

## INSIGHTS INTO E-SHOPPER PREFERENCES: AN INTERACTIVE VISUAL ANALYTICS OF MALAYSIAN ONLINE SHOPPING BEHAVIOURAL DATA

Siti Sakira Kamaruddin <sup>1</sup>  
Nurul Ain Mustakim<sup>2</sup>  
Maslina Abdul Aziz<sup>3</sup>  
Shuzlina Abdul-Rahman <sup>4</sup>

<sup>1</sup> School of Computing, Universiti Utara Malaysia, 06100 Sintok, Kedah Malaysia,  
(E-mail: [sakira@uum.edu.my](mailto:sakira@uum.edu.my))

<sup>2</sup> Faculty Bussiness & Management, Universiti Teknologi Mara, Melaka Malaysia,  
(E-mail: [ainmustakim@uitm.edu.my](mailto:ainmustakim@uitm.edu.my))

<sup>3</sup> Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Selangor Malaysia,  
(E-mail: [maslina\\_aziz@uitm.edu.my](mailto:maslina_aziz@uitm.edu.my))

<sup>4</sup> Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Selangor Malaysia,  
(E-mail: [shuzlina@uitm.edu.my](mailto:shuzlina@uitm.edu.my))

### Article history

**Received date** : 11-7-2025  
**Revised date** : 12-7-2025  
**Accepted date** : 7-9-2025  
**Published date** : 25-9-2025

### To cite this document:

Kamaruddin, S. S., Mustakim, N. A., Abdul Aziz, M., & Abdul-Rahman, S. (2025). Insights into e-shopper preferences: Interactive visual analytics of Malaysian online shopping behavioural data. *Journal of Islamic, Social, Economics and Development (JISED)*, 10 (76), 720 – 732.

---

**Abstract:** *With advancements in internet technology, secure online payment systems, and faster delivery methods, consumers are increasingly turning to online shopping as their preferred purchasing method. To remain competitive in this global industry, businesses must understand customer behaviour to meet the growing demand for online purchasing. While previous studies on online shopping behaviour primarily focused on empirical, statistical, and regression analyses, visual analytics has received less attention. This paper addresses this gap by developing an interactive dashboard that models online shopping behavioural data. Following the data visualisation pipeline, a four-step methodology was implemented to develop the dashboard consisting of data selection and pre-processing, data transformation, visual mapping and visualisation generation. The dashboard supports tasks such as data aggregation, clustering, and filtering, offering a comprehensive view of shopping behaviours. Users can interact with various tabs to explore visualisations of product categories by income level, age, and gender, as well as average shopping time by demographics, aiding decision-makers in gaining valuable insights into e-shopper preferences.*

**Keywords:** *Visual Analytics, Online Shopping Behavior, Consumer Preferences, Interactive Dashboard, E-Commerce*

---

## Introduction

Retail e-commerce has rapidly evolved into a global industry. In 2023, global retail e-commerce sales reached 5.8 trillion U. S. dollars worldwide, with projections indicating an increase of over 39% growth in the coming years and expected to surpass 8 trillion U.S. dollars by 2027 (Chevalier, 2024). The growth in online retail is driven by advancements in internet technology, enhanced online payment security, and faster delivery methods. Additionally, the COVID-19 pandemic has significantly accelerated the demand for online shopping, further establishing it as a preferred purchasing method among consumers. Understanding customers' preferences and online behaviour has thus become crucial for businesses to remain competitive, relevant, and profitable in an increasingly digital and customer-centric marketplace.

Most of the existing research on online shopping behaviour has relied on empirical studies (Mohamad Shariff & Abd Hamid, 2021), descriptive statistical analysis (Mustakim et al., 2024), and regression modelling (Ahmad Shakir & Adzhar, 2024) to explore consumer behaviour. While these methods provide useful insights, they are limited in capturing the multi-dimensional and dynamic nature of consumer behaviour. Despite the wealth of such studies, interactive visual analytics remains underexplored as a tool to analyse online shopping patterns. This lack of effective visual analysis creates a gap in understanding how factors such as income, gender, and product categories interact in shaping consumer behaviour. Addressing this gap would benefit businesses by enabling data-driven marketing strategies, policymakers by supporting evidence-based digital economy planning, and marketers by tailoring personalised campaigns to evolve consumer needs. Visual analytics systems are essential and have been proposed for exploring complex datasets and uncovering hidden insights across various domains, including health (Jiang et al., 2024), education (Chen et al., 2024), and customer behaviour (Pragarauskaite & Dzemyda, 2012).

