

Initial Data Analysis

Initial analysis/investigation of the data set for missing or problematic data found that the 'car' column has only 108 rows with non-null values and the rest of the values are null. This is addressed by setting the null values in the 'car' column to the value 'Car', so that data analysis can be done appropriately.

Additionally, the following columns also had a few null values - Bar, CoffeeHouse, CarryAway, RestaurantLessThan20 and Restaurant20to50. These rows are dropped from the data set for analysis.

Overall Data Analysis

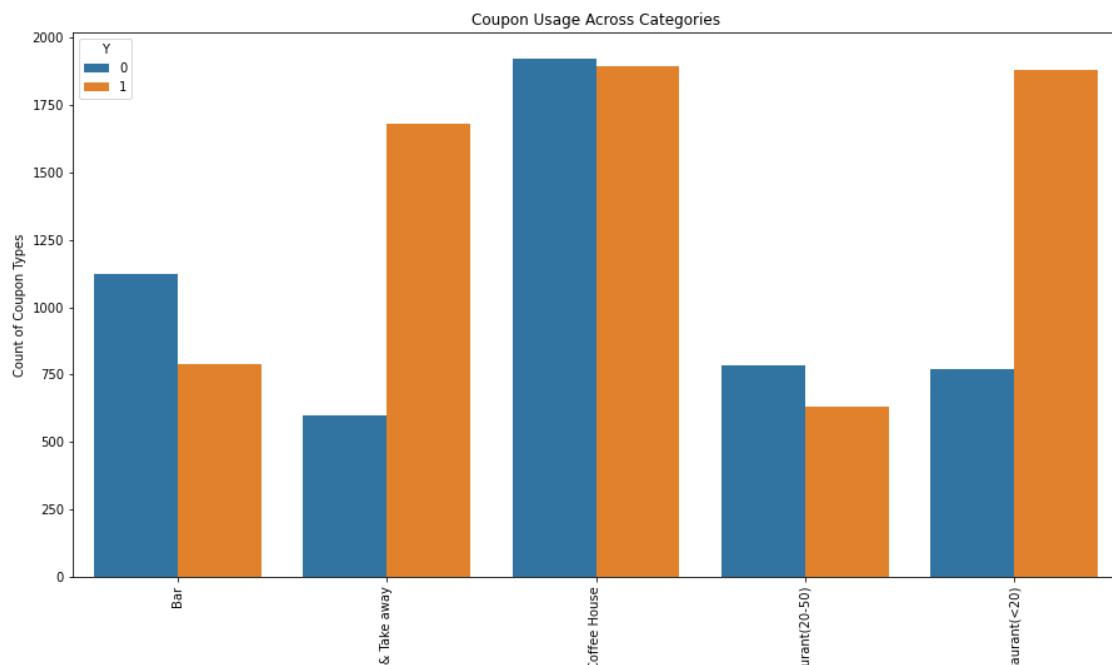
To answer the question on what proportion of the total observations chose to accept the coupon, the following is done: First, we identify the count of all rows where 'Y' == 1 (Coupon Accepted) and divide by the total count of rows in the cleaned data set.

Percent-accepted = $(\text{data.loc}[\text{data}['Y'] == 1] ['Y']. \text{count}() / (\text{data}['Y']. \text{count}())) * 100$

We find that the 56.93% of the coupons are accepted overall.

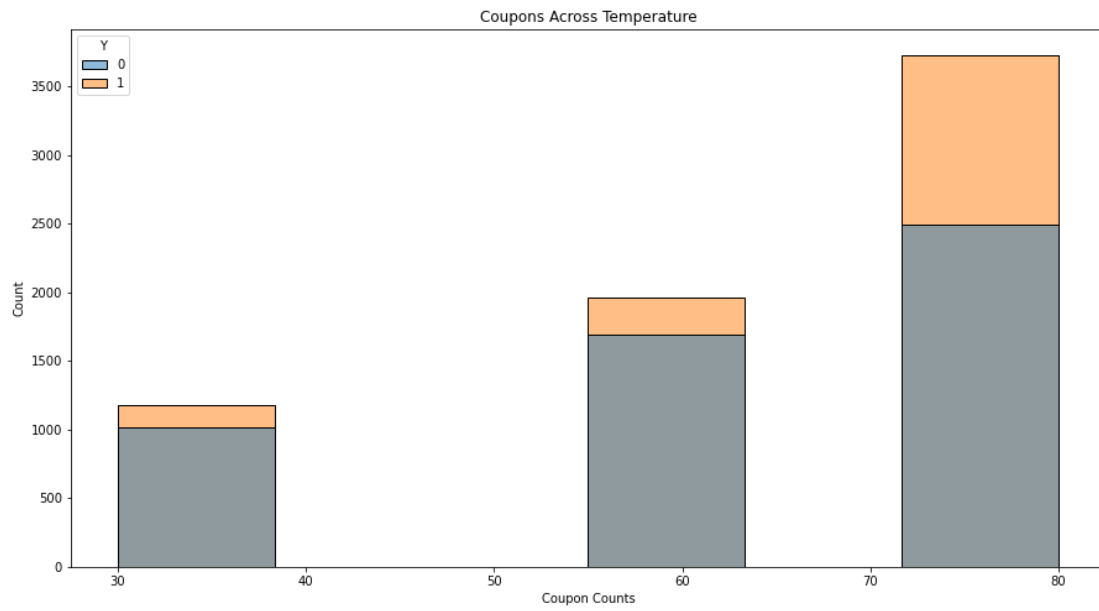
Coupon Usage Across Categories:

