Developing Behavioral Interventions and Visual Intervention Displays for a Mobile Health Application

A Major Qualifying Project
Submitted to the Faculty of the

Worcester Polytechnic Institute

In partial fulfillment of the requirements for the
Degree of Bachelor of Science

By

Michael Harney, CS
Qiaoyu Liao, CS

May 1, 2017

Approved by:

Professor Carolina Ruiz, Advisor
Professor Bengisu Tulu, Co-Advisor

Abstract

This MQP focused on developing user interfaces that are effective in visualizing data in mobile health apps, developing novel ways to deliver data driven interventions and discovering patterns from mobile health use data through the use of association rules.

Table of Contents

List of Figures

List of Tables

1. Introduction

According to the Center for Disease Control and Prevention, 36.5% of US adults are obese(Center for Disease Control and Prevention, 2016). Obesity is an epidemic spreading across our country as a result of eating too much, exercising too little and the convenience of high-portion, high-calorie fast foods (Stewart, 2006b). According to Marion Nestle, PhD, MPH, chair of the department of nutrition and food studies at New York University, corporate competition encourages higher portion sizes, which results in overeating, and the spread of obesity (Stewart, 2006a).

But how do you counteract such a widespread problem?

The National Heart, Lung and Blood Institute believes that watching portion size is a key component in the fight against obesity(National Heart, Lung, and Blood Institute, 2012), which supports Nestle's claims, in addition to healthy exercise and dietary habits. Also, the American Obesity Treatment Association(American Obesity Treatment Association, 2011) believes that behavior therapy is a very useful and viable method to implement those attitudes.

This is not new territory. There is a plethora of dietary apps on the market. They mostly take in users' calorie counts to tell them when they are overeating. But this has three major drawbacks. Firstly, the user does not want to input data to the application, particularly of their failures. Secondly, the apps focus too much attention on the 'when' instead of the 'why'. Thirdly, 'overeating' may be different or subjective based on the person. For example, if someone is trying to accumulate

muscle mass, eating four chicken legs might not be considered overeating, but eating eight Oreos might.

In response, a previous MQP group developed the SlipBuddy app, an Android application that records simple data like sleep and stress, and overeating episodes to find patterns in overeating. Overeating episodes are recorded when the user presses the in-app 'Oops' button after an overeating episode. The app takes this data, mines for patterns with machine learning, and presents a rough pattern to the user, so they can self-correct their own behavior and prevent future overeating episodes.

This MQP has been geared towards moving SlipBuddy another step further in its evolution. Whereas the last MQP built the app's foundation, this one focuses on improving the user experience. Our group worked on developing interfaces that enable visualization of the overeating data, providing visual cues for interventions and further analyzing the pilot data collected during 2016 summer.

2. Background

SlipBuddy is an application designed to help users identify and counteract their own overeating habits. This section will serve as a review of what came before our project, both the research and the justification. However, as this project is the continuation of a past MQP, this section will be two-fold. Firstly, it will review the foundation of research supporting the initial project. Secondly, it will explain the ground covered during the last MQP.

2.1 Current Dietary App Market

There is a significant number of dietary apps on the mobile market. While it is difficult to come up with an exact figure on their popularity, physical self-improvement applications have certainly established themselves as a popular genre in the mobile application market. However while these apps are flashy and consistently reviewed on several websites, more official sources have higher doubts of their quality. The Diabetes Prevention Program (DPP) conducted a study with the top on hundred free and non-free weight-loss applications from both the Android and iOS markets(Direito et al., 2014). The purpose of the study was to decide whether there was ample usage of behavioral strategies within these weight-loss applications. The study concluded that there is a harrowing vacancy in the market for applications that properly use effective behavioral strategies to reform behavior (which as we have covered in the Introduction, is an important method for losing weight). A similar study was conducted with the top forty applications, with very similar results(Direito et al., 2014)å. We also conducted our

9

own study (see Findings 4.3) where we evaluated popular and well-researched apps for their data mining and behavioral modification capabilities. Like the DPP's study, we did not find anything substantial within that realm.

2.2 Why SlipBuddy works in this market

Fortunately, smartphones and behavioral modification synchronize very well with each other. The primary challenge of behavioral modification is the collecting, crunching, and upkeep of each users' data, which cell phones, with their convenience and processing power, excel at. In *Psychology & Health*, Susan Michie details forty strategies for creating effective behavioral change (Michie, van Stralen, & West, 2011). Most of her techniques discuss when and how the information should be presented, all of which result in the following general summary. First, the user must understand their current circumstances, the consequences of their actions, and a general idea of the 'norm' of others. Next, the user is prompted to set goals and plan future steps. Finally, the guide (in this case, the app) provides the user with information as to the progression of their goals.

In theory, the purpose of SlipBuddy is to guide users away from overeating habits, and towards their goals. While the collecting, crunching, and upkeep of users' data was present, the guidance was often not. Users receive a message when they are likely to overeat, however the messages were customized by the programmers for each unique situation, and as such were not scalable as a model for message generation. In addition, there was not any infrastructure for goal-setting. Users could record their weights in-app, but any goals would have to be entirely internal. In short, while the

application worked for the data, it did not guide users in the direction it needed to. This purpose makes up the crux of our MQP and the MQP that preceded ours. We strove to make SlipBuddy more usable and functional for its users, as well as expanding its use of data to other areas (such as intervention timing). We will cover our process to approaching these challenges in the Methodology section.

2.3 Decision Trees

The primary method utilized by the SlipBuddy application to generate interventions is the decision tree method. Decision trees are graphs that use branching methods to illustrate every possible outcome of a decision. In **Figure 1** below, the population of the Titanic is analyzed through a decision tree (Milborrow, 2011) to determine which combinations of attributes resulted in the deaths of its passengers (where sibsp stands for 'spouses or siblings aboard').
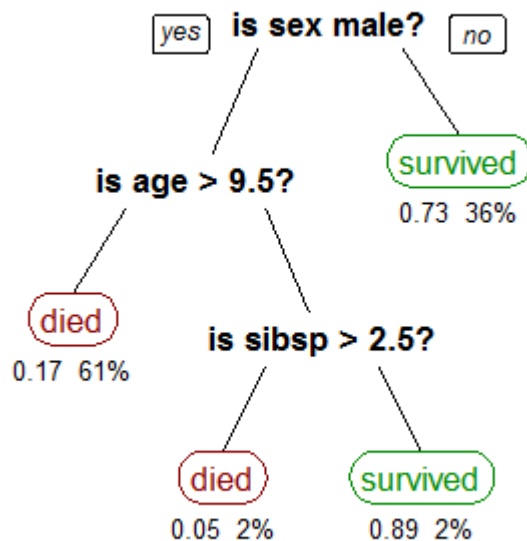
Figure 1: Depiction of Decision Tree for Titanic survivors

11

The task of the decision tree is to determine with maximum certainty which combinations of data can most determine the target outcome(Kingsford & Salzberg, 2008). In the above example, the outcome is whether passengers survived. In SlipBuddy, the outcome is whether its user overate.

Based on various user inputs during various times of day, SlipBuddy constructs a decision tree like the one above to determine which conditions result in each user overeating or not. In this way, we generate sets of conditions that are likely to result in overeating episodes. We use these conditions to generate interventions, to warn users when these conditions are met.

2.4 Association Rules and Frequent Itemsets

A new method used in this project to find relationships among certain behaviors that play a role in overeating is association rule mining. Originally association rules are defined as follows: let I be a set of binary attributes called items and D be a dataset consisting of a set of transaction. Each transaction contains a subset of items in I. An association rule is defined as $X \Rightarrow Y$ where $X,Y \subseteq I$ (Agrawal, Imieliński, & Swami, 1993).

Different metrics are used to evaluate the quality and significance of an association rules. Among them, the most common ones are support and confidence. The support of a rule $X \Rightarrow Y$ measures the relative frequency with which X and Y appear together in transactions in D: support($X \Rightarrow Y$) = number of transactions that contain X and Y divided by the number of transactions in D. The confidence of a rule $X \Rightarrow Y$ measures the likelihood that a transaction in D contains Y given that it contains X: Prob(Y|X) = support(X & Y) /support(X) x support(Y). The support of an itemset X is

12

the number of transaction in D that contain X divided by the number of transactions in D.

The traditional association rule mining task is defined as follows: Given a dataset D of transactions and a minimum support, minsupp, and a minimum confidence, minconf, find all association rules $X \Rightarrow Y$ in D such that the support and confidence of the rule are no lower than minsupp and minconf, respectively.

An algorithm called the Apriori algorithm was proposed to tackle this problem in(Agrawal & Srikant, 1994). There are two steps in finding the rules in this algorithm. First, the minimum support threshold is applied to find all frequent itemsets in a database. A frequent itemset is a set of items X whose support is no lower than minsupp. Then, the minimum confidence constraint is applied to all possible rules that can be constructed from these frequent itemsets to keep only rules that satisfy minconff.

In this project, we used Python Orange package's Apriori algorithm to mine association rules.

2.5 Previous Participation Analysis

While the previous MQP had 16 participants total that engaged in the user study last year, it is hard to conclude whether certain participants used the application actively or not. In order to understand the participant rate, we did an analysis on the data we collected from last year's user study.

The user study is divided into two parts, Phase 1 and Phase 2. Each phase is roughly a month long, with participants often having different start dates for the same phase. Phase 1 and 2 are usually two to three months apart. This was done to give participants

time to rest in between phases. Of all participants, two participants dropped after Phase 1. The remaining 14 participants finished the study with varying levels of participation.

The participation data are divided into two categories, check-ins and episodes. Check-ins are data collected during the day to monitor the participants' statuses. Three check-ins occur throughout the day, during the morning, afternoon and evenings. They are available for participants to fill in throughout the day and the next day. The second category is episodes. Episodes are what we often referred to as "Oops incidents" - when the participants overeat, they are advised to record it using the "Oops button" within the application. A participant may have multiple Episodes each day. Using these two data sets (Check-ins & Episodes), we analyzed the participation rates for every participant.

2.5.1 Check-ins Participation Analysis

Check-ins data are collected during the day to monitor each user's stress levels, as well as their sleep hours, hunger levels and weight. They are clear indicators of whether or not a user used the application during that day, since check-ins are required daily for every user. We analyzed the participation of users on a daily basis compared between phases and times.

Phase 1 has a solid participation rate. With all 16 participants, the average Phase 1 length is 36.81 days with an average of 90.27% participation rate in check-ins. And among all of them, 6 participants used the check-in function every single day, 7 have used the function in more than 85% of the days assigned. The remaining 3 have the

participation rate of 56.8%, 76% and 68.8%, including one participant that dropped out of the study after Phase 1.

Compared to Phase 1, Phase 2's participation dropped significantly. With all 14 participants in Phase 2, the average length is 42.14 days and the average participation rate is 63.30%, a significant drop from Phase 1. None of the participants used the application for check-in every single day. 6 participants have more than 85% days checked, 1 has more than 75%, the rest, 7 participants have less than 50% days checked. Using 85% as a threshold, Phase 1 has 13 out of 16 participants reach that goal, but Phase 2 has only 6 out of 14 participants. Overall, with two phase combined, the average length of use is 73.69 days and the average participation rate is 77.10%.

Table 1: Participation Rates by Phase

Individual Participants	Phase 1 rate	Phase 2 rate	Overall rate
Participant 1	56.82%	n/a	56.82%
Participant 2	91.67%	94.74%	93.24%
Participant 3	100.00%	98.15%	98.68%
Participant 4	100.00%	95.45%	97.53%
Participant 5	95.83%	75.47%	81.82%
Participant 6	100.00%	97.44%	98.67%
Participant 7	76.00%	87.80%	81.32%
Participant 8	97.06%	n/a	97.06%
Participant 9	85.29%	32.56%	55.84%
Participant 10	94.74%	21.95%	56.96%
Participant 11	92.11%	47.22%	70.27%
Participant 12	68.75%	29.73%	51.76%
Participant 13	86.05%	45.95%	67.50%
Participant 14	100.00%	27.45%	61.86%
Participant 15	100.00%	47.22%	73.61%
Participant 16	100.00%	85.00%	90.48%
AVERAGE	90.27%	63.30%	77.09%

As mentioned above, as long as a day has one check-in, the day is counted as participated. However, in one day, we have three types of check-ins: morning, afternoon and evening. Each of them has different level of participation rate that are aggregated using the number of that specific check-ins compared to the length of the study to find out the rate of each person, then aggregate the number to get the average participation rate for each kind of check-ins. In Phase 1, morning check-ins have an average participation rate of 89.6%, afternoon check-ins have an average rate of 85.6% and those in the evening have an average rate of 76.2%. In Phase 2, morning check-ins have an average participation rate of 60.9%, afternoon check-ins have an average rate of 56.5% and those in the evening have an average rate of 51.8%. Clearly, the participation rate of those in the mornings is the highest, which is close to the overall participation rate. Afternoon and evening rate dropped down gradually, with afternoon is around 4.2% less than the morning and evening less than afternoon.

