

# Black Friday Data Analysis

## The Outliers



Kristee Song  
Lara Vartanian  
Lari Boymoushakian



# Black Friday Death Count

12

DEATHS

117

INJURIES



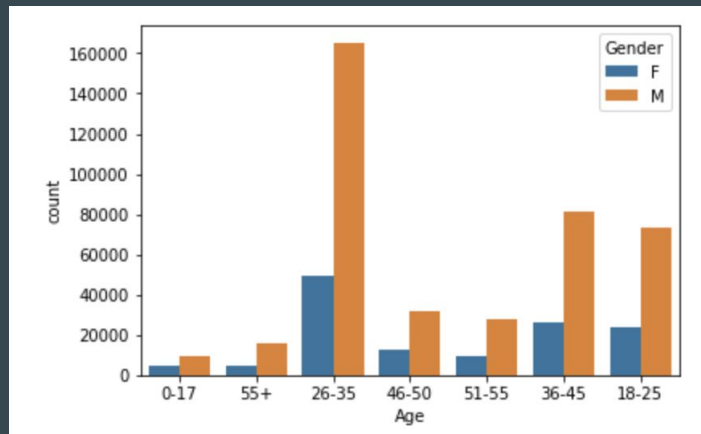
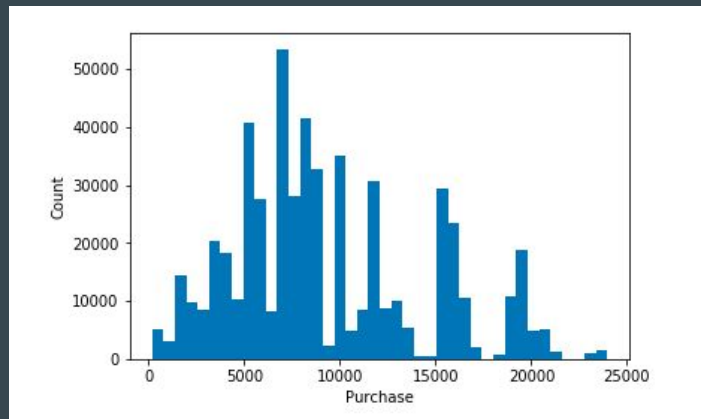


# Background Information

- User\_ID
- Product\_ID
- Gender
- Age (in bins)
- Occupation
- City\_Category (A, B, or C)
- Stay\_In\_Current\_City\_Years
- Marital\_Status
- Product\_Category\_1
- Product\_Category\_2 (optional)
- Product\_Category\_3 (optional)
- Purchase

# EDA

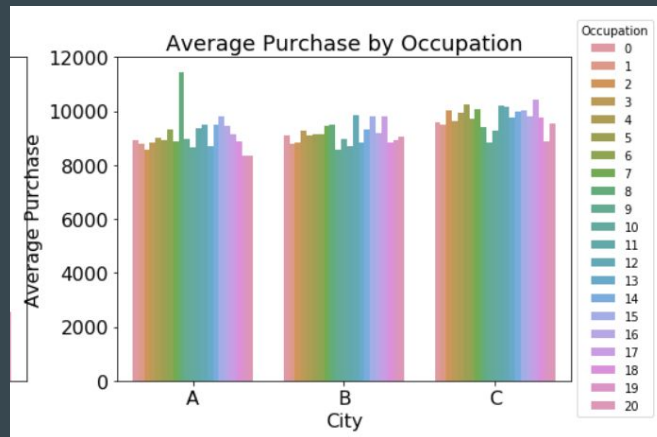
- Distribution of purchase amount is slightly right skewed, but approximately normal.
- We are interested in predicting each of the variables given all other variables.
- Gender and age seem to be a significant factor in determining purchase amount



