Exploratory Analysis on Market Basket Data using Network Visualization

Henrique L. S. Gino[1], Diogenes S. Pedro[1], Jean R. Ponciano[1], Claudio D. G. Linhares[1], Agma J. M. Traina[1]

[1]Institute of Mathematics and Computer Sciences, University of São Paulo, São Carlos, Brazil

{henriqueshibata, diogenes.pedro, jeanponciano, claudiodgl}@usp.br
agma@icmc.usp.br

Abstract. *Market basket analysis is a powerful technique for understanding customer behavior and optimizing business strategies based on that understanding. Market-based analysis over time using visualization techniques can provide insights into market trends and relations, simplify complex data, and communicate insights effectively, which can help organizations make more informed decisions. This paper leverages a dataset focused on the users' incomes and temporal aspects of market purchases. We modeled this dataset as three distinct temporal networks and performed an exploratory evaluation identifying patterns and anomalies in the data. More specifically, we identified groups of related products, indicating thematic purchases, and evaluated the impact of demographic factors, such as income, on customer spending.*

1. Introduction

Market basket analysis is a data analysis technique used to identify the relationships between frequently purchased products or items by customers [Osadchiy et al. 2019]. Market basket analysis is based on the idea that customers who buy a particular item are also likely to buy other items. This approach is commonly used in marketing to examine customer buying patterns by identifying associations among various items that customers place in their shopping baskets [Arboleda et al. 2022].

There are several ways to explore market baskets, such as statistical analysis, predictive modeling, association rules, and visualization techniques. It is also possible to model this type of data using graphs and complex networks where nodes usually represent products, and the edges depict the correlation between them [Estrada 2015]. This way, the temporal behavior of purchases could be analyzed by mapping the evolution over time through temporal (dynamic) networks [Holme and Saramäki 2019].

Networks are ideal for representing complex relations and are highly visual, which makes them easy to understand and interpret when using proper visualization techniques [Zoss et al. 2018]. Visualization helps to identify patterns, trends, and anomalies that may be difficult to detect with other methods [Wilke 2019]. Moreover, the identification and analysis of meaningful groups of nodes, such as the so-called communities — groups of nodes that interact more often with themselves than with other nodes —, can help understand related nodes and common behaviors [Fortunato and Newman 2022].

In this paper, we are particularly interested in analyzing whether purchase behavior changes according to household income. To investigate that, we explore market baskets using Network Visualization techniques and inspect the social network structure (e.g., the relation between products bought by customers) and the temporal evolution (e.g., trends and seasonality factors). We take advantage of an established system focused on the visual analysis of large temporal networks called LargeNetVis [Linhares et al. 2023], which is suitable for exploring patterns and anomalies using network communities. To detail our approach, we also took advantage of a dataset focused on the temporal network aspects and three household income levels.

2. Background and Related Work

This section describes the basic concepts and works related to Market Basket Analysis and Network Visualization, including the most relevant layouts for temporal networks in the scope of our study and related graph drawing techniques.

2.1. Market Basket Analysis

Many approaches can be used with Market Basket Analysis to assert better decision-making to the retailer about its buyers. The first study considering this theme date back to the early 1990s, in which data mining techniques were applied to a large database of customer transactions to determine relevant association rules between customer interactions [Agrawal et al. 1993]. An association rule is a machine-learning technique for detecting important relations between dataset variables. In the context of Market Basket Analysis, association rules are used to identify relations between products that are frequently purchased together [Arboleda et al. 2022]. They were proposed with the specific goal of discovering different association rules between items in an extensive database of transactions. The paper [Ünvan 2021] applies association rules in a static supermarket data set with 225 different products to address the market basket problem, creating rules that state the set of items most frequently purchased together with the highest confidence. The paper [Kaur and Kang 2016] employs periodic mining to enhance the prediction of future association rules and designs a suitable methodology to find outliers, better-assert customer behavior, and increase sales.

Customer purchasing data applied to association rules were employed by [Osadchiy et al. 2019] to create a recommendation system for market food items that are not inherently user behavior-based or rating-based but used transactions from a given population to build a collective model of preferences. Similarly, the work at [Rendle et al. 2010] used customer purchase data to propose a next-basket recommendation system that essentially uses sequential behavior to predict items commonly bought together. Some other works take a different direction, such as the paper [Griva et al. 2018], which considers not only the information of purchased items, but also the spatial localization of individual customers to define behavior in a store concerning aisles, path types, and time spent buying. However, association rules have some limitations in Market Basket Analysis, such as the sparsity of the data, which can be difficult to generate meaningful association rules or to achieve computational scalability for large datasets [Arboleda et al. 2022]. To avoid that, some approaches of market basket analysis focus on enhancing the association rules through visualization tools, to better select relevant information [Valle et al. 2018].

