

Customer Personality Analysis



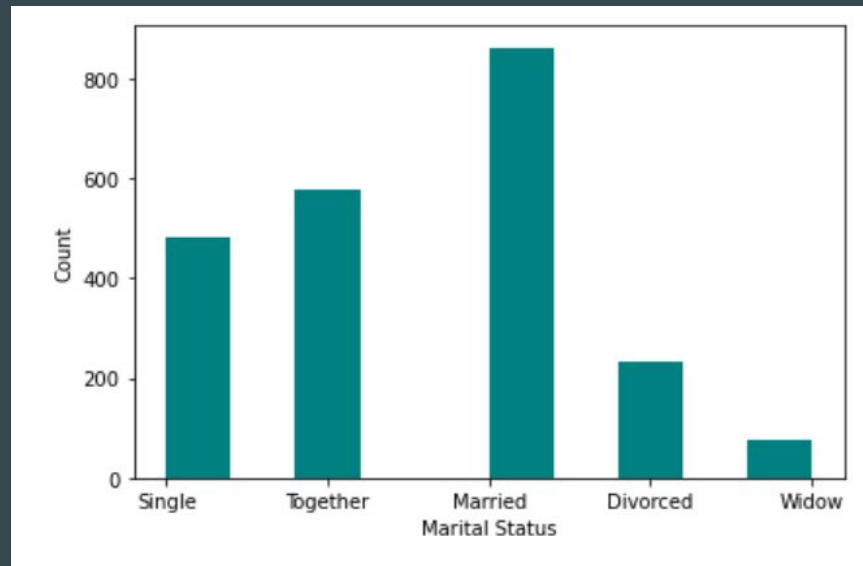
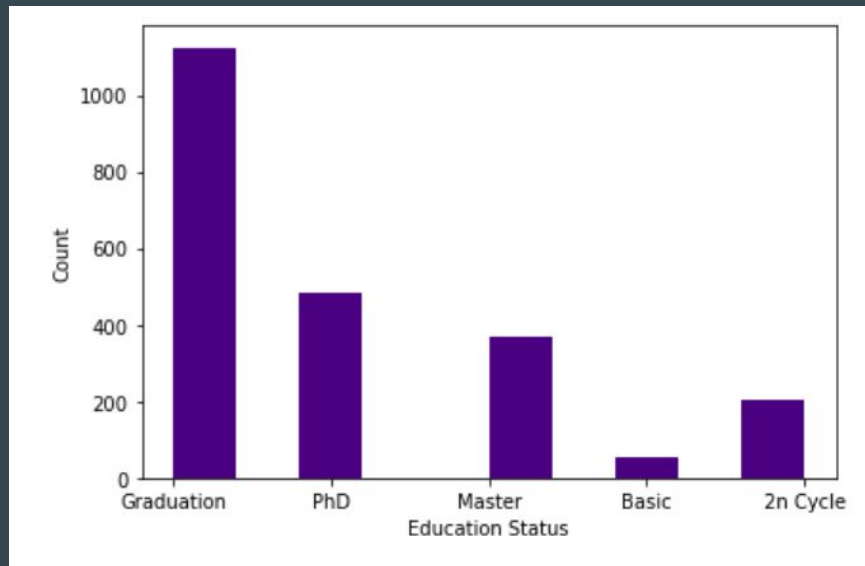
Craft a Story from a Preliminary Exploration of the Data

Data

- Sourced from Akash Patel on [kaggle](#).
- The raw dataset contains 2240 rows (representing unique customers) and 29 columns.
- Columns can be subsetted into:
 - Demographic information
 - Amount spent on types of products
 - Location of purchase (in store, on web, from catalog)
 - Promotions accepted