Table 2: Participation Rate by Phase & Time of Day

AVERAGE PARTICIPATION RATE	OVERALL	MORNING	AFTERNOON	EVENING
Phase 1	90.27%	89.60%	85.64%	76.19%
Phase 2	63.30%	60.88%	56.52%	51.79%

2.5.2 Episodes Participation Analysis

Episodes are the indicators of overeating activity. Since overeating could happen at different rates for different individuals, the degree of overeating or when the Episode occurs, do not necessarily align with the participation of the users. However, one interesting finding is that the episodes happen much less frequent than the check-ins.

16

One extreme case for example, is found in one participant in Phase 1. He/She has checked-in every day for 22 days, but only has registered 2 episodes throughout that period.

3. Methodology

3.1 Overview

Our project was focused on enhancing the SlipBuddy application. This process consisted of two principles. First, we sought to find direction by studying the application and its place in the mobile app market. Second, we improved the application by adding or improving various features such as intervention messages, and user settings.

3.2 Intervention Messages

The intervention process involved three simple phases: the 'develop' phase, where an approach was designed and implemented, the 'talk' phase, where once the approach was implemented, the team reflected on its pros and cons, and the 'revise' phase, where the approach was either readjusted to fit its flaws, or was scrapped to begin the cycle again. We cycled through this cycle four times and developed a final implementation approach.

3.3 User Settings

We wanted to provide the users with the ability to input and update their information and give them a sense of customizability within the app.

The approach we decided to take for this portion of our project was to create a selection of potential settings to implement. Then, we implemented dummy settings that fit those parameters, and after a period of time, we returned to the list of settings

and re-evaluated which settings fit the goals of what we were trying to achieve, and the theme of what we wanted our 'Settings' to mean.

3.4 Survey of Applications

As described in the Background section, it was important for our project to understand the current usages of behavior modification and data mining in the current market. To assess the potential need for and benefits of SlipBuddy, we reviewed several apps that were well-studied in the literature and were either extremely popular, had notable usages of behavior modification, or both.

Our approach for doing so was to scour various academic sources for reviews of apps that were noteworthy for their use of data mining or behavioral analysis. Apps that were well-researched, popular, and often both, made the cut, and only those that were noteworthy made the final reviews.

3.5 Graphing Participation

A common question that arose when people considered both our and past SlipBuddy projects is 'How is the Participation rate? Are users actively reporting their episodes in the experiment?'

To answer that question, we compiled the participation data according to the first two trial phases of the last SlipBuddy project's start and end dates. From the dates we got, we counted the number of check-ins and episodes in each phase as well as in between Phases 1 and 2. We used that data to visualize the participation of various participants. This includes a line chart detailing how the participants' check-in rates fluctuated across various phases. The charts can be found in section 5.4.

19

3.6 Improving Application Input

While the old application is effective, some of relevant information about overeating episodes is not being collected, and portions of the design are not as user-friendly as they can be. In order to improve the current application and streamline the input experience, we added and change some features in the application.

3.6.1 Quantifying Overeating

While doing the data analysis of the current Episode data we have at hand, one thing we noted is that there is no way of quantifying how much a user overeats at each episode. It could be just one more bite on a cookie, or pounds of chicken wings. To deal with this, we added one more question to the survey that is triggered after each episode. This question asks: "On a scale of 0 to 10, how much did you overeat this time?". The application's local database on the cellphone and the central database on the remote server were both rebased to include this additional data.

3.6.2 Streamline Check-ins

During our discussion, we found that the application requires too much page-flipping in order to finish the morning check-ins. Unlike afternoon and evening check-ins, those in the morning have a total of five questions asked, including 'stress level', 'hunger level', 'sleep quality', 'sleep hours' and the users' current weight. In the original application, each question is represented in one page. This results in the users clicking the "next" button 5 times to finish the check-in. This is time-consuming.

20

4. Results

<u>4.1 Intervention Messages</u>

Over the course of the project and our design, talk, revise cycles, we developed four major approaches to how we wanted to handle intervention messages. The results of those approaches are listed below.

<u>4.1.1 Intervention Messages: Sorting by Priority</u>

Our first implementation of self-generated user messages involved sorting elements of the message by their importance (height) in the decision tree. In a decision tree, the higher the element is on the tree, the more statistically important it is. For example if the intervention occurred because of conditions A (root node), B (child of node A), and C (child of node B), the message would look something like: "You tend to overeat when you A, B, and C". The textual representations of A, B, and C resemble (with a few exceptions): (over/under)+(verb)+(degree)+(time of day). The degrees ranged from 'a bit' to 'a lot' depending on how far the occurrence was from the perceived 'norm'.

Also, as the conditions were arranged in a tree-like format, we generated message combinations for each iteration downwards. As the conditions were pre-sorted in order of importance, we generated (for a three condition example) messages for condition A-B-C, A+B-C, and A+B+C. That way the user could select the 'depth' of their intervention message. This would allow them to choose a brief message (based on just condition A), or a detailed one (based on conditions A+B+C), or anywhere in between.

The benefits of this approach were that the information was very explicitly communicated to the user. It also maintained the order of the conditions' importance. The trigger conditions are pre-sorted by order of statistical importance. Keeping them in that order made the users' reasons for overeating very clear. The problem with this approach is that it became clunky. To say: "You tend to overeat when you overstress a bit in the morning, undereat a lot in the afternoon, undersleep a bit in the morning, and overstress a bit in the afternoon" is annoyingly long and stilted. Even with a small depth, the repetition of the times of day makes for annoying read-throughs. To make this message more user-friendly, we developed a second approach.

4.1.2 Intervention Messages: Sorting by Time of Day

The second approach we took was sorting by time of day rather than priority. This solved the problem of brevity found in our last approach. The benefits of this strategy were that the information was easier to read through and more user-friendly. However, the messages appearing in order of the time of day made the importance of each condition invisible to the user. Here we chose readability over small pieces of information. And while to an extent we were able to combine the two (sorting each element in each time of day by priority), it did not have the thoroughness of the previous model.

4.1.3 Pre-Emptive Warnings

In the then-current implementation, interventions were triggered when every condition was met. In other words, if you overate when you did A+B+C, it would warn you about overeating. But this was a system we could improve upon. If A, B both occurred in the

22

morning, perhaps it could warn you about doing C later, and overeating. While previous messages relied on interpreting the conditions for interventions, this new version demanded on constructing conditions differently.

When generating intervention conditions, a decision tree is built for each user. Each branch (or series of conditions) that results in a significant likelihood of overeating becomes an intervention. Every condition of that branch is listed in order, and those conditions are used to trigger interventions. For the new model, we had to add another step. We kept the standard interventions, but also broke them up into sections based on their time of day, so pre-emptive warnings would trigger in addition to the standard ones. So instead of printing out the branches, and having them become the trigger arguments for interventions, we reassembled the branches before they became passed as arguments.

The new message generation was similar in concept to the past versions. If you trigger every piece of the AM section of the original branch, it will trigger an intervention message citing the AM conditions, with a warning for the highest priority item of that original branch that occurs later on in the day (in this case, afternoon or evening). Alternatively, the AM&PM branches work the same way. (AM and PM are grouped together to utilize all data gathered by the afternoon).

This new message model was more proactive than past versions, but became programmatically complex. Through gradual design, it was not difficult to manage, however after reviewing it, we had realized how difficult it had been to make

23

'intervention messages' work, and how we still were dealing with the same fundamental issues.

The interventions were generated with the specific purpose of communicating important data to the user in an easy to understand way. However, we constantly ran into trade-offs with brevity vs complexity. Text is inherently flawed in that it cannot communicate large amounts of information in a dense, user-friendly space. To remedy this problem, we moved to our final intervention model: Graphical Interventions.

4.1.4 Intervention Messages: Graphical Interventions

Before we concluded our project, we decided to substitute user messages with a helpful graphic. While this meant that parts of our development had not come to fruition, it was a decision that was not made hastily, and supported by three primary reasons. The first was information density. User messages are flawed in that they cannot communicate complex information within a small space. Infographics, however, are much more efficient at storing dense information. The second reason was user engagement. We expect users to be more willing to observe and possibly return to a helpful infographic, than a message. The third reason was the complexity of tone. Prompting the user is difficult in that the program is fundamentally telling the user what to do. An effective program must be helpful and friendly while not coming across as annoying. And while this problem is not programmatically complex, it is one subverted by using images in place of text.

Pictured below, in **Figure 2**, we have an intervention represented graphically where an overeating episode is likely to occur if the user is hungry in the morning,

24

stressed in the afternoon, and has medium stress in the evening. With our model, the user is able to view their top intervention at any point throughout the day, so they know what to watch out for.
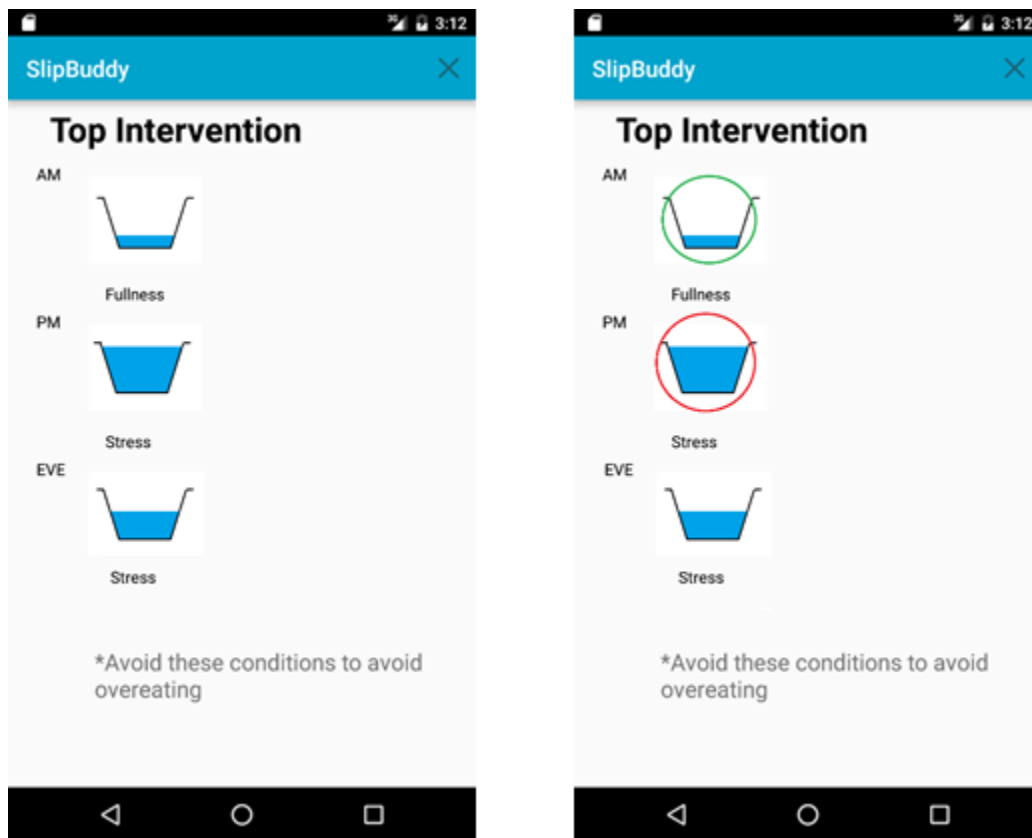
Figure 2: Graphical Intervention Model with and without checkpoints

We also implemented checkpoints into the graph. This means that as the users' day progresses through the morning, afternoon, and evening, they are able to keep track of their progress in avoiding behavior that triggers overeating. If they admit to enacting a behavior in their check-in that is associated with overeating, that behavior is circled in red in the graph. If they successfully avoid said behavior, it is circled in green.

25

In the above example, the user successfully avoided his morning hunger, but failed to avoid afternoon stress.

By enabling graph-based interventions instead of text-based ones, we are able to provide users with very information-rich experiences that would have been impossible with simple text.

