



Online Consumerism



Hypothesis

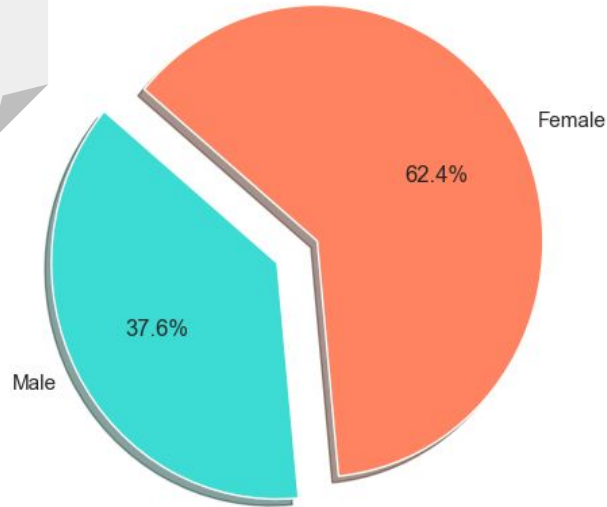
How are people spending money online?

Data cleaning

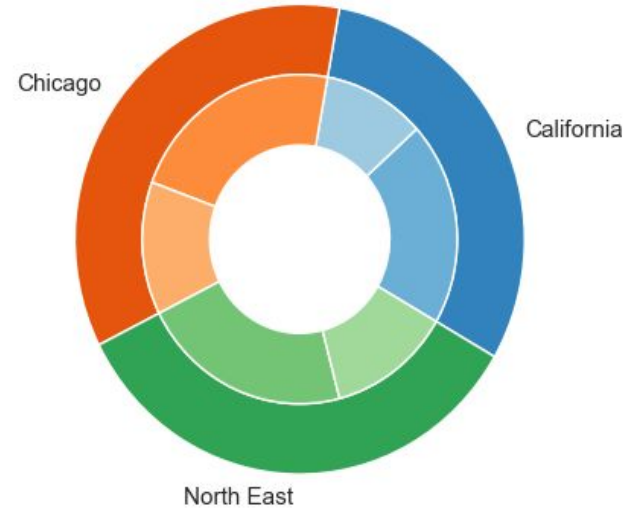
- Initially dropped two columns that didn't seem to add anything to the data
- Checked for null values and dropped them
- Combined three locations to make a north-east region to better analyze the data
- Combined the “Nest” categories to simplify analysis
- Changing the date column to datetime

Who makes up the majority of the market? Who is spending more, and what are they purchasing?

