Analyzing consumer behavior in bakery product markets: A journey through correspondence analysis and regression modeling

Michaela Soóki[1], Mária Májek[2], Zuzana Poláková[3]

Abstract: Analyzing consumer behavior in bakery product markets involves understanding the factors that influence consumers' choices and preferences when it comes to baked goods. One crucial aspect of this analysis is examining how different diseases impact the consumption of bakery products, such as celiac disease, Crohn's disease, or a low histamine diet. By incorporating these factors into consumer behavior analysis in bakery product markets, researchers can gain valuable insights into how these diseases impact individuals' choices and preferences. This knowledge can help bakeries and food manufacturers develop products that cater to specific dietary needs and expand their customer base. This study aims to investigate the consumer behavior and preferences of bakery products in Slovakia. Using a survey of 528 respondents we applied regression models and correspondence analysis to analyze the data and identify the factors influencing consumer choices and satisfaction. The results show correlations between different qualitative variables such as the consumption of bakery products the feelings arising from the consumption of bakery products and the average physical activity of the respondent. The research is supported by the VEGA project: Challenges for food security in 21st century Europe – key factors, socio-economic and environmental contexts, no. VEGA 1/0755/21.

Keywords: correspondence analysis, consumer behavior, bakery products, regression modeling
JEL Classification: D12, L66, M31, P2

1 Introduction

In the world of consumer behavior research, the allure of bakery products, with their enticing aromas and delectable flavors, has long been a subject of fascination. These products, ranging from crusty artisan bread to indulgent pastries, hold a unique place in our culinary landscape, transcending mere sustenance to become an integral part of our daily lives. However, what lies beneath the act of selecting a croissant over a muffin, or a whole-grain loaf over a baguette, is a complex web of human choices, emotions, and influences. Consumer behavior, as eloquently defined by Solomon (2019), serves as the compass guiding individuals through the labyrinth of decisions surrounding their purchases—what to buy, when to buy, where to buy, and at what cost. This intricate interplay of choices holds profound implications for marketers, managers, and policymakers alike. It is the cornerstone upon which effective strategies are designed to satisfy customer needs and desires, bolster sales and profits, and foster social welfare. Researchers have made substantial contributions to our understanding of bakery product consumption and its underlying drivers, ranging from psychology and emotions to marketing and portion control. Their work provides a rich foundation for studying the complex dynamics of consumer behavior in bakery product markets. Authors such as Brian Wansink have significantly contributed to our understanding of consumer behavior, particularly in the context of food consumption Wansink and Sobal (2007) conducted a study on the daily food decisions people overlook. In their research, Wansink, Painter, and North (2005) explored how visual cues of portion size influence food intake. Wansink's work has practical implications for understanding how individuals make decisions regarding bakery product consumption and their dietary habits. Dan Ariely's research focuses on the irrational aspects of consumer decision-making, including food choices. He has conducted experiments and studies that reveal how emotions, cognitive biases, and social factors influence our food preferences and consumption patterns. Ariely and Kreisler (2007) examined the persistence of habits despite conflicts with motives. According to Loewenstein and Ariely (2006), curiosity plays a significant role in decision-making. Ariely's work provides valuable insights into the emotional and psychological dimensions of bakery product consumption.

Barbara J. Rolls' research centres on satiety, portion control, and the factors that affect food intake. Rolls (2010) investigated the relationship between dietary energy density and energy intake. The study by Rolls and Roe (2002) focused on the effect of intragastrically infused liquid food on satiety. Rolls' findings are relevant for understanding why

[1] Slovak University of Agriculture in Nitra, Faculty of Economics and Management, Institute of Statistics, Operations Research and Mathematics, Tr. A. Hlinku 2, 949 76 Nitra Slovak republic, xkraslanova@uniag.sk
[2] Slovak University of Agriculture in Nitra, Faculty of Economics and Management, Institute of Statistics, Operations Research and Mathematics, Tr. A. Hlinku 2, 949 76 Nitra Slovak republic, xvargovam2@uniag.sk
[3] Slovak University of Agriculture in Nitra, Faculty of Economics and Management, Institute of Statistics, Operations Research and Mathematics, Tr. A. Hlinku 2, 949 76 Nitra Slovak republic, zuzana.polakova@uniag.sk

