Saint Petersburg School of Economics and Management
Department of Management

Term paper

Suchkova Angelina Sergeevna

Analysis of Promotions Based on Consumer Flows in Retail

in the field 38.03.02 «Management»

educational program «Management»

Completed by
Student of group
#BMN 196

Suchkova Angelina S.

Academic supervisor:
Doctor of Sciences,
Professor, Yuri Zelenkov

____

Saint Petersburg
2022

Abstract

Customer experience (CX) play an important role for business as it creates a competitive advantage for the company. Big media promotions are widely used for offline marketing campaigns in retail and present an effective tool to creating a satisfying consumer experience. However, despite the undeniable effectiveness of such promotions, it is rather difficult to evaluate their impact. The purpose of this work is to develop a method for evaluating the effects of promotions on customer flows and evaluate the effectiveness of various types of Big Media promotions using it.

This study proposes the method to assess the effectiveness of promotions based on the intensity of consumer flows. The first part of analysis includes clustering customers using transactional data of a Russian supermarket chain and calculation of flows between them. After that in the second part such techniques as Singular Spectrum Analysis (SSA), first differences method and trained Dynamic Bayesian Network (DBN) model are used to find causal relationships between flows coefficients and number of promotions.

As a result, the types of promotions that have the greatest impact on positive (when buyers become more active) and negative (when buying activity decreases) flows were identified. Also, assumptions have been made about consumer clusters which are most sensitive to promotions. The developed method can be used in marketing to evaluate both individual promotions and their joint effectiveness. Results of analysis provide managers a deeper understanding of changes in customer behavior and the ability to influence these changes.

Keywords: *consumer flows, clustering, efficiency of promotions*

Contents

Introduction

Customer experience (CX) plays a key role in creating competitive advantage for businesses (Lemon & Verhoef, 2016). This concept includes all interactions of the buyer with the business, which ultimately form his attitude towards the company and determine the value for it. CX can be defined as consumer responses to various stimuli. In retail, Big Media are an essential element in creating a satisfying consumer experience, they form brand awareness. Such promotions are characterized by number of outdoor advertising elements, duration, categories of goods for which they apply and the size of discounts in general. According to data from Statista, outdoor advertising was the leading segment of out-of-home (OOH) ads by spending in Russia, at over 38 billion Russian rubles in 2021. Approximately one quarter of the total spending on outdoor advertising in Russia was occupied by retail and trade services providers in 2021. Multicategory retailers who sell their goods mainly in offline stores spend a significant portion of the budget on Big Media promotions to increase brand awareness. The fact that companies spend much of their budget on promotions proves that it is possible to carefully analyze different types of promotions and evaluate their analysis on consumers.

Previous studies (Nijs et al., 2001; Ailawadi et al., 2006) have shown that not all promotions have a positive effect on retailer profits. This result is explained by high advertising costs and lower margins on advertised products. Therefore, in order to evaluate the effectiveness of promotions and the need for them, the retailer first needs to understand whether the marketing attracts profitable buyers, i.e. it is necessary to determine how different promotions affect different consumer segments (Ailawadi et al., 2009). For example, to retain loyal customers, it is necessary to periodically provide them with discounts. Promotions can be an effective tool for regulating both short-term and long-term relationships with consumers. Understanding changes in consumer behavior can help marketers design more effective promotions. Customer valuation, segmentation and the development of mixed marketing strategies help retain the customers who bring the most profit to the business (Khajvand et al., 2011). Therefore, marketers are faced with the question of grouping consumers according to some business-relevant metrics to provide not only personalized offers but also be able to optimize massive marketing campaigns in such a way as to satisfy the needs of the buyers who are most valuable to the business. This is a particular challenge for retail stores where there is a high variety of customers and satisfying the largest number of consumers is required.

In retail industry, there is the increased availability of customer transactional data that provides a good basis for analyzing consumer profiles (Heldt et al., 2021). Determining customer profiles and identifying their behavior patterns is part of Customer Relationship

Management (CRM), which is becoming increasingly important as competition increases in all areas of business. Analytical CRM allows you to get insights from customer data using such data mining and business intelligence techniques (Ngai et al., 2009). These techniques allow companies to isolate homogeneous groups of consumers using clustering techniques.

One of the most effective metrics for customer segmentation is the RFM (Recency-Frequency-Monetary) model. It allows you to segment buyers by characteristics such as recency, frequency and monetary based on their purchase history. The classical RFM model assumes that recency reflects the age of the latest transaction of customers, frequency is the number of purchases made in each period and monetary is the total amount of money spent by a buyer on goods over a certain period.

Previous papers use the RFM model and its variations for customer segmentation. The division of consumers into clusters and the assessment of buyer profiles with the identification of other significant characteristics are widely used to analyze the consumer base. However, they use a static segmentation approach that calculates metrics over a fixed period (Djurisic et al., 2020). Aggregating data and defining consumer clusters over a single period does not allow us to assess the dynamics of consumer behavior. While it is the dynamics of cluster changes that makes it possible to evaluate the effectiveness of marketing campaigns after the fact. A small number of studies assess consumer movement patterns using clustering (Abbasimehr & Bahrini, 2022), but a gap in this area is that they do not attempt to assess the influence of Big Media promotions on the intensity of consumer flows of consumers. Most works did not focus on flows between groups of consumers, thus the importance of movement between homogeneous groups was not considered. Although it is the assessment of flows which can be quantified and compared with each other, that makes it possible to evaluate the effectiveness of promotions. Previous papers have omitted to compare flow coefficients as a time series, and no one has attempted to assess their statistical significance either. This work closes this gap in marketing analytics and proposes a methodology for assessing the intensity of consumer flows and the impact of promotions on these movements of consumer cohorts. The novelty of this work lies in an analysis method that allows you to find causal relationships between the intensity of consumer flows between clusters and the number of promotions without considering other external factors.