In that direction, some studies propose the application of graphs and complex networks in Market Basket Analysis. The advantage of using graphs is that the complexity of the dataset is reduced, improving its analysis. In this case, the graph is modeled as a network of products, with each node representing a product and each edge connecting two products bought together (though an edge in a network does not imply a confirmed relation between products). The work presented at [Kafkas et al. 2021] used product networks to discover customers' purchase patterns, segmenting the products in communities through statistical inference to understand their roles in the network. Considering e-commerce environments and product networks, the article [Oestreicher-Singer et al. 2013] tested the product's influence over another product's sales, attaining a metric to estimate a product's actual value in a general product network. In the retail context, the paper [Lismont et al. 2018] analyzed product attrition with transaction data and customer-product networks, defining a strategy to identify products that would be sold significantly less. Moreover, some studies focus on transaction data through networks to create network mining techniques effective in the vast amounts of scattered data [Videla-Cavieres and Ríos 2014].

Other approaches combine multiple techniques, such as association rules, complex networks, and visualization analysis [Wu et al. 2021]. Even more, they also use the concept of network communities, which can highlight groups of products with some compatibility (e.g., chips and salsa). An objective metric of interest is used to determine the intensity of relations in such communities. Also, they use association rules networks to better detect communities whose relations are more latent to discover and explore individual hypotheses regarding products in the network [Wu et al. 2021]. Although some recent studies present network visualization techniques [Huang et al. 2019, Wu et al. 2021], none showcase market baskets or customer-based product networks with a temporal or dynamic aspect, which is essential to analyze seasonal purchases and, ultimately, increases/decreases in the purchasing power of a population. Our work aims at bridging this gap.

2.2. Network Visualization

There are several techniques (or layouts) to visualize networks. One of the most popular is the node-link diagram, which depicts nodes and edges as circles and straight lines, respectively (Fig. 1(A)) [Abdelaal et al. 2020]. Another popular technique is the matrix-based layout, in which columns and rows represent the nodes, and each marked cell represents an edge (Fig. 1(B)). There are several studies focused on graph drawing techniques to improve the readability of node-link diagrams and matrix-based representations. One famous example is the force-directed layout for node-link diagrams, which is a node positioning algorithm that considers each node as a physical object that repels other nodes and each link (i.e., edge) as a spring-like force that attracts connected nodes, improving the visualization and interpretation [Rahman et al. 2022]. Other popular graph drawing solutions for node-link diagrams include circular and tree layouts and edge bundling [Vieira et al. 2022].

Both node-link diagrams and matrix-based layouts can be used for temporal network visualization when animating them over time or presenting the network data timestamp by timestamp in side-by-side windows (a strategy called small multiples) [Abdelaal et al. 2020]. Besides these approaches, there are timeline layouts proposed specifically for temporal network visualization. Popular timeline representations include Massive Sequence View (MSV), in which each line represents a different node and each column indicates the graph activity (edges) for a specific timestamp (Fig. 1(C)) [van den Elzen et al. 2014]. Similar to MSV, the timeline technique Temporal Activity Map (TAM) focuses on the node activity using squares to represent the nodes and removing the vertical lines (edges) (Fig. 1(D)) [Linhares et al. 2020]. Popular solutions for timeline representations include node ordering [van den Elzen et al. 2014, Linhares et al. 2019] and edge sampling [Ponciano et al. 2021].

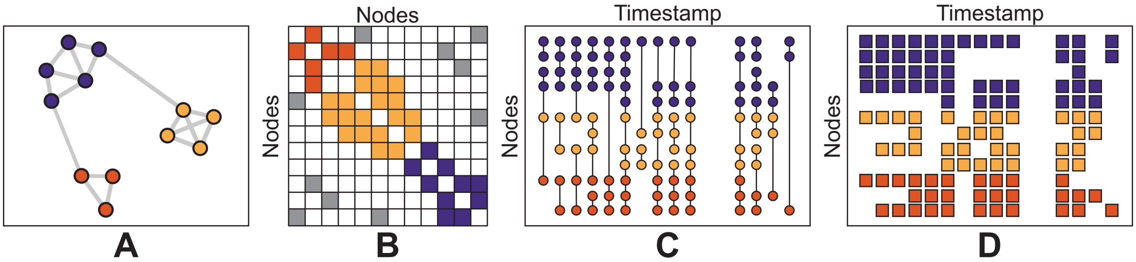

Figure 1. Techniques for network visualization: (A) Node-Link Diagram; (B) Matrix layout; (C) Massive Sequence View; (D) Temporal Activity Map.

3. Methodology

We propose a methodology to perform our market basket analysis (Fig. 2). First, we describe a literature dataset focused on household-level transactions and our data preprocessing steps (Fig. 2(A, B)), highlighting the data characteristics. Then, we model the resulting dataset as temporal networks (Fig. 2(C)), which allows a better comprehension of the product relationships and temporal evolution. After, we describe the system and network visualization techniques used to perform the intended exploratory analyses (Fig. 2(D)), to identify patterns, trends, and anomalies in the data. At last, the dataset is ready to be explored, with applications demonstrated in our case study (Fig. 2(E)), to identify groups of related products, thematic purchases, and the impact of demographic factors on customer spending.