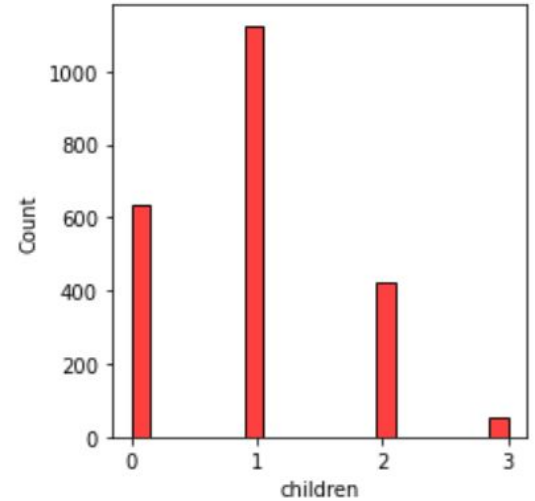
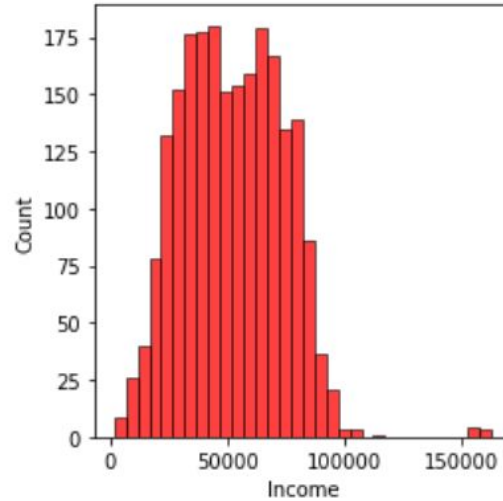
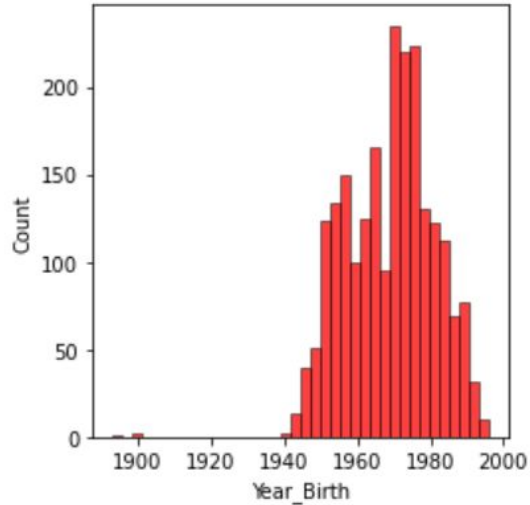
Data Distribution

Demographics



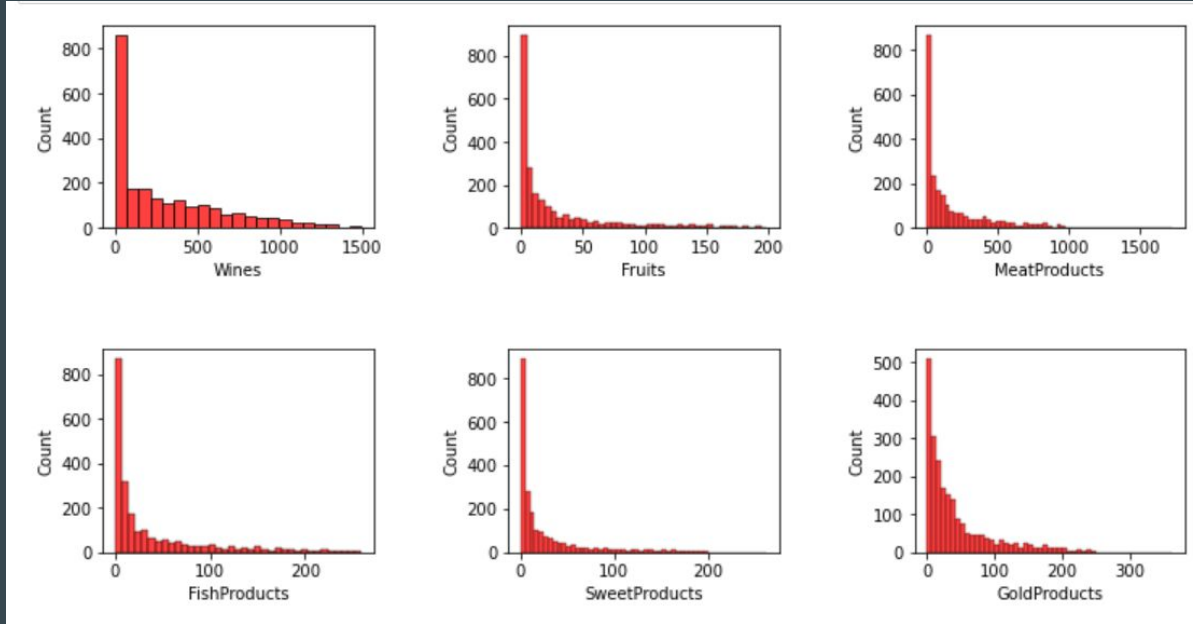
The majority of customers are college educated and/or married.