This work carried out according to the methodology of design-science proposed a framework which allows to assess the influence of Big Media promotions according to their type and number on the speed of consumer flows. Such a framework involves the following steps. The first step is the segmentation of consumers into clusters separately for each period (52 periods in total) using metrics such as frequency and monetary. The recency metric for customer characteristics does not have statistical significance due to short time periods. For segmentation, the K-means algorithm is used, which is widely used to select groups in analytics. Silhouette score is used to evaluate the quality of clustering. Next, it is needed to calculate the flows between clusters sequentially for each period. The flow coefficients between clusters represent the intensity of consumer movement. These coefficients are presented for analysis in the form of time series. Further, using Singular Spectrum Analysis (SSA), first differences method and Dynamic Bayesian Network (DBN) model we identified how different types of Big Media promotions (include billboards, facades, biweekly media, and seasonal ones) affect the change in the speed of consumer movement between clusters. We provide interpretation of model results and possible reasons that explain them.

To demonstrate practical utility the methodology described above is applied to real transactional data from a large supermarket chain from Russia. The most successful promotions are identified that can help marketers to increase the intensity of the most profitable consumer flows which in turn lead to an increase in the average check and the volume of consumer purchases.

The contributions of this research are the following: a new methodology is proposed for analyzing the dynamics of consumer transitions between clusters and assessing the impact of Big Media promotions on the intensity of these flows. The findings of this work provide retail managers with additional data on the behavior of their consumer population and the response of large segments of this population to marketing campaigns. Thus, the most effective promotions will be highlighted, and such an analysis will allow organizing more effective marketing campaigns that cause the strongest consumer flows.

Literature Review

1. Method-descriptive studies

Big data provides marketers with ample opportunity to analyze customer behavior. Data mining can be applied to identify useful customer behavior patterns from large amounts of customer and transaction data (Giudici & Passerone, 2002). Consumer behavior is widely studied using advanced clustering methods. Initially, many studies focused specifically on identifying customer segments over a certain period (Akhondzadeh-Noughabi & Albadvi, 2015). Clustering at a fixed time interval refers to static segmentation. However, this approach does not allow tracking the dynamics of cluster changes which is a serious obstacle to the analysis of customer behavior. In particular, the static approach does not allow tracking trends and using more advanced methods for predicting consumer behavior including Customer Lifetime Value (CLV) (Seret et al., 2014; Mosaddegh et al., 2021). Identification of consumer segments by periods and tracking their movement allows marketers to understand patterns and determine both internal changes that influenced the movement of consumers and external factors. K-means clustering is widely used for segmentation in marketing due to easy implementation and fast execution. This method requires a preliminary selection of the number of clusters. Since the developed framework involves first dividing buyers into periods and then clustering separately for each period, it is necessary to obtain the same number of clusters in each period, the K-means algorithm was chosen for this analysis.

Various features of consumers provide the basis for dividing into clusters, one of the approaches which represent customer characteristics is the RFM model often used in research for customer segmentation (Rygielski et al., 2002). This model is a powerful marketing tool that captures consumer value and buying habits well and is also used to predict consumer behavior and segmentation (Abbasimehr & Shabani, 2021). The traditional model uses three metrics: recency, frequency and monetary calculated based on buying history. The advantage of this model lies in the small number of metrics used which significantly speeds up the process of clustering using machine learning methods. Previous works develop improved versions of it. One of the improved versions of this model is the RdFdMd model which considers discounts on customer receipts and segments customers based on this data (Heldt et al., 2021). The mentioned study compares the segmentation efficiency using the classic RFM model and the RdFdMd model considering the discounts offered by the business. The researchers concluded that the model makes it possible to identify better consumer clusters. In the research conducted by Peker et al. (2017) an LRFMP model was developed that considers such metrics as length and periodicity to cluster buyers over a fixed period. Considering that such metrics as length, periodicity and recency when selecting clusters for short periods of time as weeks are not indicative and only worsen the quality of consumer division, in this work the RFM model is slightly simplified, and clustering is carried out only by frequency and monetary indicators.

Regarding dynamic clustering, the RFM model itself and its variations in previous studies have been used to identify customer shift patterns (Akhondzadeh-Noughabi & Albadvi, 2015) and tracked customer behavior over time by segmentation of customers (Hosseini & Shabani, 2015). However, these studies were carried out for the telecommunications and banking sectors. The retail industry has its own specifics, both when considering consumer transactions and when conducting marketing campaigns. Particularly, in retail buyers tend to go into churn when segmenting by periods, thus, when highlighting clusters, it is quite difficult to determine the overall dynamics due to the constant disappearance of those buyers who could be called regular. Due to this issue, clustering of buyers seems to be a challenging task. Therefore, in this paper the proposed methodology is designed to minimize the error in the analysis considering the characteristics of the retail industry.

Although promotions and their impact on consumer behavior have been extensively studied to develop personalized offers, there is a serious drawback in the literature since most studies do not consider the dynamics of consumer behavior and the influence of external facts on them, namely, different types of Big Media promotions carried out by marketers. Retailers often focus their efforts on retaining loyal customers and much of the work is focused on identifying the conditions for conducting a marketing campaign under which loyal consumers will remain loyal for as long as possible (Feinberg et al., 2002). Thus, when segmenting consumers by periods and tracking their dynamics, the task is to determine the effectiveness of a particular promotion for a particular segment.

The novelty and optimality of the method presented in this work is expressed, firstly, in the method of analyzing flows between clusters. Previous studies have used association rule mining (Akhondzadeh-Noughabi & Albadvi, 2015) which have also been used to identify dominant consumer behavior patterns and Hidden Markov models (HMMs) (Lemmens et al., 2012) to model consumer transactions between segments: such the studies were aimed at evaluating the evolution of the consumer's relationship with the firm. In this work coefficients of consumer flows are analyzed like time-series data. For analysis of getting coefficients of customer flows we used techniques for analysis of time series data: SSA for identifying trends and noise, first differences for normalizing distributions and DBN model for identifying causal inference between flows and promotions. The methods described above were used for analysis in such research fields such as meteorology, oceanography, medicine (Coussin, 2022) and energy consumption (Wei & Bai, 2022). Skare & Porada-Rochoń (2019) used SSA algorithm to estimate financial cycles. DBN algorithm were applied for forecasting purposes in study conducted by Shen et al. (2015) and analyzing people behavior in Li et al. (2022) study. In their study Quesada et al. (2021) use DGBN models and evaluate their ability to process time series, as a result, he concludes that these models can capture the trend of the time series with seasonality, and data noisiness does not affect the quality of the model. In marketing field these techniques were not applied that's why new methodology proposed in this study can be a basis for evaluating of efficiency of promotions in companies to develop more methods in related field.

