**Coupon Analysis Report**

The data for this assignment came from the UCI Machine Learning repository. It was a survey fielded on Amazon Mechanical Turk which offered car drivers a coupon to a local food or beverage establishment. The dataset contains the survey responses, including the type of coupon offered, demographics of the driver, and environmental conditions.

The data was analyzed to address the question, “Will a customer accept the coupon?”,, comparing customers who accepted different types of coupons with those who did not.

**Data Cleaning**

The raw data file had 12, 684 records and 26 features. There were 74 duplicate records, which I dropped. The feature “car” was almost entirely missing values, so it was dropped. Missing values in the features “CoffeeHouse”, “CarryAway”, “RestaurantLessThan20” and “Restaurant20to50” were filled with “never”. The final dataset had 12,536 records and 25 features.

**Link to Jupyter Notebook:**

**Summary of Findings**

**Types of coupons and acceptance rates**

Before I investigated individual coupons, I explored other potential influences on drivers accepting a coupon in general, such as time, temperature, direction of travel and closeness to the coupon destination.

* Coupons were offered for Restaurants $20-$50, Restaurants under $20, Bars, Coffee Houses, and Carry Out and Take Away
* Of the 12,536 drivers who were offered a coupon, 7,104 accepted it, an acceptance rate of 57%.
* Restaurant and Carry Out coupons had the highest acceptance rates
* Drivers aged 30 or under and male were more likely to accept coupons
* Drivers were more likely to accept coupons between 10am and 6pm, when they had no urgent place to go, were single, and were accompanied by adult passengers
* Driving opposite direction and distance within 15 or 25 minutes are correlated with acceptance
* No particularly strong correlation was observed between any of the variables

**Summary of findings - Bar Coupon**

We observe that no drivers who were in Farming occupation or were widowed went to bars more than once a month, so these filters did not affect the results.

The Bar coupon acceptance rate was much higher among certain groups of drivers:

* those who went to bars less frequently, fewer than 3 times a month, were 4x more likely to accept the coupon
* about 1 in 5 drivers who went to bars more than once a month and had adult passengers accepted the coupon; the rate was higher for younger drivers under 30, about 1 in 3
* about 1 in 5 drivers who frequent cheap restaurants, more than 4 times a month, and have income less than 50K accepted the coupon

We conclude that drivers who were between 25 and 30, who prefer cheaper restaurants, and who were traveling with adult passengers were likely to accept the Bar coupon.

**Summary of findings - Carry Out Coupon**

I investigated the Carryout & Takeaway coupon further to identify key characteristics of customers who accepted these coupons, as well as characteristics of the coupons that were accepted, for example, whether factors like weather, time of day, distance to destination and expiration time impacted acceptance rate.

We observe the coupon had broad appeal among drivers, ranging from never eat carryout to more than 8 times per month. The average age of drivers who accepted a Carryout coupon was 32. Drivers under 30 and over 45, and male, were more likely to accept coupons. About 43% of them had children. Single and married drivers were equally likely to accept the coupon, and lower income drivers were more likely to accept.

The acceptance rate when the coupon destination was within 15 minutes was 62%, compared to 8% when it was 25 minutes away. The acceptance rate was 24% when the coupon destination was in the same direction, but 66% when in the opposite direction. Coupons valid for one day were more likely to be accepted than those valid for only 2 hours. Drivers were more likely to accept coupons on sunny days and rainy or snowy nights.

**Next Steps and Recommendations**

This dataset contains a lot of data in string format, so a next step would be to convert these to numeric format for easier analysis. I only did a correlation heatmap for the overall dataset, but further analysis of correlations for each type of coupon could differentiate customer preferences.

In terms of analysis, clustering would be a good technique for customer segmentation, to quickly gain a clearer idea of customers attracted to each coupon and inform targeting future offers to them.