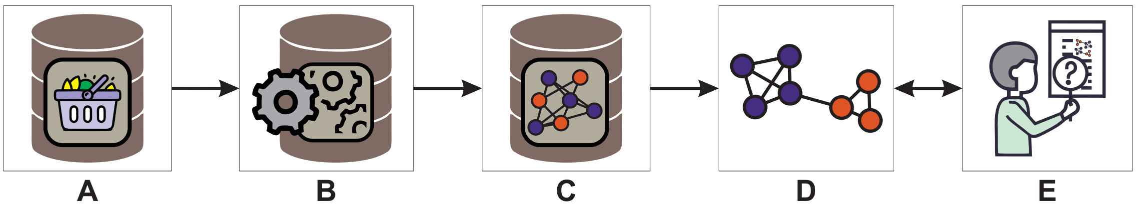

Figure 2. Workflow steps to perform exploratory analysis on market baskets: (A) Select a market basket dataset; (B) Preprocess the data; (C) Transform the resulting dataset into temporal networks; (D) Apply visualization techniques; (E) Explore the data to identify patterns, trends, and anomalies.

3.1. Dataset and data preprocessing

This work explores the Dunnhumby dataset [Gonen 2020], which contains information on household-level transactions over two years from 2500 households who are frequent shoppers at a retailer. All products from the respective households purchase are in the dataset, not limited to specific categories. Demographic information about some households are also considered, e.g., annual income, number of members, and member age.

The products are considered under three levels of detail: department, product category, and product type. The upper level of detail is the department, e.g., grocery; the next is the product category, e.g., cookies/cones; and the last is the product type, e.g., tray pack/choc chip cookies. The dataset encloses 92,353 distinct products distributed in 44 departments, 308 product categories, and 2383 product types.

Each purchase is organized with the *basket* label, categorizing groups of products bought on the same shopping trip. There is information about the cost of each product, the number of items, and whether some discount was used. There are 276,484 orders from the two years, averaging 110 orders per household and 9.4 items bought per order.

Some preprocessing was needed to properly use the Dunnhumby dataset in our study. First, since we are mainly interested in investigating whether the purchase behavior changes according to household income, we consider only family shopping trips with that (demographic) information. Given that the annual household income was initially considered under numerous classes (e.g., ‘Under 15K’, ‘15-24K’, ...), a new attribute was created to delineate a household’s income in only three categorical classes: low, middle, and high annual income. The defining ranges to create such classes were [0 to 50k – low], [$>$ 50k to 150K – middle], and [$>$ 150K – high]. We have 48% of households with low annual income, 45% middle income, and 7% with high income values.

We also concentrate our analysis on the first year of data and grocery items, which roughly represent 2/3 of products from all orders. Based on quantiles of the product price distribution, we categorize each product as having a very low, low, normal, high, or very high price (defining ranges [0 to .25 – very low], [$>$ .25 to .5 –low] [$>$ .5 to .75 – normal], [$>$ .75 to .9 – high], and [$>$ .9 to 1 – very high]). Naturally, some products could have their prices changed throughout the considered time. The most recurrent label was chosen for radical cases of a price difference that caused the same product to have different prices.

3.2. Network modeling

After preprocessing, the dataset was used to create three temporal, unweighted, and undirected networks such that nodes represent (types of) products and edges connect pairs of (types of) products bought together in the same basket. The difference between the networks relies on the household income category being analyzed. Tab. 1 presents the networks that depict purchases made by households with low, middle, and high annual incomes. Each network contains 346 days (the first year of data, as previously stated, only missing the first nineteen days in the original dataset). Note that the *high-income network* presents fewer nodes (products) and edges when compared with the others. This occurs because the original dataset presented only 7% of households with high-income values. The networks and a detailed explanation of the dataset can be accessed at `https://github.com/henrique-gino/MarketBasketDataset`.

Table 1. Three proposed networks based on household incomes.

Network	# Nodes	# Edges	# Timestamps
Low-income network	9,450	387,031	346 (days)
Middle-income network	9,859	470,224	346 (days)
High-income network	3,071	46,410	346 (days)

3.3. Visual Analysis

We have considered different interactive systems for visualizing temporal networks (e.g., DyNetVis [Linhares et al. 2020] and PaohVis [Valdivia et al. 2021]). After an initial investigation, we chose LargeNetVis [Linhares et al. 2023] because it enables effective analyses of temporal aspects and also can handle temporal networks with a few thousand nodes and edges, such as the ones described in Tab. 1, which would be difficult or infeasible with other systems. LargeNetVis's features and layouts include (i) network timeslicing (i.e., the split of the network data into time intervals of equal length) followed by community detection performed inside each timeslice using Louvain [You et al. 2022]; (ii) a node-link diagram per community that enables the visual analysis of the community's structure (Fig. 3(B)) and a summarized version of this diagram where nodes (also known as "super" nodes) represent sub-communities (Fig. 3(A)); and (iii) a Temporal Activity Map (TAM) per community (Fig. 3(C)) that enables the analysis of temporal patterns in that community (see, e.g., the three nodes with labels "High" and "Very high" that are involved in connections during six consecutive timestamps (Fig. 3(C), I) and the existence of two timestamps with no nodes (Fig. 3(C), II)).