In the realm of behavioural analysis, several studies have focused on visualising online interactions. These include systems for hierarchical user profile analysis (Nguyen et al., 2020), improving online shopping advertising (Liu et al., 2021), and web content mining (Razali et al., 2021). While these works contribute significantly to understanding user behaviour, visualisation of behavioural data is not the primary analytical method in these studies. There remains a gap in effective visualisations that can comprehensively depict and analyse consumer preferences. This paper addresses this gap by presenting a dashboard to visualise online shopping behaviour in the Malaysian context, focusing on factors such as income group, gender, time spent online, and product categories. Interactive visual analytics techniques are employed to provide deeper insights into consumer behaviour and contribute to more customised decision-making.

The following sections provide an overview of related work and details of the material and method used in the study. The visualisation results and discussion are then presented, and the paper concludes with remarks on the implications of this research, offering a structured exploration of online shopping preferences and customer behaviours.

## Related Work

Visual Analytics systems are designed to explore and understand complex datasets, offering insights that are often hidden in traditional analysis. Recent studies have demonstrated their value in health (Jiang et al., 2024), education (Chen et al., 2024) and customer behaviour (Pragarauskaite & Dzemyda, 2012). However, within the behavioural analysis domain, most studies emphasise specific aspects rather than comprehensive consumer profiling. For example, (Nguyen et al., 2020) introduced Vasabi, a system for hierarchical user profile analysis that enables detailed exploration of user behaviour across domains. Similarly, (Liu et al., 2021) present MulUBA, which applies a multi-level visual analytics approach to improve online shopping advertising, while (Razali et al., 2021) used web content mining to visualise shopping preferences. These methods, while innovative, primarily focus on advertising optimisation or trend detection rather than offering a holistic, multi-dimensional view of consumer behaviour.

In the Malaysian context, studies such as Mohamad Shariff & Abd Hamid (2021) employed empirical surveys to examine consumer behaviour during the COVID-19 pandemic, while Mustakim et al. (2024) and Yong et al. (2023) applied descriptive statistical analysis to identify factors influencing online purchasing behaviour. Ahmad Shakir & Adzhar (2024) extended this by using regression modelling to assess purchase intention predictors. Although these works provide valuable insights, they remain limited by their methodological reliance on static analysis. They do not enable interactive exploration of multi-dimensional variables such as income, gender, and product categories simultaneously.

Therefore, while prior studies advance knowledge through surveys, statistics, and regression, they underscore the need for interactive visual analytics that can integrate and dynamically visualise diverse consumer attributes. This study addresses that gap by applying visual analytics to Malaysian online shopping data, offering decision-makers deeper insights for personalised marketing, targeted policymaking, and enhanced consumer engagement.

## Material and Method

The analysis and visualisation of the online shopping behaviour data were implemented using Python, renowned for its extensive libraries and frameworks in data science and web development. The specific libraries used in developing the visualisation are Pandas for data manipulation and analysis, Plotly for creating interactive and visually appealing charts and graphs and Dash for developing an interactive dashboard. The method used in this study contains four steps as shown in Fig. 1.

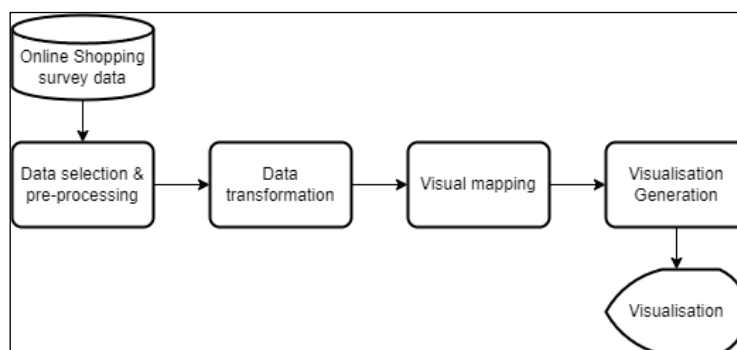


Figure 1: Method

### Data Selection and Pre-Processing

The data for this study is obtained from (Mustakim et al., 2024). They used convenience sampling to gather data from 560 respondents who met the following requirements: they had to be Malaysian citizens, older than 18 and had used an e-commerce site at least once. The obtained data was analysed to perform data selection. This process involves identifying the subset of the data that will be potentially visualised. The selected data is shown in Table 1 with data type, description and examples.