4.2 Settings

We integrated settings for Age, Height, and Ethnicity among others. We also added a Message Depth setting so users could customize the level of detail in the intervention messages. However this setting does not align with the new message model, and was thus removed. We also removed settings that took in user information, as we were not sure if that information would be utilized in the big-picture.

The remaining settings consisted of preferred check-in times for the user. This means that in future versions, the application will be able to notify the user that they should check in at a time that is convenient and customizable for them.

4.3 Survey

The results of the survey of existing weight loss apps on the market are both discouraging and heartening. The applications reviewed do not use behavior modification and data mining to the extent that SlipBuddy does. While it is unfortunate these techniques are not being utilized to their full potential, it does mean that there is a spot in the world of Health & Fitness apps for an app like SlipBuddy, which could potentially carve a niche as an application that employs data mining to a more full and thorough extent.

26

Below is the results of that survey, conducted on various Health & Fitness apps on the market which were well researched, and either very popular or notable for their usages of behavior modification (in several cases both). The simplified results of the survey can be found in the table below. Afterwards, more thorough reviews can be found.

Table 3: Survey Result

Name	Data Measured	Extent of App's Analysis	Other	Studied?
MyFitnessPal	Calories, weight, calorie composition	Graphs over time	Scans food	Yes
Jawbone Up	Sleep, calories, physical activity	Graphs over time	Plug-in for MyFitnessPal Add-on wristlet	Yes
Vegathon	Vegetable intake	Provides focused feedback and behavior modification	None	Yes
Lark	Diet, sleep, physical activity	Comparative graphs, personalized feedback on progress above or below average	No	Yes
8fit	Weight, calories, workout difficulty	Adjusts workouts based on user performance and enjoyment	No	Yes
Fooducate	Weight, calories	Daily calorie and weight counts	Rates the health value of various foods	Yes
Lose It!	Calories, weight, projected time to goal	Graphs over time, projected goal date	Calorie index of various foods	Yes

4.3.1 MyFitnesssPal
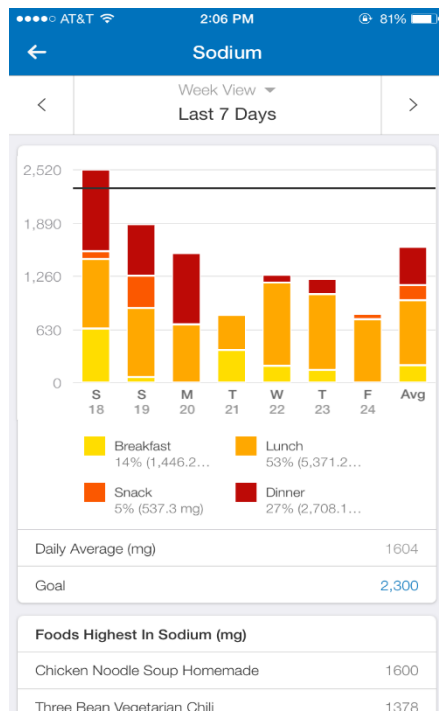
Figure 3: My Fitness Pal daily chart
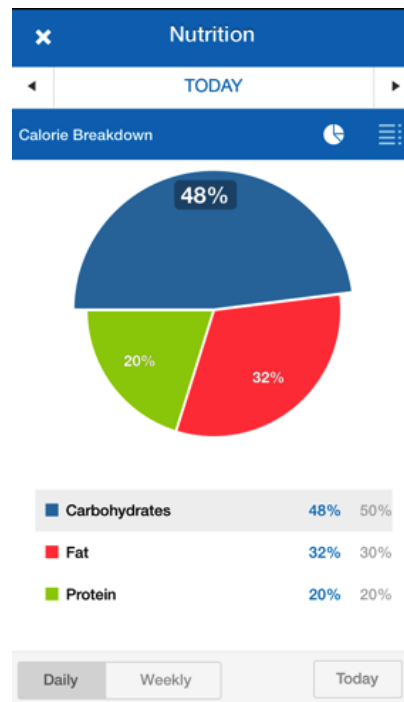
Figure 4: My Fitness Pal Nutrition chart

Data Measured: Foods you have eaten (remembers), Records meals

Tracked: Calories, weight, and calorie composition over time

Other capabilities: Scans food (determines health value of food via barcodes)

Study: (Ipjian & Johnston)

The app was recommended in a study analyzing the potential benefits of using apps to modify dietary behaviors. The study determined that MyFitnessPal did, in fact, contribute to significant success in correcting their behavior metric (sodium intake).

In terms of its data analysis, it tracks calories and weight over time, in addition to other miscellaneous diet help. In terms of data analysis and behavior modification,

28

while it does provide some graphs to analyze (pictured above), it does not provide

users with any specific paths to success aside from their own graphical analysis.

4.3.2 Jawbone up
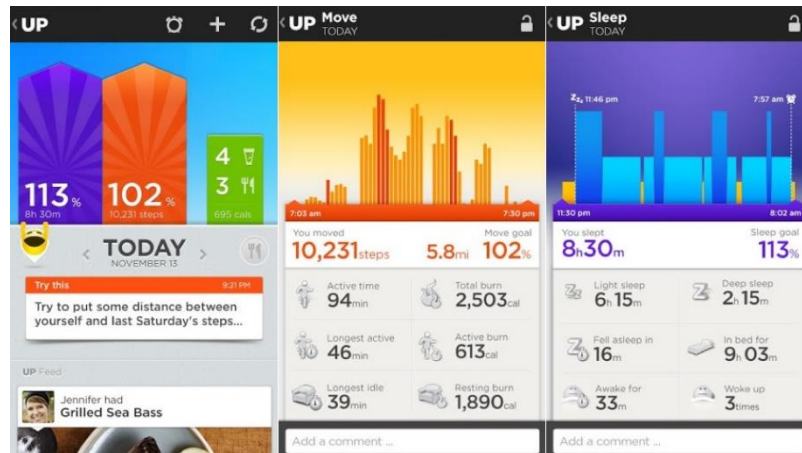
Figure 5: Jawbone Up charts

Data Measured/Tracked: Sleep, calories, and steps over time

Other capabilities: Mostly used as a helpful plug in for MyFitnessPal.

Add-ons: Wristlet (precisely measures steps)

Study:(de Zambotti, Baker, & Colrain, 2015)

This app was evaluated in a study that determined that it was able to nearly

match polysomnography in its analysis of sleep behaviors. Its most significant claim to

fame though is that of MyFitnessPal (the top fitness app on the Apple store)'s plug-ins,

this one has the most data analytics.

Overall, this app adds sleep analysis and steps taken to the MyFitnessPal model.

The metrics listed in the above graphs are also more varied. However, despite the 'Try

this' tip on the layout above, it does not provide any consistent specific guidance for

wading through the data and course-correcting.

29

4.3.3 Vegathon

Study: (Mummah, King, Gardner, & Sutton, 2016)

This app was developed by Stanford to attempt to use several behavioral modification techniques such as goal-setting, feedback, social comparison, prompts/queues, framing identity. The study is actually very thorough in its uses. However, the app is not available to the open market.

Overall, the app uses analysis and personal comparisons to try and direct behavior. It does recommend particular changes so it is helpful in that regard. Speaking in terms of product viability though, it is much more specialized in its range over other apps, as its expertise begins and ends at vegetable consumption. It is an interesting example though of the different forms behavior modification can take.

4.3.4 Lark
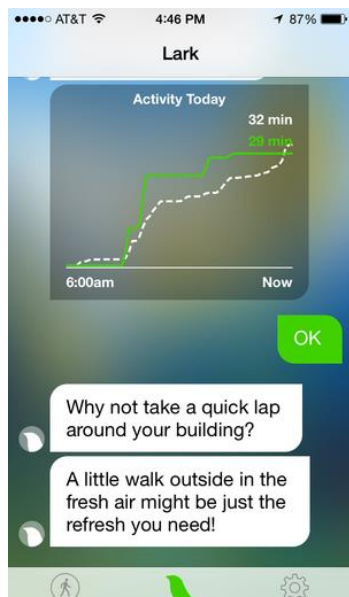
Figure 6: Lark Daily Progress Image
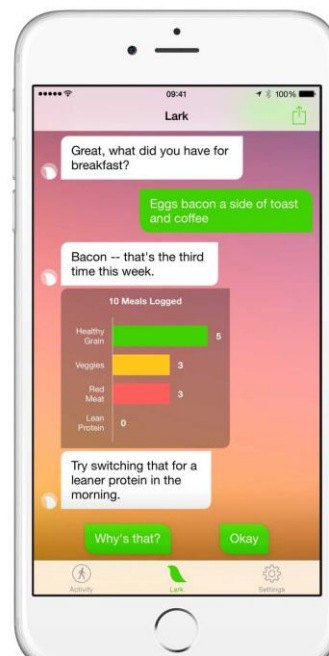
Figure 7: Lark Weekly Progress Image

30

Data Measured/Tracked: Diet, Sleep, Physical Activity

Study:(McCormick, Ask, Pilecki, & Ngo, 2015)

The app uses 'compassionate' AI to encourage users to sleep, eat, and diet better. It tracks data, such as diet and physical activity and recommends very specific changes. It feels extremely customized in the way it 'chats' with you and personalizes advice. It certainly goes the extra mile in not only representing data, but recommending dynamic changes to the user.

4.3.5 8fit
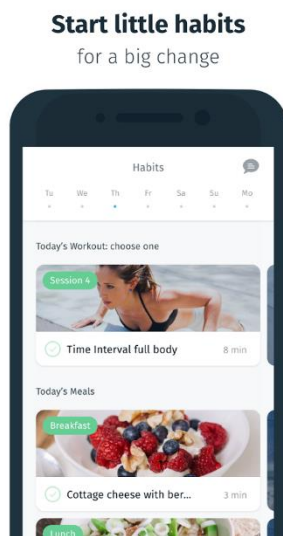
Figure 8: 8fit image - Habits
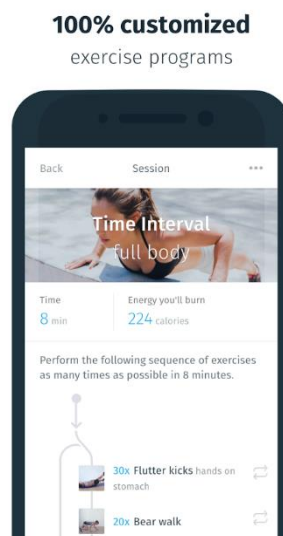
Figure 9: 8fit image - Custom
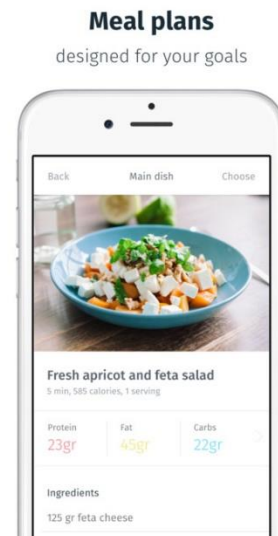
Figure 10: 8fit image - Meals

Data Measured: Physical attributes, goals.

Data Tracked: Weight, calories, workout difficulty/enjoyment.

Study: (Pan, 2016)

The app recommends very specific workouts and diets for you and your goals. Data is thorough and personalized. It takes in data such as your weight and goal (lose

weight / gain muscle), and constructs customized diets for your size and goal. Additionally, it constructs personalized workouts and amends them based on difficulty and enjoyment. This process of self-perfecting recommendations allows for a dynamic process of modifying user behavior. That being said, (unless you want the apple watch plugin), none of the data collected is visually displayed to the user. In short, it recommends specific solutions, but it is difficult to view how to get there.

4.3.6 Fooducate

Figure 11: Fooducate Food Rating   Figure 12: Fooducate Food Menu
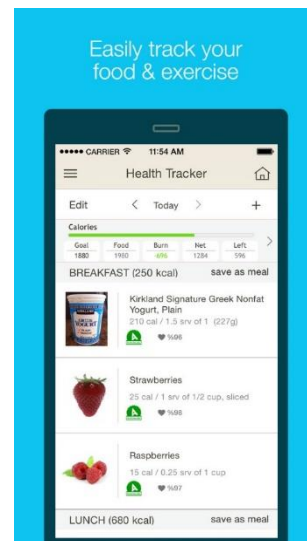
Data Measured: Calories, ideal calorie count.

Data Tracked: Foods eaten, calories consumed, and weight.