individuals may opt for certain bakery products and how these choices relate to feelings of satisfaction and fullness. David Kessler's book "The End of Overeating" (2009) discusses taking control of the insatiable American appetite. In his work, Kessler (2018) explores the concept of fast carbs and their impact on health. Investigates the factors driving overconsumption of highly palatable foods, including bakery products. He delves into the neurological and behavioral aspects of cravings and how they influence consumer behavior. Kessler's research offers insights into the motivations behind bakery product consumption, particularly in the context of indulgence and overeating. Paul Rozin's research explores food preferences, cultural influences, and the psychology of eating. His work helps unravel why individuals develop preferences for specific bakery products and how these choices relate to cultural and emotional factors. Rozin's studies shed light on the diverse range of emotions associated with bakery product consumption, from pleasure to nostalgia. Rozin and Fallon (1987) provided a perspective on disgust. Rozin (1999) discussed preadaptation and its role in pleasure. Marion Nestle's work focuses on food policy, marketing, and the influence of the food industry on consumer choices. Nestle's book "Food Politics" (2013) examines how the food industry influences nutrition and health. In her article, Nestle (2002) discusses the relationship between food marketing and childhood obesity. Chandon and Wansink (2007) proposed a psychophysical model of meal size estimation. The reduction of context effects in choice experiments was explored by Chandon and Hutchinson (2007).

In this paper, we aim to analyze the consumer behavior in the bakery products market using data from a large-scale survey conducted in Slovakia in 2023. The survey collected information on various aspects of consumer behavior, such as product attributes (e.g., taste, freshness, quality, price, etc.), consumer demographics (e.g., age, gender, income, education, etc.), purchase frequency (e.g., how often consumers buy bakery products), and purchase occasion (e.g., breakfast, lunch, snack, etc.). We use two statistical methods: correspondence analysis and regression modeling. Correspondence analysis is a technique that allows us to visualize the associations between categorical variables in a two-dimensional map (Greenacre, 2017). Regression modeling is a technique that allows us to estimate the effects of explanatory variables on a dependent variable (Wooldridge, 2016). By combining these two methods, we are able to identify the main consumer segments in the bakery products market, explore the relationships between product attributes and consumer demographics. In this study, we embark on a captivating journey through the interplay of bakery product consumption, health diagnoses, emotional states, and physical activity levels among respondents. Our research seeks to unravel the intricate web of connections that link the choices individuals make at the bakery aisle to their well-being and daily lives. The findings presented herein shed light on the nuanced dynamics that underpin these relationships, presenting valuable insights into the intricate tapestry of human behavior and its health consequences. One of the central revelations of our study lies in the profound interdependence between diagnoses related to the consumption of bakery products and the frequency of such consumption. By scrutinizing the health profiles of respondents, we uncovered patterns that illuminate how certain diagnoses may significantly influence bakery product consumption. Our exploration doesn't stop at the confines of health diagnoses; it extends into the realm of emotions. We have unearthed a fascinating connection between emotions, often rooted in bakery product consumption, and the respondents' levels of physical activity. The emotional states that individuals experience in response to their dietary preferences appear to play a pivotal role in shaping their engagement in physical activities. This intriguing finding hints at the profound psychological implications of bakery product consumption, transcending mere dietary satisfaction to influence broader aspects of daily life.

As we navigate through the intricacies of these relationships, it becomes evident that bakery products are not just items of sustenance; they are, in many ways, agents that bridge the domains of health, emotions, and physical well-being. Our study offers a holistic perspective on how choices made at the bakery aisle reverberate through the corridors of health diagnoses, emotional experiences, and daily activities.

2 Methods

When evaluating the relationship between two attributes, the investigation delves into whether the occurrence of one attribute is linked to the occurrence of another. This exploration involves the assessment of hypotheses:

- Null hypothesis: The attributes are independent.
- Alternative hypothesis: The attributes are dependent.

The assessment employs the chi-square test of independence, denoted by the formula:

$$\chi^2 = \sum_{i=1}^{m} \sum_{j=1}^{r} (E-T)^2/T \quad (1)$$

Where:

- 'm' signifies the number of rows,
- 'r' the number of columns,

- 'n' the total respondents,
- 'E' the observed frequency,
- 'T' the expected frequency,
- 'χ^2' the calculated test statistic.

