**Coupon project**

Data analysis of driver’s acceptance of online coupon offering in various conditions and situations

**Definition of “Carry Away”**

Based on the dictionary “Carry away” means “lose on self-control”.

**Use of “CarryAway” column data for data analysis for bar coupons**

I used “CarryAway” data to use for bar coupons as my understanding is driver is not allowed to drink while driving but carried away

**Strategy for data analysis**

* Load the file
* Convert to DataFrame
* Clean up the data
  + Remove duplicate
  + Impute missing values in numerical columns with the median
  + Impute missing values in categorical columns with the mode
  + Verify that there are no missing values left
  + Replace NAN with '0' (Zero)

**Formatted the columns data to suits analysis**

* “age” column Less than 25 years age as ‘less25’ and more than 25 years age ’more25’
* “age” Less than 30 years age as ‘less30’ and more than 30 years age ’more30’
* “RestaurantLessThan20 column Less” than 1 or greater than 8 as ”Less1” & “ gt8” respectively
* “bar” less than 3 and more than 3 as “less3” and “more3” respectively

**Find columns naming**

* Used “.unique” command to find columns names

**Rename columns items**

* Renamed columns entries to labels to suit for data analysis using “.isin” command

For example

data.loc[data['column\_name'].isin(['1~3', 'less1']), 'column\_name'] = 'less3'

**Sorting data and selecting columns suitable to analysis**

* Sorting data which suit analysis
* Keep columns which needed for analysis

Coupon project information

Data has 12684 indices and 26 columns

A screenshot of a computer code

Description automatically generated

Total count of coupons = 12684

A screenshot of a coupon

Description automatically generated

1. **Read in the coupons.csv file**

**A screenshot of a computer

Description automatically generated**

1. **Investigate the dataset for missing or problematic data**

**A screenshot of a computer screen

Description automatically generated**

1. **Investigate the dataset for missing or problematic data.**

**A screenshot of a computer screen

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer screen

Description automatically generated**

1. **Decide what to do about your missing data -- drop, replace, other...**

* Impute missing values in numerical columns with the median
* Impute missing values in categorical columns with the mode
* Verify that there are no missing values left
* Replace NAN with '0'

1. **Use a bar plot to visualize the coupon column.**

A graph of a bar graph

Description automatically generated

1. **Use a histogram to visualize the temperature column.**

A graph with a blue rectangle

Description automatically generated**Number of coupons accepted**

Total numbers of accepted coupons = 7210

Percentage of accepted coupons = 56%

**Investigating the Bar Coupons**

1. **Create a new DataFrame that contains just the bar coupons.**

A screenshot of a computer

Description automatically generated

1. **What proportion of bar coupons were accepted?**

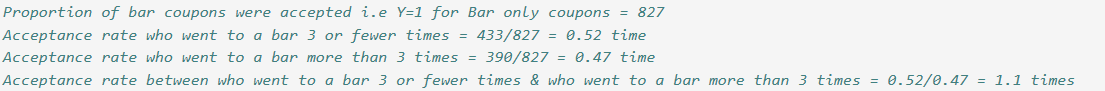
**A screenshot of a computer

Description automatically generated**

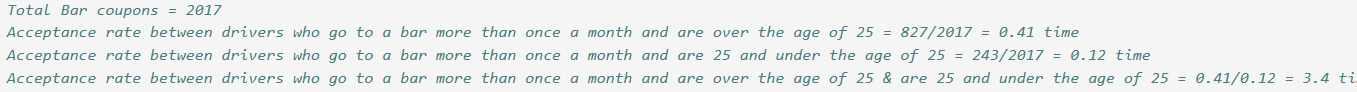
**A screenshot of a computer

Description automatically generated**

1. **Compare the acceptance rate between those who went to a bar 3 or fewer times a month to those who went more.**

****

1. **Compare the acceptance rate between drivers who go to a bar more than once a month and are over the age of 25 to the all others. Is there a difference?**



1. **Use the same process to compare the acceptance rate between drivers who go to bars more than once a month and had passengers that were not a kid and had occupations other than farming, fishing, or forestry.**

# compare the acceptance rate between drivers who go to bars more than once a month and had passengers that were not a kid and had occupations other

# than farming, fishing, or forestry.

# Proportion of bar coupons were accepted i.e Y=1 for Bar only coupons = 827

# Rate\_of\_acceptance\_with\_more3\_bar\_no\_kids\_no\_3F = 370/827 = 0.44 time

1. **Compare the acceptance rates between those drivers who:**

* go to bars more than once a month, had passengers that were not a kid, and were not widowed *OR*
* go to bars more than once a month and are under the age of 30 *OR*
* go to cheap restaurants more than 4 times a month and income is less than 50K.
* **Data analysis**# Acceptance of coupon by "Driver go to bars more than once a month, had passengers that were not a kid, and were not widowed" = 370
* # Acceptance of coupon by "Drive go to bars more than once a month and are under the age of 30" = 203
* #Acceptance rate comparison between "go to bars more than once a month, had passengers that were not a kid, and were not widowed & "go to bars more than once a month and
* #are under the age of 30" = 370/203 = 1.8 time

1. **Based on these observations, what do you hypothesize about drivers who accepted the bar coupons?**
2. Overall 56% of coupons are accepted
3. 41% acceptance rate for bar coupon i.e. Total bar coupon = 2017. Total bar coupon accepted = 827
4. Acceptance of coupon by "Driver go to bars more than once a month, had passengers that were not a kid, and were not widowed" = 370
5. Acceptance of coupon by "Drive go to bars more than once a month and are under the age of 30" = 203
6. Acceptance of coupon by "Drive go to cheap restaurants more than 4 times a month and income is less than 50K.
7. Acceptance rate comparison between "go to bars more than once a month, had passengers that were not a kid, and were not widowed & "go to bars more than once a month and

are under the age of 30" = 370/203 = 1.8 time

Number of coupons

1. Coffee House 3996
2. Restaurant(<20) 2786
3. Carry out & Take away 2393
4. Bar 2017
5. Restaurant(20-50) 1492

**Independent Investigation**

**General coupons acceptance by category**

**A graph with blue squares

Description automatically generated**

**A screenshot of a menu

Description automatically generated**

**Coffee House coupons acceptance rate**

* Total Coffee House coupons = 3996
* Accepted coupons = 1995
* Coffee House coupons acceptance rate = (1995/3996)\*100 = 49.9%

Resaving to github