Study: (Clary, Matthews, & Kestens, 2017)

Other: Rates the health value of foods (by scanning barcode)

Overall, the app's primarily used for assessing the health of foods at supermarkets. It charts your weight over time so you can assess your own progress.

And while it assesses the potential benefits of foods you scan, it does not give you any specific recommendations (aside from randomized foods to try) to improve your behavior.

4.3.7 Lose It!

Figure 13: Lose It! In-app menus

Data Measured: Calories, weight, goal-loss.

Data Tracked: Nutrient pie charts, weight over time, exercise, projected goal arrival time.

Study:(Wharton, Johnston, Cunningham, & Sterner)

Overall, the app is reportedly very effective for losing weight. Not only does it provide you with charts of weight over time, but it also provides you with organized meals to lose weight, and a projected end time for your personal goal. However, while the charts do provide helpful information, they are generic and impersonalized. In short, while the graphics and data presented are helpful, they do not provide specific avenues for change aside from the user observing the graphs presented.

4 .4 Graphing Participation

As described in the Methodology section, we decided to graph the user study's participation data over time (Phase 1, between phases, Phase 2). Below are the graphical results of our analysis. We used the D3 javascript library to visualize the data.

Figure 14 displays the trend of Episodes over time, by which we mean recorded instances of overeating.
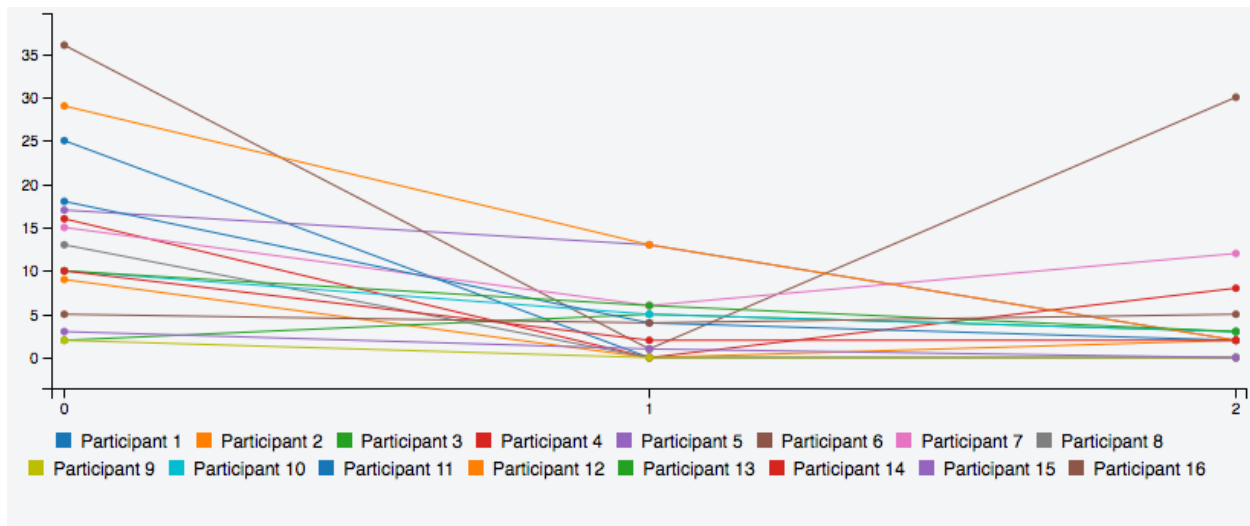
Figure 14: Graph of users' overeating episodes over time

Figure 15 presents the graph of Check-Ins over time. Check-Ins are when the user inputs data about their day into the app. Check-In frequency is a metric of user engagement.

4.5 Application History Visualization

During the summer of 2016, research students worked on developing appropriate visualization prototypes on the Android application to increase the effectiveness of using the SlipBuddy application. Four prototypes were developed, including the progress bar, overeating graph, weight over time, and the pie chart. In our project, we implemented three of them in the real application.

4.5.1 Progress Bar

It is important to keep track of users' overeating behavior. This can give users a sense of their progress over time. That is what the use of a progress bar is intended to achieve. During the summer research, three drafts of the progress bar were produced. The first prototype, shown in **Figure 16,** is a circular progress bar around the "Oops!" button, with ring decreasing as the episode number increases throughout the day. The color goes from green to yellow then to red, as the circle decreases. The second prototype, shown in **Figure 17,** is with the ring increasing as the number of episodes increases and the color goes from green to yellow then to red as this occurs. The third

35

prototype, shown in **Figure 18,** is with the ring decreasing as the number of episodes increases, but the color of the circle changes as it decreases: when it is a full circle, it means that no episode has happened yet; at that point, the circle's color is gold. When 1 to 3 episodes occur, the circles' color changes to silver. Lastly, if more than 3 episodes occur, the circle's color changes to bronze.

The third prototype comes with various limitations. The choice of three colors is based on the understanding of how we rank gold, silver and bronze. It might not be well received by people from different backgrounds. Also, the silver color is really close to the background grey color of the circle, which could lead to ambiguity. The first and second prototypes are generally the same, only with the difference of a cut off or a complete circle. We expect that most people would like to see their circle as "complete" when no episodes happen. This can encourage them to keep the circle as complete as possible. Also, the first prototype gives a clear indication of how many episodes are left within the quota. With the second, users may feel as though it resembles a task of completing 5 episodes, which is against the logic of the application as overeating is an undesired behavior.

The first prototype is used as the final implementation in the application for its intuitive design.
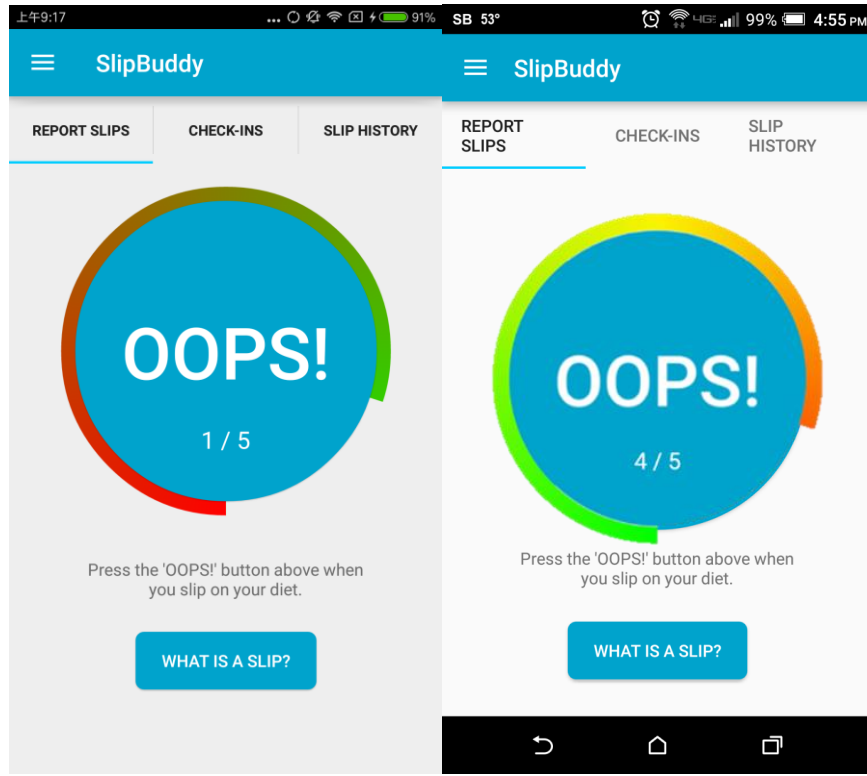
Figure 17: Slip Tracker first prototype

Figure 16: Slip Tracker second prototype
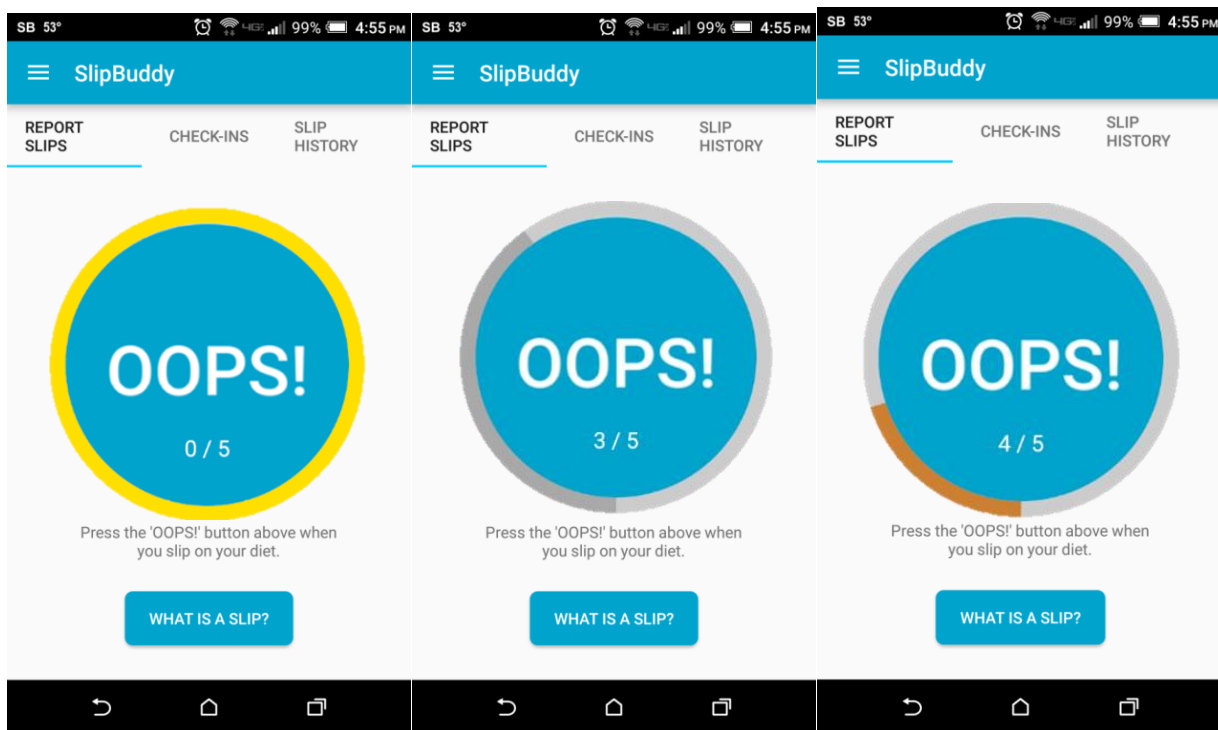
Figure 18: Slip Tracker third prototype