Using a Bar plot to visualize the 'coupon' column, it is seen that the take away coupons and the cheaper restaurant coupons have a much higher acceptance rates than the bar or expensive coupon category. The coffee house category has a fairly equal distribution of accepted and not accepted.

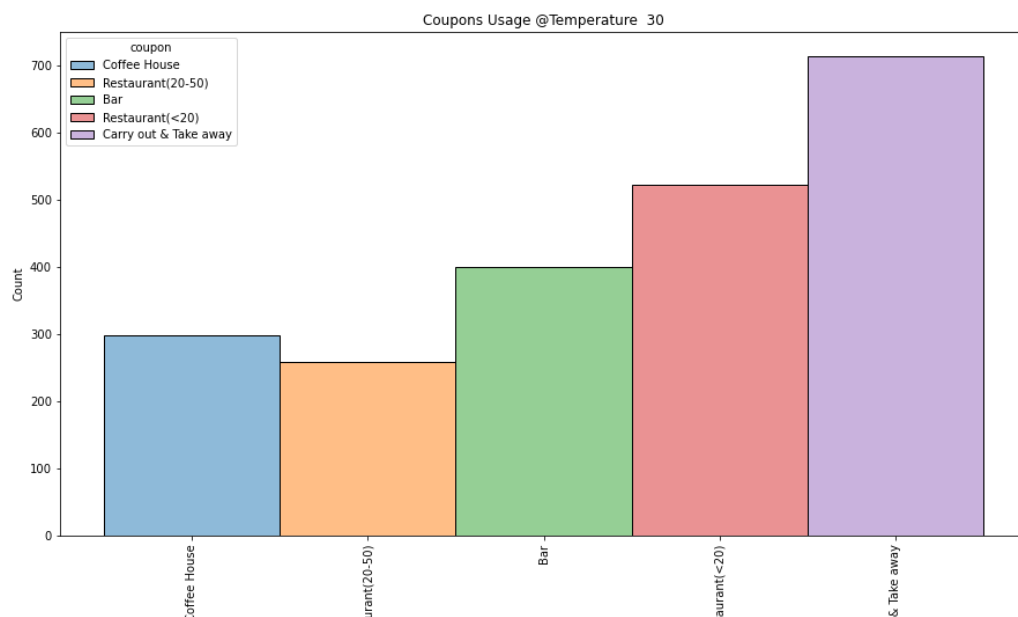


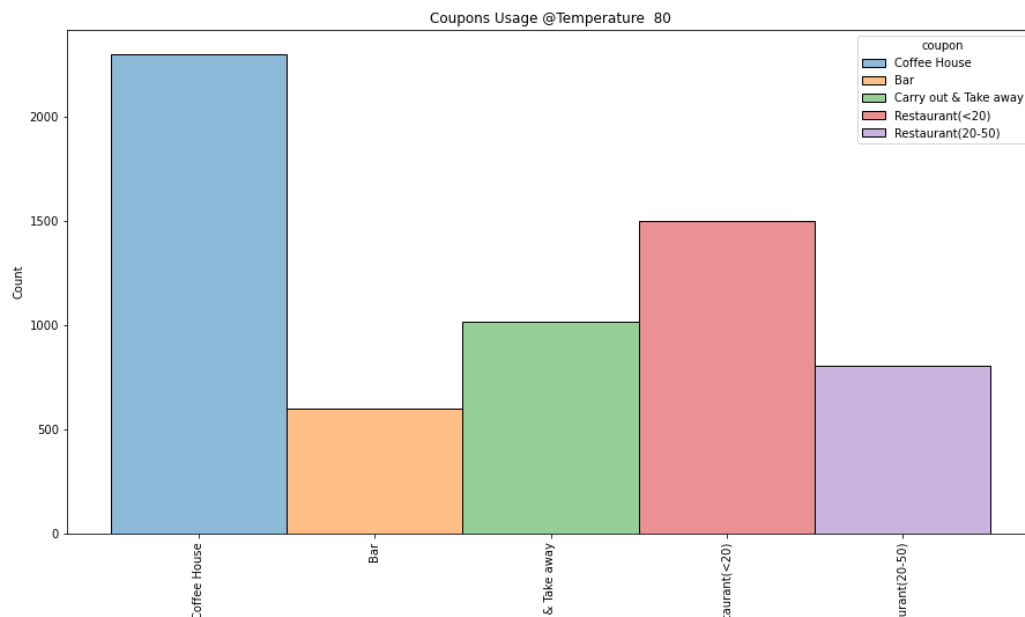
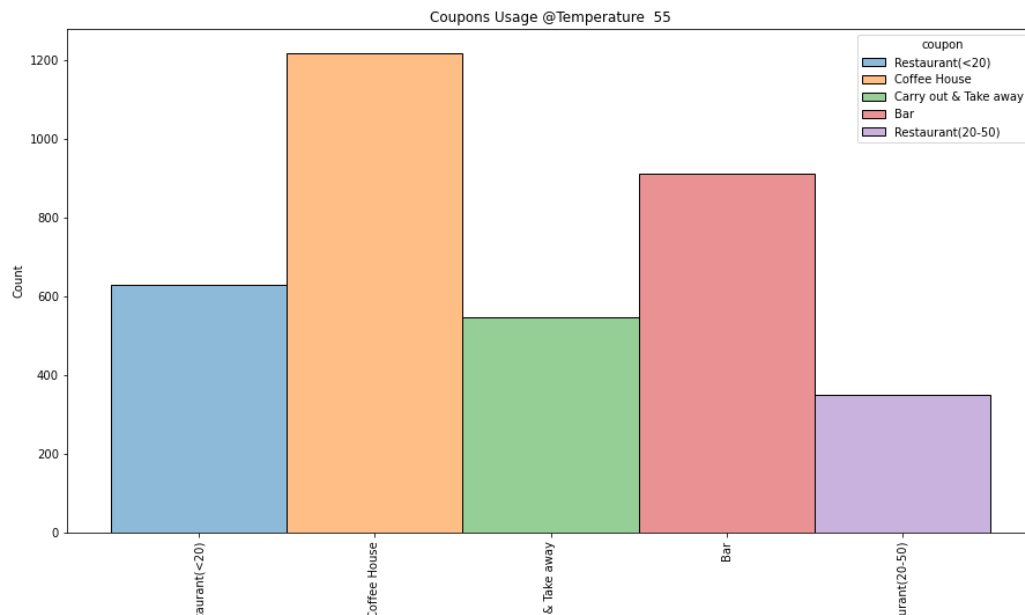
Coupon Usage Across Temperatures:

Another analysis and data plotting is done on the temperature column – essentially identify coupon acceptance across different temperatures. If you look at the histogram, the coupon acceptance rate is much higher at higher temperatures



To do a deeper analysis on the coupon usage at different temperatures, we find some interesting observations:

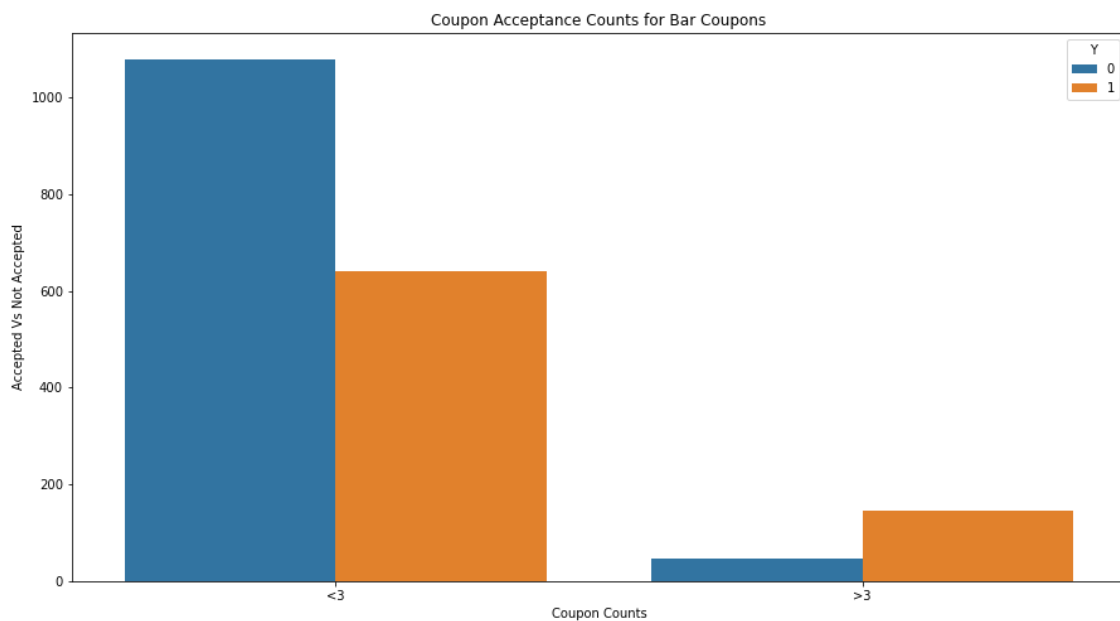
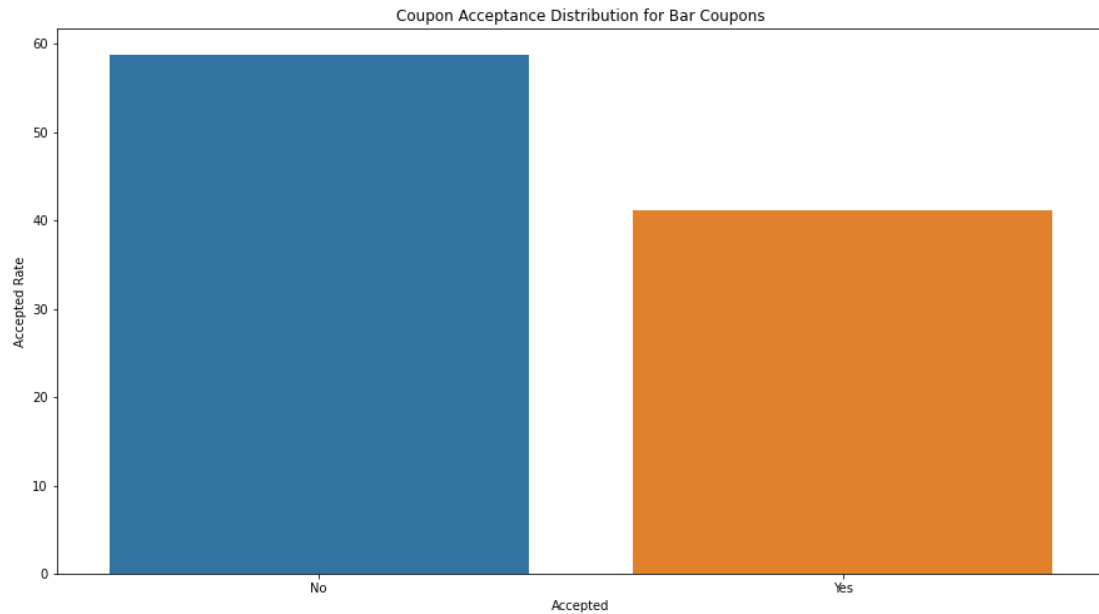




What we can see is that the total coupons for take away is highest when the temperature is 30F and the coffee house coupons are higher at higher temperatures. One would have thought that the bar coupons would be high at high temperatures but the data shows the opposite.

Deep dive into the 'Bar' coupons

A deeper analysis of the bar coupon data and plotting the same shows that while the overall acceptance rate for bar coupons is 40%, the acceptance is much higher for people who went to a bar 3 or more times a month to those who went less. The data also shows that there are far more people who go to the bar three or less times a month than those that go to a bar for more than 3 times a month.



Additional Data Analysis

It looks like the coupon acceptance rate for people who are over 25 and go to the bar more than once is 69% as compared to only 26% acceptance for the same people who go to the bar less than once. Also, for people who are below 21, the numbers are similar with 67% acceptance for people who go to the bar more than once as opposed to 38% acceptance for people who go to the bar less than once.

Also, the data shows that there is almost an equal distribution of coupon acceptance vs non acceptance for coffee house coupons when looked at in totality.

