

Group 7 Research Proposal: Online Shopping

Our dataset and why we chose it:

Kaggle datasets: Online shopping dataset

Why we want to explore this as a group:

We all enjoy shopping and do it regularly. We want to explore more about the research questions and identify trends to be better informed. Understanding how people spend money is an indicator of our economy's health.

LINK TO DATA:  [Online Shopping Dataset](https://www.kaggle.com/datasets/online-shopping-dataset)  ([kaggle.com](https://www.kaggle.com))

Research Questions:

1. Is there a difference between men and women's spending habits, e.g.: do men have fewer transactions but higher cost items vs women many small transactions with smaller transactions?
2. Shipping prices: are people more willing to spend more on delivery if the cost of the item is more?
3. Does location influence purchase frequency?
4. How do different discount percentages affect the number of repeat customers compared to other discount levels?

Inspiration - other code/articles similar to your dataset

[What our shopping choices say about the U.S. economy : NPR](https://www.npr.org/2022/08/25/1119214319/shopping-sales-walmart-amazon-inflation-economy)

<https://www.npr.org/2022/08/25/1119214319/shopping-sales-walmart-amazon-inflation-economy>

Recommend code tab on Kaggle datasets

Yes, there are many people who have had some data analytics done on the set, here are some examples:

<https://www.kaggle.com/code/jacksondivakarr/eda-and-visualizing-shopping-dataset>

<https://www.kaggle.com/code/nandhirajak/shopping-data-eda-and-model-building>

Example visualization:

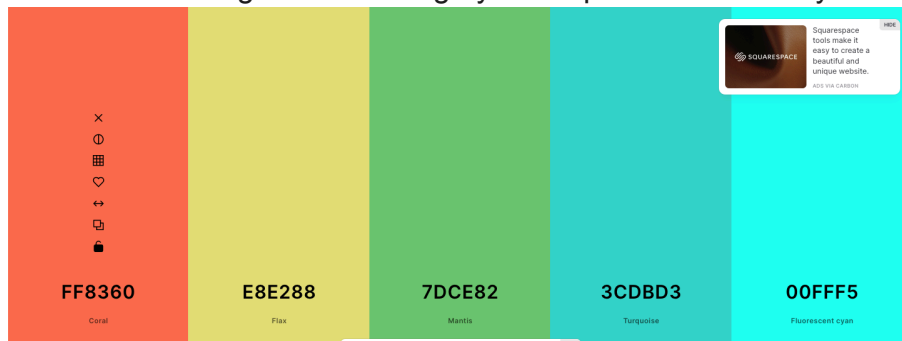
1. Question 1: double bar plot/ pie plot
2. Question 2: scatter plot
3. Question 3: bar graph
4. Question 4: scatter plot
5. Regression: could be question 2 or 3

We plan on using a regression model on question:

Question number two to see the relationship between delivery price and item price.

Color Palette:

To get people's attention, warm hues like reds, oranges and yellows are used but Cool colors are relaxing and encourage you to spend more money. Our color selection:



Roles & Responsibilities:

All of us plan on cleaning the data as a group. Then we will answer our research questions as follows:

- Alex Reid: Regression analysis
- Amber Venes: Question 1
- Abrea Jyles: Question 3
- David Carnero: Question 4
- Samantha Schutz: Question 2

We will then work on putting together a slide deck to tell our data story.

Link to GITHUB:

https://github.com/BreJyles/Project_1_group7