

Who Will Be Affected by Supermarket Health Programs? Tracking Customer Behavior Changes via Preference Modeling

Ling Luo^{1,2(✉)}, Bin Li², Shlomo Berkovsky³, Irena Koprinska¹,
and Fang Chen²

¹ School of Information Technologies, University of Sydney, Sydney,
NSW 2006, Australia

{[ling.luo](mailto:ling.luo@sydney.edu.au),[irena.koprinska](mailto:irena.koprinska@sydney.edu.au)}@sydney.edu.au

² Machine Learning Research Group, NICTA, Eveleigh, NSW 2015, Australia

{[bin.li](mailto:bin.li@nicta.com.au),[fang.chen](mailto:fang.chen@nicta.com.au)}@nicta.com.au

³ CSIRO Digital Productivity Flagship, North Ryde, NSW 2113, Australia
shlomo.berkovsky@csiro.au

Abstract. As obesity has become a worldwide problem, a number of health programs have been designed to encourage participants to maintain a healthier lifestyle. The stakeholders often desire to know how effective the programs are and how to target the right participants. Motivated by a real-life health program conducted by an Australian supermarket chain, we propose a novel method to track customer behavior changes induced by the program and investigate the program's effect on different segments of customers, split according to demographic factors like age and gender. The method: (1) derives customer preferences from the transaction data, (2) captures the customer behavior changes via a temporal model, (3) analyzes the program effectiveness on different customer segments, and (4) evaluates the program influence using a one-year data set obtained from a major Australian supermarket. Our results indicate that while overall the program had positive effect in encouraging customers to buy healthy food, its impact varied for the different customer segments. These results can inform the design of personalized health programs that target specific customers in the future and benefit more people. Our method can also be applied to other programs that use transaction data and customer profiles.

Keywords: Customer behaviors · Temporal preference modeling · Health programs · Shopping data analysis

1 Introduction

The World Health Organization (WHO) reports that in 2014 more than 1.9 billion adults were overweight, and over 600 million were obese [1]. Being overweight or obese increases the risk of cardiovascular problems, diabetes, and musculoskeletal disorders. To address the obesity problem, numerous behavior change

programs have been designed, aiming at encouraging participants to maintain a healthier lifestyle, e.g. change their diet and perform physical activity [2,3]. The program stakeholders often desire to know to what extent their program influences the participants, and more importantly, how to improve the program to benefit a broader population in the future. Therefore, mining the vast amount of participant behavior data collected by the program and understanding the behavior changes have become critical and timely research tasks.

The existing studies of health programs mainly report customer demographic statistics and customer survey results, and use statistical tests to examine the overall effect of the programs [2–4]. Although these studies can evaluate the general impact of a health program, they often overlook how the behavior of different types of participants is influenced by the program, which can potentially facilitate effective personalized programs. Thus, our aim is to explore in more depth how the participant behavior changes over time and investigate how participants from different segments are affected by the program.

We propose a systematic approach for tracking the customer behavior changes induced by the health program and evaluate how customers from different demographic segments (e.g. age and gender) are influenced by the program. The method comprises four specialized modules: we extract customer preferences from their transaction data, construct the temporal preference models, then analyze behavior change since joining the program, and most importantly, quantify the program effectiveness on different types of customers.

We evaluate our approach using a large-scale real-life health program delivered by an Australian supermarket chain. The program offered 10% discount on fresh produce to participants, in order to encourage them to eat healthier. The duration of the discount was 24 weeks, but the purchase data of the participants for the entire year was collected, which allows us to compare their behavior before, on, and after the program. We analyze the data collected by the program using the proposed method and study how the purchase behavior of different types of customers is affected. We examine four customer segmentation criteria and show that female customers, younger customers, customers who live with their family, and obese customers are more likely to be encouraged by the program. Hence, the contributions of our work are as follows:

- We construct a temporal preference model that tracks and visualizes preference changes of the program participants over time.
- We quantify the customer preference changes as well as the program effectiveness on different types of customers.
- We evaluate the effectiveness of an Australian health program. Our method and results can be used to inform future personalized health programs.

