfortunate for health scores, Figure 2-4 shows that most of the recipes, according to the FSA standards, were in the red and amber score in the category of fat, saturated fat and sodium and just scored higher in sugar. The analysis shows that most of the recipes turn out to be unhealthy, with only a few taken from this recipe plat-form considered healthy by the WHO and FSA criteria. The findings were, among others, that through bookmarking and ratings, it is possible to tell which recipes users preferred; the re-sults being that "popular and highly-rated recipes are the ones which are the least healthy" (Trattner & Elsweiler, 2017). People were less likely to choose and interact with recipes that are deemed healthy and therefore most likely consume and cook unhealthier recipes (Trattner & Elsweiler, 2017).

The same health criteria are employed in a paper by Trattner, Elsweiler and Howard (2017) where nutritional properties of a recipe platform, ready meals from a supermarket and alternatively of a TV chef are compared. Howard et al. (2012) had previously analysed the nutritional properties of recipes used by TV chefs, whereas the research above describes how the Internet sourced recipes are analysed. The study found that out of the three compared, Internet recipes were the least healthy (Trattner et al., 2017a).

Kusmierczyk and Nørvåg (2016) explored the possibility of title words of a recipe and their nutritional values showing a relationship. For their experiment as well, Allrecipes served as a provider for 58 thousand recipes that had sufficient nutrient information. The titles then were filtered to leave the researchers with 4,679 words that were statistically analysed based on their unique nutritious value. Words and nutrients, as well as nutrients among themselves showed a correlation. Another experiment in their study looks at the title words of a recipe and tries therefore to predict its nutritional value with methods such as linear regression or gradient boosted regression trees. When, in addition to fat, sugars and sodium, other nutrients were known, the calories could be determined precisely in this experiment (Kusmierczyk & Nørvåg, 2016).

By comparing the "nutritional properties and the healthiness of uploaded and bookmarked recipes" from a pool of selected users and later on also studying how hobbies or cooking interests play a role in those choices, Trattner et al. (2017c) once more make use of the platform Allrecipes and come to insightful conclusions (Trattner et al., 2017c). Bookmarked recipes show less healthy nutritional properties than the ones that are uploaded by the same user, which may or may not be an indicator of wanting to portray a wholesome lifestyle in the online community. On the recipe platform, the users have the option to present specific cooking interests or hobbies on their profile. A pattern can be seen in regard to a correlation of cooking interests to recipe healthiness. According to the health scores by WHO and FSA, which are mentioned in a previous study above, the cooking interest "Kids" had extremely unhealthy scores for both standards. "'Vegetarian', 'Middle-Eastern', 'Indian' and 'Mediterranean'" and also "Healthy" are the cuisines that score the healthiest within the standards set by those two organisations. Fewer trends can be seen in regard to the hobbies, but some can be observed. According to the research, "'Biking', 'Hiking' and 'Boating' are associated with lower intake of energy, fat and carbs", "'Hunting' and 'Fishing' score high on protein and sodium" and "'feminine' hobbies such as sewing are associated with high fat, sugar and carbs, which is associated with baking" (Trattner et al., 2017c).

Not only do people with different hobbies or interests have divergent recipe preferences, but there are also many prejudices about gender in cooking. Women and men sometimes are

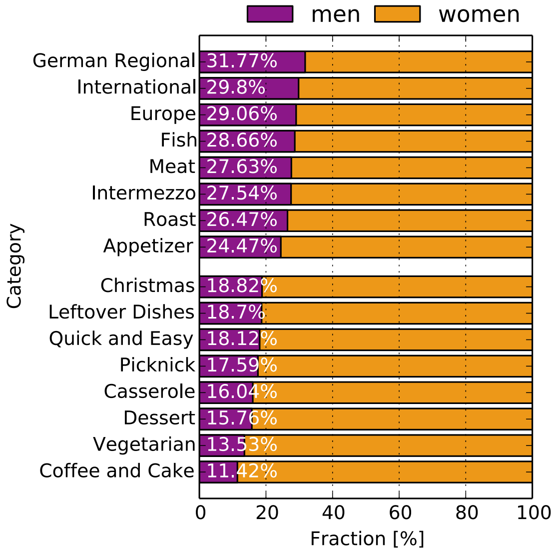


FIGURE 2-5 POPULAR CATEGORIES WITH THE HIGHEST AND LOWEST PERCENTAGES OF RECIPES PROVIDED BY MEN

Source: Rokicki et al., 2016

Post	Post tags	Canonical name(s)	Energy	Sugar	Fat	Chol.	Fiber	Protein
	butter, cakes, peanut, jelly, kellyloucakes, tea, decorating, cupcake	butter, cake, peanut, jelly, tea, cupcake	436.26	95.9	25.3	256	3.08	9.401
	healthyfood, meal, goodfood, foodgasm, carrots, vitamin, cucumber, veggies, foodisfuel, corns, lime, beetroot, nofilter, salad, potato, instafood, eatcleanmenu, apples, rich, fruits	carrot, cucumber, corn, lime, beetroot, salad, potato, apple	206.01	81.1	16.8	86	79	25.4
	strawberry, strudel, dessert, sweet_taste, jar, pastry	strawberry, strudel, pastry	322.91	21.3	80.2	114	10.2	4.202
	cajun, instagram, monday, foodporn, food, bestoftheday, instagramhub, oregano, cook, breakfast, instagood, instadaily, ig, bread, tomato, organic, iggers, egg, iphoneonly, iphonesia, morning, yum, iphone, fresh, spice	oregano, bread, tomato, egg	205.97	43.9	73.6	230	42.5	84.63
	food, dessert, lovelife, chocolate, cookie, delicious, tasty, raspberry, yummy, dough, dinner, pudding, treat, loveit, epic, pizza	chocolate, cookie, raspberry, pudding, dough, pizza	371.63	23.7	15.7	221	35.2	5.952

FIGURE 2-6 EXAMPLE INSTAGRAM POSTS WITH THEIR TAGS, MATCHING CANONICAL FOOD NAMES, AND THEIR DERIVED NUTRITIONAL PROFILE

Source: De Choudhury et al., 2016

linked to certain cooking behaviour, which in reality may or may not be true. Some hypotheses that research from Rokicki et al. (2016) analyses are “men are better cooks, men cook for impressing, women cook sweet dishes and men meat dishes, women use spices more subtly, men use more gadgets, and men are more innovative” (Rokicki et al., 2016). After crawling about 400 thousand online recipes, the insights of the study reveal that there is a difference in cooking behaviour between men and women when looking at nutritional facts. One example shows that men use more complex recipes and take longer, while women tend to cook less meat and use fewer spices. A classification experiment which was conducted shows that one can determine the gender based on how spices are used, the use of gadgets and the food type, where women tend to cook sweeter and men tend to have a higher meat affinity. Figure 2-5 shows that desserts, coffee and cake, for instance, are cooked a lot more by women than men (Rokicki et al., 2016).

Analysing the users of the social media tool Instagram, De Choudhury et al. (2016) conducted research on food deserts. Food deserts, as explained by the United States Department of Agriculture's Economic Research Service, are areas in the US that do not have sufficient access to transportation and suffer from a lack of retailers providing fresh and healthy groceries at a fair price. The language on Instagram was examined by assigning nutritional values to certain food-related words. Also, the approximate location of food deserts and non-food desert counterparts was identified. Examples of Instagram posts are portrayed in Figure 2-6. The first research question compares those two locations, and observes that high calorific food-related words, like "hamburger" are common in food deserts. Words like "spinach" are more usual in non-food deserts. By investigating the nutritional values, the second research question answers whether or not the dietary choices in food deserts seem to be less nutritional than in their counterparts. The areas examined consumed a higher sugar, fat and cholesterol level, but do not necessarily consume more or less calories. Another research question shows the result that ingestion language was an indicator of the dietary habits. The ingestion language, as well as the socioeconomic attributes and food deprivation status, can reveal with an accuracy of more than 80% if a tract is likely to be food deserts. The study determined that traces left by Instagram users are important to reveal dietary patterns and trends (De Choudhury et al., 2016).

Web usage logs can show patterns in online behaviour with food as well. The study by West et al. (2013) made use of "anonymised logs of URLs visited by users who consented to provide interaction data" (West et al., 2013). From the logs inspected, the researchers goal is to predict what kind of foods people consume. A first finding is that people shift their nutritional interests in the period of holidays to an unhealthier food consumption. They also found that people dieting had changed their search queries to certain words that revealed their weight loss interests, but also found that this shifted back to older habits after a certain number of weeks. A third result of this study shows a correlation in recipes high in sodium to hospital admissions in a certain time period in the state of Washington D.C, which shows that nutrition does correlate with health and online data could prove such a correlation. Those three findings indicate that online activity can effectively be used to predict physical conditions of a population (West et al., 2013).

Obesity prevalence is a factor where data is available all throughout the United States per county. This data from the CDC (2016) can determine in which counties obesity rates are high.

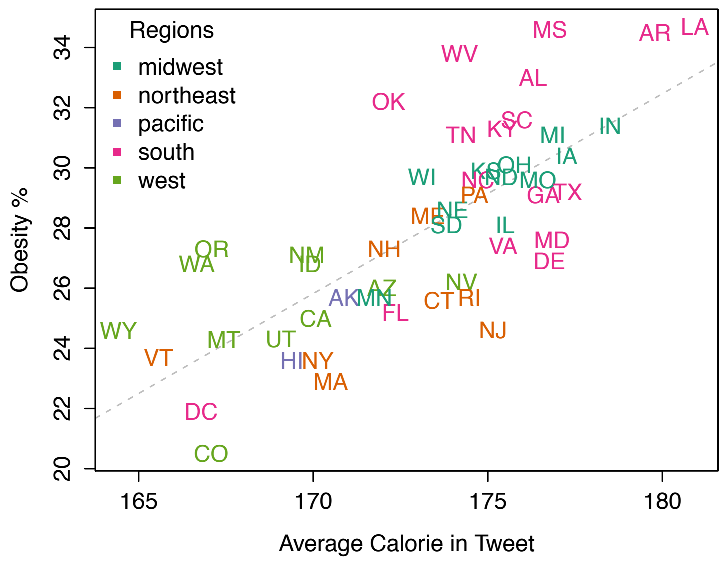


FIGURE 2-7 CALORIC VALUE OF FOODS MENTIONED IN TWEETS VERSUS OBESITY RATES.

Source: Abbar et al., 2015

A paper in which health trends can be seen throughout states in America by Trattner et al. (2017b) has measured "online activity through bookmarking to monitor obesity prevalence in the United States" (Trattner et al., 2017b). In their work, bookmarking activities of users are measured and subsequently correlated with the obesity prevalence rates in each state, an approach which is similar to the one this thesis uses in the last research question. They worked with the macro-nutritional properties of recipes that were bookmarked. The data set used for this was again Allrecipes, one of the most popular food platforms in the United States of America (USA). A Web crawler was used in order to determine which recipes were bookmarked by which kind of users, and the data set included "17,817,462 recipe bookmarks from 144,839 users" (Trattner et al., 2017b). The health criteria applied to the recipes was, as mentioned above, again taken from the FSA. Interestingly, the correlation analysis that was done between obesity levels and the FSA score with all the nutritional variables, "Fat, Saturated Fat, Sugar and Sodium" combined, shows that there is a correlation between the two variables. In addition to that, "temporal, geographical and nutritional relationships" can be observed throughout the entire data set (Trattner et al., 2017b). Bookmarking activities were able to explain the obesity levels in most of the states. The findings of Trattner et al. (2017b) again confirmed that online traces, in this case the bookmarking activities, can show significant correspondence with health data available.

Abbar et al. (2015) also investigated Twitter with their study of nutritional data called "You tweet what you eat". 210,000 users were observed to see whether their interests, demographics, and social networks could be linked to their dining experience shared on twitter. This was done through a set of keywords that were included in the collected 892.000 tweets. The food-related keywords were given nutritional properties from Internet sites that determined the dishes average calories. The food keywords, having a certain calorific value, was then correlated with the obesity rates from the CDC in order to spot a pattern. As visualised in a scatterplot in Figure 2-7, the Pearson correlation value of 0.77 showed that the food tweeted about can, in actuality, show a correlation in calories and obesity prevalence. One of their other findings was that areas with people who had a higher education level have a healthier diet, which in this case means less calories (Abbar et al., 2015).

A study by Mejova et al. (2015) also confirmed that Instagram and Foursquare, which is a location-based service whose API (Application Programming Interface) Instagram uses, can be used successfully to find a connection in user's behaviour and health data. Their findings present that there is a "relationship between small businesses and local foods with obesity, with these restaurants getting more attention on these social media" and also that when analysing Instagram pictures, the unhealthy dietary choices seemed to be socially preferred as seen through likes and comments on those pictures (Mejova et al., 2015).

Also using the recipe platform Allrecipes, Said and Bellogín (2009) tracked health in the USA per county through recipe interactions online, focusing on obesity prevalence. With a dataset of 54 thousand recipes and 8400 ingredients, their method is to take a combination of ingredients that were most commonly used and track which geographical region used a certain combination of ingredients the most. Taking into account the five counties with a low obesity prevalence and five counties with a high obesity prevalence, the research showed that the 20 ingredients used in high obesity prevalent counties, compared to the corresponding percentage in counties that have a lower obesity prevalence, do portray a difference in online interaction from users. It can be understood that it is possible to know, based on recipe interaction, if the user is from a high-risk (poor health) county or not (Said & Bellogín, 2009).

2.4 Exploiting Search Engines and Social Media to Monitor Epidemiological Patterns

One way to get an insight into health trends is online advertising through search engines, which can be used to make predictions and promote health matters. One particular paper by Yom-Tov et al. (2016), for instance, has its focal point on "Antismoking Advertising to Promote Smoking Cessation". Their research strives to test web-based advertisements in order to determine how best to promote quitting smoking. The mechanism they used was the Bing Ads system, where 10 advertisements were placed randomly that were created by a public health professional. Subsequently, the participant's post-advertisement behaviour was analysed to observe whether they followed up on smoking cessation activities, such as searching for information regarding this topic. The methodology follows a similar approach as the one in this thesis, by creating advertisements with different properties in order to measure clicks on each. The promotions contained various titles, a matching body and a link to an URL. Targeted were those people living in the United States who used the Bing search engine. As in most search

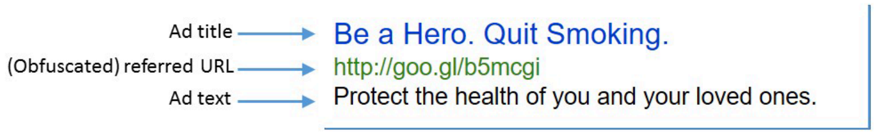


FIGURE 2-8 SAMPLE ADVERTISEMENT PROMOTING SMOKING CESSATION AIMED AT USERS OF THE BING SEARCH ENGINE.
Source: Yom-Tov et al. (2016)

engine advertising methods, the promotions appeared when users searched terms such as “smoking” or cigarettes”, or even more specifically “smoking causes black lungs”. Bing randomly generates which advertisement is shown to which individual, meaning the conductors of the study had no control over that. An example of the advertisements is portrayed in Figure 2-8. Each type of search term was then given a number to represent them, such as “0” representing generic terms such as “smoking” and “cigarettes”. A Cox proportional hazards model was used to see how likely it was that users would search for anti-smoking terms after having exposed to the advertisements. Outcomes showed which location of advertisements had the highest likelihood of generating subsequent IQSS (intention to quit smoking search). The top right of the page advertisements, for instance, were twice as likely to lead to this. Older people, according to the authors, were also more likely to follow up their search on quitting smoking. The results also show that content affects which gender and age respond. In one case, men’s likelihood of responding to empowering content was higher, while women responded well to health-related advertisements. In general, the study finds that targeting advertisements can improve effectiveness and some alterations may improve general public health of people. One limitation is that an actual change in smoking behaviour cannot be confirmed by simply placing advertisements. Other qualitative studies may help to address this limitation. This study is an important indicator of how digital traces can be used in order to examine health related data (Yom-Tov et al., 2016).

Similar to an above-mentioned study, web-based advertising is also used in a study by Yom-Tov et al. (2018). They investigated how to induce behavioural change in people making use of pro-anorexia Web content, as the issue of this disorder is becoming more predominant over the years. For this purpose, advertisements have been placed in order to examine if it can bring about behavioural change in those users through a randomised control trial. As with the advertisements in the study mentioned above, the authors placed advertisements targeting people searching for pro-anorexia content, which then lead to three websites that were randomly chosen. Those people’s behaviour post-advertisement exposure was again monitored. The location of the study was the United States of America. People that were targeted searched for terms like “Thinspo” or “Anorexia”. As the search engine works with a bidding system, bids for placing the advertisement randomly in the search engine in their case were placed between zero and US $0.99 for each keyword. Each ad expressed an attribute, which were rated by people from a crowd source and experts as well. The results show that the ads were clicked 217 times with a CTR of 0.85%, though it varied according to each advertisement. It showed that advertisements do have an effect on the search behaviour of individuals after they have seen them. The more people saw the ads, the higher the likelihood of them searching for treatment options was. People referred to a particular website, “MyProAna”, showed reduced self-harm and anorexia interests after being exposed to the promotion. The control group, on the other hand, increased this behaviour, which proves that advertisements in this direction can have positive health effects on people. The main limitation this study mentioned as well is that real life behavioural change cannot be predicted by those methods. However, effective advertising can change online search behaviour (Yom-Tov et al., 2018).

Nowadays, people leave traces in the Internet on almost every site they access. Many cases in literature have therefore shown that digital activities can tell about and even estimate public health issues. Research done by Ginsberg et al. (2009), for instance, uses search engine queries in order to predict influenza outbreaks. Multiple queries through a popular search engine, namely Google, were correlated with visits to physicians in the respective region of the United States. The level of weekly influenza activity could accurately be estimated by using this method, hence also making detection of epidemics possible (Ginsberg et al., 2009).

Not only search engines are used to monitor health. Social media networks are gaining popularity in research as well. The definition of a social network is a “network of individuals (such as friends, acquaintances and co-workers) connected by interpersonal relationships.” (Merriam Webster Online, 2018). As seen above, researchers in all kind of health fields are now looking into online tools in order to promote a behavioural change that induces health-aware actions. Further studies focus on specific social media tools to monitor and track health across countries.

Social media is a widely used tool nowadays, and includes platforms where people communicate with each other, such as Facebook, Twitter, Instagram and many others. An example for one model that portrays why people use social media sites is called “Use and Gratification Model” (Kamal et al., 2010). In this model, the motivators for users are, for instance, entertainment, social enhancement, connectivity and convenience. Additionally, users also interact on social media because they want to get information, provide information or even self-discovery (Kamal et al., 2010). Self-discovery and getting information may be important when contemplating personal health management. A theory on how to influence long term health changes, called the “social cognitive theory”, demonstrates that certain determinants, like the “ability to perform the behaviour needed to influence outcome” can influence health behaviour (Kamal et al., 2010).

Twitter, as well as other social media sites, is not only popular with users but is also gaining popularity for researchers. By inspecting the traces on social media, De Choudhury et al. (2013) try to detect the illness major depressive disorder in individuals. Through crowd sourcing, the assessments of patients diagnosed with the disorder were collected. The so called “crowd workers” took questionnaires to determine their depression level. 243 male and 233 female participants then allowed access to their Twitter feeds and were selected for the study. Furthermore, the participants data was collected for up to one year prior to the depression onset. As depressive behaviour, according to literature, often manifests itself by users being active at night, the researchers determined a day and night window for the activity on Twitter, and subsequently created an “insomnia index” for each user. Egocentric network measures, emotional state of users, linguistic style and depression language were all measures used in order to determine two types of classes, which include the depressed and non-depressed. Symptoms of depressed users were “lowered social activity, greater negative emotion, high self-attentional focus, increased relational and medicinal concerns, and heightened expression of religious thoughts” (De Choudhury et al., 2013). As a result, they proposed a model to predict depressive behaviour before the actual onset. Based on this, it would be possible to implement personalised alerts or information through specified systems (De Choudhury et al., 2013).

A similar paper has investigated the social media tool Twitter and its ability to estimate tobacco use, where topic modelling is used in order to track smoking in the United States (Prier et al., 2011).

Yet another study that makes use of Twitter to predict population characteristics was carried out by Fried et al. (2014). The study aims at demonstrating that food-related language, in this case hashtags used in tweets, can give insights into populations. In a period of about nine months, 3.5 million tweets were collected. The implemented prediction tasks are able to predict locations and also health signs in those geographic areas, such as diabetes and obesity rates. Diabetes could be predicted with a 68% accuracy, and obesity with an accuracy of 80%. Also, political interests can be predicted. This information may successively be used for a targeted marketing approach in the fields of health or others (Fried et al., 2014).

Another digital tool, namely Instagram, which is now owned by Facebook, has as well shown to be effective when tracing health patterns. In the study “Social Media Image Analysis for Public Health”, the authors Garimella et al. (2016) found that it is possible to get insights into county’s health data through user-provided and machine-generated tags. This means that images that are posted on Instagram can very specifically tell which population groups are affected by certain problems and predict a pattern in those (Garimella et al., 2016).

As seen above, advertising through social media platforms has gained popularity only in recent years. Facebook in particular lets users interact with advertisements and even “like” or “share” them with friends. They can now actively interact with what they see on the website. Research by Dehghani and Tumer (2015) indicates that advertisements can enhance brand image and equity, while also increasing the chance of consumers wanting to purchase an advertised product. The social media advertising platform is seen as a more fashionable way to promote products or brands, mostly because of the customisability (Dehghani & Tumer, 2015).

Although plenty of social media tools exist nowadays, Facebook has been growing as a social media platform ever since it was created. According to the platform, about two billion people use Facebook monthly (Facebook, 2018e). While formerly, most of the advertisements (Ads) were made through television or billboards, online marketing has increased in the past years. Eventually, Facebook made the decision to include a way to advertise on their platform as well, which is why the Facebook advertising tool was created. The platform prides itself on being able to target the "right people, capture their attention and get results". Brands and businesses are easily able to use the tool through the Facebook marketing API. All they have to do is set up a page for their business. A Facebook page is a site on Facebook with the main purpose to inform users about the business. Brands are advised to frequently update the content in order to connect with their online community. Facebook advertisements can target an audience "based on demographics, behaviours or contact information", they can take on different formats to be eye-catching, and are also able to work on multiple devices with any connection speed (Facebook, 2018e).

By investigating the efficiency of Facebook advertising in the Slovak market, Vejacka (2012) has found that the number of users and quality of data that can be collected through Facebook is enormous, with 300 million daily active users and over 900 million objects (groups, pages and events) in the year of 2012. While their study focuses on the former Facebook advertisements that were run on the side of a Facebook page, the findings are still relevant for today's marketing purposes. Google's tool "Adwords" proved to be more efficient in marketing than the former Facebook marketing tool, whereas Facebook was promising for reaching younger customers, creating targeted options and easy feedback collection (Vejacka, 2012).

Recent research by Fatehkia et al. (2018) makes use of Facebook audience estimates to predict population trends. There is no doubt that the connectivity through information and communication technologies brings about many benefits. An important developmental goal is, however, to create equal accessibility to both men and women. This paper aims to measure the global gender gap of mobile and Internet access through Facebook audience estimates. Those estimates are accessible to anyone who has a Facebook account. The research makes use of an offline dataset as well, which indicates the gender gaps from surveys. The ordinary least squares method is used to predict the outcomes and adjusted R squared, Pearson's correlation and errors are used to evaluate the performance of different models. With a correlation coefficient of 0.83, the outcome shows that Facebook estimates strongly correlate with the offline dataset's numbers, which means that they can be used accurately to represent the gender gap. An analysis on prediction capability is done on models using online data only, a mixture of online-and offline data and solely offline data. The findings here show that the strongest predictive model is the one combining online and offline data sources. However, the paper demonstrated that estimates taken from Facebook can valuably prove to be able to monitor the digital gender gap. Those predictions can also be made more frequently than with using offline sources (Fatehkia et al., 2018).

Another paper about online health monitoring also makes use of the audience estimates that Facebook provides. Mejova et al. (2018) recalls that Facebook has been used previously for recruitment of people, and that social media can be used as a tool to track health. Facebook advertisements make it possible to reach a wide range of people on the platform. The paper aims at making a connection with real-world health statistics to Facebook interests. Data from Facebook was collected from interests that were related to health conditions such as diabetes, food sensibilities, alcoholism and obesity. Some marker interests were taken for each health condition, for example “Alcoholics Anonymous” was an interest in alcohol. In addition to interests representing those health conditions, a placebo interest was taken. This placebo interest should have no relation to any of the conditions. “Fitness and Wellness" served as a baseline interest. Public health data, such as reports from the CDC and Census Bureau, served as an indicator of general health in the 50 states in the US where audiences were measured. Using Pearson’s correlation, the Facebook indices were correlated with the health indices with a significance level of 0.05. Results showed that for alcohol, some interests like “alcoholism awareness” are positively related with actual health statistics, but “Alcoholics Anonymous” is negatively related. They found that the r values are, however, similar to the placebo interest’s r value. Obesity and diabetes correlations were stronger, showing a correlation of r=0.74 for the “plus-size clothing” interest and real-world health statistics on obesity. Diabetes awareness also strongly correlated with US diabetes statistics. A linear regression model was also used to predict real-world health statistics. This model shows an adjusted “R2 of .533 for modelling Alcoholism, .712 for Obesity, and .790 for Diabetes”. With introducing demographic and financial information as the control variable, the models had an even better performance. Mejova et al. (2018) also looked at relationships between demographics and interests, which lead to the finding that plus size clothing is most popular with the African American community. Further interesting findings relate even more demographic variables to Facebook interests. They also ask themselves the question whether it is possible to understand why certain people have certain Facebook interests, which involves understanding Facebook’s algorithm, which is not public. So, although there are limitations to using audience estimates, and one cannot solely rely on them. What is an issue, for example, is the temporal difference between health statistics and data retrieved from the social media platform. Another limitation is that marker interests can be known and explored even more thoroughly, and not the whole population uses Facebook. To conclude, they state that the way to use and analyse data in this study introduces ways to design health-risk surveillance or health recruitment (Mejova et al., 2018).

Tobacco Use	
Smoking	30,000,000
Tobacco	20,000,000
Tobacco smoking	11,000,000
Lung cancer awareness	6,200,000
Cigarette	29,000,000
Hookah	10,000,000
Smoking cessation	7,500,000
union of all	77,000,000
Obesity	
Bariatrics	2,400,000
Obesity awareness	58,000,000
Plus-size clothing	29,000,000
Weight loss (Fitness And wellness)	81,000,000
Dieting	218,000,000
union of all	286,000,000
Diabetes	
Gestational diabetes	1,400,000
Insulin index	250,000
Insulin resistance awareness	1,700,000
Diabetes mellitus awareness	55,000,000
Diabetes mellitus type 1 awareness	3,200,000
Diabetes mellitus type 2 awareness	5,300,000
Diabetic diet	4,200,000
Diabetic hypoglycemia	280,000
Managing diabetes	960,000
union of all	60,000,000
Placebos/normalizers	
Facebook	863,000,000
Reading or Entertainment or Technology	1,278,000,000
Health Care	145,000,000
Fitness & Wellness	714,000,000

FIGURE 2-9 FACEBOOK MARKER INTERESTS FOR TRACKING TOBACCO USE, OBESITY, AND DIABETES, ALONG WITH PLACEBO INTERESTS

Source: Araújo et al., 2017

Making use of data from the Facebook advertising platform once more, another research effort demonstrates how one can calculate the demographic proportion of a population that is aware of schizophrenia. Saha et al. (2017) constructed an index that measured the awareness of the psychological disease schizophrenia and analysed it based on “US states”, “gender”, “age”, “ethnic affinity”, and also “education level”. The index shows that 1.03 percent of the population had a schizophrenia-related interest, with differences existing across all variables explored. The study portrays that Facebook advertisement audiences can be used to estimate interests in certain populations according to demographics (Saha et al., 2017).