4.5.2 Overeating Graph

The overeating graph displays how many episodes are in each level for each data

point using a bar chart. There are a total of three data types available: Stress Level,
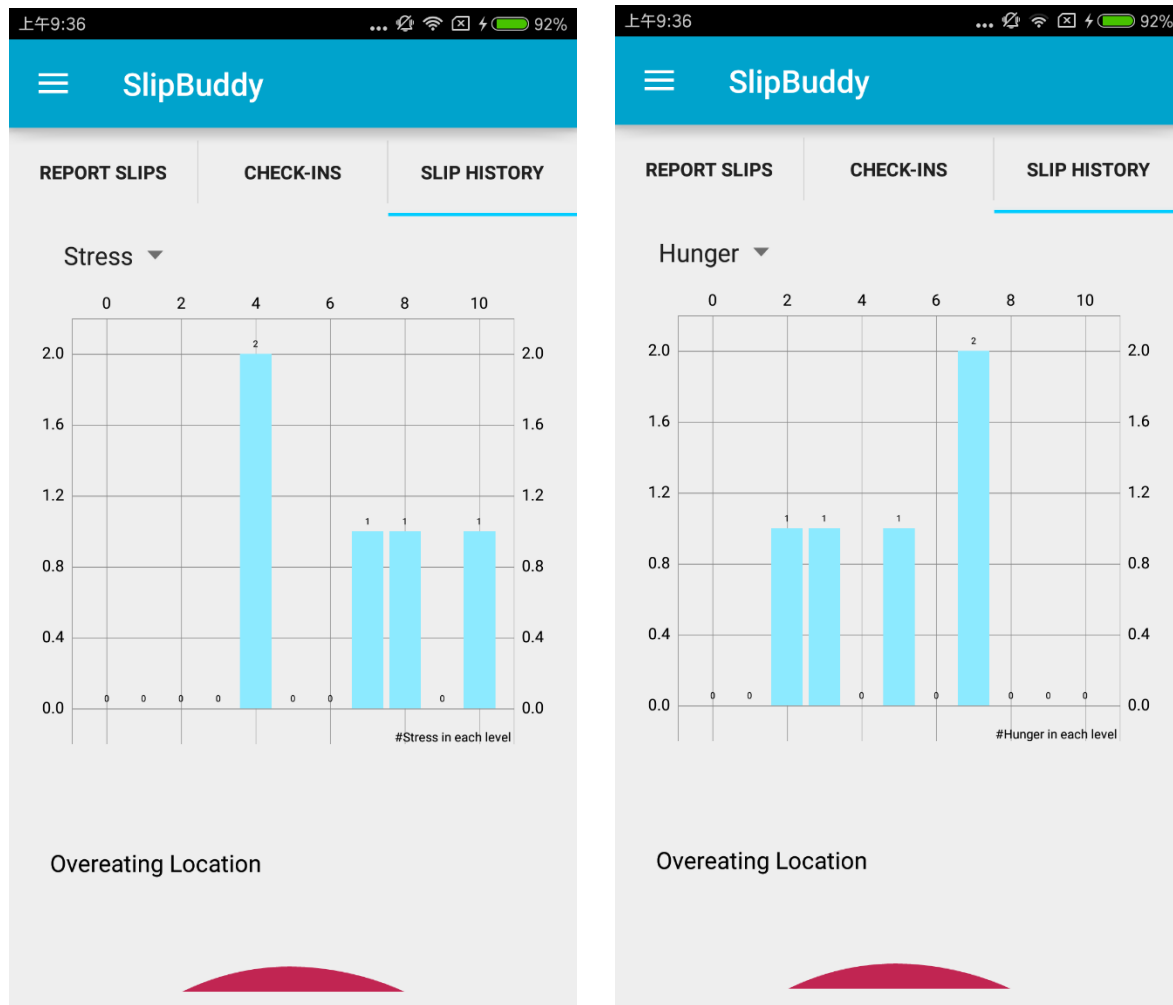
Figure 19: Data Charts for Hunger and Stress

Hunger Level and Overeating Level. The implementation used MP Android Chart

library, a chart library for android in order to build the bar chart. A spinner at the top

left of the page is the picker for selecting what data to show. The data gathered are from

the application SQLite database, using the Java activity to filter and organize the

number, then finally showing in the application.

38

4.5.3 Pie Charts

Two pie charts are implemented in the new application. One is the percentage pie chart, illustrated in **Figure 20** Right, for location where episode happened. The other is the percentage pie chart for Activity, illustrated in **Figure 20** Left, when episode happened. As with the Overeating Graph, both charts are implemented using MP Android Chart library. The data gathered are also from the application's SQLite database, followed by using a function to process and map the numbers to each arc of the pie.
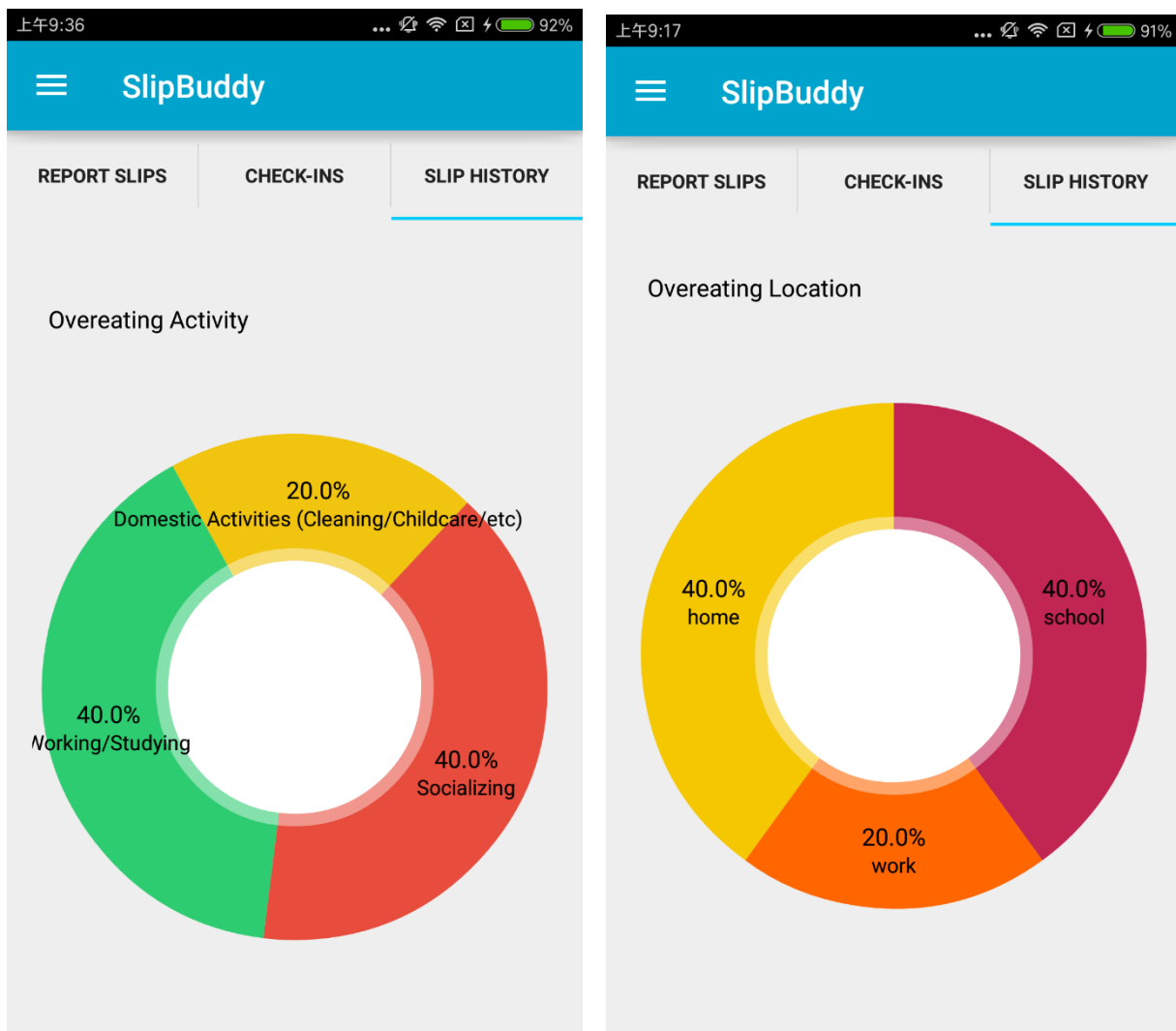
Figure 20: Overeating Location and Activity Chart

4.6 Analysis on Collected Data

During the 2016 user study with 16 participants, data were collected using the Android application as well as individual interviews. More specifically, three sets of data were collected. First is the episode data, which user reported when they overeat. This data contains user id, what activity they were doing when they overate, which meal it was, stress level, hunger level, location and when did it happen. Second is the check-in data, which user input during the morning, afternoon and evening check-ins every day to monitor their status. This data contains user id, check-in type, stress level, hunger level, rested level, sleep hour of that day, weight, as well as time and date. Third is the demographics data. This includes user id, gender, age, marital status, education level, ethnicity, work status, number of household as well as annual household income. In the past summer, a group of NSF-REU students worked on combining those data to get an overall dataset on which data mining techniques can be performed. They combined the dataset on a per-day basis, where each entry has all the check-in data as well as a Boolean attribute representing whether an overeating episode happened that day. Then decision tree method was performed to construct the attribute tree to predict future overeating outcome in a certain day.

However, the previous model was limited because it can only use the attributes in check-in data to predict whether a participant would overeat on a certain day. The attributes in episode data are not included in the analysis, while the episode data are the only set of data that is directly linked to overeating habits. We wanted to include the episode data in the data mining process. So we developed a new method that combines

40

all data of an episode basis, that is, every entry represents information associated with that overeating episode and the check-in information associate with the day that episode happened. The new dataset does not have the Boolean property for decision tree mining, as all entries in this dataset correspond to overeating episodes. We used association rules and frequent itemsets instead. This allowed us to explore the triggers and frequent attribute values associated with overeating in each participant and in all participants.

First we used the Python data analysis library, Pandas, to get the desired attributes from each of the data set, combined them into one dataset as well as convert interval data to nominal data. Then we used a Python data-mining package, Orange to use the standard Apriori algorithm to mine the frequent itemsets.

4.6.1 Analysis on the Overall Dataset

We mined frequent itemsets on the overall dataset combining the episode data, check-in data as well as the demographic data across all participants in the study, using minimum support of 40% and minimum confidence of 50%. **Table 4** presents the frequent single items for the overall dataset; inside parenthesis the support of the itemset is presented, that is the frequency with which this item appears in the dataset.

All the numeric attributes in the original dataset were translated into nominal attributes. We provide below the details of this translation.

Hunger Level: originally in a scale of 0-10.

0-3: very hungry
4-6: moderate full
7-10: very full

Stress Level: originally in a scale of 0-10.

       0-3: low

       4-6: medium

       7-10: high

Sleep Level: originally in a scale of 0-10.

       0-3: not rested

       4-6: somewhat rested

       7-10: well rested

Time: originally in Unix Timestamp, indicates exact time.

       0AM-4AM: midnight

       4AM-8AM: morning

       8AM-12AM: late morning

       12AM-4PM: afternoon

       4PM-8PM: late afternoon

       8PM-12PM: evening

Sleep Hour: originally in a scale of 0-24, which indicates how many hours of sleep the participant got that day. An average number, mean, is calculated for each participant.

       less than mean-1.5 hours: less than usual

       between mean-1.5 hours and mean+1.5 hours: usual

       more than mean+1.5 hours: more than usual

Weight Change: original value represents weight in pound. Weight is recorded daily and each day's weight is compared to the previously recorded weight of the same participant.

       Today's weight < previous weight reported: loss

       Today's weight = previous weight reported: same

       Today's weight > previous weight reported: gain

Table 4: One Level Frequent Itemsets for Overall Data across Participants

Frequent Items. Number inside (.) is the support value of the item.
(0.80) Ethnicity=Caucasian; non-Hispanic White
(0.72) AM Hunger=moderate full
(0.72) Marital Status=Married
(0.72) If employed how would you best classify your position=Professional managerial administrative
(0.66) EVE Time=evening
(0.65) sleep hour=usual
(0.61) AM Time=late morning
(0.56) stress=low
(0.56) PM Time=afternoon
(0.56) EVE Stress=low
(0.55) Work Status=Employed full-time
(0.55) Average estimated household annual income=Over $75 001
(0.54) AM Stress=low
(0.53) hunger=moderate full
(0.51) PM Stress=low
(0.51) weight change=same
(0.51) Gender=Female
(0.50) Education (Indicate highest achieved)=Bachelors degree (e.g. B.A. B.S.)
(0.49) AM Sleep=well rested
(0.49) Gender=Male
(0.46) Number of persons living in your current household=two point zero
(0.42) eating=SNACK

Some of the interesting factors in **Table 4** are as follows: people who are married, who work as professional managerial administrative, live in a household of 2, as well as have an annual household income of over $75001 tend to overeat. People usually are not stressed, are not hungry or uncomfortably full when they overeat. Most of the time, they overeat on snacks. **Table 5** presents all the item sets that contain stress=low:

Table 5: Frequent Itemsets that contain stress=low

(0.56) stress=low
(0.40) stress=low && AM Stress=low

43

(0.41) stress=low && AM Hunger=moderate full
(0.40) stress=low && PM Stress=low
(0.42) stress=low && EVE Stress=low
(0.45) stress=low && Marital Status=Married
(0.41) stress=low && Marital Status=Married && Ethnicity=Caucasian; non-Hispanic White
(0.49) stress=low && Ethnicity=Caucasian; non-Hispanic White
(0.42) stress=low && If employed how would you best classify your position=Professional managerial administrative

The frequent itemsets in Table 5 contain more than one item in the overall data set, and in this case we illustrate associations of low stress with other attributes. Low stress appears related to overall daily low stress, married, Caucasian and administrative people.

In general, there are clusters that clearly represent the associations. Most of the items are associated with married, Caucasian and professional managerial administrative. Table 6 shows our results after we excluded these two overly frequent factors in our data sample.