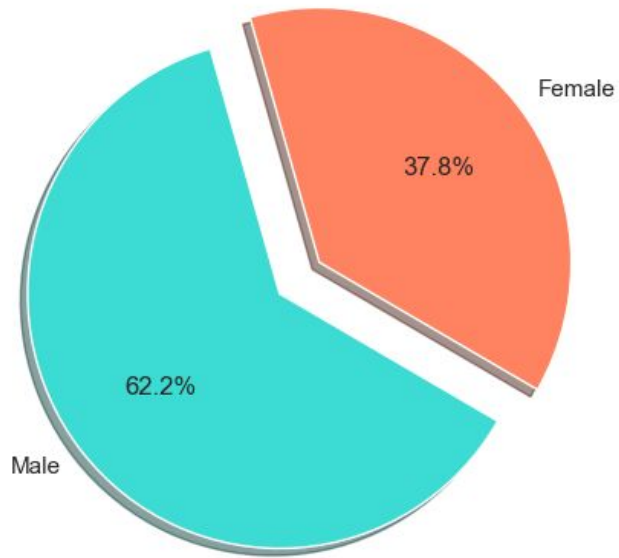
Overall Gender of Online Shoppers



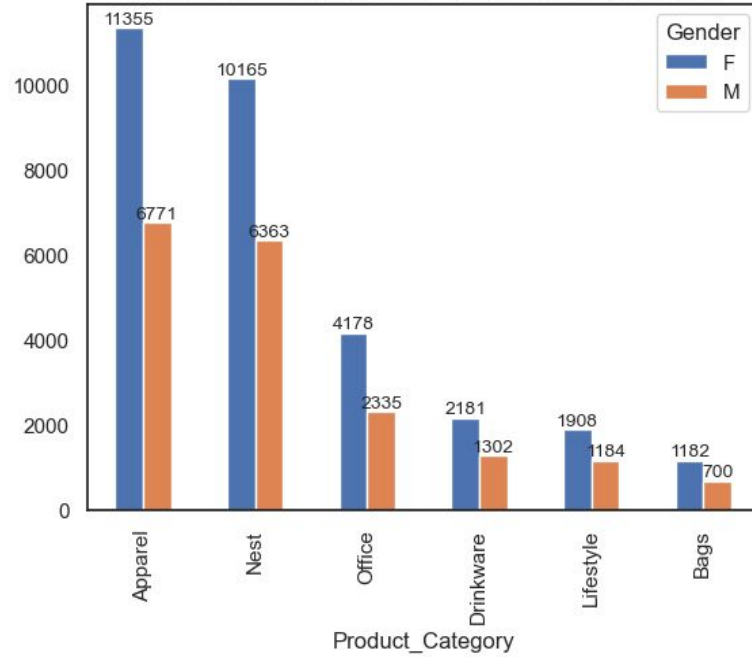
Female vs Male by Location



Total Amount Spent by Gender



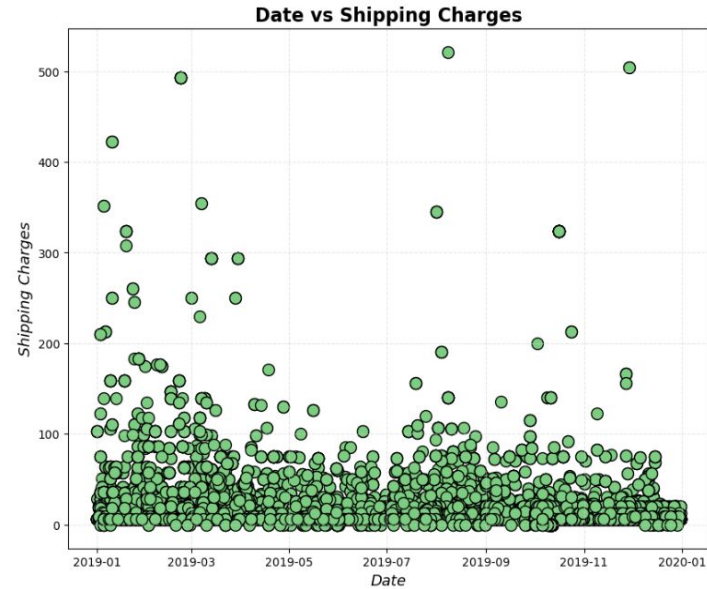
Gender-Based Analysis of Product Purchases



Question: Is there a relationship between cost of an item and the delivery charges? Date and shipping?