2 Related Work

Behavior analytics has been recognized as an indispensable part of business intelligence [5,6]. Understanding customer behavior changes allows various stakeholders to monitor dynamic business environment and evaluate their policies and

marketing campaigns [7,8]. A health program is a specific type of campaign, which promotes a healthier lifestyle and behavior [2].

The effectiveness of a program is usually evaluated by investigating customer behavior changes, which can be categorized into two groups: (1) *incremental* approach, that continuously adjusts the model with new transactions [9], and (2) *direct* approach, that models the behavior data at different time periods to identify differences [7,10]. Rule-based methods like association rules [7] and decision trees [10], are frequently used due to their easy interpretation. The incremental approach is sensitive to noise, while the direct approach can be too coarse-grained to reflect the temporal dynamics. Our goal is to design a model that facilitates direct comparison across multiple time periods – instead of just before and after the program – and track changes in customer preferences.

Temporal collaborative filtering (CF) techniques are powerful tools for analyzing patterns of customer preference over time. TimeSVD++ [11] introduces a time-dependent factor into each user-feature for modeling customer preference changes on the items. In [12], preference changes are analyzed via Bayesian tensor factorization, where the tensor is a three-dimensional array of user-item-time tuples. In contrast, [13] considers temporal dependence in Bayesian matrix factorization to model the frequency of preference changes for different types of customers. As the base underpinning for our temporal customer preference model, RMGM-OT [14] takes the advantage of probabilistic topic models to explicitly model the customer preference distributions over item groups; in addition, the method can illustrate how the customer preferences drift over time.

3 Methodology

This section introduces the proposed method for measuring the effectiveness of a health program. As shown in the flow chart in Fig. 1, our method consists of four modules: (1) extracting customer preferences from the transaction data, (2) constructing temporal model for customer preferences, (3) analyzing preference changes over time, and (4) evaluating program influence on different types of customers. The method can visualize customer preferences and provide program analytics as the output.

3.1 Extracting Customer Preferences

The first module extracts the customer preferences from the transaction data. Our program data consists of two parts: (1) transaction data, i.e. purchase records, of 931 participants captured through loyalty cards between 1st January and 31st December, and (2) self-reported survey data addressing their demographic and health information. The original transaction data set covers over 35,600 items from 200 categories. As we are interested in food and drinks relevant to the health program, the data set was reduced to 3,394 items from 24 categories, ranging from vegetables and fruits to snacks and soft drinks. Hence, we used 884 out of the 931 customers, who had a sufficient range of categories

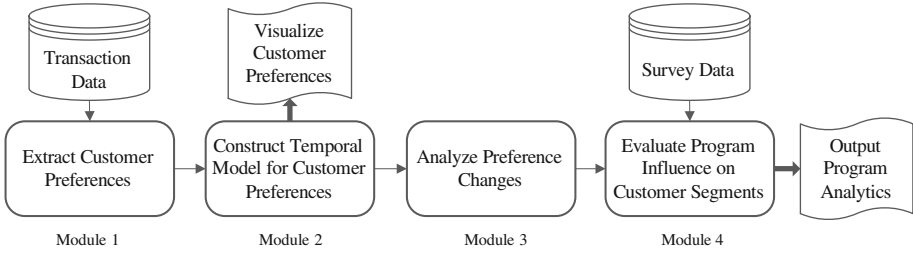


Fig. 1. Flow chart of the proposed method.

and number of purchased items in their transaction data. Specifically, each purchase record shows the customer ID, item ID, item metadata (name, code, and category), purchased quantity and the time-stamp.

Given the transaction data set comprising N customers and M items, each transaction record is a tuple $(u_i, v_j, quantity, time)$, which indicates that customer u_i purchased $quantity$ of item v_j at $time$. We represent all the transaction tuples as a tensor, $\mathcal{X} \in \{1, \dots, R\}^{N \times M \times T}$, where each element \mathcal{X}_{ijt} denotes u_i 's preference in R levels on v_j at $t \in \{1, \dots, T\}$. The whole time span of the transaction data is evenly partitioned into T periods, and the time index t is determined by the period into which $time$ falls. For example, if each time period corresponds to one month, $t = 1$ indicates January transactions. Then, we denote \mathcal{X} as T preference matrices $\{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(T)}\}$ corresponding to the T periods; and each $\mathbf{X}^{(t)}$ contains the preference information of N customers for M items in time period t . The series of preference matrices $\{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(T)}\}$ can inform the temporal analysis of customer preference changes.