Table 6: Frequent Items for Overall Data across Participants

First item	The rest of the items in the itemset				
Low stress	AM/PM/EVE low stress				
AM low stress	PM low stress				
PM low stress	EVE low stress				
EVE low stress	administrative				
Moderate Hunger	AM moderate hunger				
AM moderate hunger	EVE low stress	male	usual sleep hour	full-time	household over $75000
Female	administrative				
Bachelor degrees	administrative				
Full-time	usual sleep hour				
Administrative	usual sleep hour				

4.6.2 Analysis on Individual Participants' Datasets

We also performed association rule mining on each individual participant'
dataset. The tables below are the key item sets found for some of the participants. The
minimum support was set to 0.4 for each run. We picked the most significant itemsets
from the result to include in the tables below. The full results are given in **Appendix A**.

Note that tables are not provided for all of the participants. For participants 3, 6
and 10, there are too many repetitions in the data entries so that most of the item sets
have a support of 0.9 or more. For participants 9 and 15, there are too few data entries to
generate item sets.

Table 7 illustrates that participant 1 tends to overeat when working/studying as
well as eating breakfast and not hungry or uncomfortably full at the time. This

participant also tends to overeat when stress is medium, when at home and has gained weight that day.

Table 7: Frequent Itemsets for Participant 1

(0.40) doing=WORKING_STUDYING	eating=BREAKFAST	hunger=moderate full
(0.40) doing=WORKING_STUDYING	stress=medium	hunger=moderate full
(0.44) eating=BREAKFAST	AM Hunger=moderate full	
(0.48) location=at home		
(0.52) weight change=gain		

Table 8 illustrates that participant 2 tends to overeat in days when in the morning the stress level is medium, not that well-rested but not hungry; or when in the afternoon stress level is medium.

Table 8: Frequent Itemsets for Participant 2

(0.73) AM Stress=medium	AM Hunger=moderate full	AM Sleep=somewhat rested	sleephour=usual
(0.91) PM Stress=medium			

Table 9 illustrates that participant 4 tends to overeat when not hungry or full, when the location is at home, or in the evening. It is clear that at home and in the evening are triggers for this participant.

46

Table 9: Frequent Itemsets for Participant 4

(0.61) hunger=moderate full	AM Hunger=moderate full	PM Stress=medium	EVE Stress=medium
(0.63) location=Home	AM Hunger=moderate full		
(0.71) Episode Time=evening			
(0.63) AM Stress=medium	PM Stress=medium		

Table 10 illustrates that participant 5 tends to overeat at home.

Table 10: Frequent Itemsets for Participant 5

(0.56) location=home	
(0.59) AM Stress=low	
(0.50) AM Sleep=well rested	EVE Stress=low

Table 11 illustrates that participant 7 tends to overeat when working/studying, eating snack, very hungry or in the late afternoon.

Table 11: Frequent Itemsets for Participant 7

(0.45) doing=WORKING_STUDYING	
(0.48) eating=SNACK	
(0.52) hunger=very hungry	

(0.48) weight change=same	Sleep hour=usual
(0.42) Episode Time=late afternoon	

Table 12 illustrates that participant 8 tends to overeat when eating snack or when stress is low throughout the day.

Table 12: Frequent Itemsets for Participant 8

(0.64) eating=SNACK	PM Stress=low	sleephour=usual	
(0.64) stress=low	AM Stress=low	PM Stress=low	EVE Stress=low

Table 13 illustrates that participant 11 tends to overeat when eating dinner, when feeling very hungry at home, when stress is high and very hungry, or when not well-rested.

Table 13: Frequent Itemsets for Participant 11

(0.50) eating=DINNER	
(0.42) hunger=very hungry	location=home
(0.42) AM Stress=high	AM Hunger=very hungry
(0.46) AM Sleep=not rested	

Table 14 illustrates that participant 12 tends to overeat when socializing, eating snack, at home or in the evening. The evening time and home location is related to eating snack.

Table 14: Frequent Itemsets for Participant 12

(0.43) doing=SOCIALIZING	
(0.59) eating=SNACK	Episode Time=evening
(0.43) eating=SNACK	location=Home
(0.46) eating=SNACK	weightchange=same
(0.59) location=Home	
(0.74) Episode Time=evening	

Table 15 illustrates that participant 14 tends to overeat when watching TV, very hungry, at River Street, when the stress is high or in the late afternoon.

Table 15: Frequent Itemsets for Participant 14

(0.43) doing=SCREEN_TIME			
(0.43) stress=medium			
(0.43) hunger=very hungry	Episode Time=late afternoon		
(0.43) location=River Street	AM Hunger=moderate full	AM Sleep=well rested	Episode Time=late afternoon
(0.43) PM Stress=high			
(0.57) Episode Time=late afternoon			

49

5. Conclusion

This MQP enhanced some of the functionality of the SlipBuddy app developed by a previous MQP. The app was missing features such as 'Settings', which makes the app more welcoming and customizable. Also graphical representations were added to the app to provide a more intuitive way of conveying information and feedback to the users. Some aspects of the app could still manage to be refined. Visualizations of data over time would help the app display more complex information in a simple manner.

This MQP also utilized association rules to analyze what triggers overeating behavior. Python was used to mine frequent itemsets for the overall dataset across all participants and each participant's dataset. This method could still be refined by a better implementation to filter out the unrelated triggers. Also, the individual results could be combined with demographics information for further analysis.

Obesity is a prominent problem worldwide and will continue to be for some time. However, through learning, improving, and studying the proven techniques of behavior modification, health apps like SlipBuddy can help take obesity a step back.

References:

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *ACM SIGMOD Record, 22*(2), 207-216. doi:10.1145/170036.170072

Agrawal, R., & Srikant, R. (1994). *Fast algorithms for mining association rules.* Paper presented at the Proc. 20th int. conf. very large data bases, VLDB.

American Obesity Treatment Association. (2011). Behavioral habits influence obesity. Retrieved from http://www.americanobesity.org/behavior.htm

Center for Disease Control and Prevention. (2016). Adult Obesity Facts.   Retrieved from https://www.cdc.gov/obesity/data/adult.html

Clary, C., Matthews, S. A., & Kestens, Y. (2017). Between exposure, access and use: Reconsidering foodscape influences on dietary behaviours. *Health & Place, 44*, 1-7. doi:http://doi.org/10.1016/j.healthplace.2016.12.005

de Zambotti, M., Baker, F. C., & Colrain, I. M. (2015). Validation of Sleep-Tracking Technology Compared with Polysomnography in Adolescents. *Sleep, 38*(9), 1461-1468. doi:10.5665/sleep.4990

Direito, A., Pfaeffli Dale, L., Shields, E., Dobson, R., Whittaker, R., & Maddison, R. (2014). Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? *BMC Public Health, 14*(1), 646. doi:10.1186/1471-2458-14-646

Ipjian, M. L., & Johnston, C. S. Smartphone technology facilitates dietary change in healthy adults. *Nutrition, 33*, 343-347. doi:10.1016/j.nut.2016.08.003

Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature biotechnology, 26*(9), 1011-1013. doi:10.1038/nbt0908-1011

McCormick, J., Ask, J. A., Pilecki, M., & Ngo, S. (2015). *Case Study: Win Customers With Mobile Analytics.* Retrieved from Forrester: https://www.forrester.com/report/Case+Study+Win+Customers+With+Mobile+Analytics/-/E-RES122203

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science, 6*(1), 42. doi:10.1186/1748-5908-6-42

Mummah, S. A., King, A. C., Gardner, C. D., & Sutton, S. (2016). Iterative development of Vegethon: a theory-based mobile app intervention to increase vegetable consumption. *International Journal of Behavioral Nutrition and Physical Activity, 13*(1), 90. doi:10.1186/s12966-016-0400-z

National Heart, Lung, and Blood Institute. (2012). Overweight and Obesity.   Retrieved from https://www.nhlbi.nih.gov/health/health-topics/topics/obe

Pan, E. (2016). *How 8fit Used Outbrain to Increase App Downloads and Grow Their Mobile User Base by 80% with Relevant Content* (Vol. 2017). Outbrain.

Stewart, J. (2006a). A Conversation with Marion Nestle: Straight Talk About Obesity, Nutrition, and Food Policy. *The Permanente Journal, 10*(3), 62-65.

Stewart, J. (2006b). A Conversation with Marion Nestle: Straight Talk About Obesity, Nutrition, and Food Policy. *The Permanente Journal*.

Wharton, C. M., Johnston, C. S., Cunningham, B. K., & Sterner, D. Dietary Self-Monitoring, But Not Dietary Quality, Improves With Use of Smartphone App Technology in an 8-Week Weight Loss Trial. *Journal of Nutrition Education and Behavior, 46*(5), 440-444. doi:10.1016/j.jneb.2014.04.291

Appendix A: Full Association Rule Mining Results

Below is the full list of frequent itemsets obtained. The minimum support used was 0.4. The support of each itemset is provided within parentheses. AM, PM, and EVE Time denote the time when the participants recorded their morning, afternoon, and evening check-ins, respectively.

Frequent Itemsets obtained from the overall dataset consisting of all participants:
(0.42) eating=SNACK
(0.56) stress=low
(0.40) stress=low && AM Stress=low
(0.41) stress=low && AM Hunger=moderate full
(0.40) stress=low && PM Stress=low
(0.42) stress=low && EVE Stress=low
(0.45) stress=low && Marital Status=Married
(0.41) stress=low && Marital Status=Married && Ethnicity=Caucasian; non-Hispanic White
(0.49) stress=low && Ethnicity=Caucasian; non-Hispanic White
(0.42) stress=low && If employed how would you best classify your position=Professional managerial administrative
(0.53) hunger=moderate full
(0.41) hunger=moderate full && AM Hunger=moderate full
(0.43) hunger=moderate full && Marital Status=Married
(0.41) hunger=moderate full && Ethnicity=Caucasian; non-Hispanic White
(0.54) AM Stress=low
(0.41) AM Stress=low && PM Stress=low
(0.45) AM Stress=low && Marital Status=Married
(0.43) AM Stress=low && Marital Status=Married && Ethnicity=Caucasian; non-Hispanic White
(0.50) AM Stress=low && Ethnicity=Caucasian; non-Hispanic White
(0.40) AM Stress=low && If employed how would you best classify your position=Professional managerial administrative
(0.72) AM Hunger=moderate full
(0.45) AM Hunger=moderate full && AM Time=late morning
(0.43) AM Hunger=moderate full && PM Time=afternoon
(0.42) AM Hunger=moderate full && EVE Stress=low
(0.50) AM Hunger=moderate full && EVE Time=evening
(0.41) AM Hunger=moderate full && EVE Time=evening && Ethnicity=Caucasian; non-Hispanic White
(0.42) AM Hunger=moderate full && Gender=Male

(0.55) AM Hunger=moderate full && Marital Status=Married
(0.46) AM Hunger=moderate full && Marital Status=Married && Ethnicity=Caucasian; non-Hispanic White
(0.42) AM Hunger=moderate full && Marital Status=Married && If employed how would you best classify your position=Professional managerial administrative
(0.58) AM Hunger=moderate full && Ethnicity=Caucasian; non-Hispanic White
(0.42) AM Hunger=moderate full && Ethnicity=Caucasian; non-Hispanic White && If employed how would you best classify your position=Professional managerial administrative
(0.42) AM Hunger=moderate full && Work Status=Employed full-time
(0.53) AM Hunger=moderate full && If employed how would you best classify your position=Professional managerial administrative
(0.46) AM Hunger=moderate full && Average estimated household annual income=Over $75 001
(0.50) AM Hunger=moderate full && sleephour=usual
(0.49) AM Sleep=well rested
(0.43) AM Sleep=well rested && Ethnicity=Caucasian; non-Hispanic White
(0.61) AM Time=late morning
(0.42) AM Time=late morning && EVE Stress=low
(0.45) AM Time=late morning && EVE Time=evening
(0.40) AM Time=late morning && Marital Status=Married
(0.51) AM Time=late morning && Ethnicity=Caucasian; non-Hispanic White
(0.44) AM Time=late morning && If employed how would you best classify your position=Professional managerial administrative
(0.41) AM Time=late morning && sleephour=usual
(0.51) PM Stress=low
(0.40) PM Stress=low && EVE Stress=low
(0.41) PM Stress=low && Marital Status=Married
(0.40) PM Stress=low && Marital Status=Married && Ethnicity=Caucasian; non-Hispanic White
(0.48) PM Stress=low && Ethnicity=Caucasian; non-Hispanic White
(0.56) PM Time=afternoon
(0.43) PM Time=afternoon && EVE Time=evening
(0.48) PM Time=afternoon && Ethnicity=Caucasian; non-Hispanic White
(0.41) PM Time=afternoon && If employed how would you best classify your position=Professional managerial administrative
(0.56) EVE Stress=low
(0.45) EVE Stress=low && EVE Time=evening
(0.41) EVE Stress=low && EVE Time=evening && Ethnicity=Caucasian; non-Hispanic White
(0.44) EVE Stress=low && Marital Status=Married
(0.43) EVE Stress=low && Marital Status=Married && Ethnicity=Caucasian; non-Hispanic White