To gauge the dependency strength, two coefficients are utilized:

Pearson's coefficient (C), computed as:

$$C = \sqrt{(\chi^2/(n+\chi^2))} \quad (2)$$

Where:

- 'χ^2' represents the test statistic,
- 'n' denotes the total respondents.

Cramer's V coefficient (V), calculated as:

$$V = \sqrt{(\chi^2/(n(\min((m,r)-1))))} \quad (3)$$

Where:

- 'χ^2' represents the test statistic,
- 'n' denotes the total respondents.
- 'm' signifies the number of rows,
- 'r' the number of columns,

Cramer's V factors in the number of rows and columns, overcoming Pearson's coefficient limitation influenced by the number of groups. Both coefficients usually range from <0.1 (weak dependence) to >0.1 (stronger dependence), with values closer to 1 indicating a stronger attribute relationship.

Notably, a commonly chosen significance level (alpha) for hypothesis testing is 0.05.

Regression models serve to elucidate the link between a dependent variable and one or more independent variables. The general linear regression model follows the format:

$$y = \beta0 + \beta1x1 + \beta2x2 + ...+ \beta nxn + \epsilon \quad (4)$$

Where:

- 'y' represents the dependent variable.
- 'x1, x2,..., xn' denote the independent variables.
- 'β0, β1,..., βn' signify the regression coefficients.
- 'ϵ' denotes the error term.

Correspondence Analysis (CA) emerges as a valuable method for exploring connections between categorical variables. It finds applications in diverse domains like market research and ecology, organizing data in contingency tables. CA provides a visual representation, revealing patterns, groups, and interdependencies among variables.

The method transforms the contingency table into a lower-dimensional space. Each category from the rows and columns becomes individual points, with distances between them signifying association strength. This graphical method yields insights into categorical data relationships.

The general formula for correspondence analysis is:

$$X = UDV^T \quad (5)$$

This formula represents the contingency table's decomposition into its elements:

- 'X': The observed frequencies in the contingency table.
- 'U': A matrix of row scores indicating relationships between rows.
- 'V': A matrix of column scores indicating relationships between columns.

- 'D': A diagonal matrix of singular values reflecting relationship strength.
- 'T': Denotes the transpose of a matrix.

By employing SVD on the standardized residuals of the contingency table and decomposing it into 'U', 'D', and 'V' matrices, one gains insights into category relationships. Visualizing these relationships through biplots using row and column scores helps depict connections graphically.
Description of Available Data:

The survey responses encompass a blend of categorical and numerical variables crucial for our analyses. Categorical variables include information on the frequency of bakery product consumption ('regularly', 'sometimes', 'non'), specific diagnosed conditions that might influence this consumption ('celiac', 'milk allergy', 'low histamine', 'other'), as well as emotions experienced during consumption ('satisfaction', 'nostalgia', 'guilt', 'joy', etc.).

Moreover, we possess numerical data such as stress levels on a scale from 1 to 10 and monthly expenditures on bakery products in euros, ranging from 0 to 149 €.

Objective of Analyses:

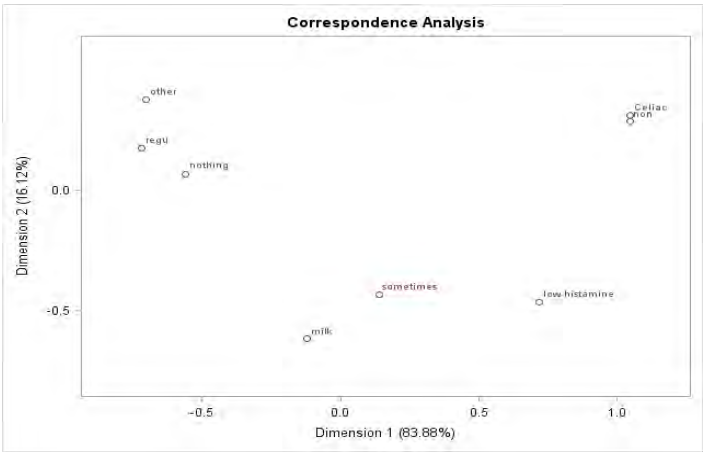
The analyses aim to identify relationships among multiple variables. We seek to elucidate connections between diagnosed conditions and consumption frequency, as well as emotions experienced during consumption. Furthermore, we explore how these diagnosed conditions influence respondents' emotional experiences during bakery product consumption. There's a specific focus on understanding the relationship between stress levels and monthly expenditures on bakery products among individuals with celiac disease.