Demographics (cont.)



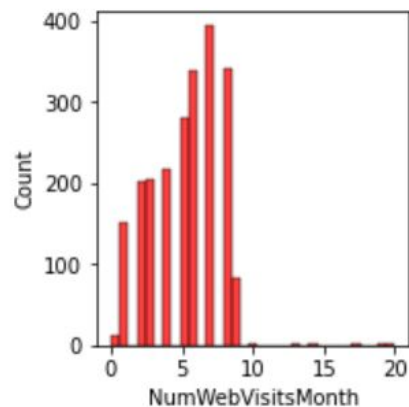
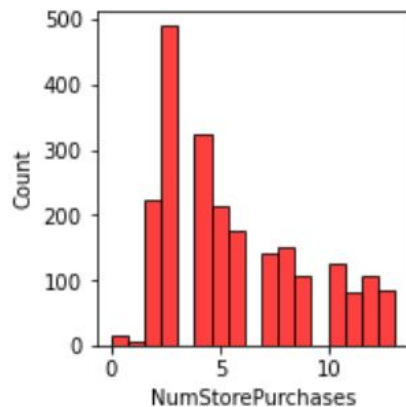
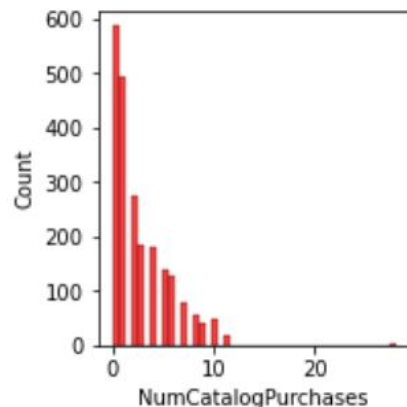
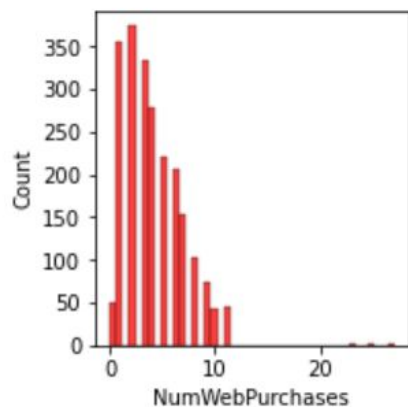
Salary is skewed to the right. Most customers seem to not have children, with no family exceeding 3 children

Amount Spent on Products



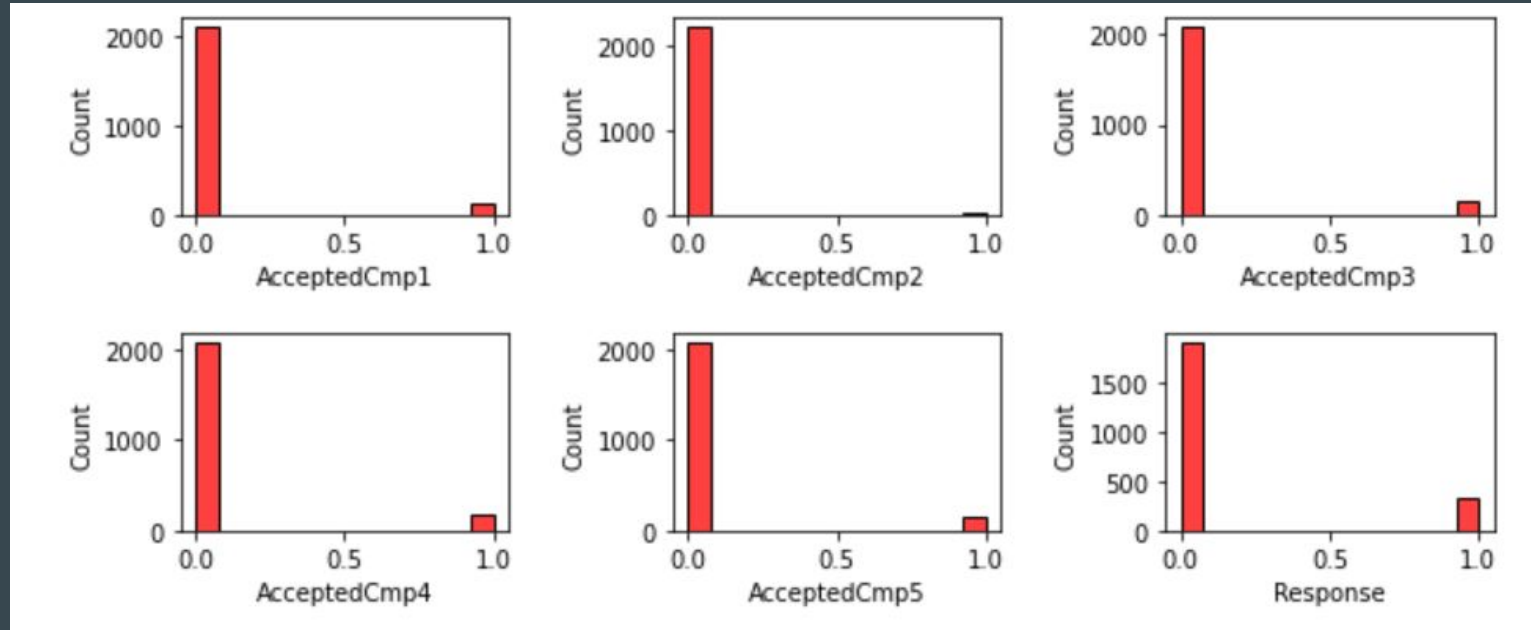
Overall, amount spent on all categories of product is skewed to the right. Most customers tend to spend on the lower end.