2. Analysis of promotions and customer behavior

Promotions can be defined as time-limited marketing strategies, implemented to directly influence customers' purchasing decisions. Much research has been done on how promotions affect brand sales and profits (Kopalle et al., 1999; Srinivasan et al., 2004). However, the effect that promotions give to retailer revenue and profits has been much less explored. In their study, Srinivasan et al. (2004) notes that promotions predominantly generate revenue for the manufacturer, while the effect on the seller is not so clear. In a study by Ailawadi et al. (2006) found that more than half of the advertising campaigns carried out do not bring profit to the retailer. Nijs et al. (2001) and Srinivasan et al. (2004) examined the long-term effects of promotions on retailers and concluded that they do not have a significant positive impact on revenue in the long run. Some research suggests that promotions do not have sustainable effects (Steenkamp et al., 2005). Thus, there is a need to constantly monitor the effectiveness of promotions, not in terms of income generated, but in terms of their impact on consumers. Therefore, in recent years, researchers have shown more interest in analyzing consumer behavior and their responses to various promotions.

Consumer reactions to business actions are part of the customer experience. CX is a broad concept that describes all interactions between a consumer and a business. In retail, relevant interactions are both product or company related information faced by customers and various in-store promotional activities, product assortments, retail atmosphere (Li et al., 2021). Satisfying CX can result in an increase of such characteristics as customer spending, shopping duration, and purchase frequency (Bleier et al., 2019; Chalil et al., 2020). With the development of various communication channels, there are more and more opportunities for contact with the consumer through the media, and retail faces the task of choosing the most effective of them.

In this area, most of the research on the effectiveness of promotions has been aimed at identifying elements and their parameters that attract the most consumer attention, and therefore are most visible, better remembered and, as a result, more effective (Narasimhan et al., 1996;

Bijmolt et al., 2005). Many studies prove that the way promos are designed strongly influences consumer behavior (Ailawadi et al., 2009).

As for the study of the effectiveness of different types of promotions, Inman et al. (1997) conducted an analysis for promotions with quantity limits, multiple unit promotions, and bonus packs, during which they found that the presence of any limit positively affects the effectiveness of the promotion. Also, Manning & Sprott (2007) conducted a study on the quantities of goods mentioned in a promotion. They found that stocks with noticeably higher amounts mentioned (e.g., 8 or 20, not 2 or 4) have a significant impact on the consumer. Thus, it can be concluded that important behavioral mechanisms are involved in promotions (Ailawadi et al., 2009).

Another potential factor that may determine the effectiveness of a promotion is promotion sensitivity across different consumer segments (Zhang & Wedel, 2009). Such sensitivity can be justified by the characteristics of individual segments, namely, customers who visit stores more often have a greater sensitivity to promotions because they have more opportunities to learn about them than those consumers who rarely visit the store and do not see in-store promotions (Vakratsas & Bass, 2002; Arce-Urriza et al., 2017). Applebaum and Spears (1950) and Andreeva et al. (2010) conducted that frequent customer exposes to sales promotions and marketing initiatives in a greater way that an infrequent one.

Some studies have looked at the behavioral patterns of occasional consumers and concluded that their reaction to promotions is stronger than that of regular consumers (Bawa & Ghosh, 1990). However, Vakratsas & Bass (2002) clarify that the effect of promotions on occasional shoppers also depends on the category of product being discounted. For example, occasional consumers do not respond as effectively to promotions for consumer goods, because often miss these promotions. For goods that are not for daily use, such buyers react more strongly than regular ones, because. use the stock as an opportunity to profitably buy goods.

Despite a fairly large number of studies on the impact of promotions, not a single work is aimed at determining the effect of outdoor promotions on different consumer segments. Thus, this work fills this gap in the literature, which will complement the existing data and provide a more complete picture of consumer reactions to marketing campaigns. The results of this work will allow marketers to better understand the impact of promotions on different consumer segments.

Methodology

This section presents the detailed description of proposed methodology which allows to assess the influence of number of promotions on the speed of flows between clusters. For the analysis two datasets were used which were presented for analysis at the hackathon from the Lenta company. The data covers the period from September 2019 to September 2020. The analysis was carried out using the Python language, since the data contain many observations (more than 20 million). The first dataset includes the receipt data of a random sample of supermarket chain customers.

The following variables were selected for the analysis:

Name	**Definition**
Client_id	Customer ID
Day	Date of purchase
Check_id	Check ID
Num_sales	Number of sales of products
Selling_price	Selling price of products

Table 1. Variables for consumer description.

The second dataset contains data on Big Media promotions of various types (billboards, facades, biweekly and seasonal catalogs) and their timing. Billboards are a type of out-of-home advertising, which is most often located in specially constructed structures. Facades are located on the walls of buildings. Two-week promos, as the name suggests, have a very limited duration, while seasonal promos offer promotions for seasonal products and the longest of the listed types. In Table 2 description of initial variables are given:

Name	**Definition**
Promo_type	Type of promotion
Offer_id	Promotion ID
Sku	Product ID
Start_date	Promotion start date
End_date	Promotion start date

Table 2. Variables for promos description.

At the stage of data preparation, we left only buyers who made more than 2 purchases per year - thus, we remove buyers from the analysis "one-time". Outliers were also removed from the data (where the check amount exceeds the 98% percentile) to improve the quality of clustering.

The proposed framework involves the following steps:

1. Clustering procedure

1.1 Periods for clustering

It was decided to carry out clustering by weeks, i.e., each week is a separate period in which clusters of consumers active during this period are independently identified. Thus, 52 periods will be considered and coefficients for 51 flows will be calculated.

1.2 RFM metrics

Indicators of the traditional RFM model: recency, frequency and monetary. When clusters are identified for short consecutive periods of time, the recency metric loses its value, therefore, in this work, only frequency and monetary metrics calculated for each customer active in a certain period were used to identify customer behavioral patterns.

1.3 Defining the number of clusters

For clustering, the K-means++ algorithm was used. This algorithm assumes a preliminary designation of the number of allocated clusters, so the silhouette score was used to evaluate the quality of clustering. Data has outliers so clustering has been improved by using Local Outlier Factor (LOF) technique to remove them.