For these analyses, we plan to employ tools such as the chi-square test of independence for categorical variables and regression analysis to understand relationships between numerical variables.

3 Research results

"In the initial segment of this research effort, centered around the survey questions on the frequency of bakery product consumption (mark in the Figure 1 'regu' for regularly, 'sometimes' for sometimes, 'non' for no consumption) and the presence of specific diagnoses that might affect such consumption ('nothing' for I don´t have any, 'other' for other than mentioned, 'milk' for any milk allergy or lactose intolerance , 'low histamine' for histamine sensitivity, 'celiac' for celiac disease), our primary focus was to explore the intricate relationship, or potential lack thereof, between these aspects. Following rigorous chi-square analysis, our calculations uncovered a statistically significant association between these variables, supported by a test statistic (311.25) surpassing the critical value (18.307). Hence, it is reasonable to posit that the existence or absence of specific diagnoses correlates distinctly and significantly with the frequency of bakery product consumption. Indeed, the examination of this correlation has yielded noteworthy outcomes. Our findings indicate a level of dependence between these variables that ranges from moderate to robust, as substantiated by both the Pearson coefficient, registering at an impressive 0.6086, and the Cramer-V coefficient, standing at a substantial 0.5424. These coefficients collectively underscore the strength of the relationship between the variables under scrutiny, underscoring the significance of the observed dependencies.

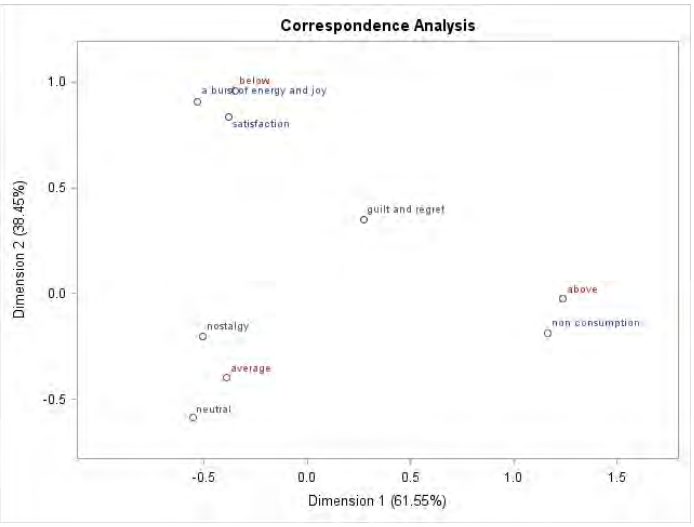
Figure 1 Relation between frequency of consumption and diagnosis



Source: Own processing

As depicted in Figure 1, our analysis of Slovak respondents afflicted with celiac disease reveals a conspicuous pattern: they conscientiously abstain from the consumption of bakery products. This notable trend may be attributed to several compelling factors. Firstly, it is plausible that newly diagnosed individuals lack comprehensive guidance regarding dietary restrictions and allowances, potentially resulting in cautious avoidance of bakery items to avert any inadvertent consumption of gluten. Additionally, the propensity to eschew bakery products may be influenced by the higher costs and limited availability of gluten-free alternatives vis-à-vis their conventional counterparts. Conversely, our survey findings indicate a contrasting consumption behavior among respondents devoid of diagnoses impacting bakery product consumption, as well as those possessing diagnoses not explicitly accounted for within the survey options. Evidently, these individuals exhibit a proclivity for regular bakery product consumption. Intriguingly, within the category of "occasional" consumption, two distinct diagnoses emerge—namely, milk allergy (inclusive of lactose intolerance) and hypersensitivity to histamine. This dichotomy warrants closer examination, as it implies that not all individuals adhering to a low-histamine diet necessarily exhibit hypersensitivity to components found in bakery products. Furthermore, it underscores the intricate interplay between dietary restrictions and the presence or absence of specific ingredients, such as milk or lactose, within bakery items. In our comprehensive investigation of the Slovak respondents within the bakery products market, we also delved into additional variables encompassing average physical activity levels and the emotional states of respondents during bakery product consumption. These variables enrich our understanding of the multifaceted dynamics at play and their influence on consumer behavior within this market segment.

