

Exploratory Data Analysis: Ride-Sharing Trips

Data Preprocessing and EDA
Telling the Story Behind 500,000 Rides

Presented by:

Malak Ahmed Saber

Marawan Abbas Mohamed

Dataset Overview



500,000 ride
records

26 features

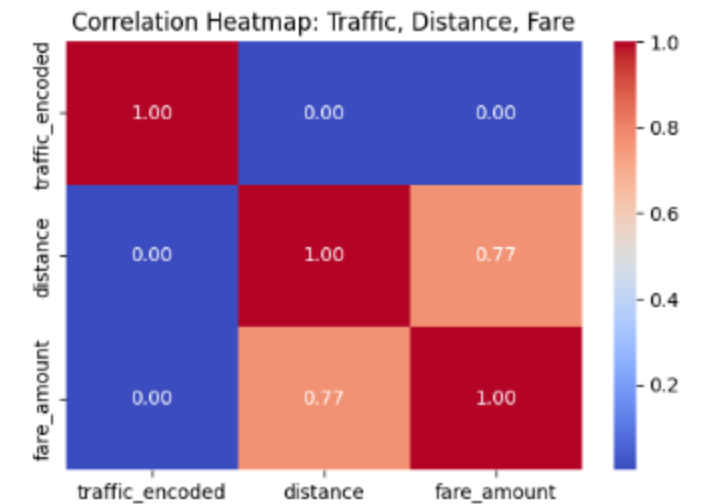
Includes time,
traffic, weather,
and geo data

No Duplicated
values

Data Preprocessing

- Converted 'pickup_date' from object to date time
- Null Values
 - Detected Nulls in 6 features (5 in each)
 - Displayed 5 nulls of first feature which appeared to affect all the other features to be null so only 5 records had 6 null features
 - Decided that the best way was to fill with mean of each feature
- Converted Categorical Columns to 'category'
- Outliers removed to improve data quality and ensure more robust analysis.
- Found Correlation between the most important features :
 - Traffic Condition
 - Fare Amount
 - Distance

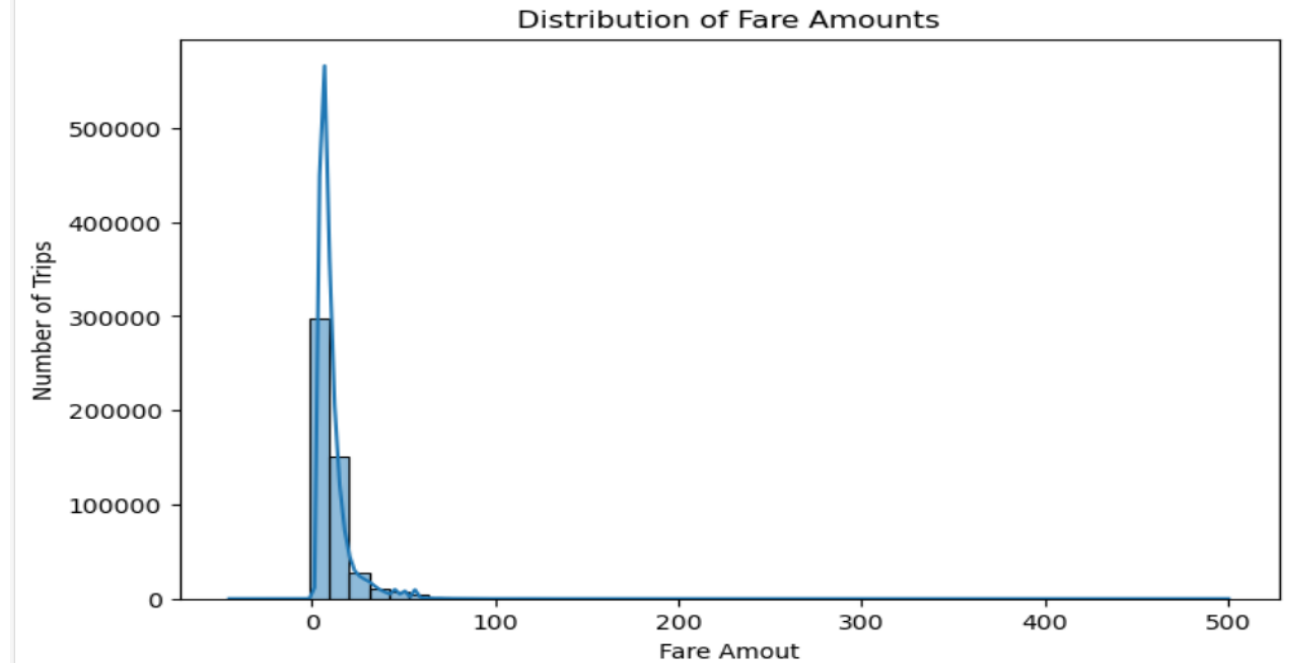
	distance	fare_amount
Traffic Condition		
Congested Traffic	20.044778	11.390527
Dense Traffic	18.739325	11.368120
Flow Traffic	19.621328	11.316381