Facebook Ads are also used when Araújo et al. (2017) completed a study that uses a similar method of research as the last research question of this thesis. Their research focused on “global lifestyle disease surveillance”, such as obesity, smoking and diabetes, and used the audiences of Facebook advertisements in their method of research. The paper introduces the interests of Facebook users as well as their age and gender. They measure those interests and correlate it with health data, which is done across 47 different countries. One interest is “obesity awareness”, which subsequently targets every Facebook user who is interested in this topic. They also used placebo interests that should not show a meaningful relationship between both variables, the WHO health data and Facebook users. Interests chosen are shown in Figure 2-9. Findings included that the strongest correlation is between the interests “Fitness & Wellness” and the WHO health data for “obesity”. However, placebo interests performed almost as good for all correlations. Gender based results showed that women often showed a correlation with the obesity and diabetes interests, but also the placebo interests. Men were more interested in tobacco. The age analysis also shows a difference in smoking interests of young people in comparison to the older population. In general, their findings show that “within-country statistics are more statistically separable than statistics across countries”. Their concluding sentences state that the Facebook API should be used with caution when examining social factors. They state that they hope their work will “encourage future efforts to use our methodology to gather user interest from the Facebook Ads for other applications and scenarios”, which this thesis will focus on as well (Araújo et al., 2017).

Zagheni et al. (2017) address the issues of looking at demographic variables through Facebook Advertising. While they monitor stocks of migrants, their approach can be used on various other causes as well. For their purposes, they used the target category “Expats (Mexico)”, of which Facebook predicted monthly users of 8.4 million that are active on the platform. Facebook also estimates a total of 202 million expat users, which is not too far from the actual data which indicates that there are 244 million of them globally in the year 2015, according to the American Community Survey. One finding includes that” Facebook data overestimate migration stocks for younger age groups and underestimate the stocks for older age groups”. In general, they came to the conclusion that using Facebook advertising is a relevant factor in estimating demographic variables, as it performed very well in estimating migrant data throughout the USA (Zagheni et al., 2017).

Using the Facebook marketing API, Dubois et al. (2017) have investigated migrant assimilation in Germany. Assimilation in this case means “the cultural absorption of a minority group into the main cultural body”, as defined by the Collins dictionary (Collins Dictionary, 2018). The estimates that were collected from the adverts platform estimated a particular number of Arabic speaking people in Germany between a certain age, who were also interested in football (Dubois et al., 2017). As an example, they provided the comparison of people from German origin interested in a football league compared to people that are Arabic expats, also interested in the same league. Many interests were investigated in order to see how those differentiate between the two ethnic groups, and eventually also other migrant populations. Their findings show that European migrants have a higher assimilation score than Turkish-speaking or Arabic-speaking migrants. In the sub-sections of their research, findings show that men are more assimilated than women, and university graduates are more assimilated than non-graduates. Young people also seem to be more assimilated than older ones. Although there were some limitations related to the type of data and their methods, they mention that there is great potential for further research (Dubois et al., 2017).

Lastly, Facebook also serves as a tool to predict population statistics. In the case of Chunara et al. (2013), the interests of users could successfully give insight into obesity prevalence in the United States. The obesity data was, among others, taken from the CDC. A cross-sectional study looked at the relationship and predictive performance of the variable. They used activities that implied to be either positively or negatively related to obesity, for instance doing sports in comparison to watching television. Users that had interests which were activity-related turned out to have an about 12% lower predicted obesity rate. The people that indicated an interest in television had a 3.9% higher obesity rate, whereas this rate increases in the measured city of New York, where it equalled 27.5%. In their conclusion they state that more research on the online social environment, including Facebook, is needed to make an appropriate resolution about obesity rates and health interventions (Chunara et al., 2013).

Not only does Facebook help to predict health patterns, but even health organisations are using the social media tool to promote their goals. A paper by Park et al. (2011) shows that those organisations especially use the free tools by the social media platform and some advertising techniques but could make use of more advertising options in order to promote their organisation (Park et al., 2011). Those options include paid advertising in order to reach a broader audience.

The fact that Facebook advertisements make it easy to target a specific audience also makes it prone to malicious activity. Not just health organisations can take advantage of this tool, but also institutions that want to promote the consumption of alcohol in young adults. Michaelidou and Moraes' (2016) qualitative study focuses on 18 to 24-year-old young adults and found that the low prices and the sales promotion of alcohol through Facebook leads to more consumption of the studied population. This can be particularly dangerous as the online advertisements are even more engaging than offline marketing. This study proves as an example that Facebook advertising can also be effective when promoting unhealthy behaviour (Michaelidou & Moraes, 2016).

2.5 Summary, Differences to Previous Work & Contributions

The literature shows that firstly, the current health status of the population is critical, and individuals carry on cooking and eating an inadequate diet. Secondly, health can be improved by implementing solutions online. Studies on online recipes show what is considered healthy and what can be done to nudge people into the direction of consuming healthier foods, therefore positively influence behaviour. There is also proof that food interactions through the Web can predict health statistics, like obesity rates. Other research portrays the ability to predict the healthiness of populations based on traces left behind on social media and search engines. Social media, including Facebook, can be used to infer health statistics and predict the behaviour of individuals while linking it to certain activities as well.

What has not been done yet is to investigate how the promotion of online recipes, classified as healthy by certain standards, through social media can affect a user's choice in how to cook. The research on Facebook advertisements has shown a lack in knowledge about how people make their health choices in regard to cooking. Measuring and improving recipe choices through Facebook advertisements is research that adds important information to the goal of implementing general health advancement in populations. A study on Facebook advertisements and the promotion of healthy behaviour can show if it is possible to measure people's food choice, how to influence them and whom interacts with health promotions the most. The method and research questions are addressing this lack of research in the field with the questions below.

3 METHODOLOGY

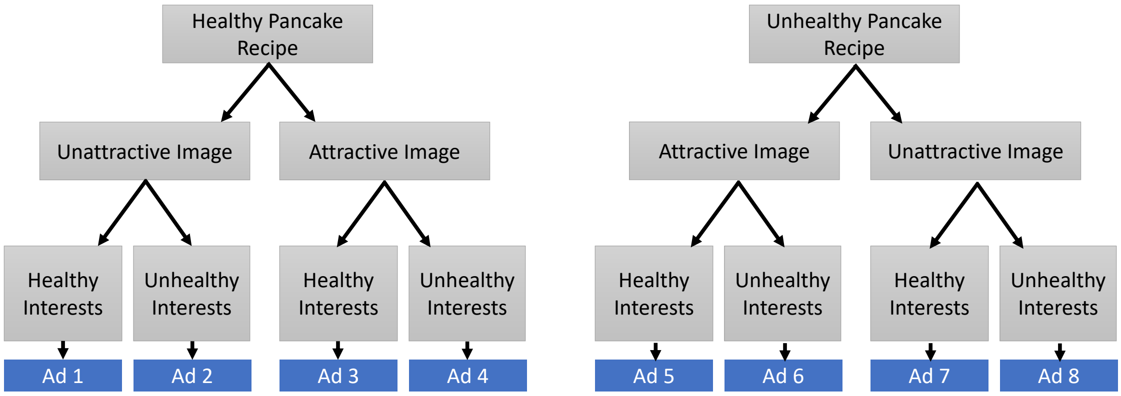


FIGURE 3-1 DIAGRAM OF THE ADVERTISEMENT STRUCTURE FOR RECIPE PROMOTIONS ON FACEBOOK

The literature in the previous chapter shows an existing lack of research where promotion of healthy eating is done via social media sites. It is clear that malnutrition is a cause to be addressed, as multiple authors already aimed at solving this issue. This thesis focuses on making clear how users interact with food promotion they are exposed to, in this case recipes. The main goal is to learn how certain factors influence the user's decision to interact with a recipe. The factors include recipe healthiness, image attractiveness, user interests and state healthiness. Apart from this, user characteristics that respond to promotions can also be identified. Figure 3-1 shows a diagram of the structure the eight advertisements have. Firstly, two recipes with different health criteria are advertised. Those each have one attractive and additionally one unattractive image. Each of the advertisements is then targeted to a different user interest group, one of them having healthy and the other unhealthy lifestyle interests.

The following sections cover how and why those advertisements were created. They also explain how the goal, to identify the factors influencing users in making a health-related decision, is worked towards. Firstly, an explanation of why this methodology was selected is provided. Afterwards, the data collection and the fitting statistical analyses will be provided.

3.1 Selection of Recipes

While the first research question covers general reactions to advertisements, the second more specific research question focuses on recipe healthiness. Trattner and Elsweiler (2017) came to the conclusion that in the internet, healthy recipes tend to be cooked less often than unhealthy ones. This is why this thesis posts advertisements to both a healthy and an unhealthy recipe in order to see if their conclusion is also confirmed in a social media setting. Monitoring

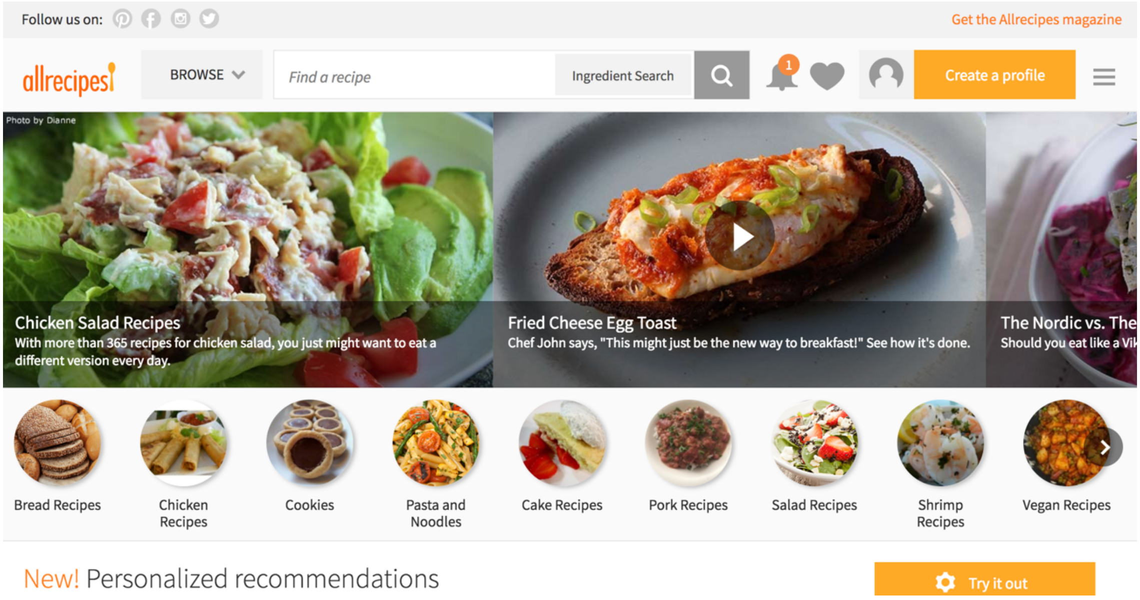


FIGURE 3-2 HOMEPAGE OF THE ALLRECIPES WEBSITE

Source: Allrecipes.com, 2018a

which recipe is clicked more often can later on help in knowing how to promote a healthy recipe. The answer to the research question then reveals if a difference between interactions on the healthy and unhealthy recipe exists. However, first a recipe needs to be selected which is afterwards pictured in the promotions.

As used by other authors before, the recipe platform Allrecipes is taken as a source to look for one healthy and one unhealthy recipe. With over 85 million users, Allrecipes serves as an online cooking platform calling themselves the “original and largest food-focused social network created for cooks by cooks” (Allrecipes.com, 2017). This platform makes it easy for online users to engage and contribute to recipe collections, as well as rate their favourite recipes online. Personalised recommendations let each user be treated individually by the website in order to appeal to the different tastes. The platform also gives an insight into how often a recipe was cooked, the average sentiment of this recipe and the popularity. Figure 3-2 shows the homepage of the platform.

For choosing a recipe from the platform, the World Health Organisation’s standards for healthy food serves as an indicator of the healthiness of a recipe in this thesis. A brief explanation of the classification system judged by WHO standards will help understand why recipes are given a certain number as a rating. Trattner and Elsweiler (2017) implemented an approach by which the 7 most important macro-nutrients, such as fiber, sodium, carbohydrates, proteins, sugars, fats, and saturated fats are within a certain range for each recipe. The ranges can be classified with numbers, starting at 0 and ending with 7. The number 0 means that the standards set by the WHO are not at all fulfilled, and 7 means that all standards are met (Trattner & Elsweiler, 2017).

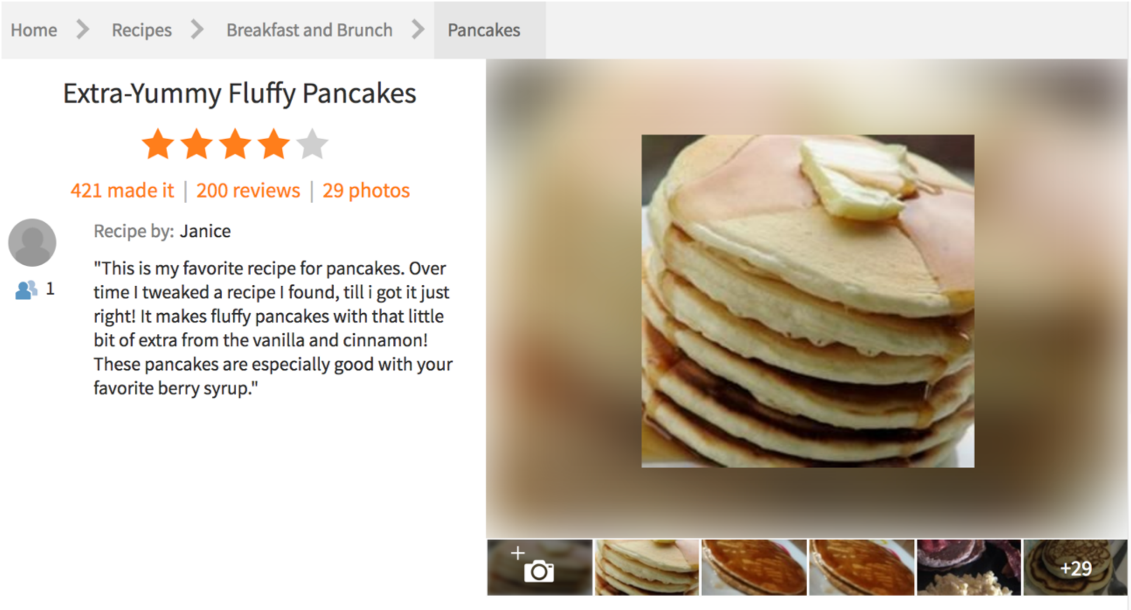


FIGURE 3-3 UNHEALTHY PANCAKE RECIPE WITH RATING, REVIEWS AND PHOTOS

Source: Allrecipes.com, 2018b

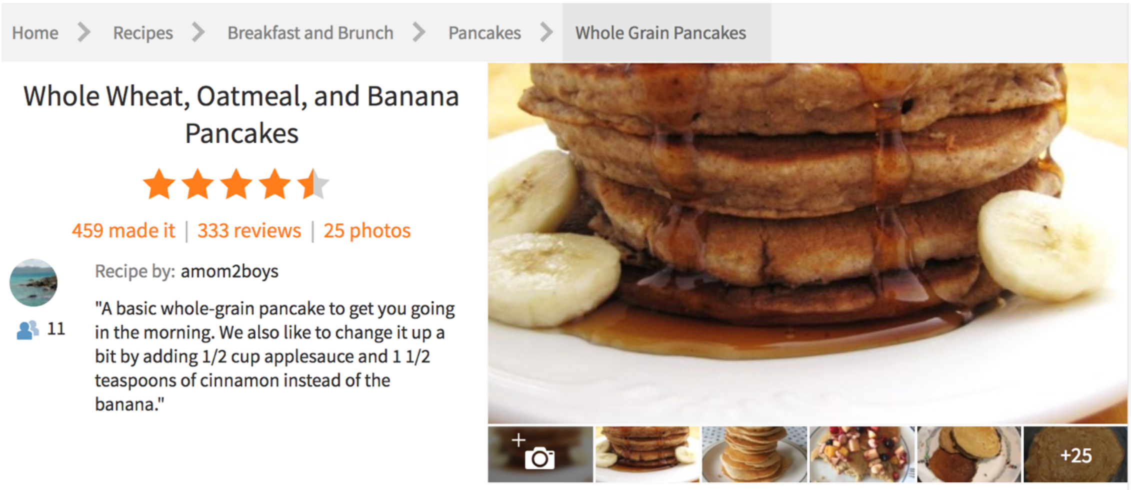


FIGURE 3-4 HEALTHY PANCAKE RECIPE WITH RATING, REVIEWS AND PHOTOS

Source: Allrecipes.com, 2018c

For knowing which recipe is classified as "healthy" and which is "unhealthy, the database from Trattner and Elsweiler (2017) with over 60 thousand recipes was used, which also show the matching health score to each recipe (Trattner & Elsweiler, 2017). The basic criteria for selecting a recipe was that a similar amount of people cooked them and that they had a similar average rating. This can be seen on the recipe website. However, the main criterion is that health scores need to be different. While the advertisements could be targeted to a great number of recipes, this thesis chooses one food in particular, namely pancakes. In contrast to other foods, this dish is vegetarian-friendly as well. According to an article in Daily Mail, research that focused on a sample population of 1,300 men and women in the USA has revealed that pancakes

recipe_id	kcal	protein	carbohydrates	fat	who_score
http://allrecipes.com/recipe/extra-yummy-fluffy-pancakes/detail.aspx	211	5.42023	26.1332	9.38861	4
http://allrecipes.com/recipe/whole-wheat-oatmeal-and-banana-pancakes/detail.aspx	187	6.18768	30.7697	4.78139	6

TABLE 3-1 TABLE OF NUTRITIONAL PROPERTIES OF HEALTHY & UNHEALTHY PANCAKES WITH THE MATCHING HEALTH SCORE
Source: Trattner & Elsweiler, 2017

are the 15th most popular dish (Peppers, 2014). Another website called "The Top Tens" lets users continuously vote on subjects. On the list "Top Ten Favourite Foods", Pancakes were currently voted as the 27th favourite food, with several quotes from users demonstrating the reasons (The Top Tens, 2005). Another site called "The Daily Press" also features pancakes as the sixth most popular breakfast food amongst Americans (Cahill, 2018). A poll made by ABC news also found that pancakes are among the most popular breakfast foods, but also found out that people eating breakfast are likely to be older rather than young (Langer, 2005). After looking through the recipes in the data set, two recipes for pancakes that had different health scores were selected. Those recipes were then used as the target website for the promotions on Facebook. The screenshot in Figure 3-3 shows how the recipe for the unhealthy pancakes looks like, while in Figure 3-4 the healthy pancake recipe is pictured. With pancakes being amongst the most popular foods, those two recipes were chosen as they have a different health score but other similar properties. As seen in Figure 3-3 and Figure 3-4, both imply that about 400 people cooked the recipe. Normalised to 100g, one pancake recipe has a calorie count of 211, while the other one is lower at 187 calories per serving. Although the healthy pancakes contain more carbohydrates, they also contain less fat than the unhealthy pancakes. Table 3-1 refers to the nutritional data of the pancake recipes. This gives them a health score of 4 for the "Extra-Yummy Fluffy Pancakes", and respectively 6 for the "Whole Wheat, Oat-meal, and Banana Pancakes" (Trattner & Elsweiler, 2017). 4 means that the first recipe is in the average health range and is rather unhealthy. The recipe with the rating 6 is therefore an extremely healthy recipe, where the cook can benefit from maximum nutrition. Most recipes rated with 7 are recipes such as flavoured water and others, which hardly introduce calories into the diet. The Table 3-1 shows a short summary of the different characteristics.

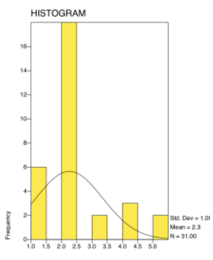
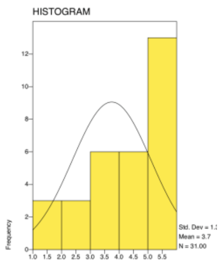
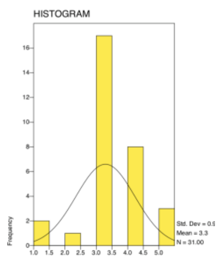
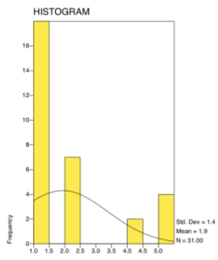
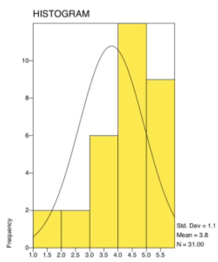
Image 1	Image 2	Image 3	**Image 4**	**Image 5**
				
				
(M = 2.26, SD = 1.09)	(M = 3.74, SD = 1.37)	(M = 3.29, SD = .94)	**(M = 1.94, SD = 1.44)**	**(M = 3.77, SD = 1.15)**

TABLE 3-2 MEANS AND STANDARD DEVIATION FOR RATINGS OF IMAGES ON "HEALTHY PANCAKES" FROM THE SURVEY

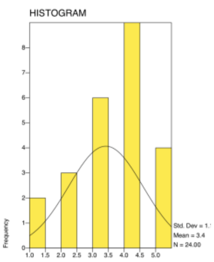
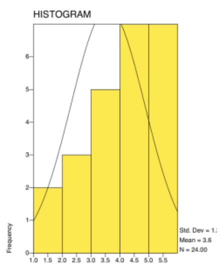
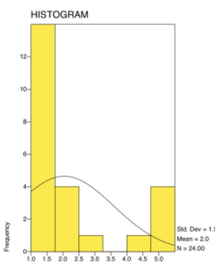
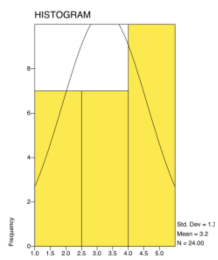
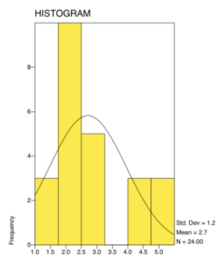
Image 1	**Image 2**	**Image 3**	Image 4	Image 5
				
				
(M = 3.42, SD = 1.18)	**(M = 3.58, SD = 1.28)**	**(M = 2.04, SD = 1.55)**	(M = 3.25, SD = 1.36)	(M = 2.71, SD = 1.23)

TABLE 3-3 MEANS AND STANDARD DEVIATION FOR RATINGS OF IMAGES ON "UNHEALTHY PANCAKES" FROM THE SURVEY

3.2 Selection of Images

Another research question in this thesis focuses on whether the image used in promoting a recipe plays a role in user interactions. Elsweiler et al. (2017) have previously investigated which factors play a role in recipe selection for users. One influential factor in their study is image, which is why two different images have been chosen to determine if image attractiveness influences recipe choice for users.

As seen in the figures above, multiple images are shown on the Allrecipes site for both pancakes. Both recipes feature a set of slightly over 20 images. In order to be able to promote the recipes on Facebook, one image has to be chosen. For this, a survey was conducted. The survey included five randomly selected images of both recipes. Those images were gathered in the survey. Respondents were asked to rate the images for each recipe separately from 1 to 5, with 1 being the “least appealing” image to 5 being the “most appealing” image of pancakes.

The survey was distributed through Facebook and e-mail from Wednesday, 11th April 2018 until Sunday, 15th April 2018 with 31 respondents for the first set of images, and 24 respondents for the second set of images. The survey, which can be looked at in Appendix 1, took approximately 1 minute to complete. The end result has shown that the images in Table 3-2 and 3-3 were chosen to be the most and least appealing pancake photos. All the images that were being rated can be found in Table 3-2 and 3-3, as well as the outcome of the survey. According to the mean value from all the rating in the survey, the healthy recipe concluded Image 5 to be the most appealing, and Image 4 the least appealing. The unhealthy recipe had Image 2 as the most appealing, whereas Image 3 was the least appealing. The winning images for most and, respectively, least appealing are pictured in Table 3-2 and Table 3-3 and emphasized in bold.

3.3 Selection of Target Group

As Trattner et al. (2017c) show that interests in cuisines and outdoor activities like biking, for instance, can be linked to certain recipe properties, the second research question addresses whether interests on Facebook influence the interactions with recipes. Different user interests can potentially have an effect on how many people click on which advertisement, for instance people interested in healthy activities might click on the healthy recipe more often.

Targeting on Facebook advertisements makes it possible to target users the way an advertiser wants to. A population sample for the advertisements placed on Facebook is employed for this experiment. This population sample is made up of Facebook users, as only those can be targeted on Facebook. The tool makes it possible to target a variety of users, which can be either all users of Facebook, or a narrowed audience. When placing advertisements, the targeting leaves advertisers a wide range of options on who to target. In the case of this thesis, two target groups have been identified. One target group represents a set of “healthy” individuals,

Interest	Audience Size
Soft drinks	162,271,430
Domino's Pizza	33,330,640
McDonald’s	153,519,150
TV game shows	111,199,740
Chocolate	293,016,710
TV talkshows	64,933,280
TV	768,151,450
TV reality shows	64,933,280
TV comedies	32,894,080
Pizza Hut	44,847,400
Sugar	177,039,110
Fast food restaurants	34,315,305
Burger King	44,351,500
KFC	81,586,240
Wendy’s	16,492,680
Fast casual restaurants	32,211,405
Pizza	247,919,350

TABLE 3-4 UNHEALTHY INTERESTS ON FACEBOOK AND THEIR ESIMATED AUDIENCE SIZE

whereas the other target group includes “unhealthy” individuals. Both groups are constituted of all genders, including a male and a female audience. The ages targeted contain all ages from 18 to over 65-year-olds. The last research question addresses those characteristics, such as age and gender, which one can gain insight to after advertisements have been run. It is imperative to know who interacts with such recipe advertisements, in order to know who to target in the future. The only difference between the two target groups for the advertisements in this thesis is the interests they have. According to an article from 2016 that appeared in The Verge, a

Facebook spokesperson has explained that “interests are formulated algorithmically” by the platform and represent what certain sets of users seem to be interested in (Havlak & Abelson, 2016). As an example, the interest “Justin Bieber” has an audience of 186,828,600 people that seem to be interested in the artist, at the time of the article in (Havlak & Abelson, 2016). Based on activities and behaviour that unhealthy people tend to have, as well as interests of healthy people, the target groups for the thesis comprise the interests summarised below.