# EDA

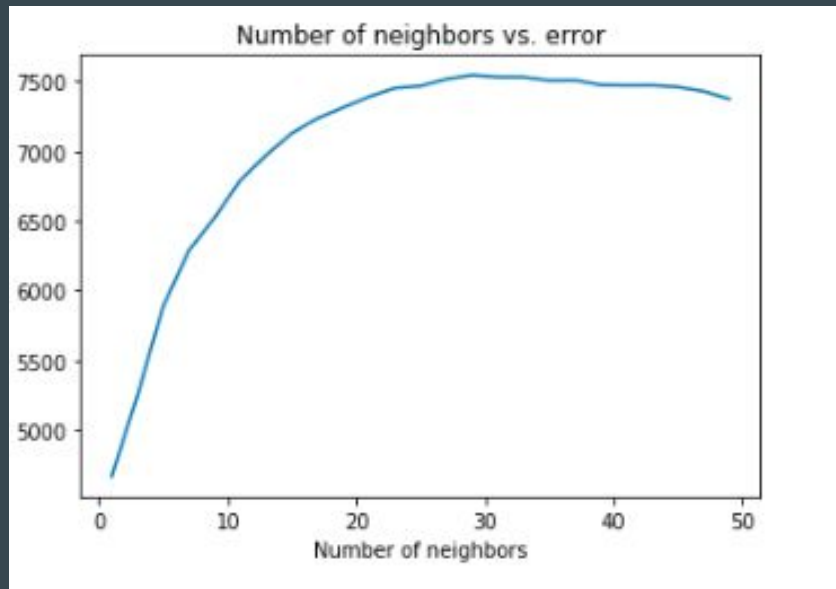
## Interesting findings from Data Exploration

- There are many repeating customers
- More male customers than female customers
- More single shoppers across gender
- Biggest chunk of customers from age group (26-35) and one age group above (36-45) and from one age group below (18-25)
- There isn't much variation in purchase amount across the age groups
- The 3 cities have similar occupation distributions which reflected similar total purchase amount across the 3 cities
- One outlier that stood out was Occupation 8 from City A bought on average more than Occupation 8 from City B and C.
- Most purchases were from customers who've resided for 1 year in the city.
- Product ID 1000 of category 18 in Product Category 1 had the highest purchase



# Modeling

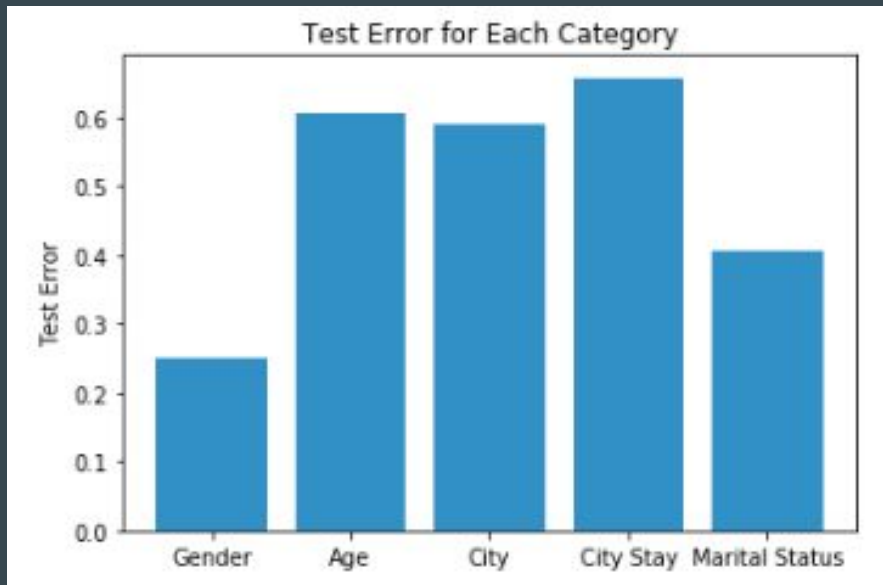
- KNeighborsClassifier was used to predict purchase amount.
- The best performance gave an RMSE of 4666 on the test data set.
- Standard deviation of the values in the Purchase variable was 4981.





# Modeling

- We also tested performance in predicting values from the categorical variables.
- The classifier did the best on the binary variables
  - ~75% accuracy classifying gender
  - ~60% accuracy predicting marital status



# Relatable Articles

According to a research paper done by Esther Swilley and Ronald E, Black Friday is more likely to increase retail sales by gender.

According to An Exploratory Investigation of Black Friday Consumption rituals by Jane Boyd Thomas, consumption rituals by the customer leads to repeated visitation to Black Friday stores than just a one-time shop, which could explain the repeated observations in our data.