Location of Purchase



Most purchases and web visits tend to skew right as well

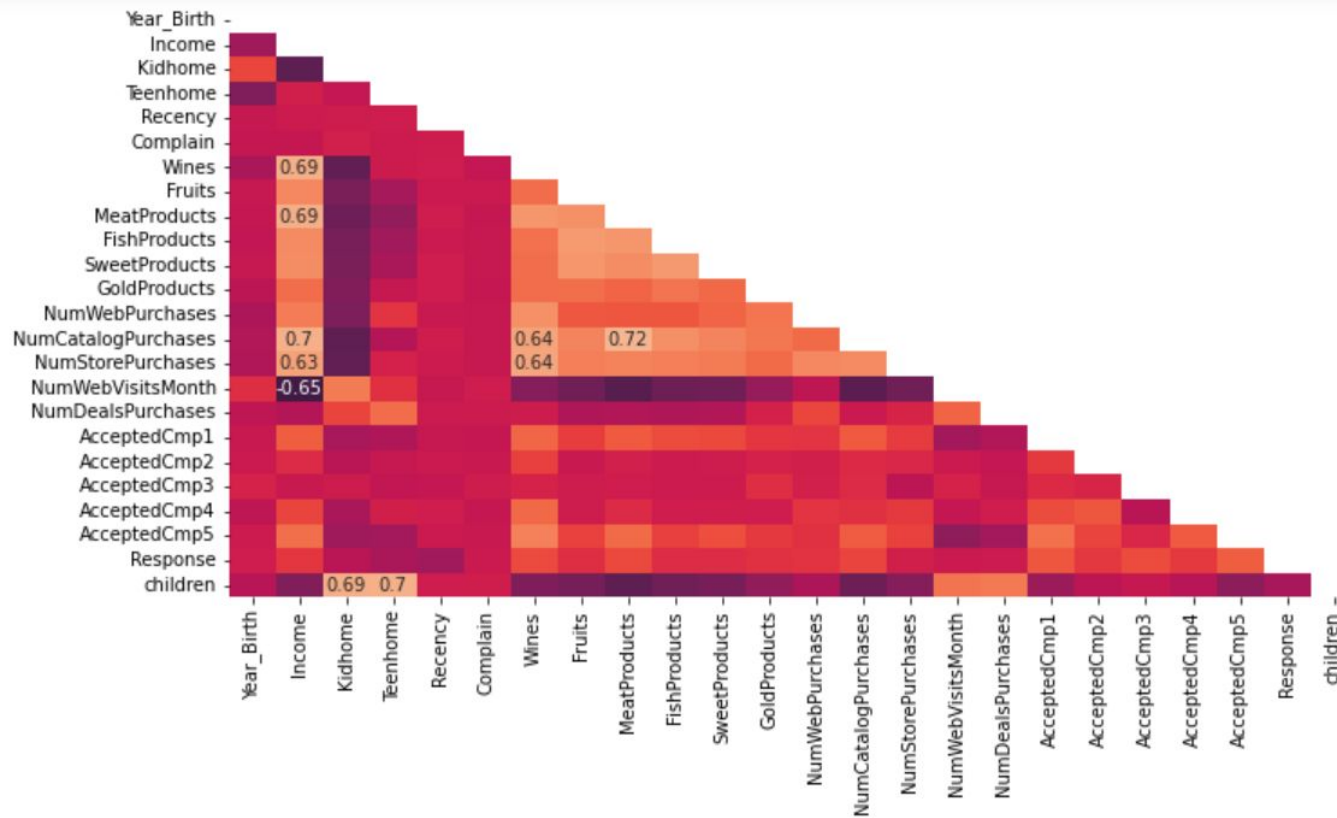
Offers Accepted



Most promotion offers are not accepted.