3.3.1 Unhealthy Interests

The interests chosen for this target group contains: Soft drinks, Domino's Pizza, Gaming, McDonald's, TV game shows, Chocolate, TV talkshows, TV, TV reality shows, TV comedies, Pizza Hut, Sugar, Fast food restaurants, Burger King, KFC, Wendy's, Fast casual restaurants, Pizza, Video games, Plus-size clothing and Fast food. The potential reach of this audience is estimated by Facebook to 130,000,000 people. Facebook defines the audience size as “Your audience selection is fairly broad” (Facebook, 2018c). Interest audience sizes can be seen in Table 3-4.

Most interests either represent people that are interested in certain kinds of food, or sedentary behaviour. One article proves that the time spent in front of a television does indeed present a positive correlation to obesity and also type 2 diabetes (Hu et al., 2003). Another study conducted in Spain has the same findings, which show that the factor playing a role in obesity patterns in grown up individuals is the time they spend watching television (Vioque et al., 2000). However, not only television watching is responsible for high obesity rates. Fast food and unhealthy food also often shows to be influencing obesity in adults and children. Jeffrey et al. have reported that how often people eat at so called “fast food restaurants” shows a positive association to the body mass index of a person (Jeffrey et al., 2006). According to Business Insider, the chain restaurants Burger King, McDonald’s, Domino’s Pizza, Pizza Hut, Wendy’s, KFC and other restaurants are among the most popular fast food chains in the USA (Fitzpatrick, 2015). Also sugar sweetened drinks count as one of the main factors of obesity, with persons increasing the likelihood to becoming obese by 1.6 times each time they consume a soft drink per day (Apovian, 2004). Plus size clothing as an interest also is a clear indicator of defining people who weigh more than average, which Yom Tov et al. also indicated in their study (Yom Tov et al., n.d.).

3.3.2 Healthy Interests

The interests chosen for this target group contain: Meditation, Physical fitness, Yoga, Running, Weight training, Bodybuilding, Physical exercise or Sports and outdoors. The potential reach of this audience is estimated by Facebook to 135,000,000 people, which is similar to the previous target group size. It is also defined “fairly broad” by Facebook (Facebook, 2018c). Table 3-5 shows the audience sizes of each interest.

Interest	Audience Size
Bodybuilding	120,707,353
Meditation	107,635,550
Physical Exercise	360,551,190
Physical Fitness	308,458,240
Running	137,898,108
Weight Training	78,922,754
Yoga	180,002,746
Sports & Outdoor	2,557,456,948

TABLE 3-5 HEALTHY INTERESTS ON FACEBOOK AND THEIR ESTIMATED AUDIENCE SIZE

Studies prove that keeping fit is a consequential part of living a healthy lifestyle. A study conducted in 1994 mentions that even low amounts of fitness related activities are enough to reap their weight and also health related benefits (Grilo, 1994). A study conducted by the Center for Diseases Control also demonstrates that physical activity is essential to preventing obesity and chronic diseases, such as diabetes or heart diseases (Centers for Disease Control, 2003). As “many proteins produced by skeletal muscle are dependent upon contraction”, Pedersen and Febbraio also mention that inactivity of those muscles leads to chronic diseases, which implies that activity of those muscles can prevent those diseases (Pedersen & Febbraio, 2012).

3.3.3 Interest Targeting Options

In addition to those interests, the advertising platform offers an option called “expand interests”. According to the platform, “this option lets Facebook automatically expand the interests in your detailed targeting criteria if there's a chance to reach more people likely to take the action you're optimising for” (Facebook, 2018c). For purposes of the following experiment, the option to expand interests is not chosen, because otherwise it would not be clearly measurable anymore where the users that clicked on advertisements come from. It is also possible in a target audience to exclude people that an advertisement should not target. Adding connections is also an option, where the advertiser has the possibility to promote their advertisement to people who have expressed interest in the page before, people that like the page or also friends of people that like the page. Individuals who are already customers of the brand or have previously purchased their products can also be targeted. This option is also not placed, as the page is newly created, which will be explained in the next section where the research instrument is described in detail.

3.4 Selection of Study Site

One research question which is answered in this thesis is whether or not state healthiness plays a role in the selection of recipe advertisements. As the advertisements are all run in specific locations, it is later on important to look at whether or not connections of recipe promotion properties to healthiness in a location can be made. Said and Bellogín (2009), for instance, display where online users come from and whether those regions are high or low health counties. Many other authors have also investigated the link between obesity and user behaviour in the Internet. Trattner et al. (2017b) used correlation analysis between bookmarking and obesity prevalence rates. Chunara et al. (2013) looked at a correlation between Facebook interests with obesity rates, and Abbar et al. (2015) also investigated Twitter data in correlation with obesity rates and found a high correlation between the two. Lastly, Fried et al. (2014) also found that Twitter data can predict obesity and diabetes rates. The last question in this thesis therefore addresses the link between obesity and diabetes prevalence and advertisement interactions, for which a target state is needed.

The study site of this research is also the United States of America, which is a country situated in the North of America. With a population of 325,719,178 as of June 2017, it is one of the biggest countries in the world, which is made up of 50 individual states who each have their own governmental jurisdiction system (U.S. Census Bureau, 2018). According to data from the Census website, the population split of male and female is made up of 50,8% female and respectively 49.2% male residents. Although ethnicities differ among states, 76.9 % are white. The rest is made up of groups such as Black or African Americans with 13.3 %, Asians about 5.7% and other groups that make up for the other percent (U.S. Census Bureau, 2018). The average person per household is 2,6 people. 63.1 % of the total population are, at the time of the report, in the labour force (U.S. Census Bureau, 2018). As this thesis focuses on research through Facebook, the 50 states were chosen as the study site because firstly, the country of origin of this platform is the United States of America. It was launched in the year 2004 by a sophomore student called Mark Zuckerberg and his friends at the site of Harvard University, in Massachusetts (Carlson, 2010). After India, the USA have the second highest amount of Facebook users in the world, with almost 250 million users (Statista, 2018a). The USA, in comparison to the country India, present data and statistics by state openly available to the all users on the Internet, which is explained in detail below. Demographic statistics can be accessed at any time. In addition to this, Facebook can target all states in one advertisement. Impressions, click-through-rate and other measurements generated through advertisements can be separately analysed on a state level. Additionally, the online source U.S. Census Bureau (2018) shows interesting trends that are visible in the target country, for instance, that most states

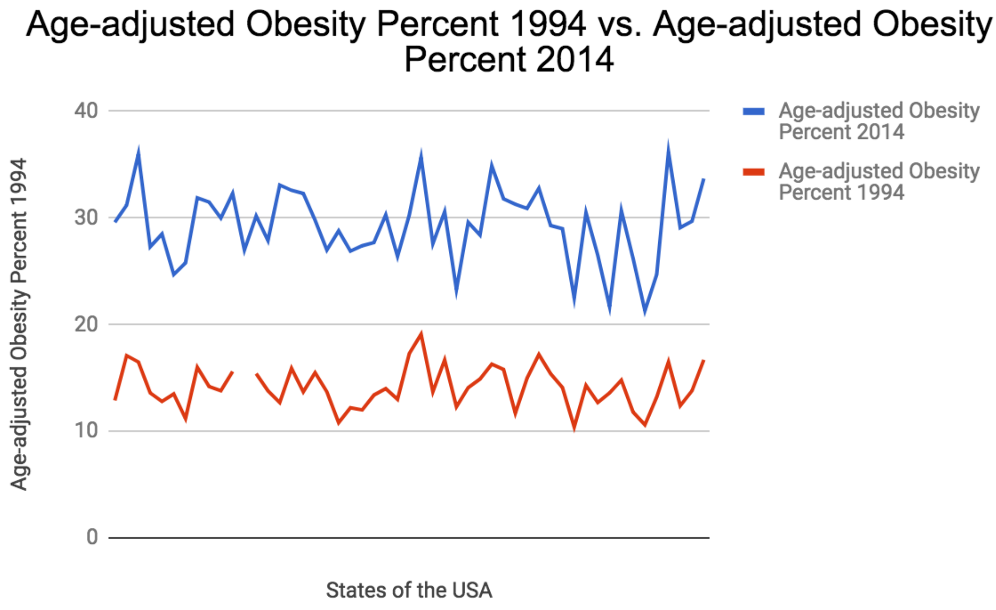


FIGURE 3-5 INCREASING OBESITY RATES IN THE STATES OF THE USA FROM 1994 TO 2014

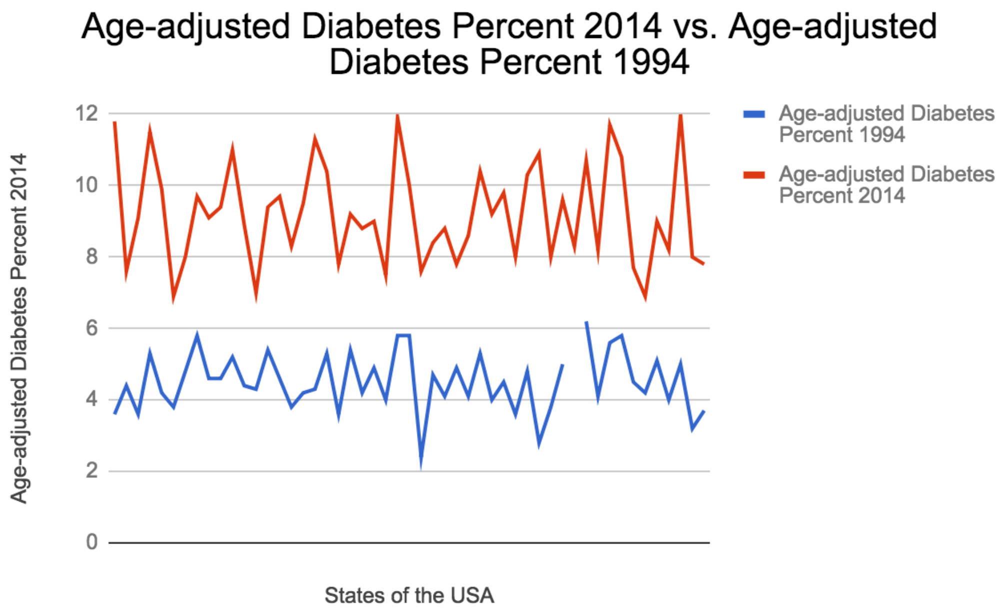


FIGURE 3-6 INCREASING DIABETES RATES IN THE STATES OF THE USA FROM 1994 TO 2014

that have the lowest obesity prevalence have more foreign-born residents (U.S. Census Bureau, 2018). It can also be seen that education level is significantly higher in states with low obesity prevalence, whereas in states of more obesity the education decreases (Census, 2018). Obesity seems to also depend on income, as it can be seen that higher obesity rates occur in countries with lower income. States with high obesity rates tend to earn about 40,000 per household, and states with low obesity rates earn notably more as median household income.

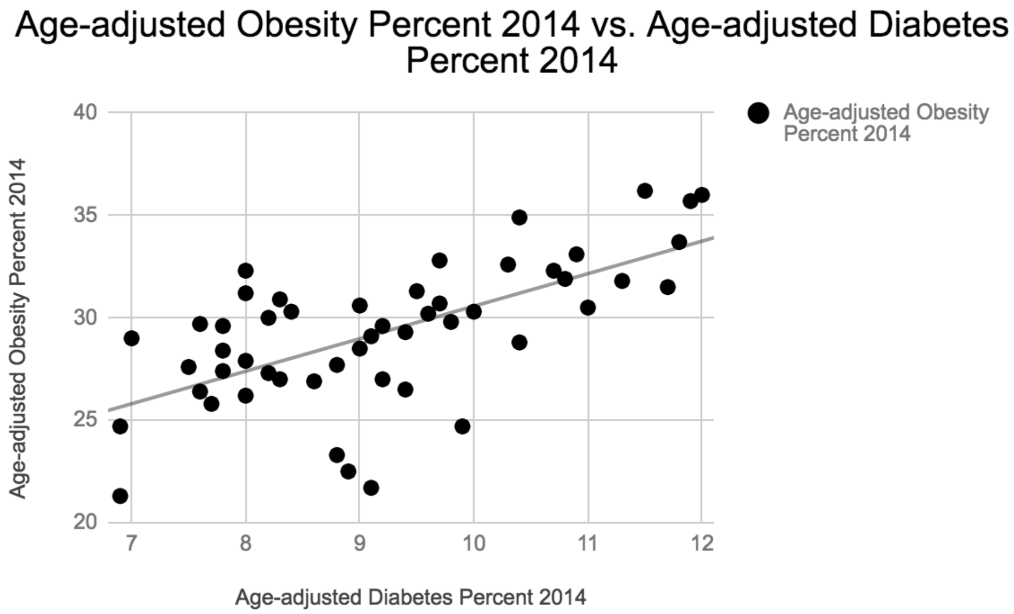


FIGURE 3-7 SCATTERPLOT OF DIABETES VS. OBESITY PREVALENCE RATES IN THE STATES OF THE USA, 2014

A direct relationship between income and obesity prevalence can therefore be seen. More specific data is portrayed in Appendix 2.

One important aspect this thesis, as well as many other research efforts, focuses on is to see how recipe interactions can be explained by obesity and diabetes rates. As the country targeted is the USA, a lot of data is available on a state level. This data comes from the CDC (Center of Disease Control). The Center of Disease Control is a federal agency who's goal it is to promote health and protect US citizens from diseases, security or health threats (Centers for Disease Control and Prevention, 2014). They own a website which makes public statistical reports and data available to every individual. Such reports include, for instance, data on obesity and diabetes prevalence in each state of the USA. This thesis makes use of the website's reports as the main source of obesity and diabetes statistics. Data is available from 1994 until the year 2014, where obesity prevalence estimates, in addition to diabetes prevalence estimates are given for each state. It is possible to see the significant difference between the most recent data of 2014 and the data 20 years earlier, namely 1994 (Centers for Disease Control and Prevention, 2016). Both prevalence rates have significantly increased over the years, which can be seen in Figure 3-5 and Figure 3-6.

As diabetes, as well as obesity rates have risen over the past years, it is possible to determine a rather strong correlation between the two. In 1994, as well as in 2014 the two rates, although both have increased, show a direct relationship. The significant correlation and the corresponding scatterplot are shown in Figure 3-7.

The states with the highest obesity rates in 2014 are Alabama, Louisiana, Mississippi, West Virginia and Arkansas. Almost all of the mentioned states also are part of the 5 states with the highest diabetes prevalence, except for Louisiana. The states with the lowest obesity rates in 2014 are Colorado, Hawaii, Massachusetts, California and Vermont. Colorado and Vermont are also the states with the lowest diabetes prevalence (Centers for Disease Control and Prevention, 2016). Those obesity and diabetes prevalence rates play an important role in RQ5, as correlations between the state's healthiness and recipe advertisement interactions are looked at.

3.5 Research instrument

When conducting research, it is possible to use a variety of tools and methods to collect and analyse data. The literature chapter manifests that for working with web-based tools, a quantitative study is most common. In the context of health trends, Trattner et al. (2017c) and Yom-Tov et al. (2018) recently conducted research that focused on using Web tools like bookmarking and Bing advertisements for data collection. The method to collect data through Facebook is chosen for this thesis because Facebook is a new and up-to-date tool that has been used previously for the purpose of measuring health standards by Mejova et al. (2018) and Araújo et al. (2017), who examined Facebook interests. However, actual advertisements were not yet used in order to gain insight into the subject of healthy eating. This is why this thesis pursues the goal of collecting data through the Facebook advertising API. This method is used in order to collect clicks from a target audience. In this case, the population of the United States and the two target audiences mentioned above lay the foundation of the advertisements. As seen in papers by De Choudhury et al. (2013), Garimella et al. (2016) and many others, trends and patterns in populations can be analysed and predicted when using social media. Next to Instagram, Twitter or Pinterest, Facebook is one of the biggest social media platforms. The advertising API makes it easy for businesses to advertise to an audience that they can individually target.

There are three main aspects that creators of ads should focus on while creating an ad, including the target audience, ad formats and later the ad reporting tool. The target audience of an ad determines which kind of demographics one wants to reach with an ad, for instance "age, gender, relationship status, education, workplace, job titles and more". The location can also be determined, which can vary from a country or a city to even a small county. Interests, such as hobbies or the entertainment of users, are also part of the target audience. Certain behaviours like "purchasing behaviours, device usage and other activities" can also be used to target an audience. A business can also specifically target people who are already a customer. "Lookalike audiences" can help businesses target users that are similar to their customer base (Facebook, 2018c). The sampling procedures above show which audience is targeted in this experiment. Ad formats are the second part of creating an advertisement, which are called the "Ad Creative". Creators of the ads can use an eye-catching photo, a video, use multiple photos and

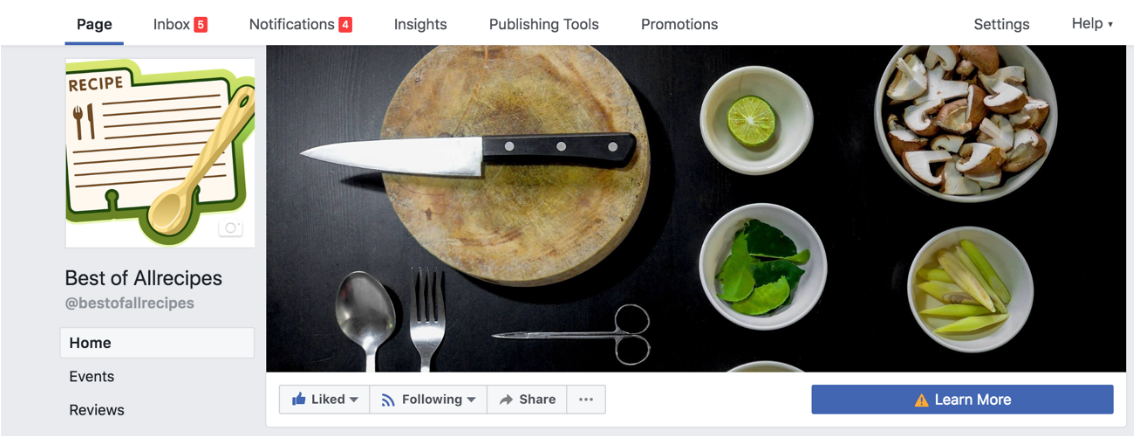


FIGURE 3-8 FACEBOOK PAGE "BEST OF ALLRECIPES" LAYOUT THAT WAS USED FOR RECIPE ADVERTISEMENTS

Source: Facebook, 2018

more. It is also possible to include the "messenger" symbol in the ad so that customers have no difficulties communicating with the business. The types of formats can be lead ads, dynamic ads with many products or link ads. With link ads, the advertising format provides a direct link to the website promoted (Facebook, 2018c). The ad reporting tool is the third aspect, that is helpful for assessing the performance of the promotion. Those insights, displayed in a tool called the "Ad manager" or the "Ads Insights API", are organised by objective. "Reach, frequency, targeting and cross-device performance" are some of the insights that are displayed there. While there are insights available for off-Facebook campaigns and partners, this thesis will focus on Facebook only with the Ad Manager tool. This tool helps the creator when "creating ads, managing when and where they'll run, and tracking how well campaigns are performing" (Facebook, 2018g). After having placed an advertisement through Facebook, it is trackable and also manageable through the Ad Manager. This means that users have the possibility to edit, and also change their advertisements. Budget, audience and placement options are some of the characteristics that can be edited. Pausing a campaign, copying or re-launching it are also possibilities. The reporting tool is the one showing whether an advertisement has served its purpose or not. It is able to spot trends over time and therefore facilitating it for a business to reach their goals (Facebook, 2018g). RQ1 focuses on the numbers the Ad Manager shows when a campaign is finished. The sections below explain how the advertisements for this thesis were created and later on analysed.

3.5.1 Facebook Page

Firstly, in order to set up advertisements and promote content to the Facebook community, it is important to own a page on the platform. A Facebook page is a "public profile created by

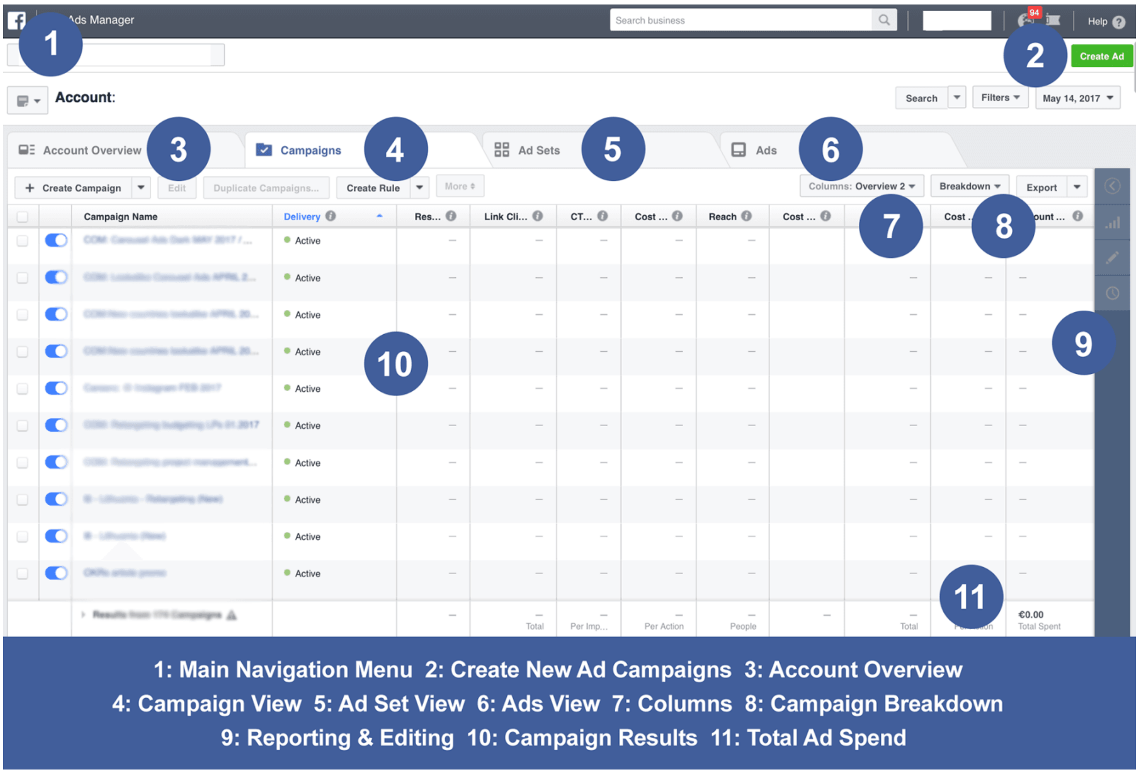


FIGURE 3-9 FACEBOOK ADS MANAGER WHICH WAS USED TO COLLECT DATA FROM THE EIGHT PLACED ADVERTISEMENTS

Source: AdEspresso Inc., 2018

businesses, organizations, celebrities and anyone seeking to promote themselves publicly through social media" (Techopedia Inc., 2018). A page can generate followers, which are people that press the button "like" on the page. Through this page, it is then possible to share posts with followers, which can range from pictures and videos to certain websites or even events. For this experiment, a page with the name "Best of Allrecipes" is created. The image below shows the layout of the page, with the images that are used taken from Google Images which were labelled "for reuse with modification" (Pixabay,2018). The layout of the page is pictured in Figure 3.8. After launching the page, Facebook gives the opportunity to advertise the page and promote posts made with it. This can be done through a shortcut on the page directly, or in the previously mentioned Ad Manager. This thesis makes use of the Ad Manager to create and advertise the recipes. Figure 3-9 below describes parts of the Ad Manager which are relevant. There are three main aspects to creating advertisements, namely the campaign, the ad sets within the campaign and the individual advertisements.

3.5.2 Campaign

Creating an advertisement first begins with starting a new campaign. The campaign is named by the creator, and usually gives insight into the purpose of the advertisement. Facebook offers a variety of goals for advertisers, such as generating likes for a business page, creating sales on a website or generating traffic to a website. This experiment focuses on the goal of generating traffic to the websites of recipes chosen above. The target website is identified in

Unhealthy Interest Group	Healthy Interest Group
Potential Audience: Potential Reach 130,000,000 people **Audience Details:** ▪ Location: ∘ United States ▪ Age: ∘ 18 - 65+ ▪ People Who Match: ∘ Interests: Soft drinks, Domino's Pizza, McDonald's, TV game shows, Chocolate, TV talkshows, TV, TV reality shows, TV comedies, Pizza Hut, Sugar, Fast food restaurants, Burger King, KFC, Wendy's, Fast casual restaurants or Pizza ▪ Interest expansion: ∘ Off	**Potential Audience:** Potential Reach 135,000,000 people **Audience Details:** ▪ Location: ∘ United States ▪ Age: ∘ 18 - 65+ ▪ People Who Match: ∘ Interests: Volleyball, Meditation, Baseball, Association football (Soccer), Auto racing, College football, Swimming, Physical fitness, Yoga, Skiing, Triathlons, Basketball, American football, Tennis, Running, Marathons, Weight training, Golf, Snowboarding, Bodybuilding or Physical exercise ▪ Interest expansion: ∘ Off

TABLE 3-6 TARGET AUDIENCES FOR HEALTHY VS. UNHEALTHY INTERESTS

Source: Facebook Ad Manager, 2018

the “ads” section of this chapter. A campaign can include numerous ads. In the case of this experiment, eight different ads are included that are explained in detail below. The campaign launched with a spending limit of 500 Euros.

3.5.3 Ad Sets

Each ad set features a different target audience and budget. Subsequently, the ad sets have the possibility to contain multiple individual advertisements. Eight ad sets are created, because it is crucial that the same budget is assigned to each advertisement. The ad sets contain two different target audience groups. Four of the ad sets are targeted to the “Healthy Interests” group, and four ad sets are targeted to the “Unhealthy Interests” group. Other options that are included in this part of the ad manager are to launch the advertisements on Instagram. Since this thesis focuses on Facebook advertising, the option to show the ads on Instagram was not chosen. Ad sets contain audiences that look as seen in Figure 3-10.

In order to get meaningful results, creators have the option to let Facebook determine how much they want to pay for generating a click. Another option is to manually enter the maximum cost-per-click (CPC) price that they are willing to pay. When selecting the manual option, it is possible to generate a larger number of clicks. This is why a maximum CPC was set at 0.15 Euros, in order to generate a high number of clicks on all advertisements. The same budget of spending a maximum of 10 Euros per day was set for each ad set. Facebook decides itself whether an advertisement is performing well, and the full amount will be spent, or only part of it. It is only possible for the creator to set a maximum spending limit.