In our study, the transaction data is partitioned monthly, so the dimension of \mathcal{X} is 844 customers \times 3394 items \times 12 months. Each element \mathcal{X}_{ijt} is a nominal preference value based on the aggregated amount of item v_j bought in month t . In more detail, the sum of amounts bought by customer u_i in period t is computed and discretized into 5 levels $\{1, 2, 3, 4, 5\}$ item-wisely. For a certain item, we sort all $N \times T$ monthly sum values in ascending order; the value smaller than the first 5-quantile (i.e. ranked within the first 20%) becomes 1 – the lowest preference level, and the value greater than the first 5-quantile but smaller than the second 5-quantile (i.e. ranked within 20%–40%) becomes 2, and so on. If u_i did not purchase v_j in month t , then \mathcal{X}_{ijt} is a missing value.

3.2 Constructing Temporal Model for Customer Preferences

The key component of investigating how the health program influences customer behavior is to build a temporal model for customer preferences on all item categories. However, the item-level preference matrix $\mathbf{X}^{(t)}$ can be very sparse in real-life cases, and the missing values do not necessarily mean the lowest preference level in that period (it is common for customers not to buy certain items at certain periods). Therefore, temporal CF techniques can be exploited to

estimate and smooth customer preferences across the time periods. In our study, we are particularly interested in customer preference changes at the category level, rather than at individual item level. Considering the above requirements, we adapt the temporal CF method RMGM-OT [14] to our problem setting to capture the temporal dynamics of the customer preference for item categories.

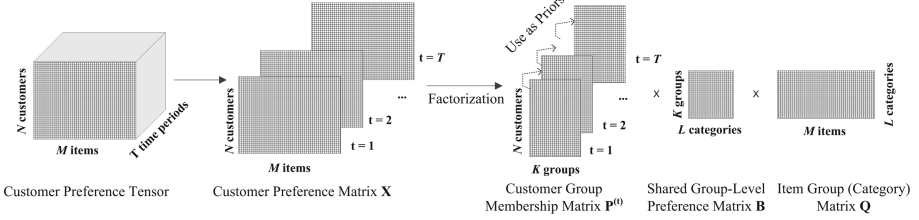


Fig. 2. Illustration of factorizing customer preference tensor \mathcal{X}

The preference matrix \mathbf{X} can be factorized by $\hat{\mathbf{X}} = \mathbf{P}\mathbf{B}\mathbf{Q}^\top$, as schematically shown in Fig. 2. In the setting of customer preference analysis, the above factorization results in K latent customer groups and L latent item groups. \mathbf{B} is a $K \times L$ group-level preference matrix, which represents the preferences of K customer groups for L item groups. $\mathbf{P} \in [0, 1]^{N \times K}$, where each row \mathbf{p}_i can be interpreted as u_i 's membership distribution over K customer groups, and $\sum_k \mathbf{p}_{ik} = 1$ (soft-membership). $\mathbf{Q} \in \{0, 1\}^{M \times L}$ represents the membership information of M items over L item groups. It is worth noting that, in our study, we adopt predefined categories of supermarket products, e.g. vegetables, fruits, and soft drinks, to define the item groups. Thus, v_j belongs to only one item category: $\mathbf{q}_{jl} = 1$ if l is the predefined category; otherwise, $\mathbf{q}_{jl} = 0$.

By taking the temporal domain into consideration, the varying customer preference over time can be modeled based on $\{\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(T)}\}$. We can obtain customer-group membership matrix $\mathbf{P}^{(t)}$ for each time period, while the group-level preference matrix \mathbf{B} is shared across all the time periods and the item-group membership matrix \mathbf{Q} is predefined. We further assume that the customer-group membership satisfies the Markov property, i.e. the state at t depends on the previous state at $t - 1$. Considering the empirical Bayes approach, we can simply use the preceding customer-group membership $\mathbf{P}^{(t-1)}$ as the prior distribution of the current customer-group membership $\mathbf{P}^{(t)}$ in the model (shown by the dotted arrows in Fig. 2).