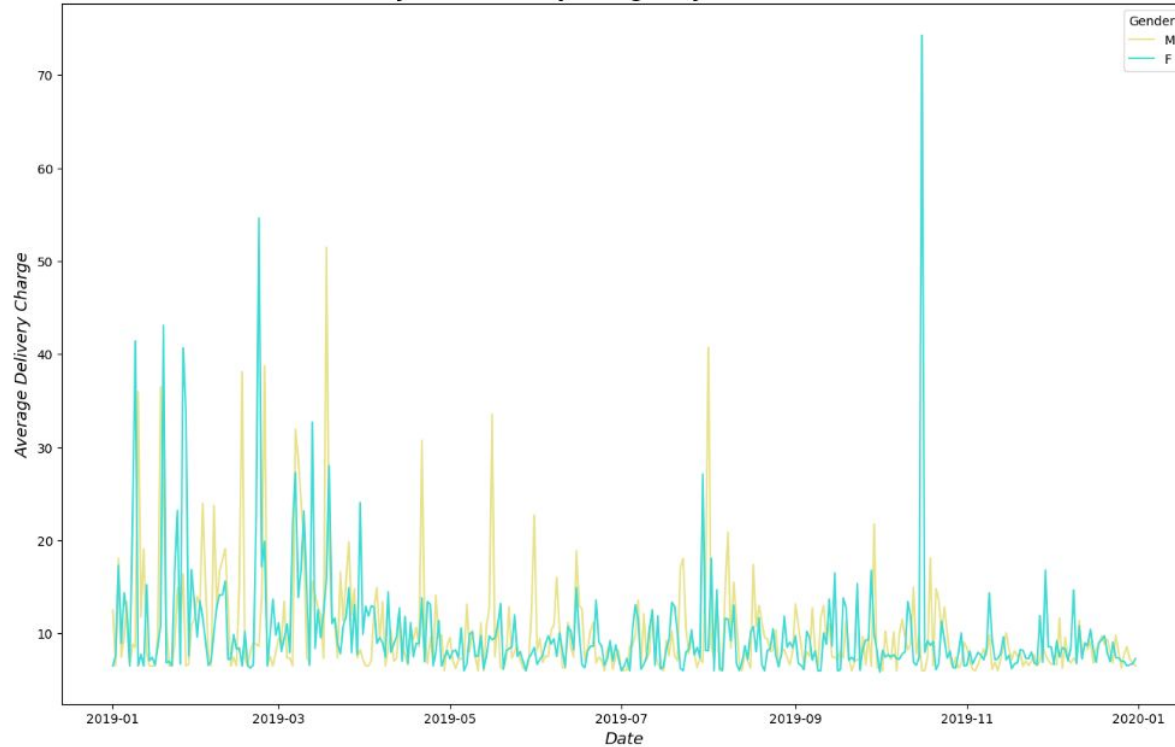


The average price of the item and its impact on the cost of delivery people are willing to spend.



The time of year and the impact it has on the cost of delivery people are willing to spend.

Analysis of Delivery Charges by Date and Gender



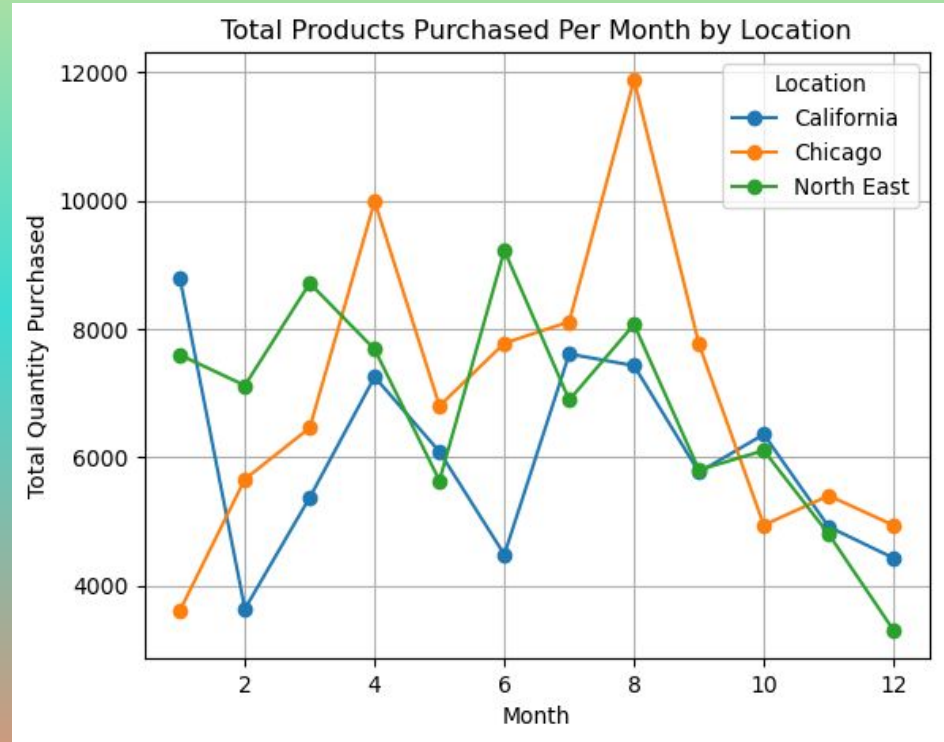
Additionally: Are the delivery charge outliers explained by a certain time of year between gender?