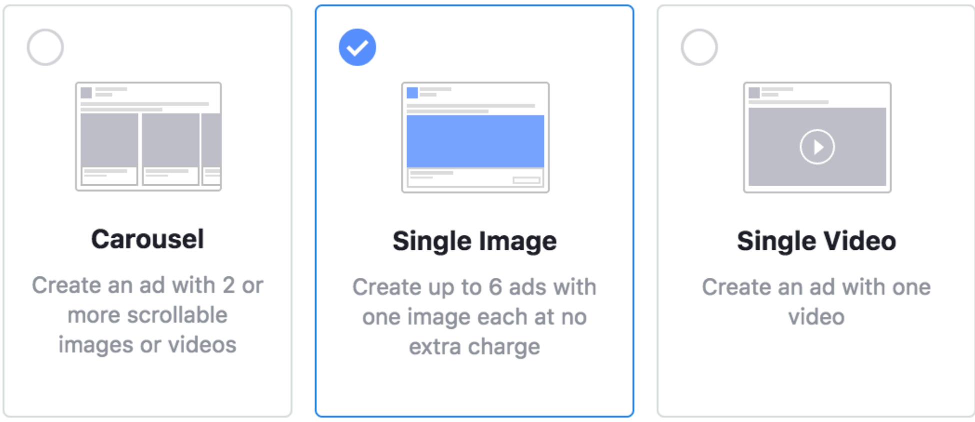


FIGURE 3-10 CREATION OF AN ADVERTISEMENT – CHOOSING OF THE AD CREATIVE

Source: Facebook Ad Manager, 2018

3.5.4 Ads

Ads are the final individual advertisements that targeted users then see in their Newsfeed. From the eight ad sets mentioned above, each feature one advertisement. Those advertisements then include the different target groups determined in the ad sets and additionally link to the different recipes and images. Firstly, the layout of the advertisement is chosen. This can include either one image or video, multiple images, collection of different items or other variations. Figure 3-11 shows three options for the creator, whereas many more are possible.

As the survey indicated which images individuals rate as appealing, compared to unappealing, the advertisements in this experiment contain the first option of an ad with only one image. The images selected through the survey are chosen for each advertisement. Then, creators choose the link of their advertisement. This can either be an event created on Facebook, or an external link. For each individual ad, one of the two recipe URLs was selected, healthy and unhealthy. Another option is to edit the text and headline of the advertisement. When entering the Recipe URL, a text from the platform Allrecipes, which perfectly describes the recipe, automatically shows up in the description box. This text was used for the experiment. The headline was altered to either "Fluffy Pancake Recipe" or "Healthy Pancake Recipe", as the title of the text has a limit of 40 characters. An Example of the advertising structure for both Recipes, healthy and unhealthy, can be seen in Figure 3-1.

Table 3-6 includes the details of each advertisement, with the title of the recipe, the image used for promoting it and the interest group it is targeted to. The daily budget for each ad is also included, which is spent until the campaign limit of 500 Euros is reached. The estimated daily results show an approximation of the reach and clicks calculated by Facebook, with a warning that "results are likely to differ from estimates" and that they "have limited data available to calculate this estimate, so estimates may be less accurate" (Facebook, 2018c).

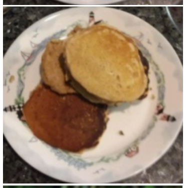
	Recipe	Image	Interest	Maximum Daily Budget	Estimated daily results
Advertisement 1	Healthy Pancakes		Interest Group 1 (Healthy)	10€/Day	Reach 340 - 2,100 Link Clicks 42 - 260
Advertisement 2	Healthy Pancakes		Interest Group 2 (Un-healthy)	10€/Day	Reach 820 - 4,000 Link Clicks 31 - 190
Advertisement 3	Healthy Pancakes		Interest Group 1 (Healthy)	10€/Day	Reach 1,100 - 4,600 Link Clicks 31 - 190
Advertisement 4	Healthy Pancakes		Interest Group 2 (Un-healthy)	10€/Day	Reach 380 - 2,400 Link Clicks 36 - 230
Advertisement 5	Unhealthy Pancakes		Interest Group 1 (Healthy)	10€/Day	Reach 830 - 4,100 Link Clicks 34 - 210
Advertisement 6	Unhealthy Pancakes		Interest Group 2 (Un-healthy)	10€/Day	Reach 950 - 3,900 Link Clicks 40 - 200
Advertisement 7	Unhealthy Pancakes		Interest Group 1 (Healthy)	10€/Day	Reach 910 - 4,300 Link Clicks 35 - 200
Advertisement 8	Unhealthy Pancakes		Interest Group 2 (Un-healthy)	10€/Day	Reach 870 - 3,900 Link Clicks 38 - 230

TABLE 3-7 EIGHT FACEBOOK ADVERTISEMENTS AND THEIR PROPERTIES

After the campaign was activated, it was active from 04/21/2018 at 3:31am until 05/01/2018 at 5:16am, which is equal to the time where the set budget limit of 500 Euros had been reached. Facebook's algorithm decides how and when to spend the money. Appendix 3 shows the advertisements as they appeared to the targeted users.

After the advertisements have run for a about a week, the outcome is tested with various quantitative measures, including t-test and Mann-Whitney U, as well as Kruskal-Wallis and ANOVA. Those are tools commonly used to show differences between chosen variables. Correlation analysis is also used for this thesis, as other authors like Chunara et al. (2013) have successfully before made use of this method to determine a relationship between obesity prevalence and data collected online.

3.6 Statistical Analysis

To confirm to what extent certain factors like healthiness, image, interests and state healthiness have on reactions to advertisements, statistical analyses can be employed. At first, Facebook Ad Manager reports the results of the advertisements with a short and simple display about the outcome metrics. The data collected through the Ad Manager can afterwards be exported to the statistical software SPSS, where further analyses are conducted.

The general ad performance can be analysed through the Ad Manager. A variety of performance measures is shown through this interface. The most important metrics to be analysed statistically are listed, with the possibility to rank by outcome. Metrics that play a role in this thesis are the following (Facebook, 2018c):

- Results: "The number of times your ad achieved an outcome, based on the objective and settings you selected."
- Reach: "The number of people who saw your ads at least once. Reach is different from impressions, which may include multiple views of your ads by the same people. (This metric is estimated.)"
- Impressions: "The number of times your ads were on screen."
- Cost per Result: "The average cost per result from your ads."
- Amount Spent: "The estimated total amount of money you've spent on your campaign, ad set or ad during its schedule. (This metric is estimated.)"
- Unique Link Clicks: "The number of people who performed a link click. (This metric is estimated.)"
- CTR: "The percentage of times people saw your ad and performed a click."

	2 Groups		>2 Groups	
	Independent	Related	Independent	Related
non-parametric	Mann-Whitney U-Test	Wilcoxon Test	Kruskal-Wallis Test	Friedman Test
parametric	t-Test	Paired Samples t-Test	ANOVA	Repeated Measures ANOVA

FIGURE 3-11 POSSIBLE OPTIONS FOR GROUP COMPARISON

Source: Ponocny & Weismayer, 2016b

This thesis will focus on the measurements Impressions and click-through rate (CTR), as this is the most accurate measurement of how many users have interacted with an advertisement, compared to seen it. After obtaining the results from the Ad Manager, the data is exported to SPSS where statistical analyses focus on answering each research question. As the research is quantitative, testing for group differences is the main goal for most research questions, except for one question which will focus on correlation analysis. When testing for group differences, depending on how many groups there are and whether or not the data is parametric, a statistical test can be selected.

Figure 3-12 by Ponocny and Weismayer (2016b) shows the selection procedure of which statistical test can be deemed accurate for a hypothesis. To begin with, the author has to determine which differences are of interest. Therefore, a variable of interest is chosen. There is one dependent and one independent variable. The dependent variable is the exploratory variable, whereas the independent variable is the explanatory variable which helps to explain the dependent variable. For this thesis, the variables will be determined by each research question separately. Then a hypothesis is made, which can be either one or two tailed. A one tailed hypothesis suggests that the author already has an idea of the direction of differences. For example, when determining which group of tourists spends more time in a city, where indicators already suggest that one of the groups spends significantly more time there. A two tailed hypothesis implies that the author does not know the direction of the hypothesis, in which case it is assumed that either group of tourists can spend more time in a city. This thesis makes use of two-tailed testing only. After knowing if it is a one or two tailed question, a hypothesis can be made. H0, which is the null hypothesis, always indicates that there is no difference between the two variables of interest. H1 therefore is the alternative hypothesis, where a significant difference between the variables can be established. Subsequently, the author has to establish the number of groups in the dataset. There are either two or more than two groups compared, as seen above, which then can be independent or related. When the groups are independent, they are compared, whereas when the groups are related, variables in those groups are compared, for example a "happiness rating before and after vacation" (Ponocny & Weismayer, 2016b). After having considered the conditions above, Ponocny and Weismayer (2016b) suggest determining whether a variable is normally distributed or not. The question to ask is whether the dataset can be normally distributed, which can be answered in two ways. The first option is to create histograms of the variable. If the histograms appear to be bell shaped, an assumption is that the data is normally distributed. To be certain, the second step is to run a One-Sample Kolmogorov-Smirnov test in SPSS. The null hypothesis for this test is that there is no deviation from normal distribution. This means that if the p-value is above 0.05, there is no significant deviation. It can be assumed that the data is normally distributed. If the p-value is below 0.05, the data is not normally distributed (Ponocny & Weismayer, 2016b).

Following all the guidelines above, the right test for analysing data can be selected. This thesis consists of data where the groups are always independent, and never related. Therefore, it will make use of the Independent t-test, Mann-Whitney U-Test, ANOVA and Kruskal-Wallis test. For the research questions that focus on more than two groups, such as examining the variables age, gender and region and their effect on advertisement interactions, an ANOVA test is used if the data is normally distributed. ANOVA stands for "Analysis of Variance". It is a "statistical technique that assesses potential differences in a scale-level dependent variable by a nominal-level variable having 2 or more categories" (Statistics Solutions, 2018a). If the above-mentioned tests all result in the conclusion that the variable is not normally distributed, the Kruskal-Wallis test is performed. This test is an alternative to ANOVA when the "assumptions of one-way ANOVA are not met" (Statistics Solutions, 2018c). If there are significant differences, post-hoc tests can identify where they lie. The research questions focused on analysing the influence that recipe healthiness, image and interest group have on interactions contain only two groups. The independent variables being the factors, and the dependent variable always being clicks or impressions. Depending on whether the data is normally distributed or not, an independent t-test or a Mann-Whitney U test will show whether there is a difference in the CTR and Impressions in, for instance, the healthy compared to the unhealthy recipe advertisements. If the tests all result in the conclusion that the variable is not normally distributed, an independent samples t-test is performed. This test is a version of the t-tests, in which means of two groups are compared with normally distributed data. It is an "analysis of dependence" (Statistics Solutions, 2018d). If the tests all result in the conclusion that the variable is not normally distributed, a Mann-Whitney U test is performed. This is the non-parametric version of the independent t-test. For this test, "results are presented in group rank differences rather than group mean differences" (Statistics Solutions, 2018b).

One research question focuses on correlation analysis in order to see if a correlation between two groups is present. In case of this thesis, the groups consist of the dependent variable, clicks or impressions. The independent variable is obesity and diabetes prevalence in the states of the USA. The testing is done between the variables in order to see if a significant correlation exists. Among different correlation analysis methods, Pearson’s correlation is chosen. According to the Merriam Webster dictionary, a correlation is “a relation existing between phenomena or things or between mathematical or statistical variables which tend to vary, be associated, or occur together in a way not expected on the basis of chance alone” (Merriam Webster, 2018).

Ponocny and Weismayer (2016a) explain that a correlation always lies between -1 and +1 and can never be a value above or beyond that. The formula in Figure 3-13 shows the calculation of a correlation Pearson’s correlation. A scatterplot represents a correlation with dots, and always indicates how the correlation looks like. When a large x value corresponds to a large y value, statisticians talk about a strong positive correlation, as well as when small x values correspond to small y values. The scatterplot indicates the direction, in this case it looks „slim” with the regression line ascending. On the other hand, when there is a strong negative correlation, a large x values corresponds to a small y value and contrarily the same. The closer the correlation value, r, is to -1 or 1, the stronger the relationship is. In a scatterplot, all points would lie on one line in this case. If the correlation value is close or equal to 0, it means that the variables show no linear relationship. In this case, the scatterplot would show a regression line that looks horizontal. If a correlation analysis is significant, then a relationship between both variables is proven. To start a correlation analysis, two variables are defined. A decision between one or two-tailed testing is made, depending on whether it is believed that the hypothesis will have a certain direction or not. Then, the analysis is performed with the chosen p-value is performed (Ponocny & Weismayer, 2016a).

Each analysis needs a measurement of when the results are statistically significant, which is the p-value. Thisted (1998) explains that “p-values exceeding 0.05 (one in twenty) just aren’t strong enough to be the sole evidence that two treatments being studied really differ in their effect” (Thisted, 1998). This is why this thesis employs a p-value of 0.05 for all statistical tests mentioned below.

4 RESULTS

	Ad 1	Ad 2	Ad 3	Ad 4	Ad 5	Ad 6	Ad 7	Ad 8	Total
Results	323	346	799	872	474	409	641	545	4409
Reach	5800	6987	15911	16152	8548	8022	12892	10560	63983
Impres-sions	6327	7639	16943	17625	9241	8401	13918	11252	91346
Cost per result	€0,12	€0,13	€0,10	€0,10	€0,12	€0,12	€0,12	€0,12	€0,11
Amount spent	€39,23	€43,35	€83,81	€84,45	€57,29	€50,91	€75,09	€65,87	€500
Unique Link Clicks	300	332	756	818	449	385	601	517	3450
CTR	6.27 %	5.51 %	5.93 %	6.03 %	6.51 %	6.25 %	6.52 %	6.13 %	6.14 %

TABLE 4-1 ADVERTISEMENT OUTCOME OF THE RECIPE PROMOTIONS IN TERMS OF VARIOUS INDICATORS

The previously explained methods are used to analyse the data discovered in this experiment. In this chapter, each research question is answered by means of statistical analyses, for which an alpha level of .05 is used. The research questions all contain sub-questions, which address the different types of measurements of interaction, such as impressions and CTR. Some sub-questions also examine each research question in more detail.

4.1 RQ1: What is the general response to advertisements promoting online recipes?

Table 4-1 shows the general outcome from the advertisements based on the listed indicators. As seen on the table, all advertisements were seen up to 91,346 times. Out of those, 4,409 users clicked on them at an average cost of €0.11 per result. The total click-through rate is 6.14%, which means that out of all the people that have seen the advertisements, 6.14% performed a click. The table shows that all advertisements used up a different amount of money from the total budget of €500. Ad3 and Ad4 used up the most, with approximately €84. The least amount of money was spent on Ad1. In regard to the Unique link clicks, Ad4 generated the most results out of all eight advertisements, while Ad1 generated the least. Ad4 is also the

Ranks

	Ad	N	Mean Rank
Impressions	1.00	51	139.98
	2.00	51	159.25
	3.00	51	235.52
	4.00	51	247.74
	5.00	51	197.61
	6.00	51	188.86
	7.00	51	246.73
	8.00	51	220.32
	Total	408	

TABLE 4-2 MEAN RANKS OF DIFFERENCES IN CTR BETWEEN ADS

ad with the highest amount of impressions, with it being on screen 17,625 times, compared to Ad1 being on screen 6,327 times. A correlation analysis between results and impressions shows that the two variables show a significant strong, positive correlation, $r(6) = .99$, $p < .001$. This means that results depend a lot on the number of impressions. The advertisement that generated the highest click-through-rate is Ad7, and Ad2 generated the lowest. However, there is no significant correlation between CTR and impressions, which implies that the advertisements were shown more often but still not a higher amount of people in relation to that clicked on them. Statistical analyses of the results also evaluate whether or not there is a general difference in the CTR and impressions between the eight different advertisements. Detailed statistical analyses can be found in Appendix 4.

RQ1.1 Are there differences between the ads in respect to CTR?

The independent variable, which consists of the different advertisements, is tested to see if it influences the dependent variable, CTR. The K-S test shows that the data is normally distributed and therefore a parametric test should be performed. However, after testing the homogeneity of variances in ANOVA, not all criteria to perform a one-way ANOVA is fulfilled. Therefore, a Kruskal-Wallis test is used to analyse the data. The outcome of the statistical Kruskal-Wallis test performed on the variables is non-significant. It shows that there is no significant difference between the CTR in the eight advertisements ($p = .624$). This implies that users do not have preferences for any specific type of advertisement, but the advertisement generated a similar CTR with all different properties.

RQ1.2 Are there differences between the ads in respect to impressions?

This question analyses if the independent variable, advertisements, has an effect on the dependent variable, impressions. After a significant K-S test, which calls for a non-parametric test, a Kruskal-Wallis test is performed for this question. The results show that there are significant differences between the groups analysed in terms of impressions ($H = 41.69$, $p < .01$). This means that certain advertisements got shown more often, whereas other advertisements were on screen a smaller number of times. Mean ranks, as seen in Table 4-2, show that advertisement 3,4,7 and 8 got shown the most often, which the statistics from the Ad Manager also confirm. The first two advertisements feature the healthy recipe with the appealing pictures targeted to both interest sets, and the latter feature the unhealthy recipe with the unappealing pictures targeted to both interest sets.

It can be determined from those questions that users did not have a preference in any specific advertisement based on how often they performed a click. Certain recipe advertisements, however, got shown more often than others, which included advertisements 3,4,7 and 8.

4.2 RQ2: To what extent does recipe healthiness influence the interaction of users with advertisements?

RQ2.1 Are there differences between the healthy and unhealthy recipe advertisements in terms of CTR?

Healthiness of the recipe, which is the independent variable, is analysed to examine whether or not it has an effect on the dependent variable, CTR. After the K-S test shows that the data is not normally distributed, a non-parametric Mann-Whitney U test is performed. There are no significant differences found between the healthy and unhealthy recipe in terms of CTR ($p = .183$). Facebook users therefore seem to have no preference of what kind of recipe advertisement they interact with, healthy or not. In fact, each recipe was clicked on a similar amount of times based on how often it was seen.

RQ2.2 Are there differences between the healthy and unhealthy recipe advertisements in terms of Impressions?

Healthiness serves as the independent variable, and impressions as the dependent. Similar to the analysis above, a K-S test is employed to test for the normality of distribution in the data. As it is not normally distributed, the non-parametric Mann-Whitney U test afterwards shows non-significant results ($p = .128$). Impressions also show no significant difference between the healthy and unhealthy recipe. This means that both recipes were on screen an almost equal amount of times. Facebook did not show the healthy recipe more often than the unhealthy one, or vice versa.

The data above shows that recipe healthiness plays no role in recipe selection for the users in the United States. Neither one of the recipes got shown more often, they simply were shown a similar amount of times with no difference in user reaction to it. Detailed analyses for this question can be found in Appendix 5.

4.3 RQ3: To what extent does the image used in a recipe advertisement influence the user's interaction?

Various images that were either rated appealing or unappealing were used in the promotions. The following questions answer whether there was a difference in interactions between the image types. The analysis can be found in Appendix 6.

RQ3.1 Is there a difference between the appealing and the unappealing image in terms of CTR?

For this question, the dependent variable CTR is used to test whether or not a difference can be seen in the independent variable, image attractiveness. The K-S test calls for a parametric test. An Independent Samples t-test is performed on all advertisement's data. The four images, two of them appealing and two of them unappealing, are compared by their means with this test. The results show that there is no significant difference between both types of images ($p = .308$). Thus, the image used in the advertisements did not have an effect on how often users clicked on the advertisements. Whether the advertisement promoted the recipe with an appealing or an unappealing image, both were interacted with an almost identical amount of times. The appealing, compared to the unappealing image, did not generate more clicks.

RQ3.2 Is there a difference between the appealing and the unappealing image in terms of Impressions?

As in the question above, the independent variable is image attractiveness, however the dependent variable is impressions. For this question, an analysis on the impressions for the different types of images is done which shows different results than for the CTR. The non-parametric Mann-Whitney U test shows a significant difference between the two variables ($U = 18170.00$, $p = .027$). The appealing images had a higher mean rank, implying that those images were on screen more often. This means that people saw the more appealing images more often than the unappealing ones in all ads. This could be different in each type of recipe though, so after analysing the different types of images for both recipes, the images for each recipe are also tested separately. The same variables as in the previous questions are used for the analyses, just filtered by recipe healthiness. The following questions address this subject.

RQ3.3 Is there a difference between the appealing and the unappealing image in terms of CTR on the healthy recipe?

For the healthy recipe, a parametric Independent Samples t-test show that no significant difference can be seen between the appealing and unappealing image in terms of CTR ($p = .129$). As with the analysis of overall images, the ones that linked to the healthy recipe also got

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	Unappealing	*Appealing*	*Total*	*Unappealing*	*Appealing*	*Unappealing*	*Appealing*
Impressions	204.00	204.00	408.00	191.57	217.43	39080.00	44356.00

TABLE 4-3 MEAN RANKS OF DIFFERENCES IN IMPRESSIONS BETWEEN IMAGES IN HEALTHY RECIPE

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	Appealing	*Unappealing*	*Total*	*Appealing*	*Unappealing*	*Appealing*	*Unappealing*
Impressions	102.00	102.00	204.00	91.78	113.22	9361.50	11548.50

TABLE 4-4 MEAN RANKS OF DIFFERENCES IN IMPRESSIONS BETWEEN IMAGES IN UNHEALTHY RECIPE

clicked through to the same amount of times. No matter if the image was appealing or unappealing, users still reacted the same way.

RQ3.4 Is there a difference between the appealing and the unappealing image in terms of Impressions on the healthy recipe?

As with impressions for both recipes together, a Mann-Whitney U test shows a significant difference in terms of impressions between the images in the healthy recipe (U = 2951.50, $p < .01$). The difference in mean ranks can be found in Table 4-3, which shows that the appealing image had a higher mean rank. The appealing image was shown to users a bigger number of times. The unappealing one did not appear on their screen as often in the healthy recipe.

RQ3.5 Is there a difference between the appealing and the unappealing image in terms of CTR on the unhealthy recipe?

Similar to the above tested images and CTR, there is no significant difference found between the image types of the healthy recipe in terms of CTR ($p = .996$). Users again were indifferent to the images used while promoting the unhealthy recipe.

RQ3.6 Is there a difference between the appealing and the unappealing image in terms of Impressions on the unhealthy recipe?

With the non-parametric data for impressions, a Mann-Whitney U test once more shows a significant difference between the images, however this time in the unhealthy recipe as well (U = 4108.50, $p < .01$). Mean ranks show that for this question, the unappealing image score is higher, which is shown in Table 4-4. This means that users did, in fact, get shown the less appealing image more often than the appealing one.

The above analysed data reveals that between the different types of images, CTR does not show a difference. Impressions, however, are always different for each type of image. This means Facebook does show certain images more often than others.

4.4 RQ4: To what extent do user interests play a role in interactions with the recipe advertisement?

As the advertisements were targeted to two different interest groups, users with unhealthy and users with healthy interests, results on both targeted audiences reveal if the groups show a difference, with statistical details being shown in Appendix 7.

RQ4.1 Is there a difference between the healthy and the unhealthy interest group in terms of CTR?

This question focuses on whether the independent variable, interest group, shows a difference in the dependent variable, CTR. The K-S test requires a non-parametric test again, which is why a Mann-Whitney U test shows that there is no significant difference between the healthy and unhealthy interest group in terms of CTR, with the significance level being above .05 ($p = .434$). The targeted interest groups consisted of one group that represents individuals with a healthy lifestyle, whereas the other group represented people with an unhealthy lifestyle. Both were targeted by the advertisements; however, the response rate is the same for the two groups. This implies that reactions on the advertisements are not influenced by the individual's life-style.

RQ4.2 Is there a difference between the healthy and the unhealthy interest group in terms of Impressions?

The independent variable, being the interest group, stays the same, however this question addresses the dependent variable impressions. Regarding impressions, the non-parametric Mann-Whitney U test also shows no significant differences between the two interest groups ($p = .937$). Impressions did not change based on which interest group was targeted in the advertisement. Whether it was the healthy group or the unhealthy, both were exposed to all advertisements a similar amount of times.

Results to both question show that user interest does not play a role in recipe selection. Both interest sets clicked on advertisements a similar number of times, which shows that users with a healthy interest do not necessarily interact with recipes more often. Users with unhealthy interests, however, also click on recipes a similar amount of times as unhealthy ones.

4.5 RQ5: To what extent does state healthiness play a role in the selection of recipe advertisements?

As obesity and diabetes prevalence data is available from the CDC website, this thesis also analyses whether interactions to the advertisements show a correlation to health data. The following questions are being answered to resolve whether the two variables are connected.

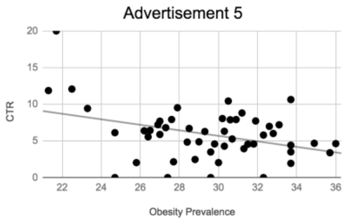
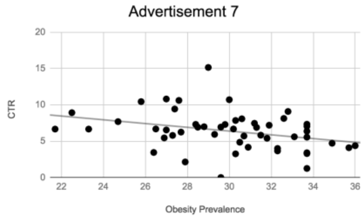
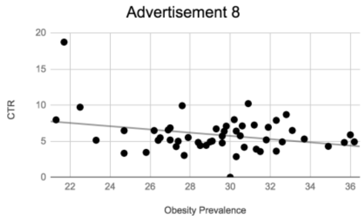
Ad 5	Ad 7	Ad 8
		
r(48) = -.41	*r*(48) = -.28	*r*(48) = -.29

FIGURE 4-1 SCATTERPLOTS OF CTR AND OBESITY PREVALENCE AND THEIR CORRESPONDING CORRELATION COEFFICIENT

Detailed correlation analyses can be found in Appendix 8. The independent variables for the following questions are the health statistics, which are diabetes prevalence rates and obesity prevalence rates. The dependent variables also vary per question, with either CTR or impressions being analysed for both independent variables.

RQ5.1 Is there a correlation of obesity prevalence and CTR on each individual advertisement?

Advertisements 1 to 4 show no significant correlations between CTR and obesity prevalence rates. Advertisement 6 also shows no significant correlation. However, Advertisements 5, 7 and 8 all show significant negative correlations for the click-through rate and the obesity rates in all states. In Advertisement 5, a rather strong significant negative correlation of $r(48) = -.41$, $p < .01$ is shown. Advertisement 7 shows a negative correlation of $r(48) = -.28$, $p = .050$, with the significance level being exactly on the edge of still being significant. Advertisement 8 also has a rather high significance value, but the correlation present is also a moderate negative one with $r(48) = -.29$, $p = .046$. All three advertisements that show a correlation between clicks and obesity prevalence link to the unhealthy recipe advertisement. As the correlation is negative, it implies that states with a low obesity prevalence responded more to the unhealthy recipes. The scatterplots in Figure 4-1 represent the correlations of Advertisement 5,7 and 8.

RQ5.2 Is there a correlation of obesity prevalence and Impressions on each individual advertisement?

No significant correlations between obesity prevalence and impressions can be observed. The amount of times the advertisements were on screen did not correlate with obesity rates in the various states.