LargeNetVis's layouts are interactive and coordinated among them so that nodes (or super nodes) selected in one of them are automatically selected in the others. Nodes can be colored according to the available metadata; in this case, super nodes are colored and labeled according to the predominant color and label among their node members. As an example, the "Low"/yellow super node in Fig. 3(A) has this color because yellow (which corresponds to the label "Low") is predominant in the corresponding sub-community (note that this sub-community contains three "Low"/yellow nodes, one "Very low"/red, and one "Normal"/white node — see Fig. 3(B)). The size of a super node and the edge thickness in the summarized node-link diagram are related to the number of nodes in the sub-community and the number of edges connecting nodes from different sub-communities, respectively (Fig. 3(A, B)). Both versions of the node-link diagram employ force-directed layouts (recall Sec. 2.2), which impacts the pattern identification. In our context, nodes close to each other are more likely to be purchased together in the same basket of a single household (see, for example, the blue nodes that refer to high- and very high-priced products in Fig. 3(B)).

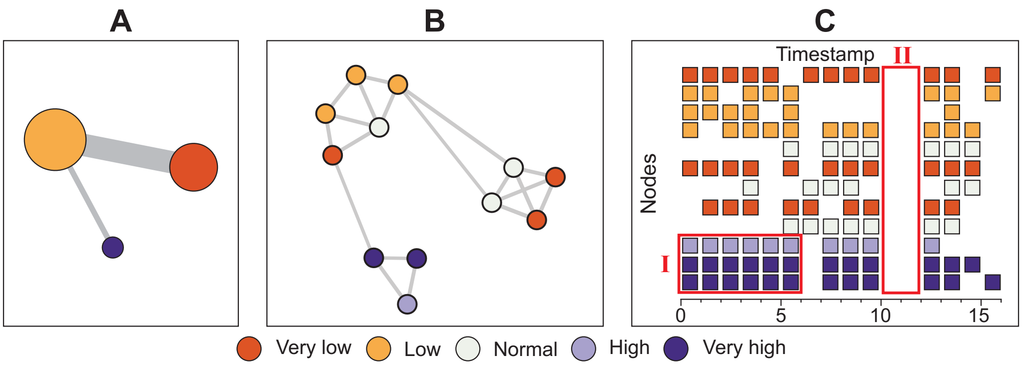

Figure 3. Community-focused layouts offered by LargeNetVis: (A) Summarized Node-link diagram; (B) Node-link diagram; (C) Temporal Activity Map.

After preliminary tests, we configured LargeNetVis so that each timeslice comprises a 31-day interval (12 timeslices). We also set the node metadata as our price categorization, using the same labels and color scale from Fig. 3.

4. Case study

In this case study, we are mainly interested in analyzing (i) if there are differences regarding the types or categories of products bought by households from different income classes and (ii) if the household income class affects the frequency in which a product is bought.

Fig. 4 illustrates five (summarized) node-link diagrams for each income class. As expected, very expensive products are typically bought by high-income households (a behavior generally found on the *high-income network* and illustrated in Fig. 4(A, 1-4)). Similarly, products categorized as having "normal" prices are mainly found on purchases from middle-income households (Fig. 4(B, 1-4)), and very cheap products are predominant in low-income households' baskets (Fig. 4(C, 1-4)).