1.4 Labelling each customer segment and value analysis

At this step each of the resulting clusters was labeled using the custom function. The main idea of determining the semantic names of clusters is to rank the clusters according to how much revenue the buyers in them bring to the store. Revenue is calculated by multiplying the frequency metric by the monetary one. Because profit is the primary measure of customer value to a business, segments have been named according to the value they provide. Thus, we divided buyers into sleeping, loyal and champions.

2. Causal inference in time-series data

2.1 Representing customer behavior as time series

The number of consumers from each cluster for each period was counted. Those consumers who were in one cluster in one month, moved to another in the next month - are considered as a flow from one cluster to another. Using pivot tables, we calculated the percentage of buyers who were in the flow relative to their number in the original cluster in the previous week. Further, the corresponding coefficients for the flows were presented in the form of time series. All flows were included in the analysis, but it was expected that some of them were affected by the number of promotions to a greater extent, others less or not at all.

2.2 Detecting trends

For preliminary analysis of time-series data we decided to use SSA method. This method allows to define the main components of the time series and suppress noise in data. Data for analysis should be a sequence of some value taken at regular intervals, in this study such data is coefficients of consumer flows while time intervals are weeks. SSA method based on the converting a one-dimensional series into a multidimensional series. After that method of principle components is applied to the resulting multivariate time series. Thus, SSA algorithm gives several components which present trends and noise data and further we continue the analysis using the trend component.

2.3 Bringing the distribution to normal

As the distributions of the trend components from the previous step usually are not normal, we took first differences of these distributions to bring them to normal (Afanasiev & Yuzbashev, 2001) using the following transformation:

$$r_{\Delta_x \Delta_y} = \frac{\sum_{i=1}^{n}\left(\Delta_{x_i} - \bar{\Delta}_x\right)\left(\Delta_{y_i} - \bar{\Delta}_y\right)}{\sqrt{\sum_{i=1}^{n}\left(\Delta_{x_i} - \bar{\Delta}_x\right)^2 * \left(\Delta_{y_i} - \bar{\Delta}_y\right)^2}}$$

where: Δ_x, Δ_y - first differences;

$\bar{\Delta}_x$, $\bar{\Delta}_y$ - average values of these differences.

2.4 Calculation of correlation between flows and number of promotions

After checking that the distributions of first differences are normal, we found the correlation coefficients between these values and number of promotions. Take that the correlation named R. Then we should check the hypothesis that correlations R are equal to zero.

$$H_0: R = 0;\ H_1: R \neq 0$$

As the own distribution of R is enough difficult, at first, we make the following transformation:

$$t = \frac{R}{\sqrt{1-R^2}}\sqrt{n-2} \sim t_{n-2}$$

Then the sample distribution of this statistic is the Student's distribution with 2 degrees of freedom. For a given significance level a, we determine the critical value t_{cr}. If $|t| > t_{cr}$, then reject H_0.

2.5 DBN model for causal inference

To define if there is a causal inference between flows and promotions, we learnt the structure of a Gaussian Bayesian dynamic network in R (natPsoho method).

The proposed methodology has not previously been used to analyze the effectiveness of promotions and assess their impact on consumer behavior, so this work makes a significant contribution to the field of marketing analytics. By following the above steps, marketing analysts will be able to evaluate the business impact of their promotions and plan the promotion calendar more effectively.

Case Study

The retail company investigated in this research is a Russian supermarket chain which operates in many cities. In this study only one region with 111 stores was chosen for analysis. Original dataset contains more than 21.4 million records where one record contains information about what product in what amount and in what price with unique check ID was purchased by a client. More detailed description of all variables is provided in Table 1 and Table 2 above.

At first from original data, we deleted customers who made only one or two purchase. Since we consider weekly dynamics of customer flows, we are more interested in constant consumers. That is also the reason why detecting outliers (people who buy too many products) should increase the accuracy of results. Such customers presumably made such purchases not for personal purposes for daily consumption but for big events or own resale. The purpose of this research is to detect influence of Big Media promotions on constant customers' activity. We believe that the purchasing activity of the above-described group of consumers is to a lesser extent due to promotions. Thus, the data for analysis contains 18.1 million records with 85 thousand customers.

After division our data into 53 weeks recency and frequency metrics were calculated for customers who were active during each period. Based on these metrics we conducted the clustering for customers from each week. The division into 3,4,5 and 6 clusters were tested, the highest average silhouette score for all periods was obtained for division into 3 clusters, the average silhouette score is 0.58+-0.02. To improve the accuracy of clustering in each period, the Label Outlier method was applied to remove outliers. The score is higher than 0.5 that means that our detection is enough correct. We got 3 clusters which were named as «sleeping», «loyal» and «champions» and those ones who didn't purchase during the period were sent to «churn» cluster. The average values for recency and monetary values for active clusters are presented on Figure 1 and in Table 3 accordingly.

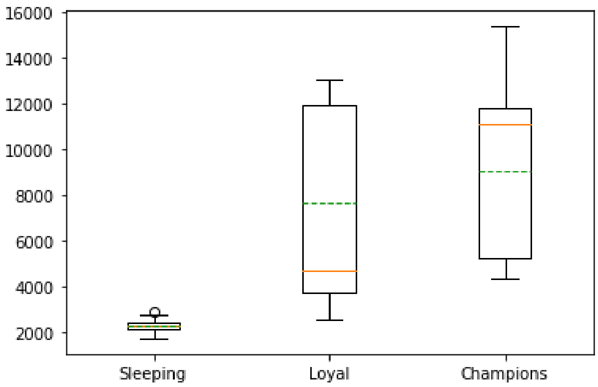

Figure 1. Distribution of monetary metric (green line illustrate average values, orange one - median value).

Statistics	Sleeping	Loyal	Champions
min	1,00	1,87	1,49
median	1,28	3,97	2,24
mean	1,31	3,09	3,52
max	1,41	5,32	5,58

Table 3. Distribution of frequency metric.

	Sleeping	Loyal	Champion	Churn
% from total	16%	3%	2%	78%
min	7180	1529	976	59165
median	13806	2686	2079	66851
mean	13702	2754	1920	66527
max	19326	4581	3058	73593

Table 3. Number of customers in clusters.