RQ5.3 Is there a correlation of diabetes prevalence and CTR on each individual advertisement?

All advertisements except for Advertisement 4 show no significant correlations of diabetes prevalence and CTR. Advertisement 7 comes close to a significant negative correlation, with $r(48) = -.27$, $p = .062$. Advertisement 4 shows a significant positive correlation of $r(49) = .36$,

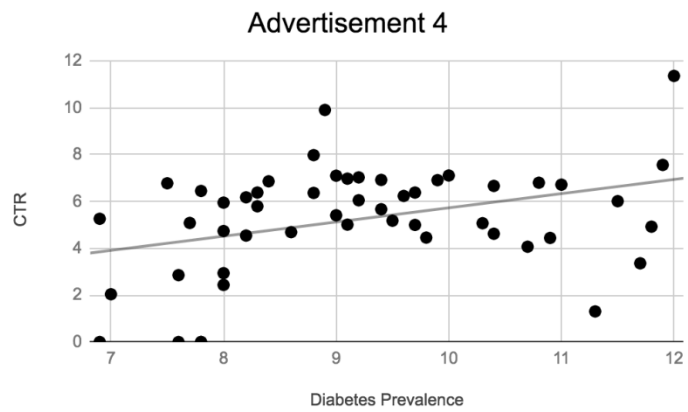
Ad 4

$r(49) = 0.36$

FIGURE 4-2 SCATTERPLOT OF CTR AND DIABETES PREVALENCE AND THE CORRESPONDING CORRELATION COEFFICIENT

$p < .01$, which is a moderate correlation between the CTR on the advertisements and diabetes prevalence in each state. The scatterplot shows the visualization of this correlation in Figure 4-2. When the click-through-rate gets smaller, diabetes prevalence also gets smaller. This advertisement was targeted to the unhealthy interest group but linked to the healthy recipe with an appealing image. The higher the diabetes rate, the more users with unhealthy interests responded, which shows that unhealthy interests and diabetes prevalence rates are connected.

RQ5.4 Is there a correlation of diabetes prevalence and Impressions on each individual advertisement?

This next question looks into the Pearson's correlations between diabetes prevalence rates and impressions in all states. Compared to the previous question, there is more than one correlation present. The first four advertisements, namely 1,2,3 and 4 all show no significant correlation between the variables. Advertisements 5, 6, 7 and 8 all show a significant correlation. All latter advertisements promote the unhealthy recipe, whereas the previous advertisements promote the healthy recipe. Advertisement 5 shows a correlation between impressions and diabetes prevalence of $r(49) = .30$, $p = .033$. Advertisement 6 also shows a moderate correlation, which is exactly the same as the previous one, $r(49) = .30$, $p = .033$. In Advertisement 7, a correlation of $r(49) = .33$, $p = .020$ is observed. Advertisement 8 again shows the same correlation value, with a different significance, $r(49) = .30$, $p = .030$. It becomes apparent that the unhealthy recipe promotions all show an almost equal correlation to diabetes prevalence. This means that the higher the amount of impressions on the unhealthy recipe promotion is, the more people from diabetes prevalent regions see the ad. The scatterplots make the data visual. As seen in Figure 4-3 and Figure 4-4, one variable can predict the other.

RQ5.5 Is there a correlation of obesity prevalence and CTR in all advertisements?

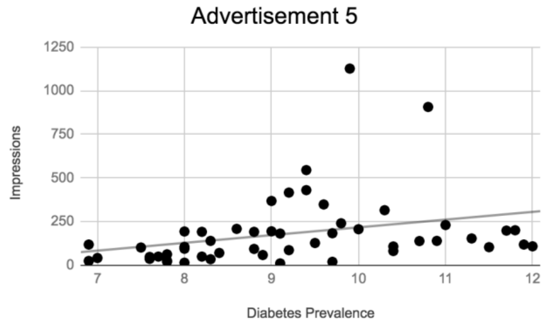
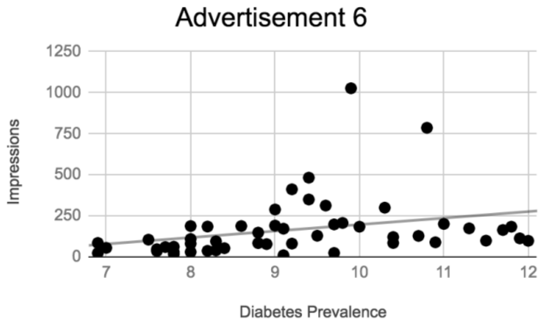
Ad 5	Ad 6
	
$r(49) = .30$	$r(49) = .30$

FIGURE 4-3 SCATTERPLOTS OF IMPRESSIONS AND DIABETES PREVALENCE FOR AD 5 & 6 AND CORRELATION COEFFICIENTS

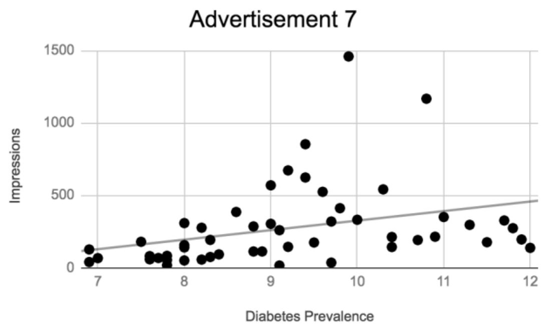
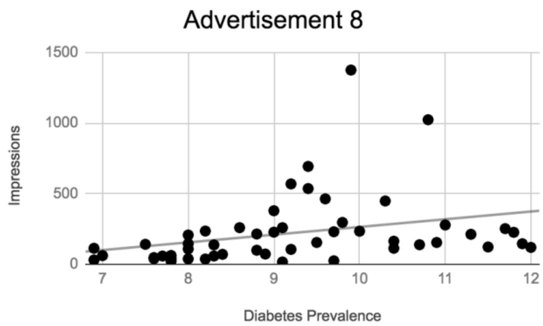
Ad 7	Ad 8
	
$r(49) = .33$	$r(49) = .30$

FIGURE 4-4 SCATTERPLOTS OF IMPRESSIONS AND DIABETES PREVALENCE FOR AD 7 & 8 AND CORRELATION COEFFICIENTS

By analysing the data of all eight different advertisements together, the Pearson's correlation value reports a significant correlation for obesity prevalence and CTR. The outcome shows a rather strong negative correlation of $r(49) = -.37$, $p < .01$. This implies that the higher the CTR, the lower the obesity prevalence of the targeted states is.

RQ5.6 Is there a correlation of diabetes prevalence and CTR in all advertisements?

In regard to diabetes in all states, there seems to be no significant correlation between CTR and diabetes prevalence, as the alpha value is too high to show significant results.

The results for this research question show that firstly, unhealthy states get shown the unhealthy recipes more often. When looking at how frequently people from those states click on the advertisements, there were no significant correlations, which means that unhealthy states do not automatically click unhealthier recipes. Also, diabetes rates tend to correlate with user interests, as results showed that users with unhealthy interests generate a higher CTR when

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
CTR	Alabama	7	3.84	1.00	.38	2.92	4.77	2.47	5.31
	Alaska	7	5.50	3.80	1.44	1.99	9.02	1.25	10.87
	Arizona	8	5.03	1.87	.66	3.46	6.59	2.92	8.19
	Arkansas	8	5.31	1.93	.68	3.70	6.93	1.94	8.33
	California	8	6.38	.77	.27	5.73	7.02	4.80	7.48
	Colorado	7	7.04	2.42	.92	4.80	9.28	4.65	11.86
	Connecticut	7	5.64	2.46	.93	3.36	7.92	2.50	9.47
	Delaware	8	6.58	5.41	1.91	2.06	11.11	.00	15.38
	District of Columbia	8	8.67	7.57	2.68	2.35	15.00	.00	20.00
	Florida	8	6.57	.93	.33	5.79	7.34	5.02	7.68
	Georgia	8	6.89	2.00	.71	5.22	8.57	4.86	10.43
	Hawaii	8	9.91	4.36	1.54	6.26	13.56	3.65	16.22
	Idaho	8	5.22	4.35	1.54	1.59	8.86	1.85	15.15
	Illinois	8	6.36	1.26	.44	5.31	7.41	4.01	8.31
	Indiana	8	6.37	2.23	.79	4.51	8.23	3.67	9.09
	Iowa	8	5.76	3.64	1.29	2.71	8.80	.00	10.64
	Kansas	8	4.32	2.21	.78	2.47	6.17	1.49	7.81
	Kentucky	8	4.44	2.05	.73	2.73	6.16	1.31	6.76
	Louisiana	8	5.14	1.24	.44	4.10	6.18	3.33	7.14
	Maine	8	7.32	2.46	.87	5.26	9.37	4.84	10.53
	Maryland	7	7.28	2.80	1.06	4.69	9.87	3.49	11.43
	Massachusetts	8	7.02	2.52	.89	4.92	9.12	2.81	10.20
	Michigan	8	6.79	1.66	.59	5.40	8.17	4.74	9.60
	Minnesota	7	7.52	2.97	1.12	4.77	10.27	2.92	11.76
	Mississippi	8	4.01	2.54	.90	1.89	6.13	.89	7.56
	Missouri	8	5.88	1.80	.64	4.38	7.39	2.73	7.85
	Montana	8	5.16	2.73	.96	2.88	7.44	.00	9.52
	Nebraska	8	6.12	3.91	1.38	2.86	9.39	2.86	14.89
	Nevada	8	4.74	2.78	.98	2.41	7.07	1.33	8.43
	New Hampshire	8	4.21	3.18	1.12	1.56	6.87	.00	9.43
	New Jersey	8	5.97	1.48	.52	4.74	7.21	3.74	8.02
	New Mexico	8	4.02	1.72	.61	2.58	5.46	1.75	6.99
	New York	8	6.48	.78	.27	5.83	7.13	5.34	7.69
	North Carolina	8	6.12	1.31	.46	5.02	7.21	4.46	7.87
	North Dakota	8	5.68	5.04	1.78	1.47	9.89	.00	16.67
	Ohio	8	6.03	1.26	.44	4.98	7.09	4.91	8.13
	Oklahoma	8	5.62	2.55	.90	3.50	7.75	1.49	10.23
	Oregon	8	6.31	3.10	1.10	3.72	8.91	2.14	10.84
	Pennsylvania	7	7.05	.76	.29	6.35	7.76	6.16	8.05
	Rhode Island	8	8.61	4.43	1.56	4.91	12.31	4.84	17.95
	South Carolina	7	4.84	1.39	.53	3.55	6.12	3.62	7.50
	South Dakota	7	3.76	4.35	1.64	-.26	7.78	.00	10.71
	Tennessee	7	4.99	1.78	.67	3.35	6.64	3.35	7.98
	Texas	8	6.41	.96	.34	5.60	7.21	4.96	7.72
	Utah	8	5.56	4.02	1.42	2.20	8.92	.00	11.54
	Vermont	7	4.17	6.82	2.58	-2.13	10.48	.00	18.18
	Virginia	8	5.57	1.40	.49	4.40	6.74	3.57	7.10
	Washington	8	6.53	1.34	.47	5.41	7.65	4.26	8.70
	West Virginia	8	7.34	2.32	.82	5.39	9.28	4.38	11.36
	Wisconsin	8	6.09	2.14	.76	4.30	7.88	2.63	8.81
	Wyoming	8	1.43	2.69	.95	-.82	3.68	.00	6.67
	Total	397	5.88	3.13	.16	5.57	6.19	.00	20.00

TABLE 4-5 DESCRIPTIVES FOR ANOVA ON CTR IN REGIONS INCLUDING MEAN AND STANDARD DEVIATION

they come from regions with a higher diabetes prevalence. However, states with a low obesity rate tend to respond more to unhealthy recipes.

4.6 RQ6: How do reactions to advertisements differ among user characteristics?

Firstly, the regions targeted by the advertisements are in the USA. One of the characteristics that Facebook analyses is how many users who live in a certain state clicked on an advertisement. The first question below looks at the regions they come from.

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean: Lower Bound	95% Confidence Interval for Mean: Upper Bound	Minimum	Maximum
Impressions	*Alabama*	8	176.25	68.58	24.25	118.92	233.58	73.00	273.00
	Alaska	8	51.13	17.59	6.22	36.42	65.83	25.00	80.00
	Arizona	8	260.38	117.22	41.44	162.38	358.37	164.00	459.00
	Arkansas	8	108.38	39.44	13.94	75.40	141.35	48.00	176.00
	California	8	1620.88	743.44	262.85	999.34	2242.41	1027.00	2903.00
	Colorado	8	136.88	65.67	23.22	81.98	191.77	84.00	247.00
	Connecticut	8	126.50	54.60	19.30	80.85	172.15	70.00	211.00
	Delaware	8	25.38	10.21	3.61	16.84	33.91	13.00	40.00
	District of Columbia	8	18.50	11.41	4.04	8.96	28.04	9.00	43.00
	Florida	8	692.13	253.69	89.69	480.04	904.21	396.00	1069.00
	Georgia	8	265.75	98.30	34.75	183.57	347.93	127.00	393.00
	Hawaii	8	83.88	33.14	11.72	56.17	111.58	37.00	137.00
	Idaho	8	58.75	22.26	7.87	40.14	77.36	28.00	98.00
	Illinois	8	566.00	262.51	92.81	346.53	785.47	302.00	970.00
	Indiana	8	198.50	73.26	25.90	137.25	259.75	90.00	319.00
	Iowa	8	124.75	38.58	13.64	92.50	157.00	78.00	192.00
	Kansas	8	143.88	47.67	16.85	104.02	183.73	67.00	208.00
	Kentucky	8	157.50	73.33	25.93	96.19	218.81	72.00	296.00
	Louisiana	8	123.63	49.02	17.33	82.65	164.60	56.00	212.00
	Maine	8	52.50	20.63	7.29	35.26	69.74	19.00	82.00
	Maryland	8	129.25	69.28	24.49	71.33	187.17	70.00	256.00
	Massachusetts	8	244.13	107.05	37.85	154.63	333.62	138.00	424.00
	Michigan	8	350.88	124.72	44.10	246.60	455.15	190.00	569.00
	Minnesota	8	126.75	38.26	13.53	94.77	158.73	73.00	179.00
	Mississippi	8	112.75	46.20	16.34	74.12	151.38	52.00	195.00
	Missouri	8	196.50	77.16	27.28	131.99	261.01	92.00	331.00
	Montana	8	33.50	12.01	4.25	23.46	43.54	20.00	58.00
	Nebraska	8	69.25	21.70	7.67	51.10	87.40	41.00	102.00
	Nevada	8	124.50	65.84	23.28	69.46	179.54	75.00	244.00
	New Hampshire	8	35.13	14.30	5.05	23.17	47.08	15.00	55.00
	New Jersey	8	356.13	204.73	72.38	184.96	527.29	187.00	682.00
	New Mexico	8	113.75	53.05	18.76	69.40	158.10	57.00	199.00
	New York	8	675.38	348.44	123.19	384.07	966.68	404.00	1305.00
	North Carolina	8	286.12	123.03	43.50	183.27	388.98	127.00	471.00
	North Dakota	8	28.00	12.48	4.41	17.57	38.43	14.00	50.00
	Ohio	8	332.75	129.66	45.84	224.35	441.15	153.00	541.00
	Oklahoma	8	127.13	48.99	17.32	86.17	168.08	67.00	214.00
	Oregon	8	132.38	47.01	16.62	93.08	171.67	81.00	205.00
	Pennsylvania	8	380.50	124.34	43.96	276.55	484.45	194.00	529.00
	Rhode Island	8	52.13	23.99	8.48	32.07	72.18	28.00	94.00
	South Carolina	8	134.13	42.85	15.15	98.30	169.95	67.00	191.00
	South Dakota	8	38.38	11.25	3.98	28.97	47.78	22.00	56.00
	Tennessee	8	193.13	79.03	27.94	127.05	259.20	94.00	326.00
	Texas	8	1279.88	699.70	247.38	694.91	1864.84	727.00	2484.00
	Utah	8	69.38	36.59	12.94	38.78	99.97	26.00	132.00
	Vermont	8	24.25	9.10	3.22	16.64	31.86	12.00	39.00
	Virginia	8	232.87	86.98	30.75	160.16	305.59	112.00	352.00
	Washington	8	240.00	78.61	27.79	174.28	305.72	161.00	356.00
	West Virginia	8	85.88	36.46	12.89	55.39	116.36	35.00	137.00
	Wisconsin	8	204.38	61.68	21.81	152.81	255.94	114.00	308.00
	Wyoming	8	14.63	5.26	1.86	10.23	19.02	7.00	21.00
	Total	408	223.83	334.04	16.54	191.32	256.34	7.00	2903.00

TABLE 4-6 DESCRIPTIVES FOR ANOVA ON IMPRESSIONS IN REGIONS INCLUDING MEAN AND STANDARD DEVIATION

RQ6.1 Is there a significant difference in CTR and the user's region in the advertisements?

The independent variable for this question is region, which comprises of the states in the USA. The dependent variable is the CTR. An ANOVA test shows that there are significant differences between the regions and the CTR, $F(50, 346) = 1.77$, $p < .01$. In Table 4-5 all values are shown, where the mean of each region can be determined. As seen in this table, not all states generated a CTR for all advertisements. Some of the states only show the number 7, which means that there was no CTR on some advertisements. "Hawaii" showed the highest click-through-rate (M = 9.91, SD = 4.36), whereas the lowest CTR is in Wyoming (M = 1.43, SD = 2.69). The

District of Columbia also shows a high CTR (*M* = 8.67, *SD* = 7.57). This means that Hawaii is the state that interacted the most with the recipes by clicking on them. Wyoming, however, did not click on the recipes very often, as seen by the mean of the CTR of *M* = 1.43, which suggests that the average rate of clicking on the advertisements compared to seeing it is quite low.

RQ6.2 Is there a significant difference in impressions and the user's region in the advertisements?

Other than in the previous question, this one tests the dependent variable impressions in the different regions, which is the independent variable. The parametric ANOVA test on Impressions shows that there are also significant differences between the regions in terms of this variable, $F(50, 357) = 23.22$, $p < .01$. Table 4-6 shows that California has the highest amount of impressions (*M* = 1620.88, *SD* = 743.44). As with the CTR, Wyoming has the least impressions (*M* = 14.63, *SD* = 5.26). Compared to the CTR, however, the mean value for impressions in the District of Columbia is very low (*M* = 18.50, *SD* = 11.41). The users that saw the advertisements the least amount of times also come from Wyoming, as did the least clicks. California got shown the advertisements most often, so many of the inhabitants of this state got exposed to the advertisements.

Apart from regions, age is another characteristic that Facebook identifies in users. The following questions examine which ages interacted the most with recipe advertisements.

RQ6.3 Is there a significant difference in CTR and the user's age in the advertisements?

Age serves as the independent variable, with the independent variable being the CTR. Following a K-S test to see if the data is normally distributed, the parametric ANOVA test shows that there are significant differences between the user's age in terms of CTR, $F(5, 40) = 25.88$, $p < .01$. The group with the highest CTR is the age group 65+ (*M* = 7.77, *SD* = .72). The lowest CTR comes from the lowest age group, which is 18-24-year-old users (*M* = 3.95, *SD* = .90). 25 to 34-year-olds have the second lowest CTR (*M* = 4.17, *SD* = .98). As seen in Table 4-7, the higher the age group is, the bigger the CTR. This means that people that are 65 years or older interacted the most with the advertisements. The younger the audience gets, the fewer of them interact with the recipe advertisements.

Since it is known that there is a difference between age groups, a post-hoc test is performed to determine between which variables the difference lies. For this, Scheffe's test is performed. The analysis shows that there are significant differences between 55-64 (*M* = 6.30, *SD* = .45) and 65+ (*M* = 7.77, *SD* = .72) to the age group of 18-24, 25-34 and 35-44. Among others, the age group 65+ is significantly different to all the other age groups. Those findings again suggest that older people interact with recipe advertisements more often, which could imply that they prefer recipe advertisements to the younger generations.

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
CTR	18-24	8	3.95	.90	.32	3.20	4.71	1.869158878505	4.791666666667
	25-34	7	4.17	.98	.37	3.26	5.07	2.787456445993	5.967078189300
	35-44	8	4.80	.80	.28	4.13	5.46	3.435804701628	5.901116427432
	45-54	8	5.10	.68	.24	4.53	5.67	4.178272980501	6.191279209829
	55-64	8	6.30	.45	.16	5.93	6.68	5.637358014304	6.847764572722
	65+	7	7.77	.72	.27	7.11	8.44	6.536541080345	8.777915632754
	Total	46	5.32	1.49	.22	4.88	5.76	1.869158878505	8.777915632754

TABLE 4-7 DESCRIPTIVES FOR ANOVA ON CTR BETWEEN AGE GROUPS INCLUDING MEAN AND STANDARD DEVIATION

Ranks

	Age	N	Mean Rank
Impressions	18-24	8	9.56
	25-34	8	12.63
	35-44	8	21.56
	45-54	8	29.63
	55-64	8	34.88
	65+	8	38.75
	Total	48	

TABLE 4-8 KRUSKAL-WALLIS RANKS ON IMPRESSIONS BETWEEN AGE GROUPS

RQ6.4 Is there a significant difference in impressions and the user's age in the advertisements?

As above, the independent variable is age. The dependent variable is impressions, with this question focusing on the difference in impressions by age. A non-parametric Kruskal-Wallis test shows that there is a significant difference between age groups in terms of impressions as well (H = 28.97, p < .01). Table 4-8 shows that for impressions, the age group of 65+ had the highest mean rank. Similar to the question before, the SPSS output shows that the higher the age group seems to be, the higher the amount of impressions on the advertisements is. To find out where the difference lies, a post hoc test is performed. In this case, 15 Mann-Whitney U tests with a Bonferroni correction to compare each group are conducted. The Bonferroni correction is done by implementing a new significance value of 0.05/15, which is p = .003. This shows that the groups 25-34 and 65+ are significantly different from each other. Also, age groups 25-34 and 55-64 are significantly different. 18-24 and 65+ show a significant difference, as well as 18-24 and 55-64. Lastly, the 18-24 and 45-54 age groups are also different from each other in terms of impressions. This shows that there are many differences between the younger and older age groups, and less between age groups that follow in rank.

Age, region and gender are the most important characteristics, so the last question analyses which gender interacted the most with the advertisements based on CTR and impressions.

RQ6.5 Is there a significant difference in CTR and the user's gender in the advertisements?

In addition to only male and female, Facebook also introduces the third "unknown" group. This is why a test for more than two groups is necessary. The independent variable is gender and

Ranks

	Gender	N	Mean Rank
CTR	1	8	18.13
	2	8	10.50
	3	8	8.88
	Total	24	

TABLE 4-9 KRUSKAL-WALLIS RANKS ON CTR BETWEEN GENDER GROUPS

the dependent variable is CTR. To see whether a difference in the dependent variable between the three groups exists, the non-parametric Kruskal-Wallis test is performed, which shows that there is a significant difference between gender in terms of CTR (H = 7.81, p = .020). This means that the click-through-rate was not equal for women and men. One of the groups clicked on the recipes more often than the other. To see where the difference lies, again a post-hoc test is performed. Three Mann Whitney U tests with a Bonferroni correction are used, with a new significance value of 0.05/3, which is p = .016. Those post-hoc tests show that there is a significant difference, (U = 8.00, p = .012), between the male and female CTR. As seen in Table 4-9, the mean rank for female with M = 18.13 is higher than the mean rank for male with M = 10.50. Male and female users did not click through to the advertisements the same amount of times. Rather, female users did more often. This implies that female users are more interested in advertisements to recipes.

RQ6.6 Is there a significant difference in impressions and the user's gender in the advertisements?

To see if the dependent variable, impressions, shows a difference in the independent variable, gender, a Kruskal-Wallis test is performed. The analysis between impressions and gender also shows a significant difference between the variables (H = 20.48, p = p < .01). This means that one of the gender groups saw the advertisements more often than the other, as they appeared on their screen a higher amount of times. As with the CTR, a post-hoc Mann-Whitney-U test with a Bonferroni correction shows that there are significant differences between all the groups tested. The same correction with a significance level of 0.016 is used for this test. Opposed to the results from the previous question, the analysis shows that significant differences between all genders are identified (U = 0.00, p < .01). The mean rank of female users is M = 20.50, whereas for males it is M = 12.50, which is shown in Table 4-10. Female users got exposed to the advertisements the highest amount of times. Male and users with unidentified gender classified as "Unknown" were exposed less often.

Results of the questions above show that users clicking on recipe promotions generally tend to be older rather than younger, with 65+ aged users responding the most. The majority of them is female. The region interacting with advertisements the least is Wyoming.

Ranks

	Gender	N	Mean Rank
Impressions	1	8	20.50
	2	8	12.50
	3	8	4.50
	Total	24	

TABLE 4-10 KRUSKAL-WALLIS RANKS ON IMPRESSIONS BETWEEN GENDER GROUPS

Hawaii and California have higher interactions with the recipe promotions. Appendix 9 shows all statistical data for this research question.

5 DISCUSSION & LIMITATIONS

5.1 Summary

The purpose of this thesis is to investigate how Facebook advertisements can be used to promote healthy eating. The objective is to answer six research questions, which focus on the factors influencing the decision of a user to click on a recipe. Factors analysed were recipe healthiness, image, interests and lastly state health statistics. Certain user characteristics also play a role in the response to recipe advertisements. The following implications summarize the findings of the results.

- Generally, an outcome of 4,409 clicks and 91,346 impressions was achieved, which resulted in a CTR of 6.14% that proves to be high for advertisements in the health industry.

- Regarding whether the healthy or the unhealthy recipe is preferred, the results show that neither of the recipes were preferred. Both generated a similar amount of impressions and CTR.

- Whether the image used while promoting an advertisement affects the user's reaction was tested as well. The clear results made apparent that CTR does not differ depending on attractiveness of the image. However, impressions did show that the more appealing images were shown to people more often in total.

- How user interests play a role in recipe selection is the second question investigated in this thesis, where no significant differences between the two targeted interest groups were seen. This implies that no matter what the target group is, CTR and impressions were the same.

- The interactions of individuals with certain recipes can to some extent explain obesity and diabetes prevalence rates. Pearson's correlation analysis showed that states with high diabetes rates get shown the unhealthy advertisements more often. Also, if a recipe is targeted to a certain group like unhealthy individuals, those are more likely to click on this promotion as well.