**Table 1: Data Description**

Attribute	Type	Description	Example
'Date'	Date	The date of the response	'2024-08-18'
'Email'	String	The email address of the respondent.	'example@example.com'
'Gender'	Categorical	The gender of the respondent.	'Female', 'Male'
'Age'	Categorical	The age range of the respondent.	'18-24', '25-34'
'Level of Education'	Categorical	The highest education level attained by the respondent.	'Bachelor's', 'Masters's'
'Ethnicity'	Categorical	The ethnicity of the respondent.	'Malay', 'Chinese'
'Monthly Income'	Categorical	The monthly income range of the respondent.	'Below RM5,000', 'RM5,000-RM10,000'
'Employment Status'	Categorical	The employment status of the respondent.	'Employed', 'Unemployed'
Current Residential State	Categorical	The current residential state of the respondent.	'Kuala Lumpur', 'Penang'
How often do you visit online shopping sites?	Categorical	Frequency of visits to online shopping sites.	'Daily', 'Weekly'
What is the average time you spend on online shopping sites daily?	Numeric	Average daily time spent on online shopping sites.	'30', '45' (in minutes)
Type of products that you usually purchase online	Categorical	Types of products usually purchased online.	Respondents choose from:  'Electronic Devices & Appliances'; 'Automotive & Motorcycles'; 'Health & Beauty'; etc.

We pre-process this data by checking for missing values by summing the total number of empty values. We found no missing values in the dataset. The next step is data transformation.

### Data Transformation

Data transformation is the process of converting and structuring data into a usable format that can be analysed to support the visualisation processes (Caroline et al., 2023). This will ensure more informative visualisation. Visualization is the practice of presenting data in a visual format, such as charts, graphs, and maps, to facilitate understanding and insights (Kharakhash, 2024). It plays a crucial role in making complex data more accessible and understandable for individuals across various domains. Data transformation is performed by calculating fields and performing aggregations to derive new metrics and insights.

First, we expand the Product Categories attribute, which contains multiple categories separated by commas. This is done by splitting it into individual rows. This is done using ``str.split`` and ``explode``, resulting in a DataFrame ``df_expanded`` where each product category is in a separate row. Next, we filter the monthly income attribute. A subset of the data is created for individuals with a monthly income of "Below RM5,000" and stored in ``df_below_5000``. Next various aggregations were performed to prepare the data for visualisation as shown in Table 2.

**Table 2: Various data aggregations performed on the dataset**

Aggregated Attribute Name	Description
<code>`income_product_counts`</code>	Count of product categories by income level
<code>`age_product_counts`</code>	Count of product categories by age group
<code>`gender_product_counts`</code>	Count of product categories by gender.
<code>`gender_counts`</code> , <code>`age_counts`</code> , <code>`income_counts`</code>	Counts of individuals by gender, age, and income level

### Visual Mapping

Visual mapping involves the ideation and definition of visualisation solutions. In this step, we define the spatial substrate (space of the visualisation), graphical elements, and graphical properties to determine how data will be visualised (Tkachev et al., 2022). All visualisations in this study are produced in 2-dimensional space. Different graphical elements such as doughnut charts, stacked bars, heatmaps, and line charts were selected based on their ability to best represent the patterns in the dataset. Graphical properties, including size, orientation, and colour, were adjusted to enhance clarity. To ensure that the dashboard communicated insights effectively, each chart type was carefully matched to the nature of the task. Table 3 presents the tasks, selected chart types, and the rationale for their use.