We adapt the collapsed Gibbs sampler used in [14] to our problem setting, where item categories \mathbf{Q} are given. In other words, item latent variables z_{ij}^v are known in advance and we only need to infer customer latent variables z_{ij}^u . The conditional distribution of z_{ij}^u for Gibbs sampling is

$$P(z_{ij}^u = k | z^{-(ij)}, l, \mathbf{X}^{(t)}) \propto \left(\frac{n_{klr}^{-(ij)} + \beta/R}{\sum_r n_{klr}^{-(ij)} + \beta} \right) \left(n_{ikt}^{-(ij)} + \lambda \mathbf{p}_{ik}^{(t-1)} \right) \quad (1)$$

where α and β are hyper-parameters, and λ is the weight of the prior knowledge. The conditional distribution is proportional to the product of two Dirichlet-multinomial distributions: the first is the proportion of preference r in customer-item joint group (k, l) and the second is the proportion of preference records falling in customer group k (for customer u_i in time period t). It is worth noting that the prior distribution of $\mathbf{p}_i^{(t)}$ is *Dirichlet*(α) for $t = 1$ and *Dirichlet*($\lambda \mathbf{p}_i^{(t-1)}$) for $t = 2, \dots, T$; so by marginalizing out $\mathbf{p}_i^{(t)}$, there exists a pseudo counting $\lambda \mathbf{p}_i^{(t-1)}$ in the second Dirichlet-multinomial distribution.

After obtaining the sample z_{ij}^u , we can estimate \mathbf{B} and $\mathbf{P}^{(t)}$ as follows

$$\mathbf{B}_{kl} = \sum_{r=1}^R r \left(\frac{n_{klr} + \beta/R}{\sum_r n_{klr} + \beta} \right), \quad \mathbf{P}_{ik}^{(t)} = \frac{n_{ikt} + \lambda \mathbf{P}_{ik}^{(t-1)}}{\sum_k n_{ikt} + \lambda \sum_k \mathbf{P}_{ik}^{(t-1)}} \quad (2)$$

where n_{klr} denotes the number of preference r in customer-item joint group (k, l) , n_{ikt} denotes the number of preference records of u_i in customer group k in time period t , and both n_{klr} and n_{ikt} are counted based on z_{ij}^u . Intuitively, $\mathbf{p}_i^{(t)} \mathbf{B} \in [0, R]^{1 \times L}$ reflects the preferences of customer u_i for L item categories in period t . Therefore, the temporal preferences of N customers can be modeled using $\{\mathbf{P}^{(1)}, \dots, \mathbf{P}^{(T)}\}$ and \mathbf{B} .

3.3 Analyzing Customer Preference Changes

Thus far, the preference change of customer u_i has been modeled by $\mathbf{p}_i^{(t)} \mathbf{B}$ for $t \in \{1, \dots, T\}$. To further understand if the change is in the direction targeted by the health program, we label each item category l as either *healthy*, *neutral* or *unhealthy*. Among the 24 food categories in our transaction data, 5 categories are labeled *healthy* (vegetables, mushrooms, fruit snacks, fruit desserts, packaged salads), 5 are labeled *unhealthy* (biscuits & cookies, chilled desserts, snacks, soft drinks, confectionery), and the remaining 14 categories are labeled as *neutral*¹. Our aim is to investigate whether the customer behavior changes due to their participation in the health program, and more specifically, whether their preference for the healthy categories increases.

Based on the category labels, we let $\mathbf{h} \in \{0, 0.5, 1\}^L$ be the healthy indicator vector, and for each category $l \in \{1, \dots, L\}$

$$\mathbf{h}_l = \begin{cases} 1 & \text{if } l \text{ is } \textit{healthy}, \\ 0.5 & \text{if } l \text{ is } \textit{neutral}, \\ 0 & \text{if } l \text{ is } \textit{unhealthy}. \end{cases} \quad (3)$$

Similarly, the indicator vector for unhealthy categories is defined as $\mathbf{1} - \mathbf{h}$, which means the value is 0 for all healthy categories, 1 for all unhealthy categories, and 0.5 for neutral categories.