- The last research question covers how the reactions to the advertisements differ among user characteristics. Users having the highest CTR were inhabitants of Hawaii, and impressions highest in California. The lowest CTR and impressions were generated in Wyoming. It also became evident that the older the users, the more they interact with advertisements promoting online recipes. Opposed to the average Facebook user

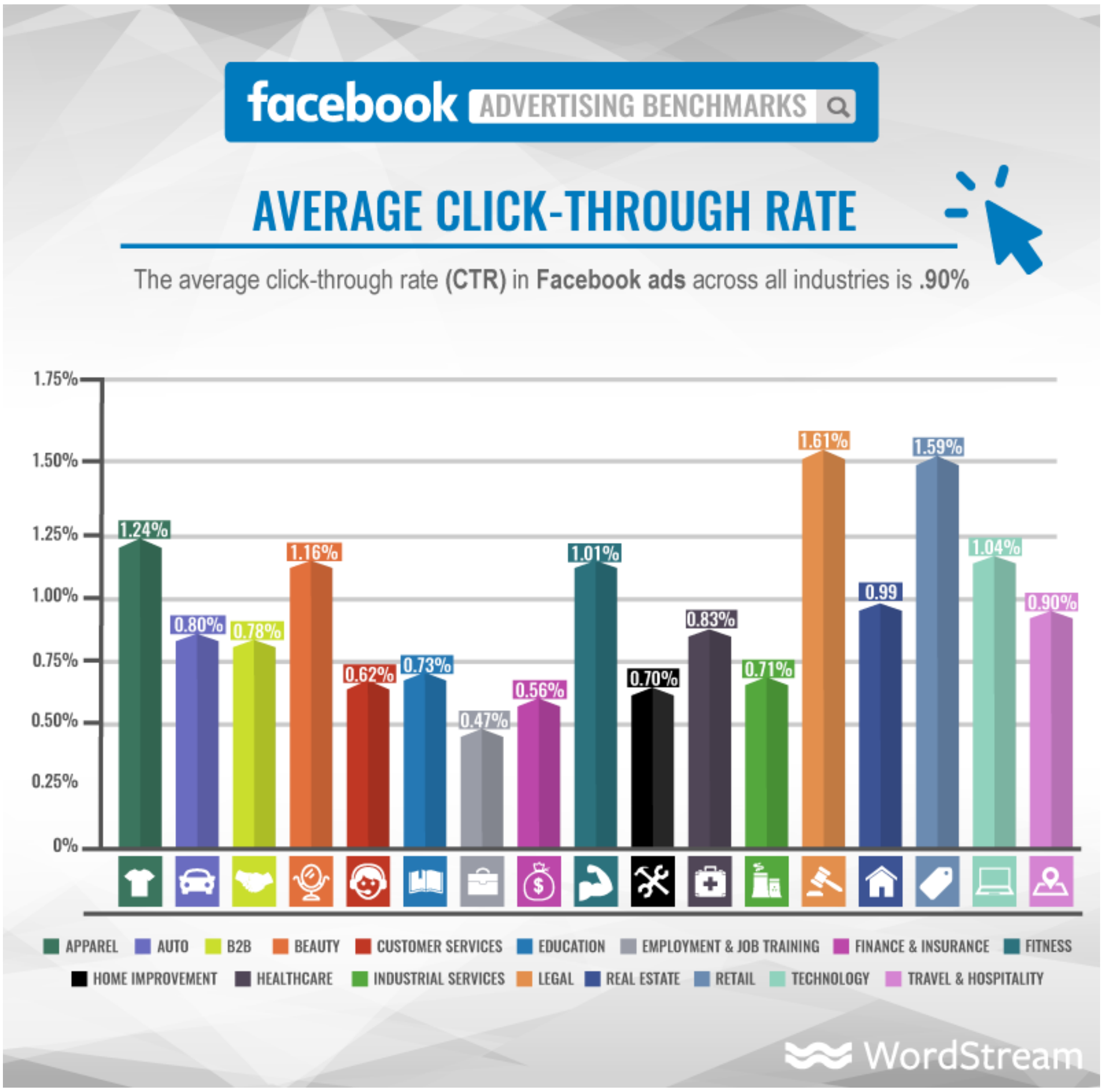


FIGURE 5-1 AVERAGE CLICK-THROUGH RATE

Source: Irvine, 2018

of age 25-34, users of ages 65+ are most likely to click and engage in the content. Females also showed more interest in the promotions, while males had fewer interactions with advertisements.

5.2 Discussion

RQ1. The general performance of the advertisements can be interpreted as a good outcome in regard to the budget. At a cost per result of about €0.11, the cost for clicks was relatively low for such a big target audience. Although a budget of €10.00 per advertisement was planned, Facebook has the right to use the money as it works best for each advertisement. An average CTR of 6.14% reveals that the generated click-through rate is high in all advertisements. Figure 5-1 portrays average click-through rates for all industries on Facebook. The results show that most money was used for Ad 3 and 4, which was the healthy recipe with the appealing image. Those were also the advertisements with most impressions. It seems like Facebook decided that they were the advertisements that most people would interact with based on image and text, and therefore the algorithm chose to show them more often. The least money was spent on Advertisement 1 and 2, which showed the unappealing image of the healthy recipe. The results therefore also were the highest for Ad 3 and Ad 4, and the lowest for Ad 1 and Ad 2. As mentioned in the results chapter, impressions and results of the promotions correlate very strongly. A logical conclusion to draw here is that the more often advertisements are seen, the higher the possibility of a click seems to be. After the Facebook Ad Manager summary of the campaign, a few statistical analyses show whether there was a difference between CTR in all of the advertisements. CTR showed no difference, but there was a difference in impressions. This could be because Facebook aims to have an alike CTR throughout a campaign if the same budget is spent, and in order to reach that CTR, some advertisements have to be shown more often. The ones shown more often afterwards generated a similar click-through-rate to the ones that did not have to be shown as often in order to get clicks.

RQ2. The second research question looked into whether or not recipe healthiness influences user interactions. The results chapter presents statistical tests on comparing the means of the healthy and unhealthy recipe, in terms of CTR and impressions. Neither of those variables display a difference between the recipes. This contradicts the finding of Trattner and Elsweiler (2017) who found that users tend to cook unhealthy recipes more than healthy ones. In this case, there was no significant difference detected. However, this might be a positive indication as it can be interpreted in a way that no matter how healthy the recipe, it will still get the same amount of clicks an advertiser pays for. For the same amount of money, either recipe can be promoted just as well. For health advocates who want to promote healthy recipes, this means it is easy to do so. With the same amount of money as any other institution, it is possible to promote healthy eating and get individuals to look into cooking a healthy dish.

RQ3. The tests on the third research question, whether image plays a role in recipe selection, show that in regard to CTR, results showed no significant difference between the appealing and the unappealing images. Impressions did show a significant difference, firstly between all images

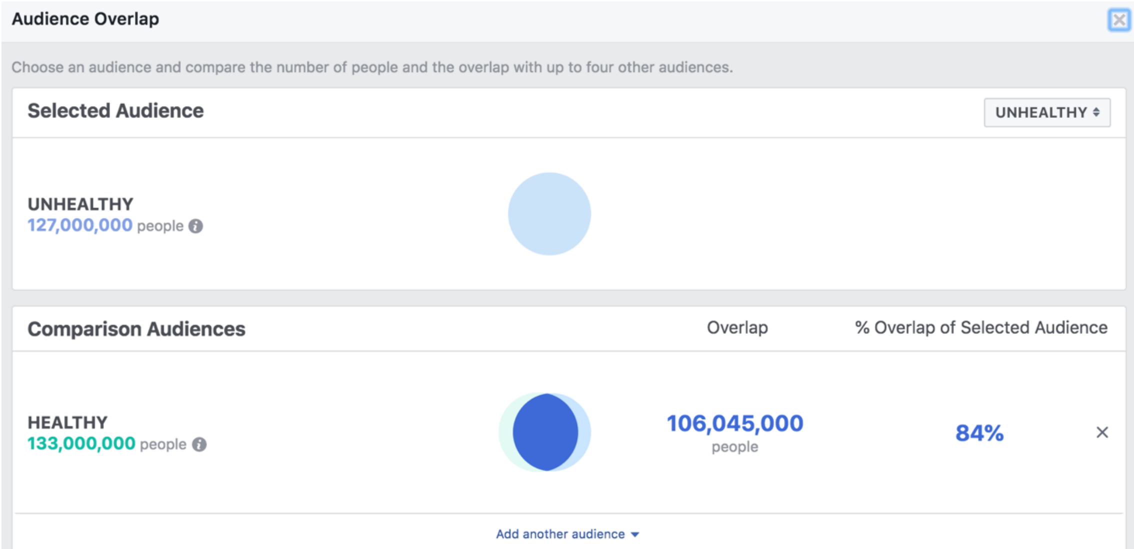


FIGURE 5-2 AUDIENCE OVERLAP

Source: Facebook, 2018a

and then between images in each individual recipe promoted. This means that the amount of times the advertisements were on screen differed between each image. For overall image comparison, the appealing image had a higher mean rank, implying that it was on screen more times than the unappealing one. Facebook sees the reactions to advertisements in the first few minutes and hours of running a promotion and adjusts the time of how often the advertisement is seen on screen. If this method of the algorithm is considered, then it is important to notice that more appealing images will show up more often in a newsfeed and thus get more exposure to users. Therefore, choosing an appealing image for an advertisement, especially for food, is important when advertising healthy recipes. It could potentially help get maximum exposure on a social media platform that is very visual-based. Yet, there is no difference in CTR in this experiment, which suggests that even if the image gets shown more often, interactions with that recipe are not guaranteed. Users might see the advertisement but not choose to click on or even cook the particular recipe, for unknown reasons. It could be that they do not like pancakes in general, or it might not be the right time to cook them. Compared to the findings of Elsweiler et al. (2017), who stated that image is one of the most important factors in recipe selection, image did not play a big role for people clicking on the advertisement. The reason why no difference could be seen in terms of CTR might be that different people find different images attractive. As the recipe choice was pancakes, the expectation for appealing pancakes can differ. The survey that asked respondents to rate images according to attractiveness was only based on 30 people which mostly are residents of Vienna, Austria. If the survey was conducted in the USA, where the advertisements were promoted, with a sample population that is higher in size, a different image could have been rated most appealing and least appealing.

RQ4. As the results highlight, there was no difference whatsoever in the interactions with the recipes according to interest group, which was the question that the fourth research question covers. It would be reasonable to say that people interested in a healthy lifestyle would click on healthy recipes more, and vice versa. However, this was not the case in this experiment. Figure 5-2 shows the audience overlap of the two targeted audience groups. As both interest groups were based in the same region and had the same age group, there was an audience overlap of 84% present. This could have affected the outcome based on interests, as some people may have been targeted by both advertisements the same. Facebook offers as a solution to either “consolidate your overlapping ad sets” or “refine your targeting”, which in the case of this thesis does not make sense as the experimental groups could not be controlled otherwise (Facebook, 2018a).

RQ5. Regarding the research question that focused on how obesity and diabetes prevalence rates influence recipe interactions, many interesting observations are made. The recipe of choice was pancakes, which is why obesity and diabetes was analysed. In the first question, if obesity prevalence affects CTR, there are some negative correlations. Those imply that as obesity rates increase, click-through rate decreases. This is a very surprising outcome, as all the advertisements with a negative correlation contained the unhealthy recipe. Advertisement 5 was targeted to the healthy interest group, as well as advertisement 7. Advertisement 8 targeted the unhealthy interest group. Advertisement 5, the appealing image targeted to the healthy group, showed the highest negative correlation of $r = -0.41$. One can assume that targeting the healthy interest group with an attractive image gets the best response of healthy people from states with low obesity rates. Healthy people seem to respond to the unhealthy recipe when targeted with it, which shows that advertising can change people’s behaviour. Although CTR showed some correlation with them, obesity prevalence did not show any correlation with impressions. Another question looks at diabetes prevalence and CTR. Only one advertisement shows a significant correlation, which was advertisement 4. The moderate correlation of $r = 0.36$ entails that diabetes rates predict CTR by 36%. This suggests that if diabetes prevalence decreases, CTR decreases as well in this promotion and vice versa. Advertisement 4, the healthy pancake recipe with the appealing image, was targeted to the unhealthy group. The higher the diabetes rates in a state, the more people in this region clicked on the recipe. This makes sense when looking at targeting, as it targeted the unhealthy interests. A compelling conclusion to draw therefore is, that if an institution or similar actors tries to target the unhealthy population with a healthy recipe, they can reach them well with the right image and targeting. The last question, whether a relationship between diabetes prevalence and impressions exists, shows the most fascinating result. All advertisements linking to the unhealthy recipe show a highly similar positive correlation of $r = 0.30$ between the variables. As diabetes increases, the amount of times the unhealthy advertisement is on screen increases. Those promotions get shown to the unhealthy population of the USA by Facebook, which is a huge disappointment. Exposing sick people to a recipe that will potentially risk their health even more is a shocking thought. As Facebook’s algorithm might base where they show advertisements on the people’s other online activity, it is possible that this is why the unhealthy recipes were shown in those areas with high diabetes prevalence the most. As seen in Michaelidou and Moraes (2016) anyone can make use of Facebook advertising and therefore it is particularly important to also promote health, rather than only unhealthy habits (Michaelidou & Moraes, 2016). This fact only leads to the conclusion that there is an even higher need of intervention in those states. Social media advertisements seem to show people what they are already interested in, rather than what should be beneficial for them and their overall health. Schäfer et al. (2017) have come to a similar conclusion and already said that recommender systems sometimes recommend items people like but may not be good for them (Schäfer et al., 2017). What they should do is to recommend items that are good for them, such as diabetes friendly food to people who prefer sweets. In their case, they choose to implement a "health recommender system" (Schäfer et al., 2017). In case of this thesis and Facebook, maybe Facebook advertisements should include such an advertisement algorithm as well.

RQ6. By answering the last research question, which addresses user characteristics, the thesis aims to see which kind of users respond to certain advertisements and how those can best be targeted. The characteristics of the target group of the thesis are region, age and gender. In regard to region, it is surprising that Hawaii was the region with the highest click-through-rate, whereas Wyoming had the least. Hawaii has a population of 1,417,710, which ranks it number 40 amongst all states of the USA. Wyoming has a population estimate of 583,334 and therefore has the least inhabitant of all states. However, both states score low amongst population statistics, so it is surprising that both have a different CTR. California had the highest amount of impressions and Wyoming again the least. This is not surprising, as California also has the biggest population out of all the states in the USA, and Wyoming the least. When looking at the age groups that interacted with the advertisements, it is surprising to see that although, according to Statista, Facebook's community is mostly between 25 and 34, the users clicking and looking at the advertisements in this experiment were mostly 65 years and older (Statista, 2018b). The age with most Facebook users, 25-34, had the second lowest CTR with M = 4.17, which implies that although most users come from this group, they do not engage with advertisements too much. The outcomes show that it is easy for advertisers to target older people, as they tend to interact with cooking advertisements more. This could be because 65+ aged people are usually already retired and may have time to cook more than younger generations. As seen in ABC news by Langer (2005) in the methodology chapter, it was identified that people eating breakfast are more likely to be old than young. Therefore, recipe advertisements could improve the health of older age groups particularly by advertising healthy recipes to them. Should one want to target young audiences, it is useful to target them in separate advertisements to get clicks from this age group alone. Regarding gender, it is important to note that mostly females saw, as well as clicked through to the advertisements. With both variables, Impressions and CTR, the female mean rank was higher than the male and the unknown one. Why this is the case might be because it is possible that more females on Facebook have cooking interests and take their recipe inspiration from social media platforms. Also, findings from

Rokicki et al. (2016) show that women are more likely to cook sweet recipes, while men go for salty ones. Taking into consideration that mostly the older population responded to advertisements, one could conclude that in older generations, the woman is more likely to cook the meals at home.

5.3 Limitations

As with any work of quantitative research, certain limitations must be considered. The Facebook Advertising API is a tool that is new to be used in research. It has many restrictive properties, such as needing a certain budget in order to get meaningful results. This is a limitation related to the sample size. As this thesis had a budget of €500, it was not possible to target a big audience of the United States, which consists of a population of about 325 million people. With a higher budget, more Facebook users can be reached, and a more indicative conclusion could be drawn. It is also possible to go into further detail on targeting and address a bigger number of different audiences. This would have meant a split of the budget into even more than eight advertisements, which cannot be done as a minimum budget for an advertisement is required by the Facebook platform. The budget could also be a reason for some outcomes to be a coincidence, even if they were statistically significant. Considering this, a larger sample size is always better because it is more representative and leads to greater statistical power.

Another limitation could be that, since the advertiser is paying Facebook by a cost-per-click method, the algorithm might intend to make the CTR similar, or close to equal, for all advertisements in a campaign if an overall budget is set. This implies that for each advertisement, regardless of the properties included, the platform will try to optimise the outcome and get the largest number of clicks per impressions possible. A solution to this could be to implement eight different campaigns and individually set up each advertisement within the campaign.

Sampling introduces another limitation. Considering that this experiment relies on the Facebook algorithm for advertisements, the conductor of the research has no control over which people get exposed to the promotions. Therefore, the experiment did not allow for guaranteed equal distribution of advertisements in all states.

Another limitation may be related to the generalisation of the findings, as this study was conducted in the USA. This means that it is not guaranteed that the findings can be applied to other populations.

One major limitation is that this experiment does not include qualitative research. People that click on recipes, no matter how healthy, are not guaranteed to cook this recipe in actuality. They may have clicked on the advertisement but might have no intention of using the recipe later on. Qualitative research, which would follow up on people that clicked the advertisements, could be a way to determine whether or not a recipe is cooked. At last, the advertisements in this thesis were not intended for a vegan community. With the growing trend of veganism, a big audience might have been excluded.

6 CONCLUSION

This research analyses the factors that influence which recipe promotion users on Facebook click on. The main goal is to learn those aspects, in order to later on successfully promote a healthy diet to individuals in need of it. The experiment revealed that healthiness of a recipe does not play a role in how many people click on it. This shows that with the right budget, such kinds of recipes and potentially others can be advertised easily, which facilitates healthy eating promotion. Image does also not influence the clicks, but the more attractive the image, the more often it is shown on screen. User interests have no effect on CTR and impressions. Some correlations are seen between health statistics of the USA and CTR or impressions. Generally, older individuals interacted with the recipes more often than young ones. Females were also more likely to click on the recipe promotions.

6.1 Contribution to knowledge

Considering the outcome of this thesis, many different stakeholders can benefit from the information obtained. First and foremost, a lot of research has already been done in the field of promoting health related goals in the Internet. This experiment suggests that doing so is possible and shows who best responds to it. Facebook advertising is a new tool that can be used for implementing research, not only in the field of health and nutrition but in various other areas of interest as well. The research above focuses on learning how factors influence recipe choice. The next step for the future is how to exploit this knowledge.

Apart from this, governmental bodies and multiple other institutions share the common goal of disease and obesity reduction, as it becomes a monetary issue when the amount of people to be treated is constantly increasing. Those institutions also benefit from the knowledge of how to promote health best.

6.2 Implications for relevant stakeholders

An implication for health advocates is that the use of Facebook advertisements can successfully attract the right target audience. It makes it possible to target very specific types of people with even more interests than mentioned in this thesis. As it is now known that older people tend to respond more to cooking advertisements, it is important that especially they are exposed to healthy recipes, rather than unhealthy ones. When wanting to reach a younger audience, the target group should consist of their age only, as that guarantees a reaction from them, rather than older individuals. Another important suggestion to health institutions is that an intervention on those social media platforms is necessary, since the people interested in unhealthy activities will keep being confronted with them unless a marketer actively targets them with health promotional subjects.

One additional implication is that other papers may use the Facebook advertising API in their studies and are able to use a similar approach as the one in this thesis to do so. Not only recipe advertisements, but advertisements of innumerable other health areas can show an insight into which people are the right ones to target, and how to do so. Knowing how to use this advertising platform to extract information and observe how users interact with information given to them uncovers a new tool that not many have made use of. Using it can be beneficial to multiple economic and charitable sectors.

6.3 Future research

As this experiment was a quantitative one, a new type of research worth considering can be done through qualitative designs. One possibility is to only target recipes to a small chosen number of people, and afterwards observe their real-life cooking behaviour by potentially conducting interviews or making observations. Interviews can identify why they decided to cook one recipe, and not the other. Another type of research can be done through the Web as well by tracking users.

If an online-cooking platform would allow a pixel on their recipe website, it is also possible to track a user on the website the advertisement leads to. That way, it is possible to observe how long people stay on the recipe homepage, if they rate the recipe, whether they reply to comments and many other factors. Web heat-mapping on those cooking websites then be interesting, as conductors of a study are able to see which areas of the site are most often scanned and hovered over by visitors.

Since this thesis exclusively focuses on the social media platform Facebook, it would be compelling to look at their partner network Instagram too. This image-based tool is growing in society and has big potential for analysing health related data. Inspecting how people respond to recipe advertisements there may also be revealing, because a study here would need to put a lot of emphasis on the images shown in the advertisements. Similar to how De Choudhury et al. (2016) and Mejova et al. (2015) used Instagram, further studies can also use the platform, however with advertisements. Outcomes could show promising implications for stakeholders, just as this thesis has.

7 BIBLIOGRAPHY

Abbar, S., Mejova, Y., & Weber, I. (2015). 'You tweet what you eat: Studying food consumption through twitter. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.' *ACM* pp. 3197-3206.

AdEspresso, Inc. (2018). '*Facebook Ads Manager Guide (How-To Set Up Your Facebook Ad Campaigns).*' Available: https://adespresso.com/guides/facebook-ads-beginner/facebook-manager-campaign-setup - retrieved on: 15 April, 2018.

Aggarwal C.C. (2016). An Introduction to Recommender Systems. In Aggarwal C.C., '*Recommender Systems*.' (pp. 1-28). Switzerland: Springer International Publishing.

Allrecipes.com (2017). '*Allrecipes Press Portal*.' Available: http://press.allrecipes.com - retrieved on: 12 April, 2018.

Allrecipes.com (2018a). '*Allrecipes.*' Available: http://allrecipes.com - retrieved on: 12 April, 2018.

Allrecipes.com (2018b). '*Extra-Yummy Fluffy Pancakes*.' Available: http://allrecipes.com/recipe/216755/extra-yummy-fluffy-pancakes - retrieved on: 2 April, 2018.

Allrecipes.com (2018c). '*Whole Wheat, Oatmeal and Banana Pancakes*.' Available: http://allrecipes.com/recipe/166601/whole-wheat-oatmeal-and-banana-pancakes - retrieved on: 2 April, 2018.

Apovian, C. M. (2004). 'Sugar-sweetened soft drinks, obesity, and type 2 diabetes.' *Jama, 292(8)* pp. 978-979.

Araújo, M., Mejova, Y., Weber, I., & Benevenuto, F. (2017). 'Using Facebook Ads Audiences for Global Lifestyle Disease Surveillance: Promises and Limitations. In Proceedings of the 2017 ACM on Web Science Conference.' *ACM* pp. 253-257.

Assimilation [Def. 3]. (2018). '*Collins Dictionary*.' Available: https://www.collinsdictionary.com/dictionary/english/assimilation - retrieved on: 12 May, 2018.

Cahill, M. (2018). '*Top 10 breakfast foods*.' Available: http://www.dailypress.com/entertainment/food/dp-toptenlist.o29-story.html - retrieved on: 25 April, 2018.

Carlson, N. (2010). '*At last — the full story of how Facebook was founded.*' Available: https://www.businessinsider.de/how-facebook-was-founded-2010-3?r=US&IR=T - retrieved on: 2 May, 2018.

Carmichael, D., & Cleave, D. (2012). 'How effective is social media advertising? A study of Facebook Social Advertisements. In Internet Technology And Secured Transactions, 2012 International Conference.*' IEEE, (1)* pp. 226-229.

Centers for Disease Control. (2003). 'Physical activity and good nutrition: essential elements to prevent chronic diseases and obesity 2003.' *Nutrition in clinical care: an official publication of Tufts University, 6(3)* pp. 135.

Centers for Disease Control and Prevention (2014). '*Mission, Role and Pledge.*' Available: https://www.cdc.gov/about/organization/mission.htm - retrieved on: 1 March, 2018.

Centers for Disease Control and Prevention (2016). '*Diabetes Home.*' Available: https://www.cdc.gov/diabetes/data/countydata/statecountyindicators.html - retrieved on: 5 March, 2018.

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

Correlation [Def. 1]. (2018). '*Merriam Webster Online*.' Available: https://www.merriam-webster.com/dictionary/correlation - retrieved on: 3 May, 2018.

Cunningham, S.J., & Bainbridge, D. (2015). 'An analysis of cooking queries; Implications for supporting leisure cooking.' *iConference 2013 Proceedings* pp. 112-123.

Dehghani, M., & Tumer, M. (2015). 'A research on effectiveness of Facebook advertising on enhancing purchase intention of consumers.' *Computers in Human Behavior, 49* pp. 597-600.

De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). 'Predicting depression via social media.' *ICWSM, 13* pp. 1-10.

Dubois, A., Zagheni, E., Garimella, K., & Weber, I. (2018). 'Studying Migrant Assimilation Through Facebook Interests.*' arXiv preprint arXiv* pp. 1801.09430.

De Choudhury, M., Sharma, S., & Kiciman, E. (2016). 'Characterizing dietary choices, nutrition, and language in food deserts via social media. In Proceedings of the 19th acm conference on computer-supported cooperative work & social computing.' *ACM* pp. 1157-1170.

Elsweiler, D., Trattner, C., & Harvey, M. (2017). 'Exploiting food choice biases for healthier recipe recommendation. In Proceedings of the 40th international acm sigir conference on research and development in information retrieval.' *ACM* pp. 575-584.

EUFIC (2006). '*The determinants of food choice*.' Available: https://www.eufic.org/en/en/healthy-living/article/the-determinants-of-food-choice - retrieved on: 6 March, 2018.

Facebook (2018a). '*About overlapping audiences.*' Available: https://www.facebook.com/business/help/1679591828938781 - retrieved on: 13 May, 2018.

Facebook (2018b). '*Ad formats.*' Available: https://www.facebook.com/business/learn/facebook-create-ad-basics - retrieved on: 13 May, 2018.

Facebook (2018c). '*Ad Manager*.' Available: https://www.facebook.com/adsmanager - retrieved on: 1 May, 2018.

Facebook (2018d). '*Advertising Insights.*' Available: https://www.facebook.com/iq/advertising-insights?ref=fbb_ads - retrieved on: 13 May, 2018.

Facebook (2018e). '*Choose your audience*.' Available: https://www.facebook.com/business/products/ads/ad-targeting - retrieved on: 13 May, 2018.

Facebook (2018f). '*Facebook Ads guide*.' Available: https://www.facebook.com/business/ads-guide - retrieved on: 13 May, 2018.

Facebook (2018g). '*Measure your ads*.' Available: https://www.facebook.com/business/learn/facebook-ads-measuring-results - retrieved on: 13 May, 2018.

Facebook Page [Def. 2]. (2018). '*Techopedia Inc*.' Available: https://www.techopedia.com/definition/15441/facebook-page - retrieved on: 1 May, 2018.

Fatehkia, M., Kashyap, R., & Weber, I. (2018). 'Using Facebook ad data to track the global digital gender gap.' *World Development, 107* pp. 189-209.

Fried, D., Surdeanu, M., Kobourov, S., Hingle, M., & Bell, D. (2014). 'Analyzing the language of food on social media. In Big Data (Big Data), 2014 IEEE International Conference.' *IEEE* pp. 778-783.

Fitzpatrick, H. (2015). '*The 20 most successful fast food chains in America*.' Available: http://www.businessinsider.com/the-20-most-successful-fast-food-chains-in-america-2015-8?IR=T - retrieved on: 11 April, 2018.