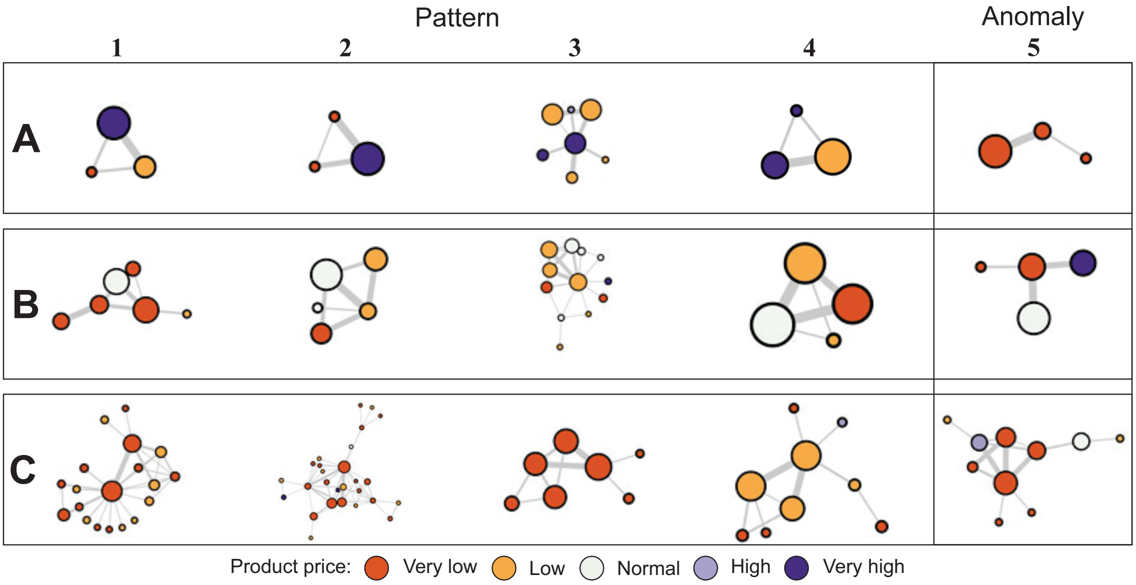

Figure 4. Summarized node-link diagrams showing communities containing (1-4) patterns or (5) anomalies: (A) High; (B) Middle; and (C) Low incomes.

It is important to highlight that exceptions from the aforementioned behavior are relevant and can also be found using the visualization. One of them is shown in Fig. 4(A, 5), where we see high-income households buying very cheap products (probably those of basic needs). Analogously, we can also observe middle- and low-income households buying more expensive products (Fig. 4(B-C, 5)). To help us understand a case where low-income households purchase high-priced products, Fig. 5(B) shows the products (nodes) that belong to the super node from Fig. 4(C, 5) with most products having high prices (see this super node highlighted in Fig. 5(A)). Besides other products, we can see that this low-income household has bought seven different types of frozen meat, beverages, and paper drinking cups. Inspired by this case and after finding other similar purchases, we could identify many thematic purchases where more expensive products are bought to be used in celebrations or other special occasions.

Thematic purchases are observed in all three networks. We can see, for example, that most purchases made by high-income households contain alcoholic beverages, mainly wine, and products often consumed with these beverages. This pattern is illustrated in Fig. 6 — which depicts the node members of the community shown in Fig. 4(A, 1) —, where we have different types of wine that were probably bought by the same household (given that they are close to each other in the layout). Seltzers are also commonly found on purchases made by households with high incomes. In the example depicted in Fig. 6, the same type of seltzer was bought on 19 days of the month (Fig. 6(B)).

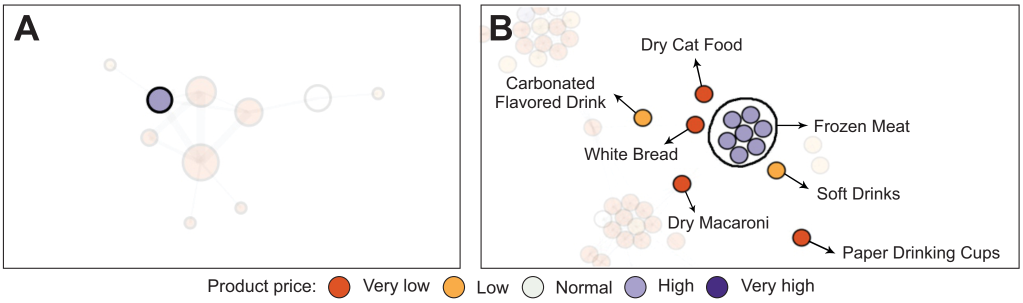

Figure 5. Analysis of a community with purchases made by a low-income household (the same community depicted in Fig. 4(C, 5)). (A) Summarized node-link diagram and (B) the corresponding node-link diagram.

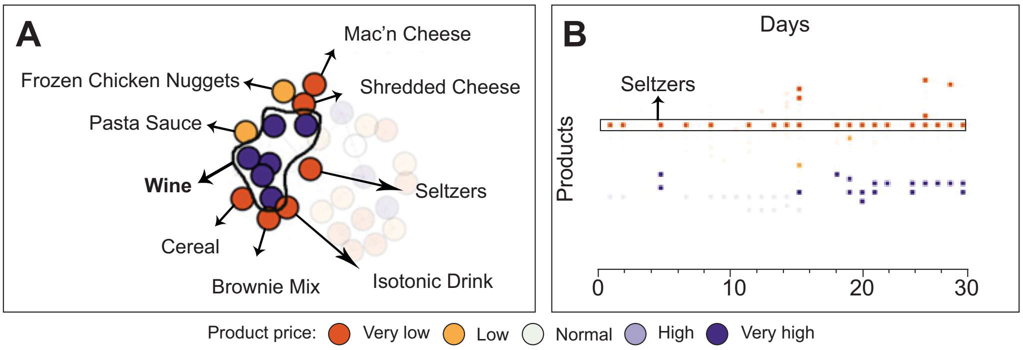

Figure 6. Analysis of a community with purchases made by high-income households (the same community depicted in Fig. 4(A, 1)). (A) Node-link diagram and (B) Temporal Activity Map (TAM).