Table 4 presents the number of people in clusters. Number of customers total in 4 clusters in each week is constant in perspective of the whole period. The biggest cluster is «churn». It is predictable since not so many people make purchases every week. The second biggest cluster is sleeping. Most customers from this cluster visit stores only once per week and their average check is slightly above 2000. The loyal customers make purchases more often (3-4 times a week) and spend about 5000-8000 rubles per time. The most profitable and spending cluster is champions who visit stores as often as loyal customers but spend significantly more (9000-11000 rubles).

Further we got the percentage amount of people who was in one cluster in the previous period and moved to another cluster (or stay in the same one) in the next week. In total we calculate 16 flows for 53 periods. The Figure 2 allows you to assess the scale of the received flows and promotions. At the first step of analysis id time-series data it is necessary to get rid of noise in data. Figure 3 illustrates different components which were detected from our data. Such decomposing was made using python method Singular Spectrum Analysis. We get 4 PCA (Principal Component Analysis) for each flow. The example on the Figure 2 shows that SSA 1 describes the trend in data while the other 3 components are Brownian motion. We continue our analysis using only SSA 1 component.

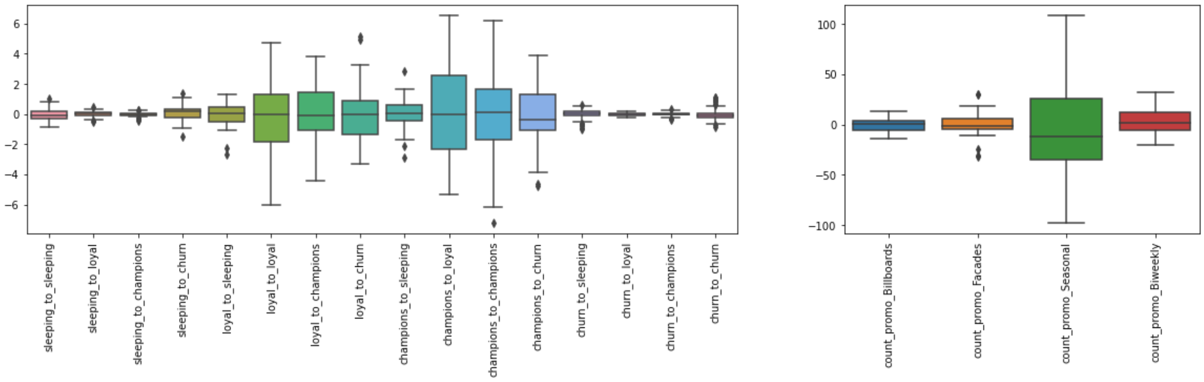

Figure 2. Distributions of flows and promos.

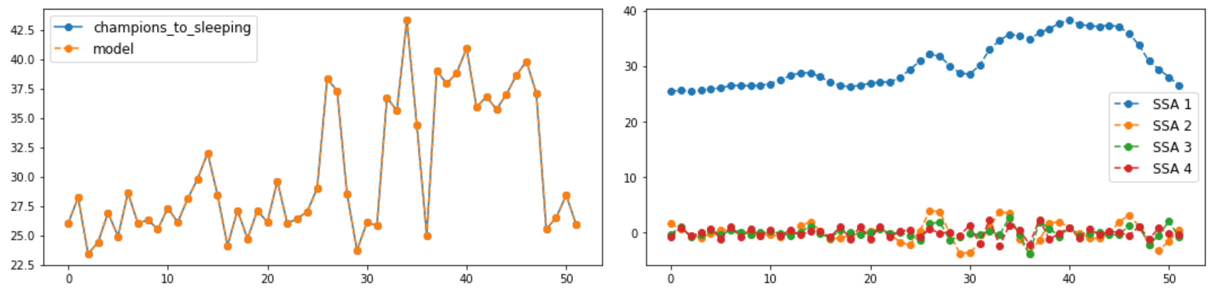

Figure 3. Example of SSA output for champions_to_sleeping flow.

Next stage is the correlation analysis to understand if any dependencies between flows and promotions. The correlation coefficients are presented in Table 5, the table shows results with values of promotions for period t-1 (as if the promotion is moved on one previous period to flows data). It is expected that the influence of Big Media on customer behavior during one previous week is the most significant. Thus, the dependence between flows and coefficients are, but they may be both correlation and causal inference.

	sleeping_to_sleeping	sleeping_to_loyal	sleeping_to_churn	loyal_to_sleeping	loyal_to_loyal	loyal_to_champions	loyal_to_churn	champions_to_sleeping	champions_to_loyal	champions_to_champions	champions_to_churn	churn_to_sleeping	churn_to_loyal	churn_to_champions	churn_to_churn
Billboards	0.27	-0.31						0.24	-0.26	0.38	-0.24			-0.29	
Facades	-0.31		0.23		0.42	-0.32			-0.42		0.31	0.51	0.29		-0.51
Seasonal	0.28														0.24
Biweekly															

Table 4. Correlation matrix between flows and promos for period t-1.

To determine whether there is a causal relationship between flows and the number of promotions of different types, we used the training of the structure of the Gaussian Baysian dynamic network. The results obtained in the form of a system of linear equations are analyzed in the next chapter.

Results

The result obtained using DBN model is a system of linear equations where the dependent variables are the 16 customer flows under consideration and the independent variables are the number of promotions of 4 types in the previous t-1 period and the preceding t-2 period (in total 8 independent variables). We analyzed the impact of the number of promotions on flows from two sides. First, we consider only the coefficients, that is, the percentage change in the number of people in the flow. Then we calculate the revenue that various flows bring and evaluate the effectiveness of promotions considering the money they bring to the business. Revenue analysis is necessary since the number of customers in the flows is different as well as the revenue, they bring is different for different clusters. Also, the number of promotions differs by their types. The above factors can affect the significance of the flow and, accordingly, the significance of a particular type of promotion.