Garimella, V. R. K., Alfayad, A., & Weber, I. (2016). 'Social media image analysis for public health. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.' *ACM* pp. 5543-5547.

Gatti, L., Guerini, M., Stock, O., & Strapparava, C. (2014). 'Sentiment variations in text for persuasion technology. In International Conference on Persuasive Technology.' *Springer, Cham* pp. 106-117.

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

Grilo, C. M. (1994). 'Physical activity and obesity.' *Biomedicine & pharmacotherapy, 48(3-4)* pp. 127-136.

Harvey, M., Ludwig, B., & Elsweiler, D. (2012). 'Learning user tastes: a first step to generating healthy meal plans. In First international workshop on recommendation technologies for lifestyle change.' *Lifestyle 2012* pp. 18.

Havlak, H. and Abelson, B. (2016) '*The definitive list of what everyone likes on Facebook.*' Available: https://www.theverge.com/2016/2/1/10872792/facebook-interests-ranked-preferred-audience-size - retrieved on: 2 May, 2018.

Howard, S., Adams, J., & White, M. (2012). 'Nutritional content of supermarket ready meals and recipes by television chefs in the United Kingdom: cross sectional study.' *BMJ, 345* pp. e7607.

Hu, F. B., Li, T. Y., Colditz, G. A., Willett, W. C., & Manson, J. E. (2003). 'Television watching and other sedentary behaviors in relation to risk of obesity and type 2 diabetes mellitus in women.' *Jama, 289(14)* pp. 1785-1791.

Irvine, M. (2018). '*Facebook Ad Benchmarks for YOUR Industry [New Data].'* Available: https://www.wordstream.com/blog/ws/2017/02/28/facebook-advertising-benchmarks - retrieved on: 3 June, 2018.

Jeffery, R. W., Baxter, J., McGuire, M., & Linde, J. (2006). 'Are fast food restaurants an environmental risk factor for obesity?' *International Journal of Behavioral Nutrition and Physical Activity, 3(1)* pp. 2.

Kamal, N., Fels, S., & Ho, K. (2010). 'Online social networks for personal informatics to promote positive health behavior. In Proceedings of second ACM SIGMM workshop on Social media.' *ACM* pp. 47-52.

Kusmierczyk, T., & Nørvåg, K. (2016). 'Online Food Recipe Title Semantics: Combining Nutrient Facts and Topics. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management.' *ACM* pp. 2013-2016.

Langer, G. (2005). '*POLL: What Americans Eat for Breakfast*.' Available: https://abcnews.go.com/GMA/PollVault/story?id=762685 - retrieved on 25 April, 2018.

Mejova, Y., Weber, I., & Fernandez-Luque, L. (2018). 'Online Health Monitoring using Facebook Advertisement Audience Estimates in the United States: Evaluation Study.' *JMIR public health and surveillance, 4(1).*

Mejova, Y., Haddadi, H., Noulas, A., & Weber, I. (2015). '# foodporn: Obesity patterns in culinary interactions. In Proceedings of the 5th International Conference on Digital Health 2015.' *ACM* pp. 51-58.

Michaelidou, N., & Moraes, C. (2016). 'How Companies Use Facebook to Promote Alcohol Brands to Young Adults. In Looking Forward, Looking Back: Drawing on the Past to Shape the Future of Marketing.*' Springer, Cham* pp. 487-490.

Müller, M., Harvey, M., Elsweiler, D., & Mika, S. (2012). 'Ingredient matching to determine the nutritional properties of internet-sourced recipes. In Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference.' *IEEE* pp. 73-80.

Murphy, C. (2015). '*Facebook Ad Targeting for Hotels*.' Available: https://www.revinate.com/blog/2015/09/facebook-ad-targeting-hotels - retrieved on: 12 June, 2018.

Ornish, D., Scherwitz, L. W., Billings, J. H., Gould, K. L., Merritt, T. A., Sparler, S., ... & Brand, R. J. (1998). 'Intensive lifestyle changes for reversal of coronary heart disease.' *Jama, 280(23)* pp. 2001-2007.

Park, H., Rodgers, S., & Stemmle, J. (2011). 'Health organizations' use of Facebook for health advertising and promotion.' *Journal of interactive advertising, 12(1)* pp. 62-77.

Pedersen, B. K., & Febbraio, M. A. (2012). 'Muscles, exercise and obesity: skeletal muscle as a secretory organ.' *Nature Reviews Endocrinology, 8(8)* pp. 457.

Peppers, M. (2014). '*America's 50 favorite foods revealed - steak and chicken top the list.*' Available: http://www.dailymail.co.uk/femail/article-2788395/pass-protein-america-s-50-favorite-foods-revealed-steak-chicken-topping-list.html - retrieved on: 25 April, 2018.

Pixabay (2018). '*Food Vegetarian Kitchen Cooking*.' Available: https://www.pexels.com/photo/food-vegetarian-kitchen-cooking-33242 - retrieved on: 11 April, 2018.

Pixabay (2018). '*Recipe Label Icon Symbol Spoon.*' Available: https://pixabay.com/en/recipe-label-icon-symbol-spoon-575434/ - retrieved on: 11 April, 2018.

Ponocny, I. and Weismayer, C. (2016a). correlation and crosstabs. Presentation presented in the course *0102 Advanced Data Analysis*, Modul University Vienna, Vienna, November.

Ponocny, I. and Weismayer, C. (2016b). Testing for Group Differences. Presentation presented in the course *0102 Advanced Data Analysis*, Modul University Vienna, Vienna, October.

Prier, K. W., Smith, M. S., Giraud-Carrier, C., & Hanson, C. L. (2011). 'Identifying health-related topics on twitter. In International conference on social computing, behavioral-cultural modeling, and prediction.' *Springer* pp. 18-25.

Rimer, B. K., & Kreuter, M. W. (2006). 'Advancing tailored health communication: A persuasion and message effects perspective.*' Journal of communication, 56(s1).*

Rokicki, M., Herder, E., Kuśmierczyk, T., & Trattner, C. (2016). 'Plate and prejudice: Gender differences in online cooking. In Proceedings of the 2016 conference on user modeling adaptation and personalization.' *ACM* pp. 207-215.

Sacks, G., Rayner, M., and Swinburn, B. (2009). 'Impact of front-of-pack 'traffic-light'nutrition labelling on consumer food purchases in the uk.' *Health promotion international, 24(4)* pp. 344–352.

Said, A., & Bellogín, A. (2014). 'You are What You Eat! Tracking Health Through Recipe Interactions.' *Rsweb@ recsys.*

Saha, K., Weber, I., Birnbaum, M. L., & De Choudhury, M. (2017). 'Characterizing Awareness of Schizophrenia Among Facebook Users by Leveraging Facebook Advertisement Estimates.*' Journal of medical Internet research, 19(5).*

Schäfer, H., Hors-Fraile, S., Karumur, R. P., Calero Valdez, A., Said, A., Torkamaan, H., ... & Trattner, C. (2017). 'Towards health (aware) recommender systems. In Proceedings of the 2017 international conference on digital health.' *ACM* pp. 157-161.

Social Nework [Def. 1]. (2018). '*Merriam Webster Online*.' Available: https://www.merriam-webster.com/dictionary/social%20network - retrieved on: 30 May, 2018.

Statista (2018a). '*Countries with the most Facebook users as of January 2018*.' Available: https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/ - retrieved on: 4 May, 2018.

Statista (2018b). '*Distribution of Facebook users in the United States as of January 2018, by age group and gender*.' Available: https://www.statista.com/statistics/187041/us-user-age-distribution-on-facebook - retrieved on: 20 May, 2018.

Statistics Solutions (2018a). '*ANOVA (Analysis of Variance)*.' Available: https://www.statisticssolutions.com/manova-analysis-anova - retrieved on: 12 May, 2018.

Statistics Solutions (2018b). '*Conduct and Interpret an Independent Sample T-Test.*' Available: https://www.statisticssolutions.com/independent-sample-t-test - retrieved on: 12 May, 2018.

Statistics Solutions (2018c). '*Kruskal-Wallis Test*.' Available: http://www.statisticssolutions.com/kruskal-wallis-test - retrieved on: 12 May, 2018.

Statistics Solutions (2018d). '*Mann-Whitney U Test*.' Available: https://www.statisticssolutions.com/mann-whitney-u-test - retrieved on: 12 May, 2018.

Thisted, R. A. (1998) 'What is a P-value?'. *The University of Chicago* pp. 1-6.

The Top Tens (2005). '*Top Ten Favourite Foods*.' Available: https://www.thetoptens.com/favourite-foods - retrieved on: 25 April, 2018.

Trattner, C., Elsweiler, D., & Howard, S. (2017a). 'Estimating the healthiness of internet recipes: a cross-sectional study.' *Frontiers in public health, 5* pp. 16.

Trattner, C., Parra, D., & Elsweiler, D. (2017b) 'Monitoring obesity prevalence in the United States through bookmarking activities in online food portals.' *PloS one, 12(6)* pp. e0179144.

Trattner, C., Rokicki, M., & Herder, E. (2017c). 'On the Relations Between Cooking Interests, Hobbies and Nutritional Values of Online Recipes: Implications for Health-Aware Recipe Recommender Systems. In Adjunct publication of the 25th conference on user modeling, adaptation and personalization.' *ACM* pp. 59-64.

Trattner, C., & Elsweiler, D. (2017) 'Investigating the healthiness of internet-sourced recipes: implications for meal planning and recommender systems. In Proceedings of the 26th international conference on world wide web.*' International World Wide Web Conferences Steering Committee* pp. 489-498.

U.S. Census Bureau (2018). '*QuickFacts UNITED STATES*.' Available: https://www.census.gov/quickfacts/fact/map/US/PST045217 - retrieved on: 4 April, 2018.

Vejacka, M. (2012). 'Facebook advertising and its efficiency on the Slovak market.*' E+ M Ekonomie a Management, (1)* pp. 116.

Vioque, J., Torres, A., & Quiles, J. (2000). 'Time spent watching television, sleep duration and obesity in adults living in Valencia, Spain.*' International journal of obesity, 24(12)* pp. 1683.

West, R., White, R. W., & Horvitz, E. (2013). 'From cookies to cooks: Insights on dietary patterns via analysis of web usage logs. In Proceedings of the 22nd international conference on World Wide Web.' *ACM* pp. 1399-1410.

World Health Organization (2003). *'Diet, nutrition, and the prevention of chronic diseases: report of a joint WHO/FAO expert consultation (Vol. 916)*.' Geneva: World Health Organization.

Yom-Tov, E., Brunstein-Klomek, A., Mandel, O., Hadas, A., & Fennig, S. (2018). 'Inducing Behavioral Change in Seekers of Pro-Anorexia Content Using Internet Advertisements: Randomized Controlled Trial.*' JMIR mental health, 5(1).*

Yom-Tov, E., Muennig, P., & El-Sayed, A. M. (2016). 'Web-based antismoking advertising to promote smoking cessation: a randomized controlled trial.' *Journal of medical Internet research, 18(11).*

Yom-Tov, E., Shembekar, J., Barclay, S., & Muennig, P. (n.d.). 'The Effectiveness of Public Health Advertisements to Promote Health: A Randomized-Controlled Trial on 794,000 Participants.' *(n.p.).*

Zagheni, E., Weber, I., & Gummadi, K. (2017). 'Leveraging Facebook's Advertising Platform to Monitor Stocks of Migrants.' *Population and Development Review, 43(4)* pp. 721-734.

1999-2018 SurveyMonkey (2018). '*Summary*.' Available: https://de.surveymonkey.com/user/sign-in/ - retrieved on: 18 April, 2018.

8 LIST OF ABBREVIATIONS

WHO	**World Health Organization**
RQ	**Research Question**
FSA	**Food Standards Agency**
CDC	**Centers for Disease Control and Prevention**
USA	**United States of America**
THC	**Tailored Health Communications**
HRS	**Health Recommender Systems**
Ads	**Advertisements**
CPC	**Cost per Click**
CTR	**Click-through-rate**

APPENDICES

Appendix 1: Survey for Recipe Image Selection

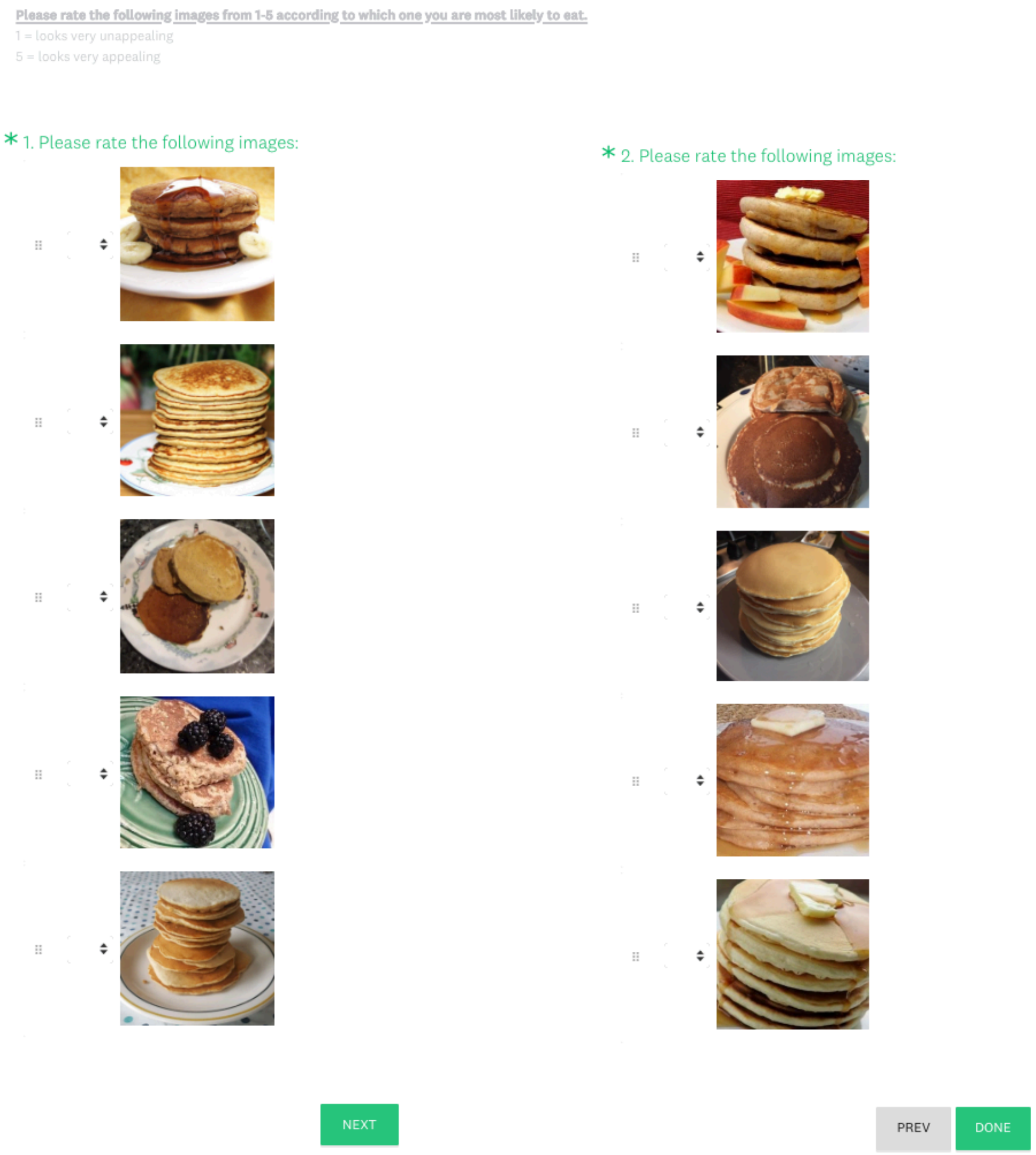


Figure A-1 Pancakes Survey distributed via Facebook as it appeared to people

Source: 1999-2018 SurveyMonkey

Appendix 2: Demographic Data for high and low obesity prevalent states

State	Age-adjusted Diabetes Percent 2014	Age-adjusted Obesity Percent 2014	Population Estimate 2014	Median household income (in 2016 dollars), 2012-2016	Foreign born persons, percent, 2012-2016	Female persons, percent, July 1, 2016, (V2016)	Bachelor's degree or higher, percent of persons age 25 years+, 2012-2016
States with the lowest obesity prevalence							
Colorado	6.9	21.3	5342311	62520.00	9.8	49.7	38.7
Hawaii	8.9	22.5	1417710	71977	17.9	49.8	31.4
Massachusetts	8.8	23.3	6757925	70954	15.7	51.5	41.2
Vermont	6.9	24.7	625665	56104	4.4	50.6	36.2
California	9.9	24.7	38701278	63783	27.0	50.3	32.0
States with the highest obesity prevalence							
Alabama	11.8	33.7	4840037	44758	3.4	51.6	24.0
Louisiana	10.4	34.9	4648797	45652	4.0	51.1	23.0
Mississippi	11.9	35.7	2988578	40528	2.3	51.5	21.0
West Virginia	12	36	1847624	42644	1.6	50.5	19.6
Arkansas	11.5	36.2	2964800	42336	4.7	50.9	21.5

TABLE A-1 DEMOGRAPHIC DATA FOR THE 5 STATES WITH THE HIGHEST AND LOWEST OBESITY RATES

Appendix 3: Advertisements as they appeared to targeted users

Ad 1 & 2

Ad 3 & 4

Ad 5 & 6

Ad 7 & 8

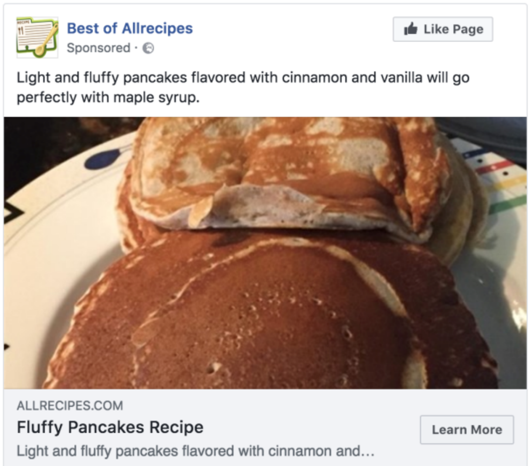


TABLE A-2 FOUR ADVERTISEMENTS AS THEY APPEARED TO THE TWO TARGETED USER GROUPS

Appendix 4: Statistical Analyses of RQ1

Ranks

	Ad	N	Mean Rank
CTR	1.00	51	203.01
	2.00	48	196.86
	3.00	47	182.70
	4.00	51	182.18
	5.00	50	201.81
	6.00	51	207.66
	7.00	50	225.32
	8.00	49	191.33
	Total	397	

Test Statistics

	CTR
Chi-Square	5.29
df	7
Asymp. Sig.	.624

TABLE A-3 KRUSKAL-WALLIS ON DIFFERENCES BETWEEN ALL ADVERTISEMENTS IN RESPECT TO CTR

Ranks

	Ad	N	Mean Rank
Impressions	1.00	51	139.98
	2.00	51	159.25
	3.00	51	235.52
	4.00	51	247.74
	5.00	51	197.61
	6.00	51	188.86
	7.00	51	246.73
	8.00	51	220.32
	Total	408	

Test Statistics

	Impressions
Chi-Square	41.69
df	7
Asymp. Sig.	.000

TABLE A-4 KRUSKAL-WALLIS ON DIFFERENCES BETWEEN ADVERTISEMENTS IN RESPECT TO IMPRESSIONS

Appendix 5: Statistical Analyses of RQ2

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	Healthy Recipe	*Unhealthy Recipe*	*Total*	*Healthy Recipe*	*Unhealthy Recipe*	*Healthy Recipe*	*Unhealthy Recipe*
CTR	197.00	200.00	397.00	191.27	206.61	37681.00	41322.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
CTR	18178.00	37681.00	-1.33	.183

TABLE A-5 MANN-WHITNEY U TEST ON DIFFERENCES BETWEEN THE HEALTHY AND UNHEALTHY RECIPE ADVERTISEMENTS IN TERMS OF CTR

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	Healthy Recipe	*Unhealthy Recipe*	*Total*	*Healthy Recipe*	*Unhealthy Recipe*	*Healthy Recipe*	*Unhealthy Recipe*
Impressions	204.00	204.00	408.00	195.62	213.38	39906.50	43529.50

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	18996.50	39906.50	-1.52	.128

TABLE A-6 MANN-WHITNEY U TEST ON DIFFERENCES BETWEEN THE HEALTHY AND UNHEALTHY RECIPE ADVERTISEMENTS IN TERMS OF IMPRESSIONS

Appendix 6: Statistical Analyses of RQ3

Group Statistics

	Image	N	Mean	Std. Deviation	S.E. Mean
CTR	Unappealing	198	6.04	3.17	.23
	Appealing	199	5.72	3.08	.22

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
CTR	Equal variances assumed	.35	.552	1.02	395.00	.308	.32	.31	-.30	.94
	Equal variances not assumed			1.02	394.57	.308	.32	.31	-.30	.94

TABLE A-7 INDEPENDENT SAMPLES T-TEST ON DIFFERENCES BETWEEN THE APPEALING AND THE UNAPPEALING IMAGE IN TERMS OF CTR

Ranks

	N			Mean Rank		Sum of Ranks	
	Unappealing	Appealing	Total	Unappealing	Appealing	Unappealing	Appealing
Impressions	204.00	204.00	408.00	191.57	217.43	39080.00	44356.00

Test Statistics

	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
Impressions	18170.00	39080.00	-2.22	.027

TABLE A-8 MANN-WHITNEY U TEST ON DIFFERENCES BETWEEN THE APPEALING AND THE UNAPPEALING IMAGE IN TERMS OF IMPRESSIONS

Group Statistics

	Image	N	Mean	Std. Deviation	S.E. Mean
CTR	Unappealing	99	5.99	3.63	.36
	Appealing	98	5.33	2.21	.22

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
CTR	Equal variances assumed	16.09	.000	1.53	195.00	.129	.65	.43	-.19	1.50
	Equal variances not assumed			1.53	162.21	.128	.65	.43	-.19	1.50

TABLE A-9 INDEPENDENT SAMPLES T-TEST ON DIFFERENCES BETWEEN THE APPEALING AND THE UNAPPEALING IMAGE IN TERMS OF CTR ON THE HEALTHY RECIPE

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	Unappealing	*Appealing*	*Total*	*Unappealing*	*Appealing*	*Unappealing*	*Appealing*
Impressions	102.00	102.00	204.00	80.44	124.56	8204.50	12705.50

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	2951.50	8204.50	-5.34	.000

TABLE A-10 MANN-WHITNEY U TEST ON DIFFERENCES BETWEEN THE APPEALING AND THE UNAPPEALING IMAGE IN TERMS OF IMPRESSIONS ON THE HEALTHY RECIPE

Group Statistics

	Image	*N*	*Mean*	*Std. Deviation*	*S.E. Mean*
CTR	Appealing	101	6.10	3.71	.37
	Unappealing	99	6.10	2.65	.27

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	*Sig.*	*t*	*df*	*Sig. (2-tailed)*	*Mean Difference*	*Std. Error Difference*	*Lower*	*Upper*
CTR	Equal variances assumed	6.75	.010	.00	198.00	.996	.00	.46	-.90	.90
	Equal variances not assumed			.00	181.00	.996	.00	.46	-.90	.90

TABLE A-11 INDEPENDENT SAMPLES T-TEST ON DIFFERENCES BETWEEN THE APPEALING AND THE UNAPPEALING IMAGE IN TERMS OF CTR ON THE UNHEALTHY RECIPE

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	Appealing	*Unappealing*	*Total*	*Appealing*	*Unappealing*	*Appealing*	*Unappealing*
Impressions	102.00	102.00	204.00	91.78	113.22	9361.50	11548.50

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	4108.50	9361.50	-2.59	.009

TABLE A-12 MANN-WHITNEY U TEST ON DIFFERENCES BETWEEN THE APPEALING AND THE UNAPPEALING IMAGE IN TERMS OF IMPRESSIONS ON THE UNHEALTHY RECIPE

Appendix 7: Statistical Analyses of RQ4

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	Healthy	*Unhealthy*	*Total*	*Healthy*	*Unhealthy*	*Healthy*	*Unhealthy*
CTR	198.00	199.00	397.00	203.52	194.50	40297.00	38706.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
CTR	18806.00	38706.00	-.78	.434

TABLE A-13 MANN-WHITNEY U TEST ON DIFFERENCES BETWEEN THE HEALTHY AND THE UNHEALTHY INTEREST GROUP IN TERMS OF CTR

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	Healthy	*Unhealthy*	*Total*	*Healthy*	*Unhealthy*	*Healthy*	*Unhealthy*
Impressions	204.00	204.00	408.00	204.96	204.04	41811.50	41624.50

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	20714.50	41624.50	-.08	.937

TABLE A-14 MANN-WHITNEY U TEST ON DIFFERENCES BETWEEN THE HEALTHY AND THE UNHEALTHY INTEREST GROUP IN TERMS OF IMPRESSIONS

Appendix 8: Statistical Analyses of RQ5

Correlations

		DiabetesPrevalence	CTR
DiabetesPrevalence	Pearson Correlation	1.00	-.02
	Sig. (2-tailed)		.872
	N	51	51
CTR	Pearson Correlation	-.02	1.00
	Sig. (2-tailed)	.872	
	N	51	51

Correlations

		DiabetesPrevalence	CTR
DiabetesPrevalence	Pearson Correlation	1.00	-.14
	Sig. (2-tailed)		.345
	N	51	48
CTR	Pearson Correlation	-.14	1.00
	Sig. (2-tailed)	.345	
	N	48	48

Correlations

		DiabetesPrevalence	CTR
DiabetesPrevalence	Pearson Correlation	1.00	.23
	Sig. (2-tailed)		.121
	N	51	47
CTR	Pearson Correlation	.23	1.00
	Sig. (2-tailed)	.121	
	N	47	47

Correlations

		DiabetesPrevalence	CTR
DiabetesPrevalence	Pearson Correlation	1.00	*.36*
	Sig. (2-tailed)		.009
	N	51	51
CTR	Pearson Correlation	*.36*	1.00
	Sig. (2-tailed)	.009	
	N	51	51

Correlations

		DiabetesPrevalence	CTR
DiabetesPrevalence	Pearson Correlation	1.00	-.06
	Sig. (2-tailed)		.670
	N	51	50
CTR	Pearson Correlation	-.06	1.00
	Sig. (2-tailed)	.670	
	N	50	50

Correlations

		DiabetesPrevalence	CTR
DiabetesPrevalence	Pearson Correlation	1.00	-.16
	Sig. (2-tailed)		.276
	N	51	51
CTR	Pearson Correlation	-.16	1.00
	Sig. (2-tailed)	.276	
	N	51	51

Correlations

		DiabetesPrevalence	*CTR*
DiabetesPrevalence	*Pearson Correlation*	1.00	-.27
	Sig. (2-tailed)		.062
	N	51	50
CTR	*Pearson Correlation*	-.27	1.00
	Sig. (2-tailed)	.062	
	N	50	50