While it is expected that different households from the same income level have different preferences and hence purchase different types of products, some of them may opt for purchasing the same products. These shared types of products are represented by *connection nodes* that link different clusters of products (baskets) in the node-link diagrams, as illustrated in Fig. 7. Surprisingly, the types of products shared by high-income households' baskets refer to products that tend to be cheaper than those shared by both middle- and low-income baskets, as we can see when comparing Figs. 7(A-C). The type of product shared by different baskets also varies according to the household income class, from some expendable products in high-income baskets to day-to-day cooking products (e.g., spices and vegetable oil) in low-income baskets. In general, most of the shared types of products found during our analysis refer to products of daily use.

Regarding the frequency in which types of products are purchased monthly, we could not establish a relation between frequency and household income class. Different shopping behaviors were found in that sense, regardless of the income. First, there are types of products that were bought on many days in a month, but belonging to baskets with few products, as illustrated in Fig. 8(A, 1), which highlights three types of products (see red boxes): a very cheap one that is bought throughout the whole month (seltzers), and two very expensive products, one that is mainly bought on the first half of the month and one mainly bought near the end of the month (two different types of wines). Similarly, there are also recurrent purchases of a product as part of baskets with several other products (see, e.g., the number of products per day in Fig. 8(A, 2)).

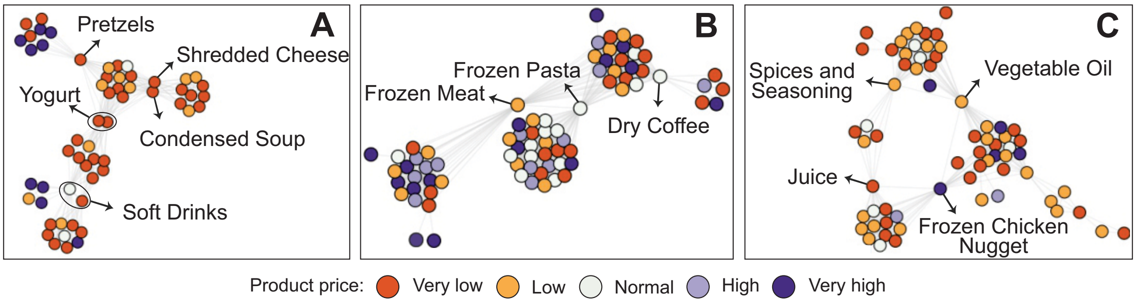

Figure 7. Analysis of *connection nodes*, i.e., products shared by different baskets. Each node-link diagram shows a community from a different network: (A) High-income; (B) Middle-income; and (C) Low-income.

Another shopping behavior refers to types of products that were bought on few and spaced days in a month, possibly indicating seasonal purchases (people buy them only on weekends, for example), as illustrated in Figs. 8(B, 1-2) (see the red boxes). Finally, there are also types of products that were bought once or a few times in a month, as illustrated in Figs. 8(C, 1) and 8(C, 2) (see the red boxes). Besides the immediate identification of seasonal purchases, the investigation of temporal shopping behaviors through the proposed visual analysis can be used to assist retailers and analysts in recognizing increases or decreases in the households' purchasing power over time, thus enabling fast and appropriate decision-making to encourage/discourage those behaviors (e.g., through marketing campaigns in e-commerce platforms).