(0.51) EVE Stress=low && Ethnicity=Caucasian; non-Hispanic White

(0.41) EVE Stress=low && If employed how would you best classify your position=Professional managerial administrative

(0.66) EVE Time=evening

(0.44) EVE Time=evening && Marital Status=Married

(0.53) EVE Time=evening && Ethnicity=Caucasian; non-Hispanic White

(0.43) EVE Time=evening && Work Status=Employed full-time

(0.42) EVE Time=evening && Work Status=Employed full-time && If employed how would you best classify your position=Professional managerial administrative

(0.48) EVE Time=evening && If employed how would you best classify your position=Professional managerial administrative

(0.46) EVE Time=evening && sleephour=usual

(0.51) weightchange=same

(0.51) Gender=Female

(0.45) Gender=Female && Ethnicity=Caucasian; non-Hispanic White

(0.44) Gender=Female && If employed how would you best classify your position=Professional managerial administrative

(0.49) Gender=Male

(0.41) Gender=Male && Marital Status=Married

(0.72) Marital Status=Married

(0.61) Marital Status=Married && Ethnicity=Caucasian; non-Hispanic White

(0.41) Marital Status=Married && Ethnicity=Caucasian; non-Hispanic White && If employed how would you best classify your position=Professional managerial administrative

(0.51) Marital Status=Married && If employed how would you best classify your position=Professional managerial administrative

(0.44) Marital Status=Married && If employed how would you best classify your position=Professional managerial administrative && Average estimated household annual income=Over $75 001

(0.42) Marital Status=Married && Number of persons living in your current household=two point zero

(0.47) Marital Status=Married && Average estimated household annual income=Over $75 001

(0.47) Marital Status=Married && sleephour=usual

(0.50) Education (Indicate highest achieved)=Bachelors degree (e.g. B.A. B.S.)

(0.43) Education (Indicate highest achieved)=Bachelors degree (e.g. B.A. B.S.) && Ethnicity=Caucasian; non-Hispanic White

(0.43) Education (Indicate highest achieved)=Bachelors degree (e.g. B.A. B.S.) && Ethnicity=Caucasian; non-Hispanic White && If employed how would you best classify your position=Professional managerial administrative

(0.50) Education (Indicate highest achieved)=Bachelors degree (e.g. B.A. B.S.) && If employed how would you best classify your position=Professional managerial administrative

(0.80) Ethnicity=Caucasian; non-Hispanic White
(0.56) Ethnicity=Caucasian; non-Hispanic White && If employed how would you best classify your position=Professional managerial administrative
(0.42) Ethnicity=Caucasian; non-Hispanic White && Average estimated household annual income=Over $75 001
(0.52) Ethnicity=Caucasian; non-Hispanic White && sleephour=usual
(0.55) Work Status=Employed full-time
(0.52) Work Status=Employed full-time && If employed how would you best classify your position=Professional managerial administrative
(0.41) Work Status=Employed full-time && sleephour=usual
(0.72) If employed how would you best classify your position=Professional managerial administrative
(0.49) If employed how would you best classify your position=Professional managerial administrative && Average estimated household annual income=Over $75 001
(0.47) If employed how would you best classify your position=Professional managerial administrative && sleephour=usual
(0.46) Number of persons living in your current household=two point zero
(0.55) Average estimated household annual income=Over $75 001
(0.65) sleephour=usual

Frequent Itemsets obtained from Participant 1's dataset:
(0.56) doing=WORKING_STUDYING
(0.40) doing=WORKING_STUDYING && eating=BREAKFAST
(0.40) doing=WORKING_STUDYING && eating=BREAKFAST && hunger=moderate full
(0.40) doing=WORKING_STUDYING && stress=medium
(0.40) doing=WORKING_STUDYING && stress=medium && hunger=moderate full
(0.44) doing=WORKING_STUDYING && hunger=moderate full
(0.40) doing=WORKING_STUDYING && AM Stress=low
(0.40) doing=WORKING_STUDYING && AM Hunger=moderate full
(0.44) doing=WORKING_STUDYING && AM Time=late morning
(0.60) eating=BREAKFAST
(0.40) eating=BREAKFAST && stress=medium
(0.48) eating=BREAKFAST && hunger=moderate full
(0.40) eating=BREAKFAST && AM Stress=low
(0.44) eating=BREAKFAST && AM Hunger=moderate full
(0.44) eating=BREAKFAST && AM Time=late morning
(0.52) stress=medium
(0.40) stress=medium && hunger=moderate full
(0.40) stress=medium && AM Stress=low
(0.40) stress=medium && AM Hunger=moderate full
(0.40) stress=medium && AM Time=late morning
(0.52) hunger=moderate full

(0.40) hunger=moderate full && AM Hunger=moderate full
(0.40) hunger=moderate full && AM Time=late morning
(0.48) location=at home
(0.68) AM Stress=low
(0.40) AM Stress=low && AM Hunger=moderate full
(0.48) AM Stress=low && AM Sleep=well rested
(0.40) AM Stress=low && AM Sleep=well rested && AM Time=late morning
(0.52) AM Stress=low && AM Time=late morning
(0.40) AM Stress=low && AM Time=late morning && PM Stress=low
(0.48) AM Stress=low && PM Stress=low
(0.64) AM Hunger=moderate full
(0.44) AM Hunger=moderate full && AM Time=late morning
(0.40) AM Hunger=moderate full && PM Time=afternoon
(0.40) AM Hunger=moderate full && PM Time=afternoon && weightchange=gain
(0.40) AM Hunger=moderate full && weightchange=gain
(0.48) AM Sleep=well rested
(0.40) AM Sleep=well rested && AM Time=late morning
(0.64) AM Time=late morning
(0.48) AM Time=late morning && PM Stress=low
(0.40) AM Time=late morning && PM Stress=low && weightchange=gain
(0.40) AM Time=late morning && weightchange=gain
(0.56) PM Stress=low
(0.40) PM Stress=low && weightchange=gain
(0.52) PM Time=afternoon
(0.44) PM Time=afternoon && weightchange=gain
(0.40) EVE Stress=low
(0.52) weightchange=gain
(0.52) sleephour=usual

Frequent Itemsets obtained from Participant 2's dataset:
(0.73) hunger=moderate full
(0.73) hunger=moderate full && AM Hunger=moderate full
(0.73) hunger=moderate full && AM Hunger=moderate full && PM Stress=medium
(0.73) hunger=moderate full && PM Stress=medium
(0.91) AM Stress=medium
(0.91) AM Stress=medium && AM Hunger=moderate full
(0.73) AM Stress=medium && AM Hunger=moderate full && AM Sleep=somewhat rested
(0.73) AM Stress=medium && AM Hunger=moderate full && AM Sleep=somewhat rested && sleephour=usual
(0.82) AM Stress=medium && AM Hunger=moderate full && PM Stress=medium
(0.73) AM Stress=medium && AM Hunger=moderate full && sleephour=usual
(0.73) AM Stress=medium && AM Sleep=somewhat rested

57

(0.73) AM Stress=medium && AM Sleep=somewhat rested && sleephour=usual

(0.82) AM Stress=medium && PM Stress=medium

(0.73) AM Stress=medium && sleephour=usual

(1.00) AM Hunger=moderate full

(0.82) AM Hunger=moderate full && AM Sleep=somewhat rested

(0.73) AM Hunger=moderate full && AM Sleep=somewhat rested && PM Stress=medium

(0.73) AM Hunger=moderate full && AM Sleep=somewhat rested && PM Stress=medium && sleephour=usual

(0.82) AM Hunger=moderate full && AM Sleep=somewhat rested && sleephour=usual

(0.91) AM Hunger=moderate full && PM Stress=medium

(0.73) AM Hunger=moderate full && PM Stress=medium && sleephour=usual

(0.82) AM Hunger=moderate full && sleephour=usual

(0.82) AM Sleep=somewhat rested

(0.73) AM Sleep=somewhat rested && PM Stress=medium

(0.73) AM Sleep=somewhat rested && PM Stress=medium && sleephour=usual

(0.82) AM Sleep=somewhat rested && sleephour=usual

(0.91) PM Stress=medium

(0.73) PM Stress=medium && sleephour=usual

(0.82) sleephour=usual

Frequent Itemsets obtained from Participant 4's dataset:

(0.74) hunger=moderate full

(0.61) hunger=moderate full && AM Stress=medium

(0.68) hunger=moderate full && AM Hunger=moderate full

(0.63) hunger=moderate full && AM Hunger=moderate full && PM Stress=medium

(0.61) hunger=moderate full && AM Hunger=moderate full && PM Stress=medium && EVE Stress=medium

(0.63) hunger=moderate full && AM Hunger=moderate full && EVE Stress=medium

(0.61) hunger=moderate full && AM Hunger=moderate full && sleephour=usual

(0.66) hunger=moderate full && PM Stress=medium

(0.63) hunger=moderate full && PM Stress=medium && EVE Stress=medium

(0.68) hunger=moderate full && EVE Stress=medium

(0.61) hunger=moderate full && sleephour=usual

(0.68) location=Home

(0.63) location=Home && AM Hunger=moderate full

(0.71) AM Stress=medium

(0.68) AM Stress=medium && AM Hunger=moderate full

(0.61) AM Stress=medium && AM Hunger=moderate full && PM Stress=medium

(0.63) AM Stress=medium && AM Hunger=moderate full && sleephour=usual

(0.63) AM Stress=medium && PM Stress=medium

(0.63) AM Stress=medium && sleephour=usual

(0.95) AM Hunger=moderate full

(0.63) AM Hunger=moderate full && AM Sleep=somewhat rested
(0.84) AM Hunger=moderate full && PM Stress=medium
(0.68) AM Hunger=moderate full && PM Stress=medium && EVE Stress=medium
(0.63) AM Hunger=moderate full && PM Stress=medium && Episdoe Time=evening
(0.68) AM Hunger=moderate full && PM Stress=medium && sleephour=usual
(0.71) AM Hunger=moderate full && EVE Stress=medium
(0.61) AM Hunger=moderate full && EVE Stress=medium && sleephour=usual
(0.61) AM Hunger=moderate full && weightchange=same
(0.71) AM Hunger=moderate full && Episdoe Time=evening
(0.79) AM Hunger=moderate full && sleephour=usual
(0.66) AM Sleep=somewhat rested
(0.61) AM Sleep=somewhat rested && PM Stress=medium
(0.87) PM Stress=medium
(0.71) PM Stress=medium && EVE Stress=medium
(0.63) PM Stress=medium && Episdoe Time=evening
(0.68) PM Stress=medium && sleephour=usual
(0.76) EVE Stress=medium
(0.61) EVE Stress=medium && sleephour=usual
(0.61) weightchange=same
(0.71) Episdoe Time=evening
(0.79) sleephour=usual

Frequent Itemsets obtained from Participant 5's dataset:
(0.50) stress=low
(0.62) hunger=moderate full
(0.56) hunger=moderate full && AM Hunger=moderate full
(0.50) hunger=moderate full && AM Hunger=moderate full && EVE Stress=low
(0.53) hunger=moderate full && AM Hunger=moderate full && EVE Time=evening
(0.53) hunger=moderate full && PM Time=afternoon
(0.50) hunger=moderate full && PM Time=afternoon && EVE Stress=low
(0.50) hunger=moderate full && PM Time=afternoon && EVE Time=evening
(0.56) hunger=moderate full && EVE Stress=low
(0.53) hunger=moderate full && EVE Stress=low && EVE Time=evening
(0.59) hunger=moderate full && EVE Time=evening
(0.56) location=home
(0.53) location=home && AM Hunger=moderate full
(0.50) location=home && AM Hunger=moderate full && EVE Time=evening
(0.53) location=home && EVE Time=evening
(0.59) AM Stress=low
(0.50) AM Stress=low && AM Hunger=moderate full
(0.50) AM Stress=low && EVE Time=evening
(0.91) AM Hunger=moderate full
(0.59) AM Hunger=moderate full && AM Sleep=well rested