¹ Although this manual labeling may be simplistic and coarse-grained, we posit that it generally reflects the accepted health perception of food categories.

For customer u_i , we compute the correlation coefficient between the customer preference $\mathbf{p}_i^{(t)}\mathbf{B}$ for the L categories and the healthy indicator \mathbf{h} , and also the correlation coefficient between $\mathbf{p}_i^{(t)}\mathbf{B}$ and the unhealthy indicator $\mathbf{1} - \mathbf{h}$. The difference between these two correlation coefficients defines the health score for customer u_i in time period t as follows

$$health_score_i^{(t)} = corr((\mathbf{p}_i^{(t)}\mathbf{B})^\top, \mathbf{h}) - corr((\mathbf{p}_i^{(t)}\mathbf{B})^\top, \mathbf{1} - \mathbf{h}) \quad (4)$$

where $corr(\mathbf{x}, \mathbf{y})$ denotes Pearson's correlation coefficient.

Given the program starting time t_p , we split the T time periods into two phases: the first includes the periods before the program $t \in \{1, \dots, t_p - 1\}$ and the second includes the remaining periods $t \in \{t_p, \dots, T\}$. The change in customer preferences towards the healthy food categories δ_i is quantified by the difference between the average $health_score$ in the second phase and the average $health_score$ in the first phase. More formally,

$$\delta_i = \frac{1}{T - t_p + 1} \sum_{t=t_p}^T health_score_i^{(t)} - \frac{1}{t_p - 1} \sum_{t=1}^{t_p-1} health_score_i^{(t)} \quad (5)$$

The positive value of δ_i indicates an increase in customer u_i 's preference towards healthy categories after joining the program, while a higher value of δ_i implies a greater change in the right direction. Therefore, δ_i is the key measure for evaluating the effect of the program on customer u_i .

3.4 Evaluating Program Influence on Customer Segments

The values of δ_i for all the customers can provide a general understanding of the health program effect. However, we are also interested in determining the types of customers that are more responsive to the program. This insightful information can inform the design of future personalized programs targeting specific customers.

Generally, the entire customer base can be segmented using a number of features, e.g. geographic, demographic or behavioral [15]. In our study, the self-reported demographic and health information collected by completing a survey is used for customer segmentation. We consider four features: *gender*, *age*, *who customers live with* (alone, with partner, with family), and *Body Mass Index* (BMI). All the customers are partitioned into S_A segments with respect to the value of an attribute A , such that the customers in segment $s \in \{1, \dots, S_A\}$ have the same value of A . If an attribute is numeric, e.g. age, its values are discretized into S_A levels.

As transaction data of customers not participating in the program is unavailable, we split the customers into *experimental group* and *control group* according to the duration of their participation in the program. Specifically, the experimental group completed two surveys – 1) at the start of the program and 2) 12 weeks after the start date, and they participated in the entire program; whereas the

control group completed only the first survey, so that they participated only in half of the program or less.

Overall, the observed changes δ_i in the behavior of individual customers are minor and hard to pick in the short period of one year. Thus, to quantify the effect of the program in customer segment s , we sort all the customers in s according to their δ_i , and measure the portion of the experimental group customers in the set of top- n customers with the highest δ_i . Intuitively, this reflect whether the experimental group customers have greater preference changes towards healthy categories than the control group customers. We define the *effectiveness* of the program for a segment s as:

$$eff_s = \frac{\sum_{n=1}^{N_s} counter_n}{(1 + N_s)N_s} \quad (6)$$

where N_s is the number of customers in s and $counter_n$ is the number of experimental group customers in top- n customers. The baseline for eff_s is 0.5 and greater eff_s values imply that the experimental group customers are ranked higher than the control group customers, indicating that the program is more effective in segment s . Importantly, eff_s is computed for each customer segment and it is used as the main metric for evaluating the effect of the health program.

