Deliverable 2: Project Proposal

Kelby Palmer

2025-09-28

Introduction

- **Background Information:** Analyzing shopping behavior is valuable because consumer decisions drive business strategy, marketing, new and old product design, etc. Understanding the different factors that influence people's purchases can help companies predict demand and tailor their offerings.
- **Purpose of the Project:** In this project, my goal is to use my dataset to analyze different demographics, purchasing habits, and trends to find info on what products are most popular, when they're being bought, and who buys them.
- **Significance:** This project is important because it can help find patterns that could help retail industries make decisions about their products, especially around marketing campaigns, inventory planning, and their seasonal sales. Data analytics helps businesses to be the most productive

Objectives

- **Primary Objective:** Find and identify relationships between demographics (age, gender, location) and purchasing behaviours (product type, purchase amount, season).
- Secondary Objectives: Find seasonal trends in purchases (holiday trends, weather trends, etc). Find out which payment methods are most popular with certain categories of purchases or in which seasons or within different demographic groups.

Methodology

- **Approach:** My dataset seems pretty organized, but I could probably clean it up further. Then, just some analysis to find trends (like the ones already said). To show what I've found, I will make graphs and charts.
- **Tools and Resources:** R, Rstudio, Excel, Google Sheets, Kaggle, GitHub
- **Timeline:**
  - Deliverable 1 – Data set of choice: Due September 21st
  - Deliverable 2 – Project Proposal: A maximum one-page document with your research motivation and description and links to the dataset you plan on using: Due on Canvas on September 28th
  - Deliverable 3 – Data cleaning/ organizing/ basic summary statistics : Due on October 19th
  - Deliverable 4 – Exploratory Data Analysis: Due on October 26th

- Deliverable 5 – Draft Final Report including Results: Due on November 9th
- Final Deliverable – Zip File with proper file structure/ GitHub Repo Link – Due November 16th
- **Final Presentation – November 17th – 21st**

Expected Outcomes

- **Deliverables:** The final outcome of the project will be a GitHub repo with everything from the project (data, visuals, analysis, etc) and then a presentation covering it all.
- **Success Criteria:** A successful project to me is a project where you know what happened and learned something. For this project, it will be successful if by the end of it I can clean, analyze, and explain the data and what it means.

Challenges and Limitations

- **Potential Challenges:** My data contains 3,900 customer records with 18 attributes that describe purchase details, shopping habits, and preferences and more. I do worry that it will have too many variables and that I won't know how to organize it or that it will just be a lot to organize. For example, the colors variable: if there's 30 different shades of each color I think I will cry.
- **Limitations:** Even though the dataset has a lot of records, it doesn't say which specific store or brand the purchases came from, which makes it harder to tie results to a certain type of store. It also lists whether shopping was in-store or online, but it's not clear if that all comes from one store or several different ones.

Conclusion

- **Summary of the Proposal:** For this project, I will analyze the Shopping Behaviours Dataset to find patterns in how different demographics shop, what products are most popular, and how trends shift across seasons. Using tools like R, Excel, and GitHub, I'll clean the dataset, create visuals, and present important insights about consumer behaviour.
- **Call to Action** I'm confident that this project will give me a better understanding of data analytics and how it can be applied to real-world situations. I'm committed to completing each deliverable on time and putting together a final project that clearly shows what the data means.

References

Kaggle Dataset: https://www.kaggle.com/datasets/zubairamuti/shopping-behaviours-dataset/data