Hypothesis and Visualization

QUESTION - DOES LOCATION INFLUENCE PURCHASE FREQUENCY?

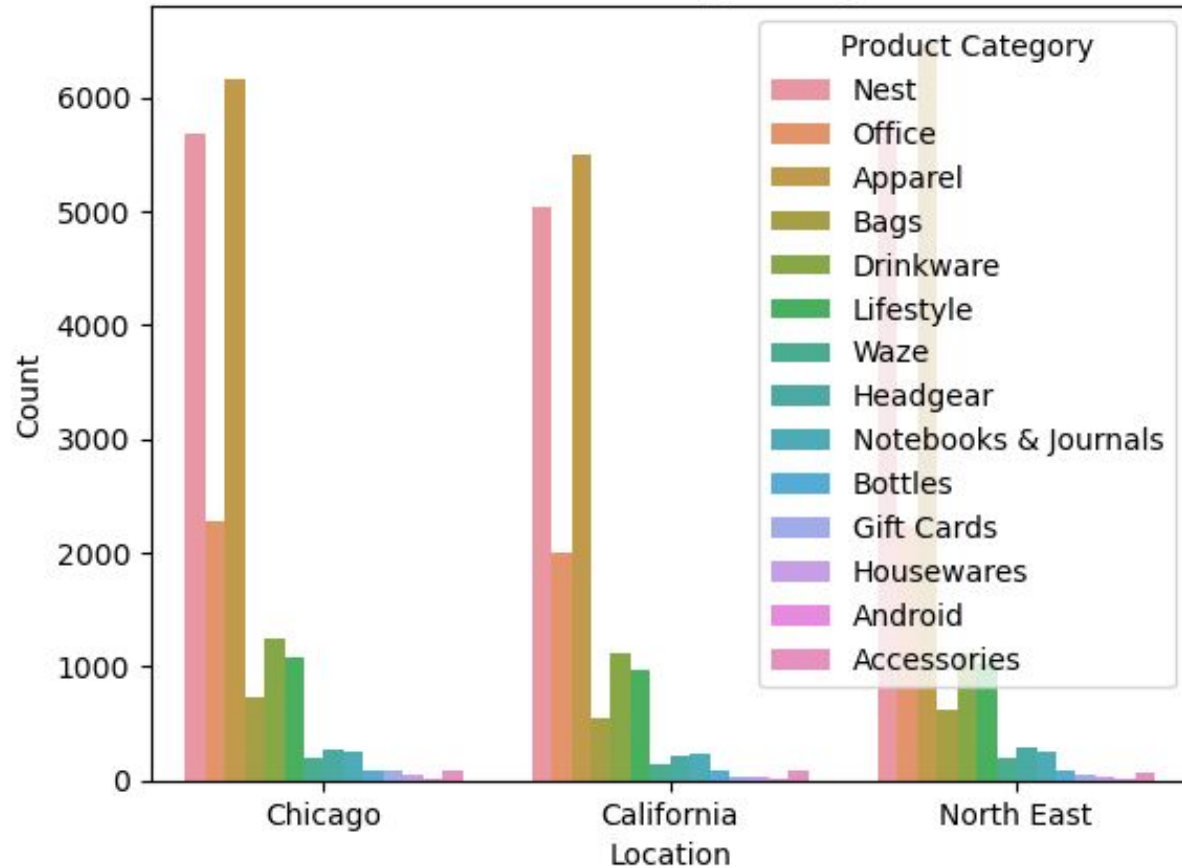
MY HYPOTHESIS - I HYPOTHESIZE THAT CALIFORNIA WILL SURPASS THE NORTHEAST DUE TO ITS LARGE POPULATION AND HIGH PROPENSITY FOR E-COMMERCE TRANSACTIONS.