Correlations

		DiabetesPrevalence	*CTR*
DiabetesPrevalence	*Pearson Correlation*	1.00	-.02
	Sig. (2-tailed)		.897
	N	51	49
CTR	*Pearson Correlation*	-.02	1.00
	Sig. (2-tailed)	.897	
	N	49	49

TABLE A-15 CORRELATION ANALYSIS ON DIABETES PREVALENCE AND CTR

Correlations

		DiabetesPrevalence	Impressions
DiabetesPrevalence	Pearson Correlation	1.00	.19
	Sig. (2-tailed)		.175
	N	51	51
Impressions	Pearson Correlation	.19	1.00
	Sig. (2-tailed)	.175	
	N	51	51

Correlations

		DiabetesPrevalence	Impressions
DiabetesPrevalence	Pearson Correlation	1.00	.19
	Sig. (2-tailed)		.188
	N	51	51
Impressions	Pearson Correlation	.19	1.00
	Sig. (2-tailed)	.188	
	N	51	51

Correlations

		DiabetesPrevalence	Impressions
DiabetesPrevalence	Pearson Correlation	1.00	.19
	Sig. (2-tailed)		.186
	N	51	51
Impressions	Pearson Correlation	.19	1.00
	Sig. (2-tailed)	.186	
	N	51	51

Correlations

		DiabetesPrevalence	Impressions
DiabetesPrevalence	Pearson Correlation	1.00	.19
	Sig. (2-tailed)		.172
	N	51	51
Impressions	Pearson Correlation	.19	1.00
	Sig. (2-tailed)	.172	
	N	51	51

Correlations

		DiabetesPrevalence	Impressions
DiabetesPrevalence	Pearson Correlation	1.00	*.30*
	Sig. (2-tailed)		.033
	N	51	51
Impressions	Pearson Correlation	*.30*	1.00
	Sig. (2-tailed)	.033	
	N	51	51

Correlations

		DiabetesPrevalence	*Impressions*
DiabetesPrevalence	*Pearson Correlation*	1.00	*.30*
	Sig. (2-tailed)		.033
	N	51	51
Impressions	*Pearson Correlation*	*.30*	1.00
	Sig. (2-tailed)	.033	
	N	51	51

Correlations

		DiabetesPrevalence	*Impressions*
DiabetesPrevalence	*Pearson Correlation*	1.00	*.33*
	Sig. (2-tailed)		.020
	N	51	51
Impressions	*Pearson Correlation*	*.33*	1.00
	Sig. (2-tailed)	.020	
	N	51	51

Correlations

		DiabetesPrevalence	*Impressions*
DiabetesPrevalence	*Pearson Correlation*	1.00	*.30*
	Sig. (2-tailed)		.030
	N	51	51
Impressions	*Pearson Correlation*	*.30*	1.00
	Sig. (2-tailed)	.030	
	N	51	51

TABLE A-16 CORRELATION ANALYSIS ON DIABETES PREVALENCE AND IMPRESSIONS

Correlations

		Impressions	*ObesityPrevalence*
Impressions	*Pearson Correlation*	1.00	-.13
	Sig. (2-tailed)		.346
	N	51	51
ObesityPrevalence	*Pearson Correlation*	-.13	1.00
	Sig. (2-tailed)	.346	
	N	51	51

Correlations

		Impressions	*ObesityPrevalence*
Impressions	*Pearson Correlation*	1.00	-.14
	Sig. (2-tailed)		.312
	N	51	51
ObesityPrevalence	*Pearson Correlation*	-.14	1.00
	Sig. (2-tailed)	.312	
	N	51	51

Correlations

		Impressions	*ObesityPrevalence*
Impressions	*Pearson Correlation*	1.00	-.14
	Sig. (2-tailed)		.336
	N	51	51
ObesityPrevalence	*Pearson Correlation*	-.14	1.00
	Sig. (2-tailed)	.336	
	N	51	51

Correlations

		Impressions	*ObesityPrevalence*
Impressions	*Pearson Correlation*	1.00	-.14
	Sig. (2-tailed)		.339
	N	51	51
ObesityPrevalence	*Pearson Correlation*	-.14	1.00
	Sig. (2-tailed)	.339	
	N	51	51

Correlations

		Impressions	ObesityPrevalence
Impressions	*Pearson Correlation*	1.00	-.02
	Sig. (2-tailed)		.878
	N	51	51
ObesityPrevalence	*Pearson Correlation*	-.02	1.00
	Sig. (2-tailed)	.878	
	N	51	51

Correlations

		Impressions	ObesityPrevalence
Impressions	*Pearson Correlation*	1.00	-.02
	Sig. (2-tailed)		.888
	N	51	51
ObesityPrevalence	*Pearson Correlation*	-.02	1.00
	Sig. (2-tailed)	.888	
	N	51	51

Correlations

		Impressions	ObesityPrevalence
Impressions	*Pearson Correlation*	1.00	.00
	Sig. (2-tailed)		.986
	N	51	51
ObesityPrevalence	*Pearson Correlation*	.00	1.00
	Sig. (2-tailed)	.986	
	N	51	51

Correlations

		Impressions	ObesityPrevalence
Impressions	*Pearson Correlation*	1.00	-.03
	Sig. (2-tailed)		.856
	N	51	51
ObesityPrevalence	*Pearson Correlation*	-.03	1.00
	Sig. (2-tailed)	.856	
	N	51	51

TABLE A-17 CORRELATION ANALYSIS ON OBESITY PREVALENCE AND IMPRESSIONS

Correlations

		ObesityPrevalence	*CTR*
ObesityPrevalence	*Pearson Correlation*	1.00	-.10
	Sig. (2-tailed)		.500
	N	51	51
CTR	*Pearson Correlation*	-.10	1.00
	Sig. (2-tailed)	.500	
	N	51	51

Correlations

		ObesityPrevalence	*CTR*
ObesityPrevalence	*Pearson Correlation*	1.00	-.03
	Sig. (2-tailed)		.815
	N	51	48
CTR	*Pearson Correlation*	-.03	1.00
	Sig. (2-tailed)	.815	
	N	48	48

Correlations

		ObesityPrevalence	*CTR*
ObesityPrevalence	*Pearson Correlation*	1.00	-.07
	Sig. (2-tailed)		.659
	N	51	47
CTR	*Pearson Correlation*	-.07	1.00
	Sig. (2-tailed)	.659	
	N	47	47

Correlations

		ObesityPrevalence	*CTR*
ObesityPrevalence	*Pearson Correlation*	1.00	.08
	Sig. (2-tailed)		.593
	N	51	51
CTR	*Pearson Correlation*	.08	1.00
	Sig. (2-tailed)	.593	
	N	51	51

Correlations

		ObesityPrevalence	*CTR*
ObesityPrevalence	*Pearson Correlation*	1.00	-.41
	Sig. (2-tailed)		.003
	N	51	50
CTR	*Pearson Correlation*	-.41	1.00
	Sig. (2-tailed)	.003	
	N	50	50

Correlations

		ObesityPrevalence	*CTR*
ObesityPrevalence	*Pearson Correlation*	1.00	-.20
	Sig. (2-tailed)		.160
	N	51	51
CTR	*Pearson Correlation*	-.20	1.00
	Sig. (2-tailed)	.160	
	N	51	51

Correlations

		ObesityPrevalence	*CTR*
ObesityPrevalence	*Pearson Correlation*	1.00	-.28
	Sig. (2-tailed)		.050
	N	51	50
CTR	*Pearson Correlation*	-.28	1.00
	Sig. (2-tailed)	.050	
	N	50	50

Correlations

		ObesityPrevalence	*CTR*
ObesityPrevalence	*Pearson Correlation*	1.00	-.29
	Sig. (2-tailed)		.046
	N	51	49
CTR	*Pearson Correlation*	-.29	1.00
	Sig. (2-tailed)	.046	
	N	49	49

TABLE A-18 CORRELATION ANALYSIS ON OBESITY PREVALENCE AND CTR

Appendix 9: Statistical Analyses of RQ6

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
CTR	Alabama	7	3.84	1.00	.38	2.92	4.77	2.47	5.31
	Alaska	7	5.50	3.80	1.44	1.99	9.02	1.25	10.87
	Arizona	8	5.03	1.87	.66	3.46	6.59	2.92	8.19
	Arkansas	8	5.31	1.93	.68	3.70	6.93	1.94	8.33
	California	8	6.38	.77	.27	5.73	7.02	4.80	7.48
	Colorado	7	7.04	2.42	.92	4.80	9.28	4.65	11.86
	Connecticut	7	5.64	2.46	.93	3.36	7.92	2.50	9.47
	Delaware	8	6.58	5.41	1.91	2.06	11.11	.00	15.38
	District of Columbia	8	8.67	7.57	2.68	2.35	15.00	.00	20.00
	Florida	8	6.57	.93	.33	5.79	7.34	5.02	7.68
	Georgia	8	6.89	2.00	.71	5.22	8.57	4.86	10.43
	Hawaii	8	9.91	4.36	1.54	6.26	13.56	3.65	16.22
	Idaho	8	5.22	4.35	1.54	1.59	8.86	1.85	15.15
	Illinois	8	6.36	1.26	.44	5.31	7.41	4.01	8.31
	Indiana	8	6.37	2.23	.79	4.51	8.23	3.67	9.09
	Iowa	8	5.76	3.64	1.29	2.71	8.80	.00	10.64
	Kansas	8	4.32	2.21	.78	2.47	6.17	1.49	7.81
	Kentucky	8	4.44	2.05	.73	2.73	6.16	1.31	6.76
	Louisiana	8	5.14	1.24	.44	4.10	6.18	3.33	7.14
	Maine	8	7.32	2.46	.87	5.26	9.37	4.84	10.53
	Maryland	7	7.28	2.80	1.06	4.69	9.87	3.49	11.43
	Massachusetts	8	7.02	2.52	.89	4.92	9.12	2.81	10.20
	Michigan	8	6.79	1.66	.59	5.40	8.17	4.74	9.60
	Minnesota	7	7.52	2.97	1.12	4.77	10.27	2.92	11.76
	Mississippi	8	4.01	2.54	.90	1.89	6.13	.89	7.56
	Missouri	8	5.88	1.80	.64	4.38	7.39	2.73	7.85
	Montana	8	5.16	2.73	.96	2.88	7.44	.00	9.52
	Nebraska	8	6.12	3.91	1.38	2.86	9.39	2.86	14.89
	Nevada	8	4.74	2.78	.98	2.41	7.07	1.33	8.43
	New Hampshire	8	4.21	3.18	1.12	1.56	6.87	.00	9.43
	New Jersey	8	5.97	1.48	.52	4.74	7.21	3.74	8.02
	New Mexico	8	4.02	1.72	.61	2.58	5.46	1.75	6.99
	New York	8	6.48	.78	.27	5.83	7.13	5.34	7.69
	North Carolina	8	6.12	1.31	.46	5.02	7.21	4.46	7.87
	North Dakota	8	5.68	5.04	1.78	1.47	9.89	.00	16.67
	Ohio	8	6.03	1.26	.44	4.98	7.09	4.91	8.13
	Oklahoma	8	5.62	2.55	.90	3.50	7.75	1.49	10.23
	Oregon	8	6.31	3.10	1.10	3.72	8.91	2.14	10.84
	Pennsylvania	7	7.05	.76	.29	6.35	7.76	6.16	8.05
	Rhode Island	8	8.61	4.43	1.56	4.91	12.31	4.84	17.95
	South Carolina	7	4.84	1.39	.53	3.55	6.12	3.62	7.50
	South Dakota	7	3.76	4.35	1.64	-.26	7.78	.00	10.71
	Tennessee	7	4.99	1.78	.67	3.35	6.64	3.35	7.98
	Texas	8	6.41	.96	.34	5.60	7.21	4.96	7.72
	Utah	8	5.56	4.02	1.42	2.20	8.92	.00	11.54
	Vermont	7	4.17	6.82	2.58	-2.13	10.48	.00	18.18
	Virginia	8	5.57	1.40	.49	4.40	6.74	3.57	7.10
	Washington	8	6.53	1.34	.47	5.41	7.65	4.26	8.70
	West Virginia	8	7.34	2.32	.82	5.39	9.28	4.38	11.36
	Wisconsin	8	6.09	2.14	.76	4.30	7.88	2.63	8.81
	Wyoming	8	1.43	2.69	.95	-.82	3.68	.00	6.67
	Total	397	5.88	3.13	.16	5.57	6.19	.00	20.00

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
CTR	Between Groups	788.05	50	15.76	1.77	.002
	Within Groups	3079.23	346	8.90		
	Total	3867.28	396			

TABLE A-19 ANOVA ON DIFFERENCES IN CTR AND THE USER'S REGION IN THE ADVERTISEMENTS

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Impressions	Alabama	8	176.25	68.58	24.25	118.92	233.58	73.00	273.00
	Alaska	8	51.13	17.59	6.22	36.42	65.83	25.00	80.00
	Arizona	8	260.38	117.22	41.44	162.38	358.37	164.00	459.00
	Arkansas	8	108.38	39.44	13.94	75.40	141.35	48.00	176.00
	California	8	1620.88	743.44	262.85	999.34	2242.41	1027.00	2903.00
	Colorado	8	136.88	65.67	23.22	81.98	191.77	84.00	247.00
	Connecticut	8	126.50	54.60	19.30	80.85	172.15	70.00	211.00
	Delaware	8	25.38	10.21	3.61	16.84	33.91	13.00	40.00
	District of Columbia	8	18.50	11.41	4.04	8.96	28.04	9.00	43.00
	Florida	8	692.13	253.69	89.69	480.04	904.21	396.00	1069.00
	Georgia	8	265.75	98.30	34.75	183.57	347.93	127.00	393.00
	Hawaii	8	83.88	33.14	11.72	56.17	111.58	37.00	137.00
	Idaho	8	58.75	22.26	7.87	40.14	77.36	28.00	98.00
	Illinois	8	566.00	262.51	92.81	346.53	785.47	302.00	970.00
	Indiana	8	198.50	73.26	25.90	137.25	259.75	90.00	319.00
	Iowa	8	124.75	38.58	13.64	92.50	157.00	78.00	192.00
	Kansas	8	143.88	47.67	16.85	104.02	183.73	67.00	208.00
	Kentucky	8	157.50	73.33	25.93	96.19	218.81	72.00	296.00
	Louisiana	8	123.63	49.02	17.33	82.65	164.60	56.00	212.00
	Maine	8	52.50	20.63	7.29	35.26	69.74	19.00	82.00
	Maryland	8	129.25	69.28	24.49	71.33	187.17	70.00	256.00
	Massachusetts	8	244.13	107.05	37.85	154.63	333.62	138.00	424.00
	Michigan	8	350.88	124.72	44.10	246.60	455.15	190.00	569.00
	Minnesota	8	126.75	38.26	13.53	94.77	158.73	73.00	179.00
	Mississippi	8	112.75	46.20	16.34	74.12	151.38	52.00	195.00
	Missouri	8	196.50	77.16	27.28	131.99	261.01	92.00	331.00
	Montana	8	33.50	12.01	4.25	23.46	43.54	20.00	58.00
	Nebraska	8	69.25	21.70	7.67	51.10	87.40	41.00	102.00
	Nevada	8	124.50	65.84	23.28	69.46	179.54	75.00	244.00
	New Hampshire	8	35.13	14.30	5.05	23.17	47.08	15.00	55.00
	New Jersey	8	356.13	204.73	72.38	184.96	527.29	187.00	682.00
	New Mexico	8	113.75	53.05	18.76	69.40	158.10	57.00	199.00
	New York	8	675.38	348.44	123.19	384.07	966.68	404.00	1305.00
	North Carolina	8	286.12	123.03	43.50	183.27	388.98	127.00	471.00
	North Dakota	8	28.00	12.48	4.41	17.57	38.43	14.00	50.00
	Ohio	8	332.75	129.66	45.84	224.35	441.15	153.00	541.00
	Oklahoma	8	127.13	48.99	17.32	86.17	168.08	67.00	214.00
	Oregon	8	132.38	47.01	16.62	93.08	171.67	81.00	205.00
	Pennsylvania	8	380.50	124.34	43.96	276.55	484.45	194.00	529.00
	Rhode Island	8	52.13	23.99	8.48	32.07	72.18	28.00	94.00
	South Carolina	8	134.13	42.85	15.15	98.30	169.95	67.00	191.00
	South Dakota	8	38.38	11.25	3.98	28.97	47.78	22.00	56.00
	Tennessee	8	193.13	79.03	27.94	127.05	259.20	94.00	326.00
	Texas	8	1279.88	699.70	247.38	694.91	1864.84	727.00	2484.00
	Utah	8	69.38	36.59	12.94	38.78	99.97	26.00	132.00
	Vermont	8	24.25	9.10	3.22	16.64	31.86	12.00	39.00
	Virginia	8	232.87	86.98	30.75	160.16	305.59	112.00	352.00
	Washington	8	240.00	78.61	27.79	174.28	305.72	161.00	356.00
	West Virginia	8	85.88	36.46	12.89	55.39	116.36	35.00	137.00
	Wisconsin	8	204.38	61.68	21.81	152.81	255.94	114.00	308.00
	Wyoming	8	14.63	5.26	1.86	10.23	19.02	7.00	21.00
	Total	408	223.83	334.04	16.54	191.32	256.34	7.00	2903.00

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Impressions	Between Groups	34731760.99	50	694635.22	23.22	.000
	Within Groups	10681561.00	357	29920.34		
	Total	45413321.99	407			

TABLE A-20 ANOVA ON DIFFERENCES IN IMPRESSIONS AND THE USER'S REGION IN THE ADVERTISEMENTS

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
CTR	18-24	8	3.95	.90	.32	3.20	4.71	1.869158878505	4.791666666667
	25-34	7	4.17	.98	.37	3.26	5.07	2.787456445993	5.967078189300
	35-44	8	4.80	.80	.28	4.13	5.46	3.435804701628	5.901116427432
	45-54	8	5.10	.68	.24	4.53	5.67	4.178272980501	6.191279209829
	55-64	8	6.30	.45	.16	5.93	6.68	5.637358014304	6.847764572722
	65+	7	7.77	.72	.27	7.11	8.44	6.536541080345	8.777915632754
	Total	46	5.32	1.49	.22	4.88	5.76	1.869158878505	8.777915632754

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
CTR	.37	5	40	.865

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
CTR	Between Groups	76.64	5	15.33	25.88	.000
	Within Groups	23.69	40	.59		
	Total	100.33	45			

TABLE A-21 ANOVA ON DIFFERENCES IN CTR AND THE USER'S AGE IN THE ADVERTISEMENT

Multiple Comparisons (CTR)

	(I) Age	(J) Age	Mean Difference (I - J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Scheffé	18-24	25-34	-.21	.40	.998	-1.61	1.18
		35-44	-.84	.38	.453	-2.19	.50
		45-54	-1.15	.38	.138	-2.50	.20
		55-64	-2.35	.38	.000	-3.70	-1.00
		65+	-3.82	.40	.000	-5.21	-2.43
	25-34	18-24	.21	.40	.998	-1.18	1.61
		35-44	-.63	.40	.775	-2.02	.76
		45-54	-.94	.40	.374	-2.33	.46
		55-64	-2.13	.40	.000	-3.53	-.74
		65+	-3.61	.41	.000	-5.05	-2.17
	35-44	18-24	.84	.38	.453	-.50	2.19
		25-34	.63	.40	.775	-.76	2.02
		45-54	-.31	.38	.986	-1.65	1.04
		55-64	-1.50	.38	.020	-2.85	-.16
		65+	-2.98	.40	.000	-4.37	-1.58
	45-54	18-24	1.15	.38	.138	-.20	2.50
		25-34	.94	.40	.374	-.46	2.33
		35-44	.31	.38	.986	-1.04	1.65
		55-64	-1.20	.38	.109	-2.55	.15
		65+	-2.67	.40	.000	-4.06	-1.28
	55-64	18-24	2.35	.38	.000	1.00	3.70
		25-34	2.13	.40	.000	.74	3.53
		35-44	1.50	.38	.020	.16	2.85
		45-54	1.20	.38	.109	-.15	2.55
		65+	-1.47	.40	.033	-2.87	-.08
	65+	18-24	3.82	.40	.000	2.43	5.21
		25-34	3.61	.41	.000	2.17	5.05
		35-44	2.98	.40	.000	1.58	4.37
		45-54	2.67	.40	.000	1.28	4.06
		55-64	1.47	.40	.033	.08	2.87

TABLE A-22 POST HOC SCHEFFE TEST ON DIFFERENCES IN CTR AND THE USER'S AGE IN ADVERTISEMENTS

Ranks

	Age	N	Mean Rank
Impressions	18-24	8	9.56
	25-34	8	12.63
	35-44	8	21.56
	45-54	8	29.63
	55-64	8	34.88
	65+	8	38.75
	Total	48	

Test Statistics

	Impressions
Chi-Square	28.97
df	5
Asymp. Sig.	.000

TABLE A-23 KRUSKAL-WALLIS TEST ON DIFFERENCES IN IMPRESSIONS AND THE USER'S AGE IN THE ADVERTISEMENTS

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	18-24	*25-34*	*Total*	*18-24*	*25-34*	*18-24*	*25-34*
Impressions	8.00	8.00	16.00	7.94	9.06	63.50	72.50

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	27.50	63.50	-.47	.636

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	18-24	*35-44*	*Total*	*18-24*	*35-44*	*18-24*	*35-44*
Impressions	8.00	8.00	16.00	6.00	11.00	48.00	88.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	12.00	48.00	-2.10	.036

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	18-24	*45-54*	*Total*	*18-24*	*45-54*	*18-24*	*45-54*
Impressions	8.00	8.00	16.00	4.63	12.38	37.00	99.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	1.00	37.00	-3.26	.001

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	18-24	*55-64*	*Total*	*18-24*	*55-64*	*18-24*	*55-64*
Impressions	8.00	8.00	16.00	4.50	12.50	36.00	100.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	.00	36.00	-3.36	.001

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	18-24	*65+*	*Total*	*18-24*	*65+*	*18-24*	*65+*
Impressions	8.00	8.00	16.00	4.50	12.50	36.00	100.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	.00	36.00	-3.36	.001

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	25-34	*35-44*	*Total*	*25-34*	*35-44*	*25-34*	*35-44*
Impressions	8.00	8.00	16.00	6.44	10.56	51.50	84.50

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	15.50	51.50	-1.73	.083

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	25-34	*45-54*	*Total*	*25-34*	*45-54*	*25-34*	*45-54*
Impressions	8.00	8.00	16.00	5.75	11.25	46.00	90.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	10.00	46.00	-2.31	.021

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	25-34	*55-64*	*Total*	*25-34*	*55-64*	*25-34*	*55-64*
Impressions	8.00	8.00	16.00	4.75	12.25	38.00	98.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	2.00	38.00	-3.15	.002

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	25-34	*65+*	*Total*	*25-34*	*65+*	*25-34*	*65+*
Impressions	8.00	8.00	16.00	4.63	12.38	37.00	99.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	1.00	37.00	-3.26	.001

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	35-44	*45-54*	*Total*	*35-44*	*45-54*	*35-44*	*45-54*
Impressions	8.00	8.00	16.00	6.13	10.88	49.00	87.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	13.00	49.00	-2.00	.046

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	35-44	*55-64*	*Total*	*35-44*	*55-64*	*35-44*	*55-64*
Impressions	8.00	8.00	16.00	6.13	10.88	49.00	87.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	13.00	49.00	-2.00	.046

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	35-44	*65+*	*Total*	*35-44*	*65+*	*35-44*	*65+*
Impressions	8.00	8.00	16.00	5.75	11.25	46.00	90.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	10.00	46.00	-2.31	.021

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	45-54	*55-64*	*Total*	*45-54*	*55-64*	*45-54*	*55-64*
Impressions	8.00	8.00	16.00	6.88	10.13	55.00	81.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	19.00	55.00	-1.37	.172

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	45-54	*65+*	*Total*	*45-54*	*65+*	*45-54*	*65+*
Impressions	8.00	8.00	16.00	6.25	10.75	50.00	86.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	14.00	50.00	-1.89	.059

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	55-64	*65+*	*Total*	*55-64*	*65+*	*55-64*	*65+*
Impressions	8.00	8.00	16.00	7.13	9.88	57.00	79.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	21.00	57.00	-1.16	.248

TABLE A-24 15 POST HOC MANN-WHITNEY U TESTS ON DIFFERENCES IN IMPRESSIONS AND THE USER'S AGE IN THE ADVERTISEMENTS

Ranks

	Gender	N	Mean Rank
CTR	1	8	18.13
	2	8	10.50
	3	8	8.88
	Total	24	

Test Statistics

	CTR
Chi-Square	7.81
df	2
Asymp. Sig.	.020

TABLE A-25 KRUSKAL-WALLIS TEST ON DIFFERENCES IN CTR AND THE USER'S GENDER IN THE ADVERTISEMENTS

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	female	*male*	*Total*	*female*	*male*	*female*	*male*
CTR	8.00	8.00	16.00	11.50	5.50	92.00	44.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
CTR	8.00	44.00	-2.52	.012

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	female	*unknown*	*Total*	*female*	*unknown*	*female*	*unknown*
CTR	8.00	8.00	16.00	11.13	5.88	89.00	47.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
CTR	11.00	47.00	-2.21	.027

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	male	*unknown*	*Total*	*male*	*unknown*	*male*	*unknown*
CTR	8.00	8.00	16.00	9.50	7.50	76.00	60.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
CTR	24.00	60.00	-.84	.401

TABLE A-26 3 POST HOC MANN-WHITNEY U TESTS ON DIFFERENCES IN CTR AND THE USER'S GENDER IN THE ADVERTISEMENTS

Ranks

	Gender	N	Mean Rank
Impressions	1	8	20.50
	2	8	12.50
	3	8	4.50
	Total	24	

Test Statistics

	Impressions
Chi-Square	20.48
df	2
Asymp. Sig.	.000

TABLE A-27 KRUSKAL-WALLIS TEST ON DIFFERENCES IN IMPRESSIONS AND THE USER'S GENDER IN THE ADVERTISEMENTS

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	female	*male*	*Total*	*female*	*male*	*female*	*male*
Impressions	8.00	8.00	16.00	12.50	4.50	100.00	36.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	.00	36.00	-3.36	.001

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	female	*unknown*	*Total*	*female*	*unknown*	*female*	*unknown*
Impressions	8.00	8.00	16.00	12.50	4.50	100.00	36.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	.00	36.00	-3.36	.001

Ranks

	N			*Mean Rank*		*Sum of Ranks*	
	male	*unknown*	*Total*	*male*	*unknown*	*male*	*unknown*
Impressions	8.00	8.00	16.00	12.50	4.50	100.00	36.00

Test Statistics

	Mann-Whitney U	*Wilcoxon W*	*Z*	*Asymp. Sig. (2-tailed)*
Impressions	.00	36.00	-3.36	.001

TABLE A-28 3 POST HOC MANN-WHITNEY U TESTS ON DIFFERENCES IN IMPRESSIONS AND THE USER'S GENDER IN THE ADVERTISEMENTS