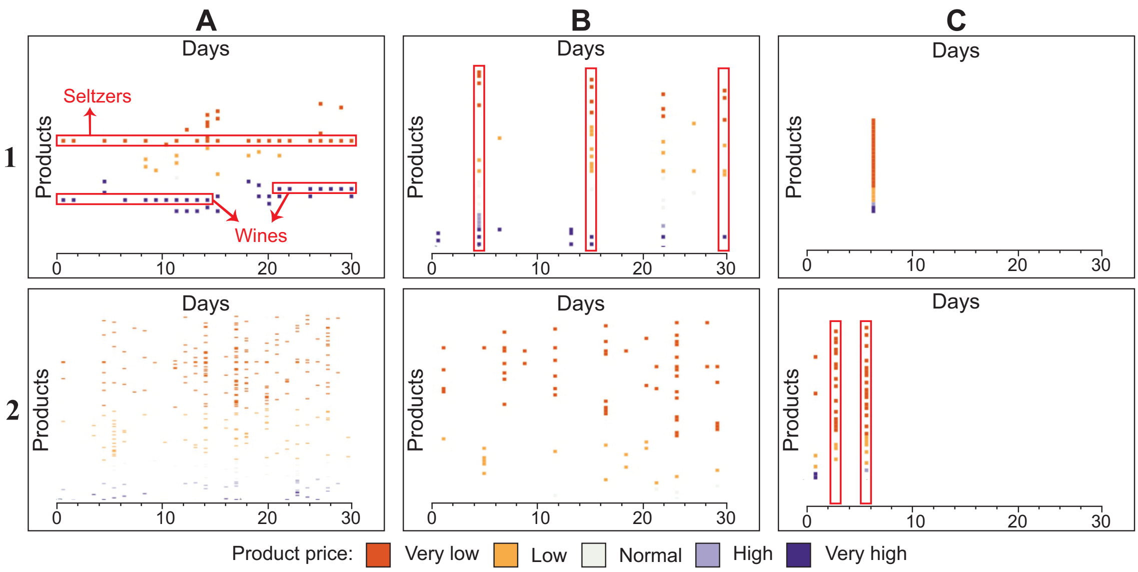

Figure 8. Temporal Activity Maps showing different temporal shopping behaviors. (A) Types of products that were bought on many days in a month but belonging to baskets with (1) few or (2) several products. (B, 1-2) Types of products that were bought on few and spaced days in a month, possibly indicating seasonal purchases. (C) Types of products that were bought (1) once or (2) a few times in a month.

5. Limitations and Future Work

Using visualization techniques for market basket analysis is valuable because it allows identifying patterns and relations in the data that might not be apparent from other analysis strategies. However, adequately validating such a visual exploration approach can be difficult to achieve. In that regard, a second iteration of this study will include validating the quality and applicability of the visualization results with objective-oriented metrics through user study settings and quantitative evaluations.

Since we analyzed a dataset potentially focused on a specific geographical region, it is hard to generalize the results of related products and customer behaviors for other regions. Moreover, our analysis also considered only individual network communities, resulting in edges across communities being lost, which could have potentially affected the findings [Ponciano et al. 2021]. Lastly, the proposed network analysis derives from a dataset with only different households' income levels relating to the prices of purchased products. Considering other factors (e.g., customers' age, number of family members, and place of residence (country, state, city)), or even utilizing different network architectures (e.g., modeling the dataset as a temporal multilayer network and visualizing it using MuxViz [De Domenico et al. 2014]) could enrich our analysis and lead to new and more assertive conclusions, and is on-par with the future iterations of this study.

6. Conclusion

Market Basket Analysis is a data analysis technique used to identify relations between items purchased by customers, and it is a powerful tool for understanding customer behavior and making data-driven decisions. In this paper, we investigated the effects of annual household income on purchase behavior. To do that, we took advantage of a dataset with market basket information that we modeled as temporal networks focused on purchases of products from the perspective of three distinct groups of household customers, characterized by their annual income. Using LargeNetVis [Linhares et al. 2023], an interactive system with layouts and features to analyze small and large temporal networks, we then performed a series of exploratory analyses that allowed us to inspect structure (identifying, e.g., purchase patterns and relations between items and users' preferences) and temporal dynamics (e.g., purchases trends and seasonalities).

Our case study described meaningful purchase behaviors identified through the proposed analysis procedure, including existing relations involving household incomes and types (and prices) of products bought, frequency of purchases over time, and the presence of thematic purchases. As demonstrated throughout the paper, our approach can greatly help domain experts (data or business analysts, among others) by leading them to new insights into the data and supporting faster and more reliable decision-making, for example, in marketing campaigns, product pricing/placement, and recommendations. As previously stated, we now intend to expand our analysis by considering other demographic factors, using other complementary tools to perform analyses, and assessing usability and usefulness through quantitative and qualitative (user study) evaluations.

Acknowledgment

This research was supported by grants #2020/10049-0, #2020/07200-9, #2022/13190-1, and #2016/17078-0 from São Paulo Research Foundation (FAPESP), by grant #2022-2094 from Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (CNPq) and by Coordenação de Aperfeicoamento de Pessoal de Nivel Superior (CAPES).

References

Abdelaal, M., Lhuillier, A., Hlawatsch, M., and Weiskopf, D. (2020). Time-aligned edge plots for dynamic graph visualization. In *2020 24th International Conference Information Visualisation (IV)*, pages 248–257.

Agrawal, R., Imieliński, T., and Swami, A. (1993). Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD international conference on Management of data*, pages 207–216.

Arboleda, F., Ortega, G., and Luna, J. (2022). Temporal visual profiling of market basket analysis. *IAENG International Journal of Computer Science*, 49(2).

De Domenico, M., Porter, M. A., and Arenas, A. (2014). MuxViz: a tool for multilayer analysis and visualization of networks. *Journal of Complex Networks*, 3(2):159–176.

Estrada, E. (2015). *Introduction to Complex Networks: Structure and Dynamics*, pages 93–131. Springer International Publishing, Cham.

Fortunato, S. and Newman, M. E. J. (2022). 20 years of network community detection. *Nature Physics*, 18(8):848–850.

Gonen, F. (2020). Dunnhumby - The Complete Journey. `https://www.kaggle.com/datasets/frtgnn/dunnhumby-the-complete-journey`. Accessed: 2023-02-23.

Griva, A., Bardaki, C., Pramatari, K., and Papakiriakopoulos, D. (2018). Retail business analytics: Customer visit segmentation using market basket data. *Expert Syst. Appl.*, 100:1–16.