For analyzing coefficients, we divided the consumer flows into 3 groups: positive flows (sleeping to loyal, sleeping to champions, loyal to champions, churn to sleeping, churn to loyal, churn to champions), negative flows (sleeping to churn, loyal to sleeping, loyal to churn, champions to loyal, champions to sleeping, champions to loyal) and neutral flows (sleeping to sleeping, loyal to loyal, champions to champions and churn to churn accordingly). For positive flows, the higher the coefficients for the promo, the more effective these promotions are for the business. Positive values indicate that the positive flow has increased due to, for example, placed billboards. For negative flows, on the contrary, the lower the coefficient for the promotion, the more it reduces the flow which is unfavorable for business. Tables 6, 7 and 8 present the coefficients of promotions for positive, negative and neutral customer flows, respectively. The columns contain information about 4 types of promotions carried out in the previous period. The values in the Intercept column are the slope coefficients, that are, independent change in the size of the flow.

Flows	Intercept	Billboards t-1	Facades t-1	Seasonal t-1	Biweekly t-1
sleeping_to_loyal	-0,0142	-0,0097	-0,0037		
sleeping_to_champions	-0,0348	-0,0082	0,0204		0,0112
loyal_to_champions	0,0301	-0,0030	-0,0212	-0,0056	
churn_to_sleeping	0,0308				
churn_to_loyal	0,0014	0,0042	0,0011	0,0001	
churn_to_champions	0,0000	-0,0053		-0,0010	0,0023

Table 5. Coefficients of promos for positive flows.

Conclusions that are drawn about the impact of promotions on positive flows (Table 6):

a. Flows from the sleeping cluster to more active clusters are reduced regardless of the number of promotions (slope coefficients for them are negative: -0.0142 and -0.0348), that is, over a prolonged period, more and more buyers are flowing out. No type of promo has a positive effect on the flow from sleeping to loyal.

b. However, facades have a strong positive effect on the transition of buyers from sleeping directly to the champions cluster which is equal 0.0204. At the same time, this type of promo reduces the flow from loyal to champions at 0.0212. It can be assumed that buyers from the sleeping cluster are overly sensitive to various discounts that can be presented on the facades, while already loyal consumers are indifferent to them. Moreover, the flow from loyal to champions by itself is constantly increasing (the slope coefficient is 0.0301), in contrast to the sleeping to champions flow (the slope coefficient is -0.0348), which once again indicates the fundamentally different behavior of the buyers of the two clusters: sleeping customers are hypersensitive to discounts and have a high risk of churn without them, loyal customers have a natural the tendency not to decrease, but to increase their purchasing activity without external incentives in the form of promotions.

c. Only biweekly promotions have an indisputable positive effect on positive flows, but this effect is not very pronounced, because coefficients are statistically significant only for 2 flows (from sleeping to champions and from churn to champions) and are only 0.0112 and 0.0023.

d. As for flows from the churn cluster, billboards and seasonal promotions have a statistically significant impact on the flow from churn to champions, however this influence is rather low and negative (-0.0053 and -0.0017 respectively). Such values can be explained by the fact that such frequent and large purchases after the absence of any buying activity are due to personal reasons, but not Big Media promotions. The transition from churn to loyal is slightly positively affected by all types of promotions placed in the previous period.

Flows	Intercept	Billboards t-1	Facades t-1	Seasonal t-1	Biweekly t-1
sleeping_to_churn	0,0050	-0,0100			0,0286
loyal_to_sleeping	0,0001		-0,0327		
loyal_to_churn	0,1492		-0,0137		
champions_to_sleeping	0,0411			0,0017	
champions_to_loyal	-0,2267	-0,2248	-0,1050	0,0030	
champions_to_churn	-0,1564				

Table 6. Coefficients of promos for negative flows.

As for negative overflows, the following insights can be distinguished (Table 7):

a. Loyal consumers flow quite strongly into churn (the slope coefficient is 0.1492) regardless of the number of promos, although the transition from loyal to sleeping is almost 0. Facades reduce these flows a little (-0.0137 for churn and -0.0327 for sleeping), but to other types of promo loyal consumers are not sensitive. We came to the same conclusion when analyzing positive flows from the loyal cluster.

b. As for the outflow from champions, flows to loyal and churn clusters are steadily declining even without promos (the slope coefficients are -0.2267 and -0.1564). But if no type of promotions can affect the outflow to churn, then the transition to loyal is noticeably reduced by billboards at 0.2248 and facades at 0.1050.

c. Based on the results of the model seasonal and biweekly promos slightly increase negative flows. These types of promotions last longer than billboards and facades, so their impact in a given short period is harder to gauge. Perhaps, the coefficients obtained are due to the specifics of these types of promotions.

Flows	Intercept	Billboards t-1	Facades t-1	Seasonal t-1	Biweekly t-1
sleeping_to_sleeping	0,0816	-0,0015	-0,0175	-0,0004	-0,0139
loyal_to_loyal	-0,0349		0,1564		
champions_to_champions	0,0068			-0,0067	
churn_to_churn	-0,0305		0,0252	0,0043	-0,0060

Table 7. Coefficients of promos for neutral flows.

Regarding neutral flows (Table 8):

a. We can note the constant growth of the sleeping cluster regardless of the number of promos (the slope coefficient is 0.0816). Such an increase may be due, among other things, to the attraction of new buyers, who at first do not have high purchasing activity.

b. Although the number of buyers who remain loyal from period to period is decreasing (the slope is -0.0349), facades significantly increase the number of such buyers (at 0.1564).

Discussions

This work addresses a CX-related issue that has not been sufficiently explored in the literature. Customer response to a retailer's marketing activities is subjective and difficult to track. This work provides information on how outdoor promotions affect the activity of various consumer segments. Since the analysis was carried out not on each individual consumer, but on selected clusters, possible subjective motives of consumers were leveled.

In previous research it was concluded that promotion sensitivity is a potential factor behind the differential promotion effect across consumer types (Zhang & Wedel, 2009). This study confirms this conclusion from various aspects. Applebaum and Spears (1950) and Andreeva et al. (2010) found that the more frequently customers visit a store, the greater their exposure to sales promotions and marketing initiatives. Buyers from frequent and champions clusters are those ones which visit store more frequent that sleeping ones. According to our results, promotions of any type have no significant positive effects on positive flows from loyal and champions clusters. That’s why we can’t provide the same findings like in previous studies. However, negative flows from champions cluster are reduced by promotions very much. Thus, we can conclude that customers who visit stores more frequently can be more affected by different promotion activities, but the influence is expressed in reducing the negative effect of their outflow on business. This a new issue, which wasn't considered in previous studies, and it presents a wide field for future research.