(0.50) AM Hunger=moderate full && AM Sleep=well rested && EVE Stress=low
(0.50) AM Hunger=moderate full && AM Sleep=well rested && EVE Time=evening
(0.72) AM Hunger=moderate full && AM Time=late morning
(0.56) AM Hunger=moderate full && AM Time=late morning && EVE Stress=low
(0.53) AM Hunger=moderate full && AM Time=late morning && EVE Stress=low && EVE Time=evening
(0.62) AM Hunger=moderate full && AM Time=late morning && EVE Time=evening
(0.56) AM Hunger=moderate full && PM Time=afternoon
(0.53) AM Hunger=moderate full && PM Time=afternoon && EVE Stress=low
(0.50) AM Hunger=moderate full && PM Time=afternoon && EVE Stress=low && EVE Time=evening
(0.53) AM Hunger=moderate full && PM Time=afternoon && EVE Time=evening
(0.69) AM Hunger=moderate full && EVE Stress=low
(0.66) AM Hunger=moderate full && EVE Stress=low && EVE Time=evening
(0.78) AM Hunger=moderate full && EVE Time=evening
(0.59) AM Sleep=well rested
(0.50) AM Sleep=well rested && EVE Stress=low
(0.50) AM Sleep=well rested && EVE Time=evening
(0.72) AM Time=late morning
(0.56) AM Time=late morning && EVE Stress=low
(0.53) AM Time=late morning && EVE Stress=low && EVE Time=evening
(0.62) AM Time=late morning && EVE Time=evening
(0.66) PM Time=afternoon
(0.62) PM Time=afternoon && EVE Stress=low
(0.59) PM Time=afternoon && EVE Stress=low && EVE Time=evening
(0.62) PM Time=afternoon && EVE Time=evening
(0.78) EVE Stress=low
(0.75) EVE Stress=low && EVE Time=evening
(0.88) EVE Time=evening

Frequent Itemsets obtained from Participant 7's dataset:
(0.45) doing=WORKING_STUDYING
(0.48) eating=SNACK
(0.42) hunger=moderate full
(0.52) hunger=very hungry
(0.48) AM Stress=low
(0.67) AM Hunger=moderate full
(0.52) AM Hunger=moderate full && AM Time=late morning
(0.45) AM Hunger=moderate full && AM Time=late morning && sleephour=usual
(0.48) AM Hunger=moderate full && PM Time=afternoon
(0.52) AM Hunger=moderate full && EVE Time=evening
(0.52) AM Hunger=moderate full && sleephour=usual
(0.48) AM Sleep=well rested

(0.64) AM Time=late morning
(0.48) AM Time=late morning && EVE Time=evening
(0.55) AM Time=late morning && sleephour=usual
(0.48) PM Stress=low
(0.61) PM Time=afternoon
(0.48) PM Time=afternoon && EVE Time=evening
(0.48) PM Time=afternoon && sleephour=usual
(0.73) EVE Time=evening
(0.42) EVE Time=evening && weightchange=same
(0.55) EVE Time=evening && sleephour=usual
(0.58) weightchange=same
(0.48) weightchange=same && sleephour=usual
(0.42) Episdoe Time=late afternoon
(0.76) sleephour=usual

Frequent Itemsets obtained from Participant 8's dataset:
(0.79) eating=SNACK
(0.64) eating=SNACK && PM Stress=low
(0.64) eating=SNACK && PM Stress=low && sleephour=usual
(0.71) eating=SNACK && sleephour=usual
(0.71) stress=low
(0.64) stress=low && AM Stress=low
(0.64) stress=low && AM Stress=low && PM Stress=low
(0.64) stress=low && AM Stress=low && PM Stress=low && EVE Stress=low
(0.64) stress=low && AM Stress=low && EVE Stress=low
(0.71) stress=low && PM Stress=low
(0.71) stress=low && PM Stress=low && EVE Stress=low
(0.64) stress=low && PM Stress=low && EVE Stress=low && sleephour=usual
(0.64) stress=low && PM Stress=low && sleephour=usual
(0.71) stress=low && EVE Stress=low
(0.64) stress=low && EVE Stress=low && sleephour=usual
(0.64) stress=low && sleephour=usual
(0.71) AM Stress=low
(0.64) AM Stress=low && PM Stress=low
(0.64) AM Stress=low && PM Stress=low && EVE Stress=low
(0.64) AM Stress=low && EVE Stress=low
(0.71) AM Hunger=moderate full
(0.64) AM Hunger=moderate full && PM Stress=low
(0.64) AM Hunger=moderate full && sleephour=usual
(0.86) PM Stress=low
(0.79) PM Stress=low && EVE Stress=low
(0.71) PM Stress=low && EVE Stress=low && sleephour=usual
(0.79) PM Stress=low && sleephour=usual

(0.64) PM Time=afternoon
(0.79) EVE Stress=low
(0.71) EVE Stress=low && sleephour=usual
(0.86) sleephour=usual

Frequent Itemsets obtained from Participant 10's dataset:
(0.50) eating=DINNER
(0.50) stress=low
(0.42) stress=low && AM Time=late morning
(0.42) stress=low && EVE Stress=low
(0.62) hunger=very hungry
(0.42) hunger=very hungry && location=home
(0.42) hunger=very hungry && location=home && AM Time=late morning
(0.46) hunger=very hungry && AM Hunger=very hungry
(0.42) hunger=very hungry && AM Hunger=very hungry && AM Time=late morning
(0.54) hunger=very hungry && AM Time=late morning
(0.42) hunger=very hungry && AM Time=late morning && EVE Time=evening
(0.42) hunger=very hungry && EVE Time=evening
(0.58) location=home
(0.46) location=home && AM Time=late morning
(0.42) location=home && EVE Time=evening
(0.46) AM Stress=high
(0.42) AM Stress=high && AM Hunger=very hungry
(0.71) AM Hunger=very hungry
(0.58) AM Hunger=very hungry && AM Time=late morning
(0.46) AM Hunger=very hungry && AM Time=late morning && EVE Time=evening
(0.42) AM Hunger=very hungry && EVE Stress=low
(0.50) AM Hunger=very hungry && EVE Time=evening
(0.46) AM Sleep=not rested
(0.75) AM Time=late morning
(0.50) AM Time=late morning && EVE Stress=low
(0.42) AM Time=late morning && EVE Stress=low && EVE Time=evening
(0.58) AM Time=late morning && EVE Time=evening
(0.46) PM Time=afternoon
(0.62) EVE Stress=low
(0.50) EVE Stress=low && EVE Time=evening
(0.71) EVE Time=evening

Frequent Itemsets obtained from Participant 11's dataset:
(0.43) doing=SOCIALIZING
(0.70) eating=SNACK
(0.41) eating=SNACK && stress=low
(0.52) eating=SNACK && hunger=moderate full

(0.43) eating=SNACK && hunger=moderate full && Episdoe Time=evening
(0.43) eating=SNACK && location=Home
(0.52) eating=SNACK && AM Hunger=moderate full
(0.46) eating=SNACK && AM Hunger=moderate full && Episdoe Time=evening
(0.46) eating=SNACK && weightchange=same
(0.41) eating=SNACK && weightchange=same && Episdoe Time=evening
(0.59) eating=SNACK && Episdoe Time=evening
(0.54) stress=low
(0.43) stress=low && AM Hunger=moderate full
(0.41) stress=low && Episdoe Time=evening
(0.74) hunger=moderate full
(0.43) hunger=moderate full && location=Home
(0.48) hunger=moderate full && AM Hunger=moderate full
(0.46) hunger=moderate full && weightchange=same
(0.54) hunger=moderate full && Episdoe Time=evening
(0.59) location=Home
(0.41) location=Home && weightchange=same
(0.41) location=Home && Episdoe Time=evening
(0.41) AM Stress=low
(0.67) AM Hunger=moderate full
(0.43) AM Hunger=moderate full && EVE Stress=low
(0.54) AM Hunger=moderate full && Episdoe Time=evening
(0.43) AM Sleep=well rested
(0.50) EVE Stress=low
(0.41) EVE Stress=low && Episdoe Time=evening
(0.65) weightchange=same
(0.52) weightchange=same && Episdoe Time=evening
(0.74) Episdoe Time=evening

Frequent Itemsets obtained from Participant 14's dataset:
(0.43) doing=SCREEN_TIME
(0.43) stress=medium
(0.43) stress=medium && AM Sleep=well rested
(0.50) hunger=very hungry
(0.43) hunger=very hungry && Episdoe Time=late afternoon
(0.43) location=River Street
(0.43) location=River Street && AM Hunger=moderate full
(0.43) location=River Street && AM Hunger=moderate full && AM Sleep=well rested
(0.43) location=River Street && AM Hunger=moderate full && AM Sleep=well rested && Episdoe Time=late afternoon
(0.43) location=River Street && AM Hunger=moderate full && Episdoe Time=late afternoon
(0.43) location=River Street && AM Sleep=well rested

(0.43) location=River Street && AM Sleep=well rested && Episdoe Time=late afternoon
(0.43) location=River Street && Episdoe Time=late afternoon
(0.57) AM Stress=medium
(0.57) AM Stress=medium && AM Sleep=well rested
(0.50) AM Stress=medium && AM Sleep=well rested && PM Time=afternoon
(0.50) AM Stress=medium && PM Time=afternoon
(0.43) AM Hunger=moderate full
(0.43) AM Hunger=moderate full && AM Sleep=well rested
(0.43) AM Hunger=moderate full && AM Sleep=well rested && Episdoe Time=late afternoon
(0.43) AM Hunger=moderate full && Episdoe Time=late afternoon
(0.79) AM Sleep=well rested
(0.43) AM Sleep=well rested && AM Time=late morning
(0.57) AM Sleep=well rested && PM Time=afternoon
(0.50) AM Sleep=well rested && EVE Time=evening
(0.50) AM Sleep=well rested && Episdoe Time=late afternoon
(0.50) AM Sleep=well rested && sleephour=usual
(0.64) AM Time=late morning
(0.43) AM Time=late morning && PM Time=afternoon
(0.43) AM Time=late morning && EVE Time=evening
(0.43) AM Time=late morning && sleephour=usual
(0.43) PM Stress=high
(0.43) PM Stress=high && sleephour=usual
(0.64) PM Time=afternoon
(0.64) EVE Time=evening
(0.43) EVE Time=evening && sleephour=usual
(0.57) Episdoe Time=late afternoon
(0.71) sleephour=usual

Frequent Itemsets obtained from Participant 16's dataset:
(0.71) stress=medium
(0.79) AM Hunger=moderate full
(0.71) AM Hunger=moderate full && AM Sleep=somewhat rested
(0.71) AM Hunger=moderate full && AM Sleep=somewhat rested && AM Time=late morning
(0.71) AM Hunger=moderate full && AM Sleep=somewhat rested && AM Time=late morning && EVE Time=evening
(0.71) AM Hunger=moderate full && AM Sleep=somewhat rested && EVE Time=evening
(0.71) AM Hunger=moderate full && AM Time=late morning
(0.71) AM Hunger=moderate full && AM Time=late morning && EVE Time=evening
(0.79) AM Hunger=moderate full && EVE Time=evening
(0.71) AM Hunger=moderate full && EVE Time=evening && weightchange=same

64

(0.71) AM Hunger=moderate full && weightchange=same
(0.79) AM Sleep=somewhat rested
(0.79) AM Sleep=somewhat rested && AM Time=late morning
(0.71) AM Sleep=somewhat rested && AM Time=late morning && EVE Time=evening
(0.71) AM Sleep=somewhat rested && AM Time=late morning && weightchange=same
(0.71) AM Sleep=somewhat rested && EVE Time=evening
(0.71) AM Sleep=somewhat rested && weightchange=same
(0.86) AM Time=late morning
(0.79) AM Time=late morning && EVE Time=evening
(0.71) AM Time=late morning && EVE Time=evening && weightchange=same
(0.79) AM Time=late morning && weightchange=same
(0.79) PM Stress=medium
(0.71) PM Stress=medium && EVE Time=evening
(0.71) PM Stress=medium && weightchange=same
(0.71) PM Time=afternoon
(0.93) EVE Time=evening
(0.86) EVE Time=evening && weightchange=same
(0.71) EVE Time=evening && weightchange=same && sleephour=usual
(0.71) EVE Time=evening && sleephour=usual
(0.93) weightchange=same
(0.79) weightchange=same && sleephour=usual
(0.79) sleephour=usual