Holme, P. and Saramäki, J. (2019). *Temporal Network Theory*. Springer Cham.

Huang, H., Yao, L., Chang, J.-S., Tsai, C.-Y., and Kuo, R. (2019). Using product network analysis to optimize product-to-shelf assignment problems. *Applied Sciences*, 9(8).

Kafkas, K., Perdahçı, Z. N., and Aydın, M. N. (2021). Discovering customer purchase patterns in product communities: an empirical study on co-purchase behavior in an online marketplace. *J. Theor. Appl. Electron. Commer. Res.*, 16(7):2965–2980.

Kaur, M. and Kang, S. (2016). Market basket analysis: Identify the changing trends of market data using association rule mining. *Procedia computer science*, 85:78–85.

Linhares, C. D. G., Ponciano, J. R., Paiva, J. G. S., Rocha, L. E. C., and Travençolo, B. A. N. (2020). DyNetVis - an interactive software to visualize structure and epidemics on temporal networks. In *2020 ASONAM Conference*, pages 933–936.

Linhares, C. D. G., Ponciano, J. R., Pedro, D. S., Rocha, L. E. C., Traina, A. J. M., and Poco, J. (2023). LargeNetVis: Visual exploration of large temporal networks based on community taxonomies. *IEEE Trans. on Vis. and Comput. Graphics*, 29(1):203–213.

Linhares, C. D. G., Ponciano, J. R., Pereira, F. S. F., Rocha, L. E. C., Paiva, J. G. S., and Travençolo, B. A. N. (2019). A scalable node ordering strategy based on community structure for enhanced temporal network visualization. *Comput. & Graph.*, 84:185 – 198.

Lismont, J., Ram, S., Vanthienen, J., Lemahieu, W., and Baesens, B. (2018). Predicting interpurchase time in a retail environment using customer-product networks: An empirical study and evaluation. *Expert Syst. Appl.*, 104:22–32.

Oestreicher-Singer, G., Libai, B., Sivan, L., Carmi, E., and Yassin, O. (2013). The network value of products. *Journal of Marketing*, 77(3):1–14.

Osadchiy, T., Poliakov, I., Olivier, P., Rowland, M., and Foster, E. (2019). Recommender system based on pairwise association rules. *Expert Syst. Appl.*, 115:535–542.

Ponciano, J. R., Linhares, C. D., Rocha, L. E., Faria, E. R., and Travençolo, B. A. (2021). A streaming edge sampling method for network visualization. *KAIS*, 63(7):1717–1743.

Rahman, M. K., Sujon, M. H., and Azad, A. (2022). Scalable force-directed graph representation learning and visualization. *KAIS*, 64(1):207–233.

Rendle, S., Freudenthaler, C., and Schmidt-Thieme, L. (2010). Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th international conference on World wide web*, pages 811–820.

Valdivia, P., Buono, P., Plaisant, C., Dufournaud, N., and Fekete, J.-D. (2021). Analyzing dynamic hypergraphs with parallel aggregated ordered hypergraph visualization. *IEEE Trans. on Vis. and Comput. Graphics*, 27(1):1–13.

Valle, M. A., Ruz, G. A., and Morrás, R. (2018). Market basket analysis: Complementing association rules with minimum spanning trees. *Expert Syst. Appl.*, 97:146–162.

van den Elzen, S., Holten, D., Blaas, J., and van Wijk, J. (2014). Dynamic network visualization with extended massive sequence views. *IEEE Trans. on Vis. and Comput. Graphics*, 20:1087–1099.

Videla-Cavieres, I. F. and Ríos, S. A. (2014). Extending market basket analysis with graph mining techniques: A real case. *Expert Syst. Appl.*, 41(4, Part 2):1928–1936.

Vieira, R. S., Do Nascimento, H. A. D., Ferreira, J. M., and Foulds, L. (2022). Clustering ensemble-based edge bundling to improve the readability of graph drawings. In *2022 26th International Conference Information Visualisation (IV)*, pages 21–26.

Wilke, C. (2019). *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. O'Reilly.

Wu, J., Wang, Y., Shafiee, S., and Zhang, D. (2021). Discovery of associated consumer demands: Construction of a co-demanded product network with community detection. *Expert Syst. Appl.*, 178:115038.

You, Y., Ren, L., Zhang, Z., Zhang, K., and Huang, J. (2022). Research on improvement of Louvain community detection algorithm. In Zhu, L., editor, *2nd AIAHPC 2022*, volume 12348. International Society for Optics and Photonics, SPIE.

Zoss, A., Maltese, A., Uzzo, S. M., and Börner, K. (2018). *Network Visualization Literacy: Novel Approaches to Measurement and Instruction*, pages 169–187. Springer International Publishing, Cham.

Ünvan, Y. A. (2021). Market basket analysis with association rules. *Communications in Statistics - Theory and Methods*, 50(7):1615–1628.