Project 1 Analysis – December 7, 2022

Data Set: Grocery Store Customer data, 2240 entries, 29 fields

1. I started by removing all of the non-returning customer entries (0 in Response column) in order to analyze the responses from the returning customers only (1 in Response column). There was a total of 334 returning customers, or 14.91%.
2. I converted Year\_Birth to Age and found no correlation between age and returning customers
3. I noticed that many of the returning customers had no or few children in their households. I combined Kidhome and Teenhome into a “Total Kids” column and found that about 50.6% of returning customers had no children and 34.73% had 1 child.
4. I then combined all the amounts the customers had spent on different types of products into a TotalProds column. I then calculated the average amount spent between returning customers with children and returning customers without children:

Average Spent by Returning Customers without children: $1448.47

Average Spent by Returning Customers with children: $515.14

This showed that Customers without children were likely to spend almost 3x more than those without children.

1. It seemed that most of the returning customers without children had incomes on the higher end of the data set.
2. I noticed that all returning customers had made purchases in the store and visited the stores website in the past month.
3. I combined all the AcceptedCmp columns into TotalAcceptedCmp and found that 56.29% of returning customers had accepted 1 or more of these offers, which did not point to any significant correlation.

**Conclusion:** Individuals with higher income, fewer or no children and those who visited the store and website are more likely to spend more and be returning customers.