Another aspect of the analysis concerns precisely the speed of consumer response to different types of promos. Arce-Urriza et al. (2017) concluded that the more often shoppers visit a store, the faster they respond to promotions. This study looked at the period of one previous week, which allows you to track only short-term effects. In the context of the duration of consumer response, the results of this work are consistent with the results obtained in past studies. Namely, more active consumers react faster to marketing activities. In this work, the response of consumers from the champions cluster to billboards and facades during the first week was revealed: the outflow from the cluster is significantly reduced. An effect on loyal consumers was also revealed: they maintain their activity in the short term. In previous works, the authors also explain this by the fact that active buyers visit stores more often and, as a result, are more aware of various promotions.

As for the occasional buyers, namely the churn cluster, the results of this study overlap with the findings of past work, but there are conflicting points. So, no type of promo affects the flow from churn to sleeping, billboards and facades slightly increase the flow from churn to loyal but reduce the flow to champions. Bawa and Ghosh (1990) and Vakratsas & Bass (2002) concluded that occasional shoppers may react strongly to some promos and increase their activity, which is what happened for churn to loyal traffic. However, the effect on buyers from the churn cluster cannot be called regular or stable, which proves the negative impact of promo on flows to champions and the lack of effect on churn to sleeping flow.

The negative effects that certain promotions produce can be explained both by the limited set of factors in the model and by the effect recorded in the study by Kopalle et al. (1999). The authors note that increased promotions can lead to lower underlying sales and increased consumer price sensitivity. Therefore, the resulting negative odds for positive flows (for example, loyal to champions) may be the result of an excessive number of different types of promotions placed.

Conclusion

The purpose of this work was to evaluate the effect of promotions on consumer flows. From the side of the method used for the analysis, the work is of value for future design-science research, as it offers a new approach for analyzing consumer behavior. The advantage of the proposed method lies in the following facts. First, the method can be used to estimate the effect of specific promotions without considering other factors. Marketers will be able to track the effect that a specific launched promotion gives, as well as determine its starting point and expected duration. Secondly, the proposed methodology makes it possible to look at consumer behavior more deeply than previous works, namely, to determine the intensity of the effect on various customer flows. In practice, this will help to track not just an increase in the activity of the buyer, but to assess the increase in activity relative to other buyers. This opportunity is especially relevant for businesses that are highly seasonal: in their case, the characteristics of all customer segments change over the period and an increase in the activity of one segment can only be assessed relative to others.

The work also contributes to behavior research in the field of management and marketing. The results of the work give marketers the ground for analysis and more correct planning of marketing activities. This study provides information that the impact of outdoor promotions on positive traffic is not as strong. However, a strong reduction in negative flows under the influence of billboards and facades makes us look at marketing from a different angle: not as a stimulating measure, but as a preventive measure, without which the purchasing activity of the most loyal consumers will decline.

The limitations of this study are mainly related to the number of periods considered and, as a result, the uncertain effect of seasonal promos. The DBN model was trained for two lags, i.e. considered the effect of a promo lasting a maximum of 2 weeks. Seasonal and bi-weekly rolls are expected to have a much longer-term effect on consumer activity than billboards and facades, so this effect has only been captured to a small extent. Thus, to obtain relevant results regarding long-term promotions in the future, it is necessary to add more periods to the model. Also, adding more features to the DBN model is a promising idea for further research. Features such as frequency and monetary will allow you to track the change in the characteristics of buyers within the flows, a larger number of lags in the model will give an idea of the delayed effects that certain types of promotions can have and will also allow you to estimate how much before the expected decrease in buying activity it is worth launching promotions. Then the refined model will give managers the opportunity to develop customer experience with much greater understanding and much less financial loss.

References

Abbasimehr, H., & Bahrini, A. (2022). An analytical framework based on the recency, frequency, and monetary model and time series clustering techniques for dynamic segmentation. *Expert Systems with Applications, 192* doi:10.1016/j.eswa.2021.116373

Abbasimehr, H., & Shabani, M. (2021). A new framework for predicting customer behavior in terms of RFM by considering the temporal aspect based on time series techniques. *Journal of Ambient Intelligence and Humanized Computing, 12*(1), 515-531. doi:10.1007/s12652-020-02015-w

Ailawadi, K. L., Beauchamp, J. P., Donthu, N., Gauri, D. K., & Shankar, V. (2009). Communication and promotion decisions in retailing: A review and directions for future research. *Journal of Retailing, 85*(1), 42-55. doi:10.1016/j.jretai.2008.11.002

Ailawadi, K. L., Harlam, B. A., César, J., & Trounce, D. (2006). Promotion profitability for a retailer: The role of promotion, brand, category, and store characteristics. *Journal of Marketing Research, 43*(4), 518-535. doi:10.1509/jmkr.43.4.518

Akhondzadeh-Noughabi, E., & Albadvi, A. (2015). Mining the dominant patterns of customer shifts between segments by using top-k and distinguishing sequential rules. *Management Decision, 53*(9), 1976-2003. doi:10.1108/MD-09-2014-0551

Arce-Urriza, M., Cebollada, J., & Tarira, M. F. (2017). The effect of price promotions on consumer shopping behavior across online and offline channels: Differences between frequent and non-frequent shoppers. *Information Systems and e-Business Management, 15*(1), 69-87. doi:10.1007/s10257-016-0310-2

Bijmolt, T. H. A., Van Heerde, H. J., & Pieters, R. G. M. (2005). New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research, 42*(2), 141-156. doi:10.1509/jmkr.42.2.141.62296

Bleier, A., Harmeling, C. M., & Palmatier, R. W. (2019). Creating effective online customer experiences. *Journal of Marketing, 83*(2), 98-119. doi:10.1177/0022242918809930

Chalil, T. M., Dahana, W. D., & Baumann, C. (2020). How do search ads induce and accelerate conversion? the moderating role of transaction experience and organizational type. *Journal of Business Research, 116,* 324-336. doi:10.1016/j.jbusres.2020.05.033

Coussin, M. (2022). Singular spectrum analysis for real-time financial cycles measurement. *Journal of International Money and Finance, 120* doi:10.1016/j.jimonfin.2021.102532