ANSWER - LOOKING AT MY DATASET MY HYPOTHESIS WAS INCORRECT. WHO WOULD'VE GUESSED?



This data visualization represents online purchase trends from the year 2019.

Count of Product Categories by Location



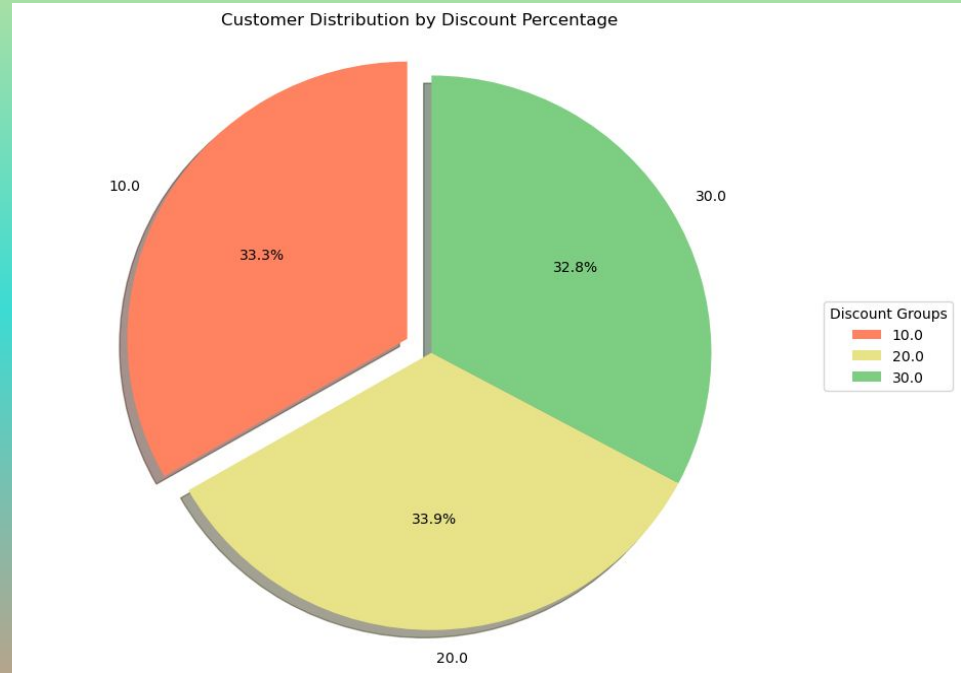
- THE DATA REVEALS A CLEAR TREND WHERE APPAREL SALES CONSISTENTLY ARE PURCHASED MORE THAN ANY OTHER PRODUCT PURCHASED.
- SMART HOME DEVICES FOLLOW CLOSELY AFTER.
- FINALLY, THE HIGHER PURCHASE OF OFFICE SUPPLIES IN CHICAGO THAN CALIFORNIA MAY INDICATE A GREATER LEVEL OF BUSINESS ACTIVITY OR OFFICE DEMAND IN CHICAGO.

Hypothesis and Visualization

QUESTION - HOW DO DIFFERENT DISCOUNT PURCHASES AFFECT THE NUMBER OF REPEAT CUSTOMERS COMPARED TO OTHER DISCOUNT LEVELS?

MY HYPOTHESIS - MY HYPOTHESIS IS THAT THE AN INCREASE IN DISCOUNT PERCENTAGE WILL LEAD TO HIGHER NUMBER OF CUSTOMERS MAKING PURCHASES.