4 Results for Our Case Study

This section presents the results for our case study on how the health program influences the behavior of the program participants. The evaluation involves 884 participants and 3,394 items from 24 categories, as described in Sect. 3.1. The experimental group comprises 190 customers who participated in the complete program, whereas the remaining 694 customers are in the control group. The program started in May (i.e. $t_p = 5$), so $t \in \{1, \dots, 4\}$ is the first phase, and $t \in \{5, \dots, 12\}$ is the second phase.

Following the four-step approach described in Sect. 3, we convert the transaction data into customer preference tensor, and construct the temporal preference model by factorizing $\hat{\mathbf{X}}^{(t)} = \mathbf{P}^{(t)}\mathbf{B}\mathbf{Q}^\top$ in each month. The parameters are configured as follows: the number of customer groups $K = 20$, the item groups L are the 24 predefined categories, the number of preference levels $R = 5$, λ is set to 10, and the hyper-parameters α and β are set to 1 as in [14].

4.1 Visualization of Customer Preference Changes

We visualize customer group membership $\mathbf{p}_i^{(t)}$ and customer preference $\mathbf{p}_i^{(t)}\mathbf{B}$ over the 12 months in Fig. 3. We select three customers who clearly demonstrate different degrees of variability in group membership and category preferences over time.

The subplots for customer group memberships are shown in the upper row of Fig. 3. Each column in a subplot indicates the mixed membership of 20 customer

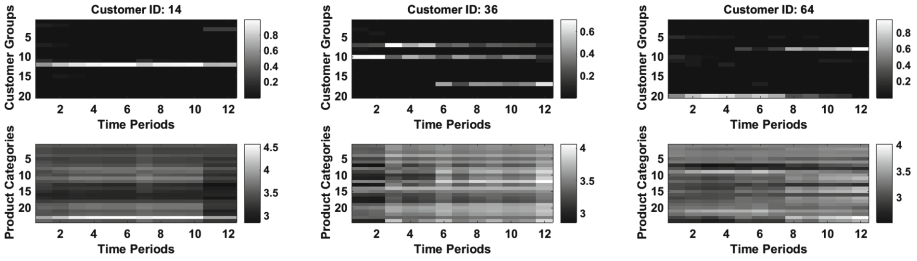


Fig. 3. Visualization of customer group membership distribution (upper row) and corresponding customer preferences for all item categories over the 12 months (bottom row). The three columns, from left to right, are for customers 14, 36, and 64.

groups in one month (sum of each column is 1). Customer 14 had a stable membership over all the time periods, consistently belonging to group 12. On the contrary, customer 36 had a mixed membership in groups 7, 10, and 17. It is interesting to note that the membership in group 17 was identified in June, after joining the program and it was observed till the end of the year. Customer 64 gradually switched from group 20 to group 8 in the middle of the year.

The subplots in the bottom row of Fig. 3 demonstrate how customer preferences for all the 24 categories change over the 12 months. Each row in a subplot reflects the fluctuations of the customer preference for a certain item category over the 12 months, and each column indicates the customer preferences for all the 24 item categories in one month. Customer 14 preferred category 23 (packaged salads) and did not like categories 13 (cheese), 16 (beef), and 17 (lamb), which implies that this customer purchased more vegetables. As for customer 36, the preference for category 20 (vegetables) was consistently high, while the preference for category 22 (fruit desserts) increased gradually, especially after June. The visualizations, such as those shown in Fig. 3, can provide an intuitive understanding of temporal preference changes of individual customers.

4.2 Program Effects for Different Types of Customers

This section quantifies customer preference changes and the program effectiveness. For the control group, the mean preference change is $\delta_i = 0.0204$, while for the experimental group the mean change is $\delta_i = 0.055$, which is more than twice as much as that of the control group. This observation generally shows that the experimental group had a greater preference change towards healthy categories than the control group.

The overall effectiveness of the program, without segmenting customers, is $eff = 0.527$. To get an insight of the fine-grained program effects, the customers are partitioned into segments according to four different criteria: *gender*, *age*, *who customers live with*, and *BMI*. For each segment, we report in Table 1 the size and the program effectiveness eff_s . The customers who had not provided their demographic and health information were excluded from this analysis.