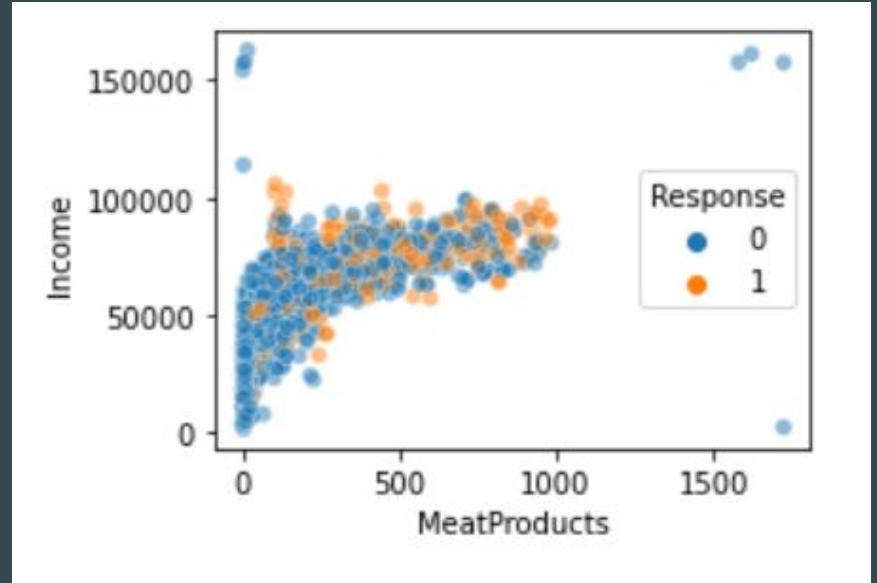
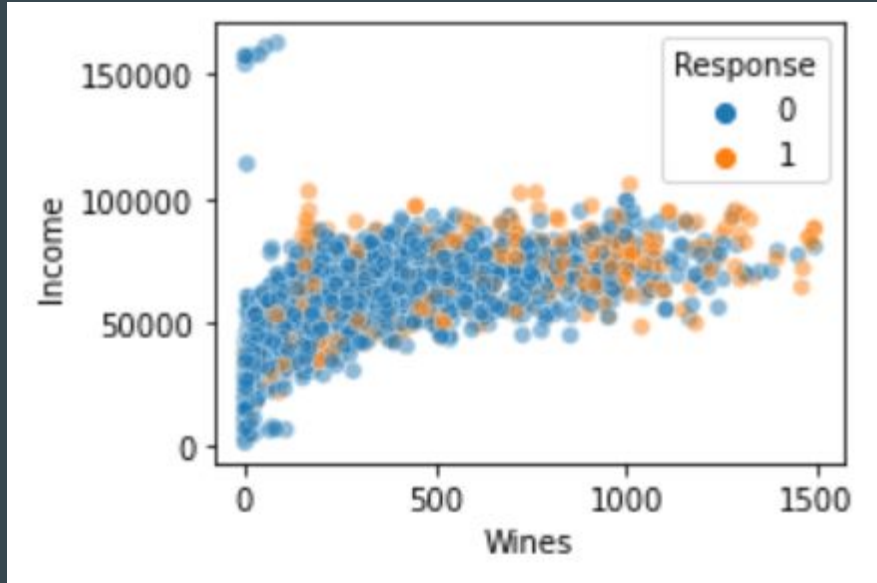
Interaction Effects

Correlating Features



Highly correlated features ($r \geq .60$) will be explored further.