**Table 3: Visual Mapping**

Tasks	Graphical elements	Graphical Properties	Rationale
<b>Gender Distribution</b>	Doughnut chart showing the distribution of shopping by gender	Different sizes and colours according to the gender category	Provides a clear visual comparison of male vs. female shoppers, easy for non-technical users to interpret.
<b>Age Group Distribution</b>	Doughnut chart showing the distribution of shopping by age group.	Different sizes and colours according to the age group category	Highlights proportions of each age segment, useful for demographic insights.
<b>Income Group Distribution</b>	Doughnut chart showing the distribution of shopping by income level.	Different sizes and colours according to the income level category	Shows proportional differences in income categories, aiding quick comparisons.
<b>Product Categories by Income Level</b>	Stacked bar chart showing product categories by income level	Different sizes and colours according to product categories	Enables simultaneous comparison of multiple product categories across income groups.
<b>Product Categories by Age</b>	Heat map of product categories by age group	Different hues according to the number of product categories by age group	Highlights intensity of product preferences across age groups, useful for spotting clusters and dominant patterns.
<b>Product Categories by Gender</b>	Heat map of product categories by gender	Different hues according to the number of product categories by gender	Provides a clear visual of gender-specific product preferences.
<b>Average Time Spent by Gender</b>	Doughnut chart showing the average time spent by gender	Different sizes and colours according to gender	Simplifies comparison of average time spent between male and female shoppers.
<b>Income Level Sorted by Average Time Spent</b>	Bar Chart showing the average time spent by income level, sorts the values and resets the index	Different sizes according to the income level category	Offers straightforward comparison of time spent shopping across income categories.
<b>Age Group by Average Time Spent</b>	Line chart showing the average time spent by age group	Different placement of points according to age group category	Shows behavioural progression and trends in shopping time across age ranges.



### Visualisation Generation

In this step, the specific projection or rendering of the visualisation objects is performed. It involves the process of generating the 2-dimensional images of the visualisation using an application program or data visualisation tool. As mentioned above in this study the Python program was used to render the visualisation objects using the *Plotly* Python library. The layout of the app is defined using the *Dash* library particularly with `'html.Div'` and `'dcc.Tabs'`, which contain multiple tabs with different visualisations. The primary tasks involve filtering data, computing averages, aggregating counts, and retrieving values for visualisation.

### Validation of Usability

To assess the dashboard's usability, a pilot testing session was conducted with a small group of users, including business students and marketing professionals. Participants were asked to perform tasks such as identifying the most active shopper demographic and comparing product preferences between genders. Feedback indicated that the use of multiple chart types improved clarity and engagement, while interactive features such as filtering and tooltips enhanced exploratory analysis. Based on this feedback, minor refinements were made to improve labelling, colour contrast, and navigation across tabs. The validated dashboard was then used to conduct the visualisation and analysis of Malaysian online shopping behaviour, as discussed in the following section.

### Visualisation and Discussion

The visualisation of the online shopping behaviour data is intricately designed to facilitate interactivity. According to (Wu & Chang, 2024), visualization, the process of transforming raw data into visually comprehensible representations is pivotal. It supports the user in selecting a specific task from a tab-style interface to accommodate diverse analytical interests. The dashboard layout and visualisations can be reconfigured by selecting different tabs which show different charts and graphs based on the underlying data. Users explore different aspects of the data through visualisations. For instance, examining product category distributions by income level, age group, and gender. These aspects of the dashboard support different interactivity such as reconfiguring and exploring.

### Visualisation Dashboard

Fig 2 shows the main interface of the visualisation dashboard. The dashboard shows tab menus for each visualisation arranged horizontally for the user to choose.

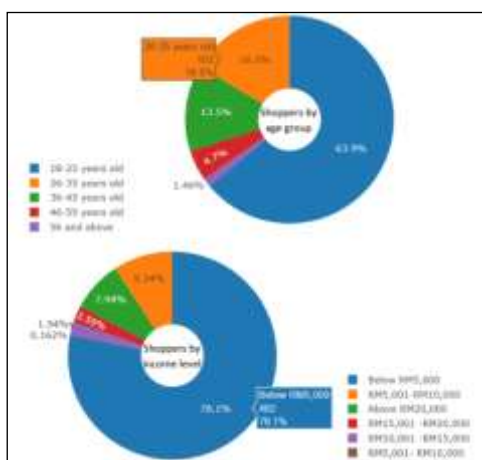


**Figure 2: Online Shopping Analysis Dashboard**

### General Shopping Behaviour Analysis

For this task, we view the shopping distributions using doughnut charts first according to gender as shown in Fig 2. where 80.4% of the shoppers are female and 19.6% are male. Next visualisation is the distribution of shoppers according to age group (Tab 2 in Fig.2) and income group (Tab 3 in Fig. 2). The visualisation is summarized as shown in Fig. 3.