Table 1. Program effectiveness for different customer segments.

Segments	Number of Customers	Number of Experimental Customers	eff _s
Male	349	72	0.4898
Female	517	118	0.5533
Age ≤30	184	30	0.5844
30 < Age ≤ 40	347	83	0.5431
40 < Age ≤ 50	232	54	0.5067
Age > 50	100	23	0.4574
Live Alone	91	21	0.4929
With Partner	344	68	0.5038
With Family	385	91	0.5555
Normal	329	73	0.5592
Overweight	303	77	0.4787
Obese	220	34	0.5705

Segmentation by Gender. As shown in the first section of Table 1, male customers account for about 40% of all the program customers and 37.8% of the experimental group. The effectiveness of the program is 0.5533 for the female customers, which is higher than 0.4898 observed for the male customers. This implies that the female customers in the experimental group were effectively motivated to purchase healthier food, while the male customers from the experimental group were less responsive to the program. Therefore, *the program was found to be more effective for female customers.*

Segmentation by Age. The reported age of the participants varied from 19 to 67. The participants are partitioned into four equal-width segments as shown in the second section of Table 1). The program effectiveness drops significantly from 0.5844 for customers younger than 30 to 0.4574 for customers older than 50. The difference between these two effectiveness scores supports that different customer segments have different responsivenesses to the program. We notice that the program effectiveness decreases gradually as the age increases. One possible reason is that it might be easier for younger customers to change their dietary habits. However, this does not imply that customers older than 50 purchase unhealthy food. On the contrary, they might purchase healthier food consistently before and after joining the program, resulting in smaller preference changes. As our effectiveness measure focuses on the “behavior changes”, the results show that *the program did not influence older customers as much as younger customers.*

Segmentation by Who Customers Live With. Based on the survey question “who you live with”, the customers are partitioned into three segments: live alone, live with partner, and live with family. The segment of “living with their family” is the largest, comprising 46.95% of all participants. The program

influence on this segment is 0.5555, which is higher than the other two segments. For the customers who live alone or with partner, the effectiveness scores are relatively low, being 0.4929 and 0.5038, respectively. Thus, from the perspective of who customers live with, the results show that *customers living with family achieved a greater preference change towards healthy food than the other customers.*

Segmentation by BMI. The BMI is derived from the height and weight of a person, and its value is classified as underweight, normal, overweight, and obese [1]. As the proportion of underweight customers is less than 1.5%, we excluded them from the analysis. It is worth noting that the overweight and obese segments together take up 61.4% of all customers, which shows the severity of the overweight problem and the necessity of encouraging people to eat healthily. The obese segment achieves the highest effectiveness 0.5705, closely followed by the normal segment with 0.5592, and leaving the overweight segment behind with 0.4787. The effectiveness scores of the obese and normal segments are close, but there is a substantial gap between these segments and the overweight segment. The results suggest that *the program was more effective for obese and normal weight customers than for overweight customers.*

Discussion. Our results clearly demonstrate that different types of customers were influenced by the program to a different extent. For the four segmentation criteria, the program was found to be more effective for female customers, younger customers, customers who live with their family, and obese or normal weight customers. However, the program effectiveness across all the segments was relatively low. There are four main possible reasons: (1) the offered 10% discount on fruits and vegetables may not be attractive enough for customers to notably change their behaviors; (2) the lack of real control group may have undermined the significance of the results, since using the customers, who participated in part of the program as the control group, may not truly represent customers not participating in the program; (3) the actual family size is unknown, so that the results may be inaccurate when comparing customers who purchased food, for example, for a family of 2 vs. for a family of 5; (4) the one-year duration of the program may not be sufficiently long to identify stable behavior changes of the customers. Despite these shortcomings of the program data, the evaluation results are encouraging and allow the health program stakeholders to get a fine-grained insight into the impact of the program. This allows tailoring or personalizing future programs, to motivate customers who are not very responsive, such as male customers, senior people and overweight customers in our case.