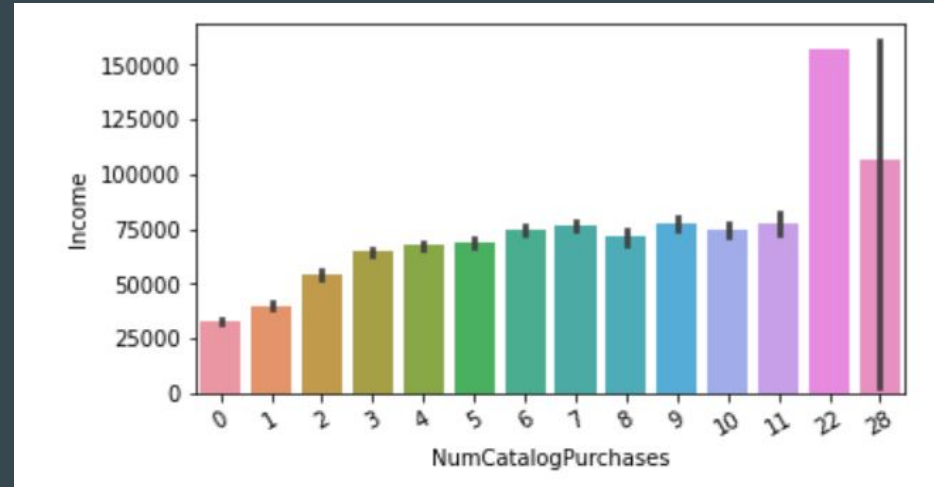
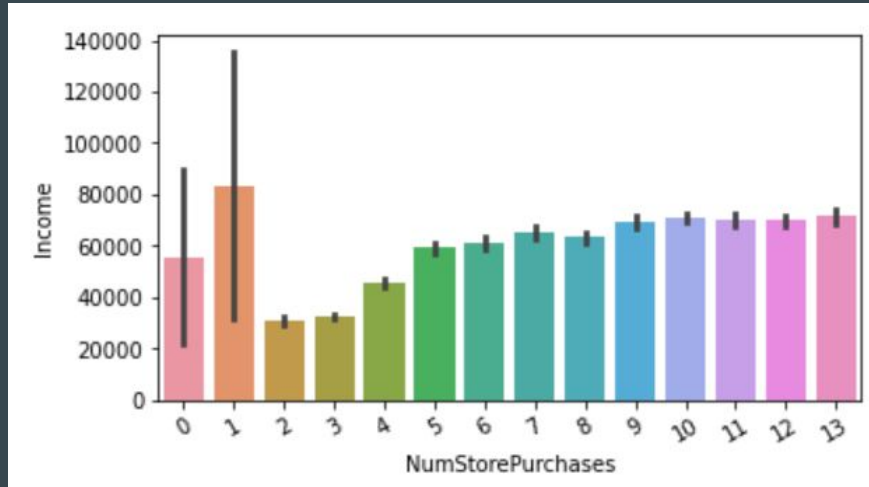
Income and Type of Purchase



Income is positively correlated with:

- Amount spent on wine ($r = .69$)
- Amount spent on meat products ($r = .69$)

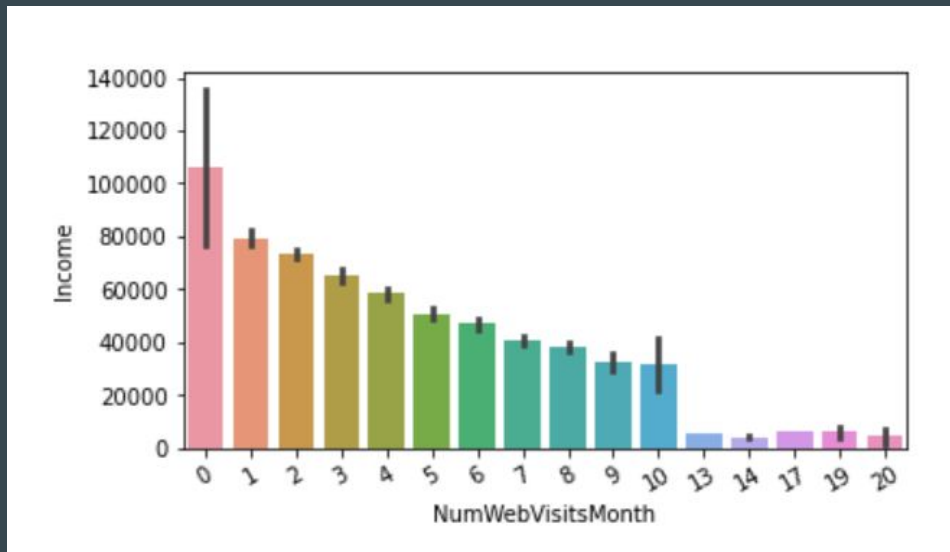
Income and Location of Purchase



Income is positively correlated with:

- Number of in-store purchases ($r = .63$)
- Number of catalog purchases ($r = .70$)

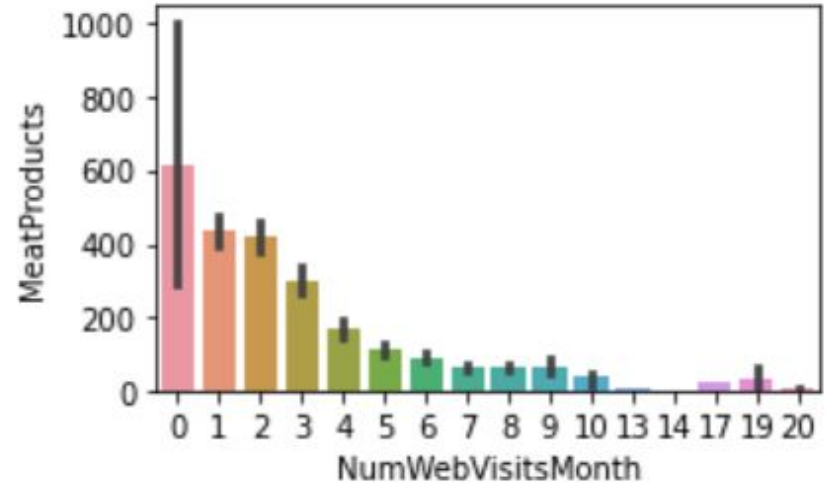
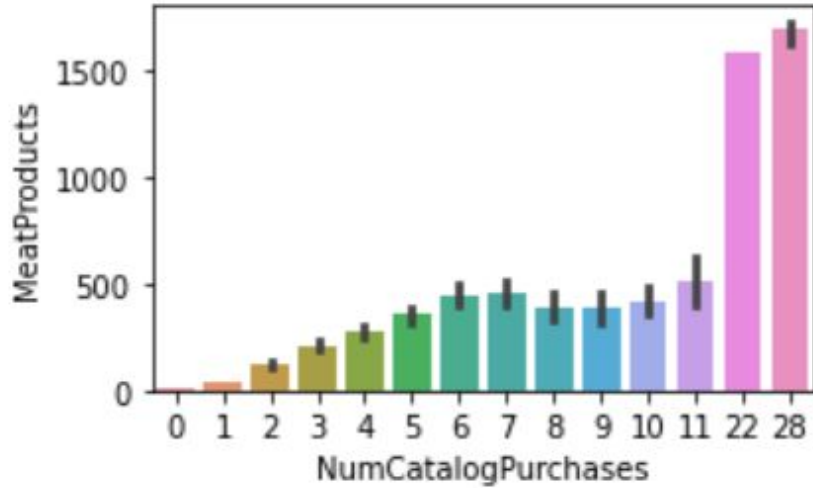
Income and Location of Purchase (cont.)



Income is negatively correlated with:

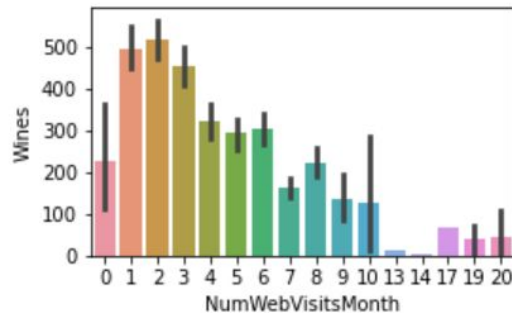
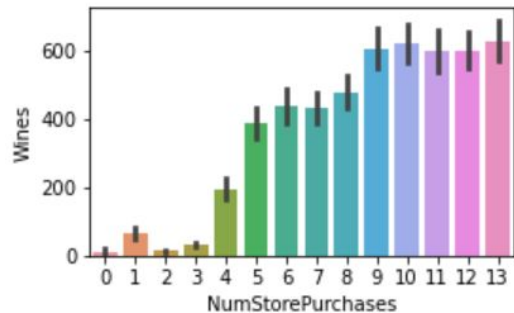
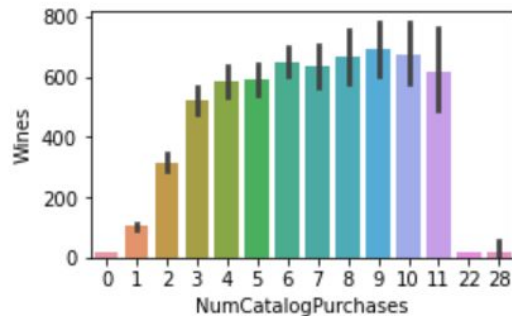
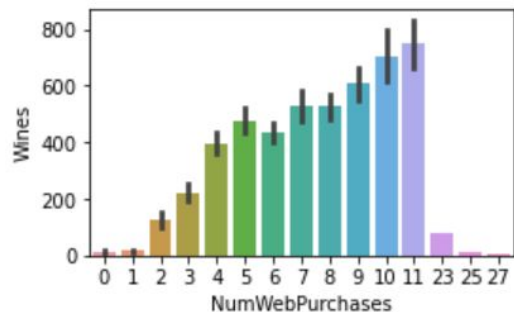
- Number of web visits per month ($r = -.65$)

Meat Products



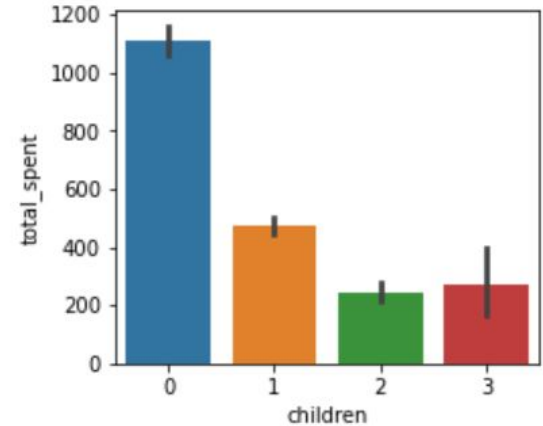
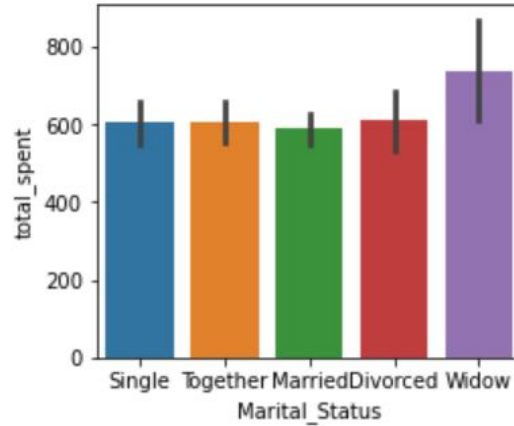
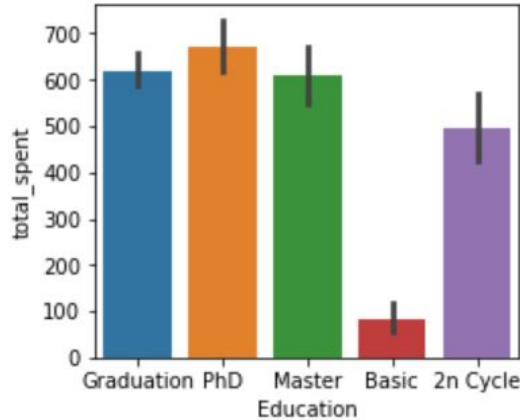
- There is a strong correlation between amount spent on meat products and number of catalog purchases ($r = .72$)
- Additionally, there seems to be a trend between meat products and number of web visits.

Wine Products



- Similarly, there seems to be a trend between amount spent on wine and location of purchase.
- The two strongest correlations, however, are:
 - Catalog Purchases ($r = .64$)
 - In Store Purchases ($r = .64$)

Total Spent Based on Demographic Information



- Those without a college degree seem to spend less per month.
- Marital status doesn't seem to have a large impact on amount spent.
- Childfree customers spend much more than those with children.

Summary

- Income seems to be the most influential feature in terms of predicting customer behavior, and will likely be the basis in forming customer segmentations.
- There seems to be some association between type of product and location of purchases. It might be possible to combine the two through some form of feature reduction.
- Demographic information, particularly whether or not a customer has children, can also be useful for forming clusters.