Figure 2 Relation between condition of respondent and their emotions



Source: Own processing

In the subsequent section of our research, we directed our focus toward the discernment of intricate relationships between the variables physical condition and an emotion accompanied by the consumption of bakery products. Employing chi-square analysis, our calculations unveiled a statistically significant association among these variables, substantiated by the test statistic (403.241) surpassing the critical value (21.026). Consequently, we posit that the emotions elicited by the consumption of bakery products are indeed correlated with the physical activity levels and overall well-being of the respondents.

Within this web of interdependence, we observe a robust connection, ranging from moderate to strong, substantiated by both the Pearson coefficient, which attains a notable value of 0.658, and the Cramer-V coefficient, standing at a substantial 0.618. Figure 2 encapsulates these associations, revealing intriguing nuances within the emotional spectrum experienced by respondents.

Notably, for respondents reporting below-average physical condition (as denoted by mark "below" in the Figure 2), the consumption of bakery products engenders a distinct sense of satisfaction. Respondents articulated this sentiment as "Eating bakery products gives me a feeling of satisfaction both physically and emotionally. It is a soothing experience." Simultaneously, they conveyed sensations of "a burst of energy and joy" after partaking in bakery products, depicted in the image as "a burst of energy and joy."

Conversely, respondents in average physical condition expressed more neutral or nostalgic emotions, describing bakery product consumption as evocative of cherished memories. For these individuals, bakery products constituted a trip down memory lane. Additionally, some respondents in average condition reported that bakery products did not elicit significant emotional reactions and were considered a routine part of their diet.

However, it is noteworthy that negative emotions, including feelings of guilt and regret tied to the fear of weight gain, transcended physical condition and activity levels. Both respondents with below-average and above-average physical activity levels harbored these concerns, underscoring the complex interplay of emotions intertwined with bakery product consumption.
In our regression analysis, we examined the relationship between the variable "stress level," which ranged from 1 to 10, and the variable "monthly household expenditures on bakery products," with expenditure amounts ranging from 0 to 149 €. Our findings indicated a positive correlation between monthly expenditures on bakery products and stress levels.

The analysis showcases an incredibly robust relationship (Multiple R = 0.9769) between the variables under study. This means that a significant proportion (around 95.44%) of the variability in stress levels among individuals with celiac disease can be accounted for by variations in their monthly expenditures on bakery products (R Square = 0.9544). The statistical significance of our regression model is highlighted by the ANOVA results. The obtained F-statistic (F = 209.23) coupled with an extremely low p-value (4.95E-08) indicates that the overall regression model is notably superior in explaining the variance in stress levels compared to a model devoid of independent variables. When the monthly household expenditure on bakery products is zero, the predicted stress level is estimated to be approximately -1.1872. While this may not hold practical significance, it forms the baseline prediction within the model. The coefficient of 0.05696 for the monthly household expenditures on bakery products signifies that for every one-unit increase in spending on bakery items, stress levels among individuals with celiac disease tend to rise by approximately 0.05696 units. The regression equation obtained was:

$$\text{Stress Level} = -1.1872 + 0.05696 \times \text{Monthly Expenditures on Bakery Products}.$$

Surprisingly, the analysis revealed a positive correlation between increased monthly expenditures on bakery products and heightened stress levels within households affected by celiac disease. This suggests that as spending on bakery items escalates among individuals with dietary restrictions, reported stress levels also demonstrate an upward trend.