Fig 3 shows two visualisations of the online shopper's characteristics according to the age group distribution and income level distribution. Through these visualisations, users gain insights into most of the shoppers are from ages between 18-25 years old and have income levels below RM5,000. This insight suggests that most respondents in the dataset are likely young individuals, possibly students, who may not yet be earning an income. The user interactions that were incorporated into the visualisations include the tooltip information when the user hovers over the specific parts of the doughnut chart. Filtering interactions were also incorporated in this visualisation where users will be able to remove some categories of age group or income level from the visualisation by clicking on the specific legends.

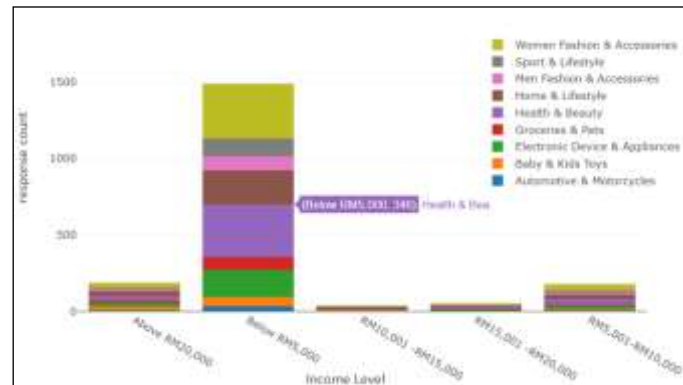


**Figure 3: Visualisation of Shoppers characteristics according to age group and income level**



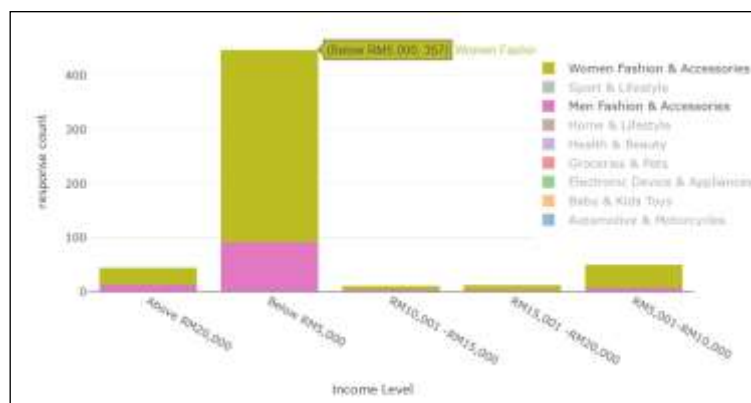
### Product Category Visualisation

This task first investigates the relationship between different income levels and the categories of products frequently purchased. This is represented in Tabs 4-6 in Fig. 2. The goal is to identify which income groups are most likely to buy certain types of products, thereby finding extreme patterns in product preferences across income groups. Fig 4 shows a stacked bar chart visualisation to depict the product category according to income level.



**Figure 4: Product Category by Income Level**

In Fig 4, when the user selects parts of the stacked bar a pop-up tooltip will display the specific information about the part selected. Additional information on the number of responses counts and product categories is displayed to enable users to detail each part of the stacked bar. As discussed in the previous visualisation users can filter out the product category for example in Fig 5 after filtering out other categories and leaving only women's fashion & accessories and men's fashion & accessories product categories, the users can compare the purchase behaviour of these two product categories.



**Figure 5: Filtering interaction to compare Product Categories**

Fig. 6 and 7 show heat map visualisations to view the product categories according to age group and gender.



Figure 6: Product Category by Age Group

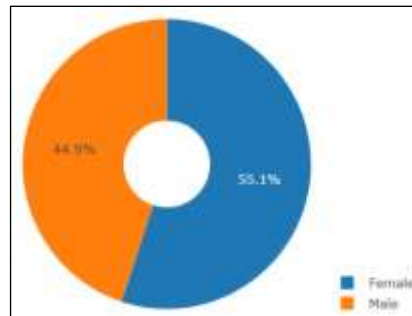


Figure 7: Product Category by Gender

The heat maps shown in Figs 6 and 7 offer another visualisation of the most popular product categories by age group and gender. The insights that can be derived from these figures is that the most popular product categories for females aged 18-25 are Women's Fashion & Accessories and Health & Beauty as indicated by darker cells in the heat maps.

