Topic domain

We aim to analyze how Uber and Lyft adjust their prices in response to weather conditions, such as rain, snow, or extreme weather, and identify the circumstances under which users prefer one service over the other. We chose this topic since we usually use Uber or Lyft to school or to the airport and we find that under different weather, for the same distance, the price would change. We want to know more about how they will change the price, how it will bring benefit to the Company, and what the impact of changing their price.

A proposed research question or hypothesis

Our research question is: Is there any relationship between ride price and weather?
We are going to use Uber and Lyft data frames as an example. While the weather is more severe and inconvenient for commuting, the price of the rides is higher. In order to validate the hypothesis, we first use Uber data and Lyft data separately. We calculate the mean price in both Uber and Lyft in different weathers in the short summary, since the mean value can represent the general value for the price of different weather. Then put it into the "Mean Price in Different Weather" plot see the distribution and discuss does the weather impacts the price. Also, we want to make a regression model to visualize the relationship between price and distance in every weather in the short summary. Assuming the distance and price have a linear relationship, after visualizing the relationship we can observe the slope of each weather. If they have a similar slope, the weather will not influence the price and if they have a different slope, the weather could impact the price.

Two features:
  ● First pair: Sean Han, Sunny Li
Features: Finding influence factors of price change of ride apps

Scenario 1: How would the time period affect Lyft's price
Given: Rush time
When: daytime between hours between 7 am-8 am and daytime between hours between 4 pm-5 pm
Then the price will increase

Scenario 2: How would the time period affect Uber's price
Given: Rush time
When: daytime between hours between 7 am-8 am and daytime between hours between 4 pm-5 pm
Then the price will increase

- Second pair: Xiaonan Yang, Siyi Jin

Feature: Finding Influence Factors of Price Change of Ride Apps
Scenario 1: How would weather affect Uber's price
Given: Bad weather
When: Humidity above 70% (Rainy)
Then: Price will increase

Scenario 2: How would weather affect Lyft's price
Given: Bad weather
When: Temperature lower than 30 F
Then: Price will increase

Data Source:
Our data Source is on Kaggle called *Uber and Lyfy Dataset Boston, MA.*
(*https://www.kaggle.com/datasets/brllrb/uber-and-lyft-dataset-boston-ma*) This dataset collects data from Uber and Lyft, and it includes all the data that we are looking for. Such as date, time, precipitation, humidity, and price

Data visualizations:
In time feature:
We will assert daytime between hours between 7 am-8 am and daytime between hours between 4 pm-5 pm as rush time in a day. We are going to make a line plot between time and price to see does the price changes during the rush time. For the second plot, we are going to make a scatterplot to see if there is any relationship between time and price. If there is a correlation, we will see an obvious straight line in the plot.

In weather feature:
We use Lyft and Uber data separately, and we are going to make two scatterplots for each ride app to see the distribution of prices in each weather. We also going to make a correlation matrix to see the relationship between weather and price. If there is a strong correlation, it will shows red and if there is a weak correlation there will be a blue in the matrix.