5 Conclusion

In this paper, we proposed an approach for tracking the customer preferences over time and evaluating the effectiveness of a health program for different types of customer segments defined by demographic and health attributes such as age, gender, living arrangements and BMI. We used data from a large-scale one-year

program conducted by an Australian supermarket, which was designed to encourage customers to build healthy dietary habits. We analyzed how customers from different segments change their preferences over time for various food categories. Overall, the results showed that the program successfully motivated customers to purchase healthier food. The segment-wise effectiveness results demonstrated that different types of customers were influenced to a different extent. We found that female customers, younger customers, customers who live with their family and obese customers were more responsive to the program than their counterparts. Our results can be used to provide guidelines to enhance future health programs, in order to target and motivate the customers who were less responsive in this program and benefit the wider society. Although our method has been designed for a health program, it is a generic method that can be applied to other programs involving transaction records and customer profiles.

References

1. World Health Organization: World health organization factsheet: obesity and overweight (2015). <http://www.who.int/mediacentre/factsheets/fs311/en/>
2. Ball, K., McNaughton, S.A., Le, H.N., Gold, L., Mhurchu, C.N., Abbott, G., Pollard, C., Crawford, D.: Influence of price discounts and skill-building strategies on purchase and consumption of healthy food and beverages: outcomes of the supermarket healthy eating for life randomized controlled trial. *Am. J. Clin. Nutr.* **101**(5), 1055–1064 (2015)
3. Berkovsky, S., Hendrie, G., Freyne, J., Noakes, M., Usic, K.: The healthieru portal for supporting behaviour change and diet programs. In: Proceedings of the 23rd Australian National Health Informatics Conference, vol. 214. IOS Press (2015)
4. Brindal, E., Hendrie, G., Freyne, J., Coombe, M., Berkovsky, S., Noakes, M.: Design and pilot results of a mobile phone weight-loss application for women starting a meal replacement programme. *J. Telemedicine Telecare* **19**(3), 166–174 (2013)
5. Sheth, J.N., Mittal, B., Newman, B.I.: *Customer Behavior: Consumer Behavior and Beyond*. Dryden press fort worth, Texas (1999)
6. Cao, L.: Behavior informatics and analytics: let behavior talk. In: Proceedings of the IEEE International Conference on Data Mining Workshops, pp. 87–96. IEEE (2008)
7. Chen, M.C., Chiu, A.L., Chang, H.H.: Mining changes in customer behavior in retail marketing. *Expert Syst. Appl.* **28**(4), 773–781 (2005)
8. Huang, C.K., Chang, T.Y., Narayanan, B.G.: Mining the change of customer behavior in dynamic markets. *Inf. Technol. Manage.* **16**(2), 117–138 (2015)
9. Masegla, F., Poncelet, P., Teisseire, M.: Incremental mining of sequential patterns in large databases. *Data Knowl. Eng.* **46**(1), 97–121 (2003)
10. Liu, B., Hsu, W., Han, H.-S., Xia, Y.: Mining changes for real-life applications. In: Kambayashi, Y., Mohania, M., Tjoa, A.M. (eds.) *DaWaK 2000*. LNCS, vol. 1874, pp. 337–346. Springer, Heidelberg (2000)
11. Koren, Y.: Collaborative filtering with temporal dynamics. *Commun. ACM* **53**(4), 89–97 (2010)
12. Xiong, L., Chen, X., Huang, T.K., Schneider, J.G., Carbonell, J.G.: Temporal collaborative filtering with bayesian probabilistic tensor factorization. In: Proceedings of the 2010 SIAM International Conference on Data Mining, vol. 10, pp. 211–222. SIAM (2010)

13. Li, R., Li, B., Jin, C., Xue, X., Zhu, X.: Tracking user-preference varying speed in collaborative filtering. In: Proceedings of the 25th AAAI Conference on Artificial Intelligence (2011)
14. Li, B., Zhu, X., Li, R., Zhang, C., Xue, X., Wu, X.: Cross-domain collaborative filtering over time. In: Proceedings of the 22nd International Joint Conference on Artificial Intelligence, pp. 2293–2298. AAAI Press (2011)
15. Baker, G.A., Burnham, T.A.: Consumer response to genetically modified foods: market segment analysis and implications for producers and policy makers. *J. Agric. Resource Econ.* **26**, 387–403 (2001)