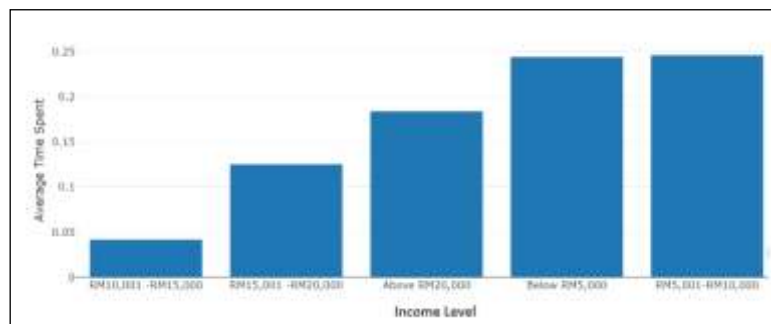
### Average Time Spent Visualisation

The original data for time spent online shopping was categorical where users choose from options such as "Less than 30 minutes", "30 minutes - 1 hour", "1 - 2 hours" or "More than 2 hours". To calculate the average time spent we map these categorical values to numeric such that "Less than 30 minutes" = 0.25 which represents 25 minutes, "30 minutes - 1 hour" = 0.75, which represents 75 minutes, "1 - 2 hours" = 1.5, which represents 1 hour 30 minutes and "More than 2 hours" = 2.5, which represents 2 hours 30 minutes. These numerical values were then averaged to visualise the respondents' average time spent shopping. Fig 8 shows the average time spent shopping according to gender.



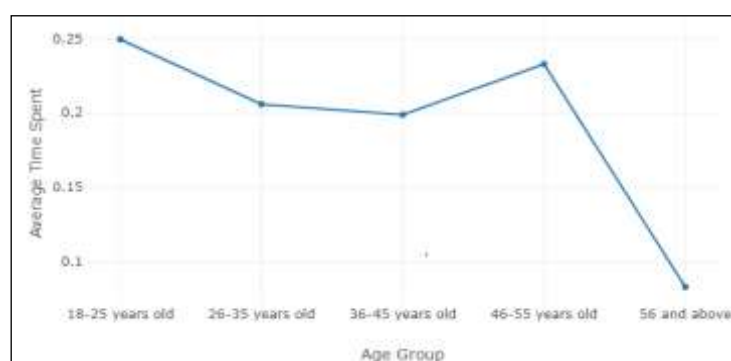
**Figure 8: Visualisation of Average Time Spent Shopping by Gender**

The original data for time spent online shopping was categorical where users choose from options such as "Less than 30 minutes", "30 minutes - 1 hour", "1 - 2 hours" or "More than 2 hours". To calculate the average time spent we map these categorical values to numeric such that "Less than 30 minutes" = 0.25 which represents 25 minutes, "30 minutes - 1 hour" = 0.75, which represents 75 minutes, "1 - 2 hours" = 1.5, which represents 1 hour 30 minutes and "More



**Figure 9: Visualisation of Average Time Spent Shopping by Income Level**

The insights that can be derived from Fig 9 are 1: income level below RM10,000 take the longest time to shop online i.e. approximately 20 minutes on average compared to other income levels and the income level of RM10,001-RM15,000 spends the least amount of time doing online shopping approximately less than 5 minutes on average. Fig 10 shows the average time spent shopping according to age group.



**Figure 10: Visualisation of Average Time Spent Shopping by Age Group**

## Conclusion and Recommendation

In this work, we presented visualisations of behavioural data on online shopping among Malaysians. The visualisation dashboard incorporates various tasks and interactions, enabling users to gain insights into shopping behaviour patterns. We first provided a general overview of shopping behaviours based on gender, income level, and age group. We then focused on identifying popular product categories across these demographics. Lastly, we explored the average time spent shopping among different groups. The tasks involved in the dashboard include data aggregation, clustering, and filtering to offer a detailed yet accessible view of online shopping behaviours. In terms of interactivity, users can navigate through different tabs to explore visualisations such as product categories by income level, age, and gender, as well as average shopping time by demographics. The dashboard also abstracts detailed data into summarised visual representations, such as doughnut charts and bar charts, allowing users to engage with the data in various ways. Future work could focus on more complex tasks, such as exploring correlations between product categories and shopping duration or conducting a deeper analysis of specific income levels or age groups to uncover their preferred product categories. Additionally, incorporating more interactive elements, such as dropdowns and sliders, could further enhance user engagement and facilitate data exploration.