ANSWER - MY HYPOTHESIS WAS INCORRECT; CUSTOMER PURCHASES REMAINED FAIRLY CONSISTENT REGARDLESS OF THE DISCOUNT PERCENTAGE. THEREFORE, I BELIEVE THAT CUSTOMERS ARE MORE FOCUSED ON THE PRODUCT ITSELF RATHER THAN THE DISCOUNT OFFERED.

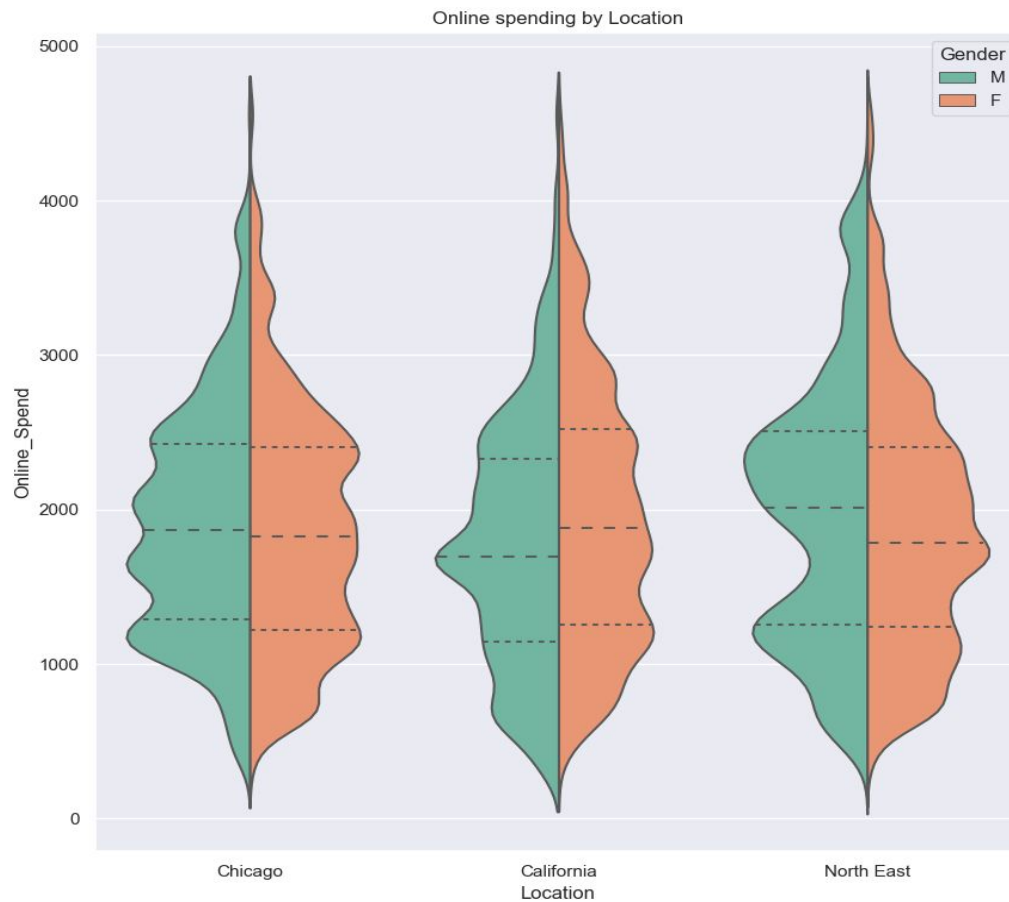


This data visualization represents discount purchase trends from the year 2019.



- Offline and online spend are the most correlated
- The GST and the average price are the least correlated

In Chicago men and women have similar online spending habits





- The r-squared value is 0.011
- 1% of the data considered to be significant

Call to action

- Given the lack of correlations so far, more research is required to determine how people are spending their money online.

Bias limitations

- The data has unknown origins
- Needs more locations to draw more meaningful insights from the data

Future work

- Compare pre pandemic, pandemic, and post pandemic data to see if trends have changed
- Use the data to identify trends to better market products



Works Cited

<https://www.kaggle.com/datasets/jacksondivakarr/online-shopping-dataset>

Questions?