Djurisic, V., Kascelan, L., Rogic, S., & Melovic, B. (2020). Bank CRM optimization using predictive classification based on the support vector machine method. *Applied Artificial Intelligence, 34*(12), 941–955. http://dx.doi.org/10.1080/08839514.2020. 1790248

Feinberg, F. M., Krishna, A., & Zhang, Z. J. (2002). Do we care what others get? A behaviorist approach to targeted promotions. *Journal of Marketing Research, 39*(3), 277-291. doi:10.1509/jmkr.39.3.277.19108

Giudici, P., & Passerone, G. (2002). Data mining of association structures to model consumer behaviour. *Computational Statistics and Data Analysis, 38*(4), 533-541. doi:10.1016/S0167-9473(01)00077-9

Heldt, R., Silveira, C. S., & Luce, F. B. (2021). Predicting customer value per product: From RFM to RFM/P. *Journal of Business Research, 127,* 444-453. doi:10.1016/j.jbusres.2019.05.001

Hosseini, M., & Shabani, M. (2015). New approach to customer segmentation based on changes in customer value. *Journal of Marketing Analytics, 3*(3), 110-121. doi:10.1057/jma.2015.10

Inman, J. J., Peter, A. C., & Raghubir, P. (1997). Framing the deal: The role of restrictions in accentuating deal value. *Journal of Consumer Research, 24*(1), 68-79. doi:10.1086/209494

Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. *Paper presented at the Procedia Computer Science, 3,* 57-63. doi:10.1016/j.procs.2010.12.011 Retrieved from www.scopus.com

Kopalle, P. K., Mela, C. F., & Marsh, L. (1999). The dynamic effect of discounting on sales: Empirical analysis and normative pricing implications. *Marketing Science, 18*(3), 317-332. doi:10.1287/mksc.18.3.317

Lemmens, A., Croux, C., & Stremersch, S. (2012). Dynamics in the international market segmentation of new product growth. *International Journal of Research in Marketing, 29*(1), 81-92. doi:10.1016/j.ijresmar.2011.06.003

Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. Journal of Marketing, 80(6), 69-96. doi:10.1509/jm.15.0420

Li, X., Dahana, W. D., Ye, Q., Peng, L., & Zhou, J. (2021). How does shopping duration evolve and influence buying behavior? the role of marketing and shopping environment. *Journal of Retailing and Consumer Services, 62* doi:10.1016/j.jretconser.2021.102607

Li, Y., Wu, G., & Liu, C. (2022). Research on intelligent recommendation model of E-commerce commodity based on feature selection and deep belief network. *Security and Communication Networks, 2022* doi:10.1155/2022/6469217

Manning, K. C., & Sprott, D. E. (2007). Multiple unit price promotions and their effects on quantity purchase intentions. *Journal of Retailing, 83*(4), 411-421. doi:10.1016/j.jretai.2007.03.011

Mosaddegh, A., Albadvi, A., Sepehri, M. M., & Teimourpour, B. (2021). Dynamics of customer segments: A predictor of customer lifetime value. *Expert Systems with Applications, 172* doi:10.1016/j.eswa.2021.114606

Narasimhan, C., Neslin, S. A., & Sen, S. K. (1996). Promotional elasticities and category characteristics. *Journal of Marketing, 60*(2), 17-30. doi:10.1177/002224299606000202

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications, 36*(2 PART 2), 2592-2602. doi:10.1016/j.eswa.2008.02.021

Nijs, V. R., Dekimpe, M. G., Steenkamp, J. -. E. M., & Hanssens, D. M. (2001). The category-demand effects of price promotions. *Marketing Science, 20*(1), 1-22. doi:10.1287/mksc.20.1.1.10197

Peker, S., Kocyigit, A., & Eren, P. E. (2017). LRFMP model for customer segmentation in the grocery retail industry: A case study. *Marketing Intelligence and Planning, 35*(4), 544-559. doi:10.1108/MIP-11-2016-0210

Quesada, D., Valverde, G., Larrañaga, P., & Bielza, C. (2021). Long-term forecasting of multivariate time series in industrial furnaces with dynamic gaussian bayesian networks. *Engineering Applications of Artificial Intelligence, 103* doi:10.1016/j.engappai.2021.104301

Rygielski, C., Wang, J. -., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in Society, 24*(4), 483-502. doi:10.1016/S0160-791X(02)00038-6

Seret, A., Vanden Broucke, S. K. L. M., Baesens, B., & Vanthienen, J. (2014). A dynamic understanding of customer behavior processes based on clustering and sequence mining. *Expert Systems with Applications, 41*(10), 4648-4657. doi:10.1016/j.eswa.2014.01.022

Shen, F., Chao, J., & Zhao, J. (2015). Forecasting exchange rate using deep belief networks and conjugate gradient method. *Neurocomputing, 167,* 243-253. doi:10.1016/j.neucom.2015.04.071

Skare, M., & Porada-Rochoń, M. (2019). Tracking financial cycles in ten transitional economies 2005–2018 using singular spectrum analysis (ssa) techniques. *Equilibrium.Quarterly Journal of Economics and Economic Policy, 14*(1), 7-29. doi:10.24136/eq.2019.001

Srinivasan, S., Pauwels, K., Hanssens, D. M., & Dekimpe, M. G. (2004). Do promotions benefit manufacturers, retailers, or both? *Management Science, 50*(5), 617-629. doi:10.1287/mnsc.1040.0225

Steenkamp, J. -. E. M., Nijs, V. R., Hanssens, D. M., & Dekimpe, M. G. (2005). Competitive reactions to advertising and promotion attacks. *Marketing Science, 24*(1), 35-54. doi:10.1287/mksc.1040.0069

Vakratsas, D., & Bass, F. M. (2002). The relationship between purchase regularity and propensity to accelarate. *Journal of Retailing, 78*(2), 119-129. doi:10.1016/S0022-4359(02)00068-4

Wei, S., & Bai, X. (2022). Multi-Step Short-Term building energy consumption forecasting based on singular spectrum analysis and hybrid neural network. *Energies, 15*(5) doi:10.3390/en15051743

Zhang, J., & Wedel, M. (2009). The effectiveness of customized promotions in online and offline stores. *Journal of Marketing Research, 46*(2), 190-206. doi:10.1509/jmkr.46.2.190