4 Conclusions

Our exploration through the complex web of bakery product consumption and its associated factors has been an illuminating journey, unraveling the intricate dynamics that govern our dietary choices and their broader implications. Beginning with the investigation into the link between bakery product consumption frequency and health diagnoses, our statistical analyses revealed a substantial and meaningful connection. This connection underscores the significant impact that health conditions have on shaping the consumption patterns of bakery products, emphasizing the critical role of informed dietary decisions, especially in the face of health challenges. Venturing deeper into the emotional realm, our exploration unearthed a fascinating correlation between the emotional responses evoked by bakery product consumption and respondents' physical activity levels and overall well-being. The diverse spectrum of emotions - from satisfaction to nostalgia and even guilt - highlights the complex interplay between dietary choices and human psychology. This exploration broadens our comprehension of bakery products, transforming them from mere sustenance to conduits of emotional experiences.

Moreover, our regression analysis introduced a compelling layer to our narrative, revealing an unexpected positive correlation between monthly expenditures on bakery products and stress levels among individuals with celiac disease. This intriguing finding beckons further inquiry into the psychological and emotional intricacies related to dietary restrictions and household spending behaviors. Understanding these correlations sheds light not only on the complex interplay between consumer behaviors and stress factors but also emphasizes the significance of considering individual dietary needs in shaping psychological well-being within household expenditure dynamics. Our comprehensive investigation has brought to light the multifaceted nature of bakery product consumption, transcending mere dietary preferences. These choices influence health outcomes, trigger emotional responses, impact physical activity, and wield influence over household budgets.

Reflecting on these insights, it becomes evident that bakery products aren't mere commodities but storytellers of human experiences, intricately woven with health, emotions, and financial implications. Our journey signifies a crucial step towards a deeper understanding of the nuanced world of bakery products and their profound impact on our lives. As we conclude this chapter of exploration, we acknowledge that our findings serve as a foundation for a more comprehensive comprehension of the captivating realm of bakery products, inviting continued inquiry into the intricate relationships they weave within the fabric of our lives.

Acknowledgement

The research presented in the article is supported by the VEGA project: Challenges for ensuring food security in Europe in the 21st century - key factors, socio-economic and environmental contexts, No VEGA 1/0755/21

References

Ariely, D., & Kreisler, A. (2007). The pull of the past: When do habits persist despite conflict with motives? *Personality and Social Psychology Bulletin*, 33(7), 917-928.

Greenacre, M. (2007*). Correspondence Analysis in Practice*. Boca Raton, FL: Chapman & Hall/CRC.

Chandon, P., & Hutchinson, J. W. (2007). The reduction of context effects in choice experiments: Experimental evidence. *Marketing Science*, 26(3), 393-402.

Chandon, P., & Wansink, B. (2007). Is obesity caused by calorie underestimation? A psychophysical model of meal size estimation. *Journal of Marketing Research*, 44(1), 84-99.

Kessler, D. A. (2009). *The end of overeating: Taking control of the insatiable American appetite.* Rodale Books.

Kessler, D. A. (2018). *Fast Carbs, Slow Carbs: The Simple Truth About Food, Weight, and Disease*. HarperOne.

Loewenstein, G., & Ariely, D. (2006). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, 132(5), 751-778.

Nestle, M. (2002). Food marketing and childhood obesity—a matter of policy. New England *Journal of Medicine*, 346(21), 1620-1622.

Nestle, M. (2013). *Food politics: How the food industry influences nutrition and health* (3rd ed.). University of California Press.

Rolls, B. J. (2010). The relationship between dietary energy density and energy intake. *Physiology & Behavior*, 100(5), 438-445.

Rolls, B. J., & Roe, L. S. (2002). Effect of the volume of liquid food infused intragastrically on satiety in women. *Physiology & Behavior*, 76(4-5), 623-631.

Rozin, P. (1999). Preadaptation and the puzzles and properties of pleasure. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*, 109-133. Russell Sage Foundation.

Rozin, P., & Fallon, A. E. (1987). A perspective on disgust. *Psychological Review*, 94(1), 23-41.

Wansink, B., & Sobal, J. (2007). Mindless eating: The 200 daily food decisions we overlook. *Environment and Behavior*, 39(1), 106-123.

Wansink, B., Painter, J. E., & North, J. (2005). Bottomless bowls: Why visual cues of portion size may influence intake. Obesity Research, 13(1), 93-100.

Wooldridge J.M. (2016). *Introductory Econometrics: A Modern Approach*. Cengage Learning.