## References

- Ahmad Shakir, A. A. A., & Adzhar, N. (2024). Incorporating Multiple Linear Regression in Analysing Factors Influencing Consumers Purchase Intention for Online Shopping in Malaysia. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 47(1), 257–267. <https://doi.org/10.37934/araset.47.1.257267>
- Caroline, Yuswardi, & Rofi'i, Y. U. (2023). Analysis of E-Commerce Purchase Patterns Using Big Data: An Integrative Approach to Understanding Consumer Behavior. *International Journal Software Engineering and Computer Science (IJSECS)*, 3(3), 352–364. <https://doi.org/10.35870/ijsecs.v3i3.1840>
- Chen, Z., Wang, J., Xia, M., Shigyo, K., Liu, D., Zhang, R., & Qu, H. (2024). *StuGPTViz: A Visual Analytics Approach to Understand Student-ChatGPT Interactions*. <http://arxiv.org/abs/2407.12423>
- Chevalier, S. (2024). *Retail e-commerce sales worldwide from 2014 to 2027*. Statista.
- Jiang, Z., Chen, H., Zhou, R., Deng, J., Zhang, X., Wang, Y., Ngai, E. C. H., & Zhao, R. (2024). *HealthPrism : A Visual Analytics System for Exploring Children ' s Physical and Mental Health Profiles with Multimodal Data Visualization for Time-Series Data Visualization for Healthcare Visualization tools are widely used in the medical field to analyze*. 30(1), 1205–1215.
- Kharakhash, O. (2024). Data Visualization: Transforming Complex Data into Actionable Insights. *Journal of Technology and Systems*, 6(3), 52–77. <https://doi.org/10.47941/jts.1911>
- Liu, S., Peng, D., Zhu, H., Wen, X., Zhang, X., Zhou, Z., & Zhu, M. (2021). MulUBA: multi-level visual analytics of user behaviors for improving online shopping advertising. *Journal of Visualization*, 24(6), 1287–1301. <https://doi.org/10.1007/s12650-021-00771-1>
- Mohamad Shariff, N. S., & Abd Hamid, N. H. I. (2021). Consumers' Buying Behavior Towards Online Shopping During The Covid-19 Pandemic: An Empirical Study In Malaysia. *Malaysian Journal of Science Health & Technology*, 7(2), 1–7. <https://doi.org/10.33102/mjosht.v7i2.164>
- Mustakim, N. A., Abdul-Rahman, S., Aziz, M. A., & Hasan, Z. (2024). A Descriptive Study

- of Factors Influencing Online Purchasing Behavior: Malaysian Consumer Perspective. *Communications in Computer and Information Science*, 2002 CCIS, 296–308. [https://doi.org/10.1007/978-981-99-9592-9\\_23](https://doi.org/10.1007/978-981-99-9592-9_23)
- Nguyen, P. H., Henkin, R., Chen, S., Andrienko, N., Andrienko, G., Thonnard, O., & Turkay, C. (2020). VASABI: Hierarchical User Profiles for Interactive Visual User Behaviour Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 26(1), 77–86. <https://doi.org/10.1109/TVCG.2019.2934609>
- Pragarauskaite, J., & Dzemyda, G. (2012). Visual decisions in the analysis of customers online shopping behavior. *Nonlinear Analysis: Modelling and Control*, 17(3), 355–368. <https://doi.org/10.15388/na.17.3.14061>
- Razali, N. F. C., Mohamad, M., Salleh, K. A., Sani, M. H. A. R., & Alfat, L. (2021). Online Shopping Preferences Visualization System Using Web Content Mining. *2021 2nd International Conference on Artificial Intelligence and Data Sciences, AiDAS 2021*, 1–6. <https://doi.org/10.1109/AiDAS53897.2021.9574386>
- Tkachev, G., Cutura, R., Sedlmair, M., Frey, S., & Ertl, T. (2022). Metaphorical Visualization: Mapping Data to Familiar Concepts. *Conference on Human Factors in Computing Systems - Proceedings*. <https://doi.org/10.1145/3491101.3516393>
- Wu, E., & Chang, R. (2024). *Design-Specific Transformations in Visualization*. <http://arxiv.org/abs/2407.06404>
- Yong, S. C. S. C., Huan, R. T., Poh, W. S., Osman, M., & Ng, D. C. W. (2023). Assessing the Factors Influencing Consumer Behaviour in E-Commerce Platforms. *International Journal of Multidisciplinary: Applied Business and Education Research*, 4(10), 3725–3735. <https://doi.org/10.11594/ijmaber.04.10.25>