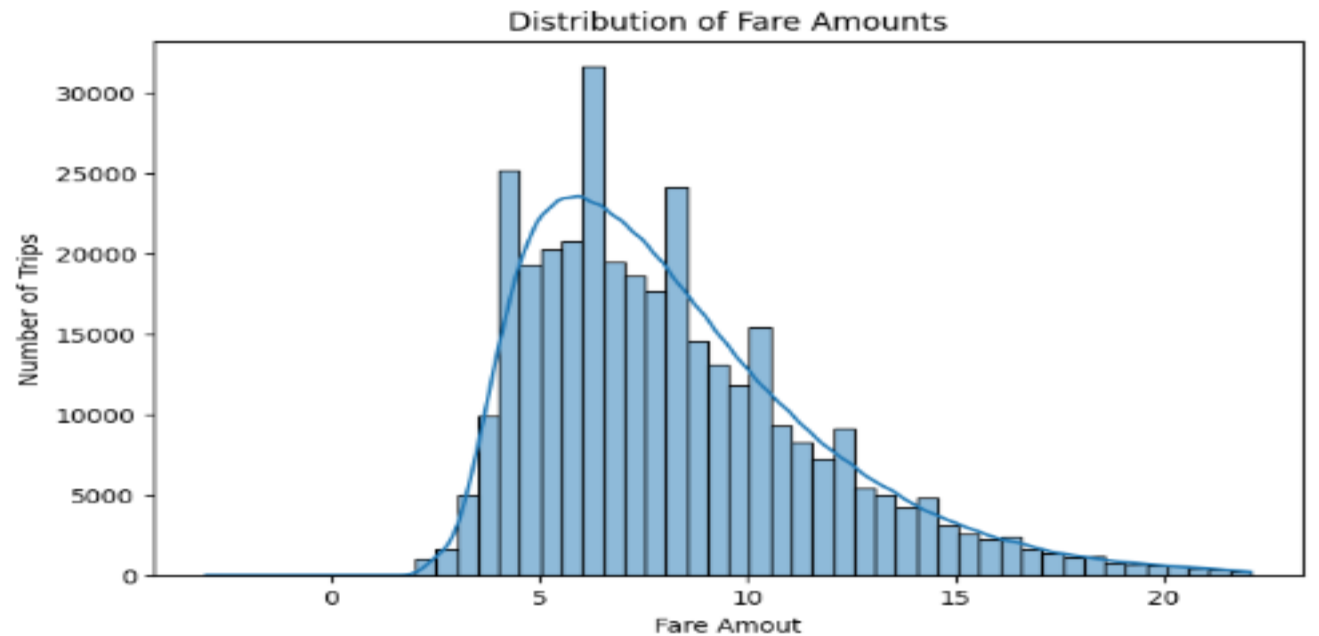
How much do we pay for a ride?

Insight : Most fares are under \$20

Before removing outliers

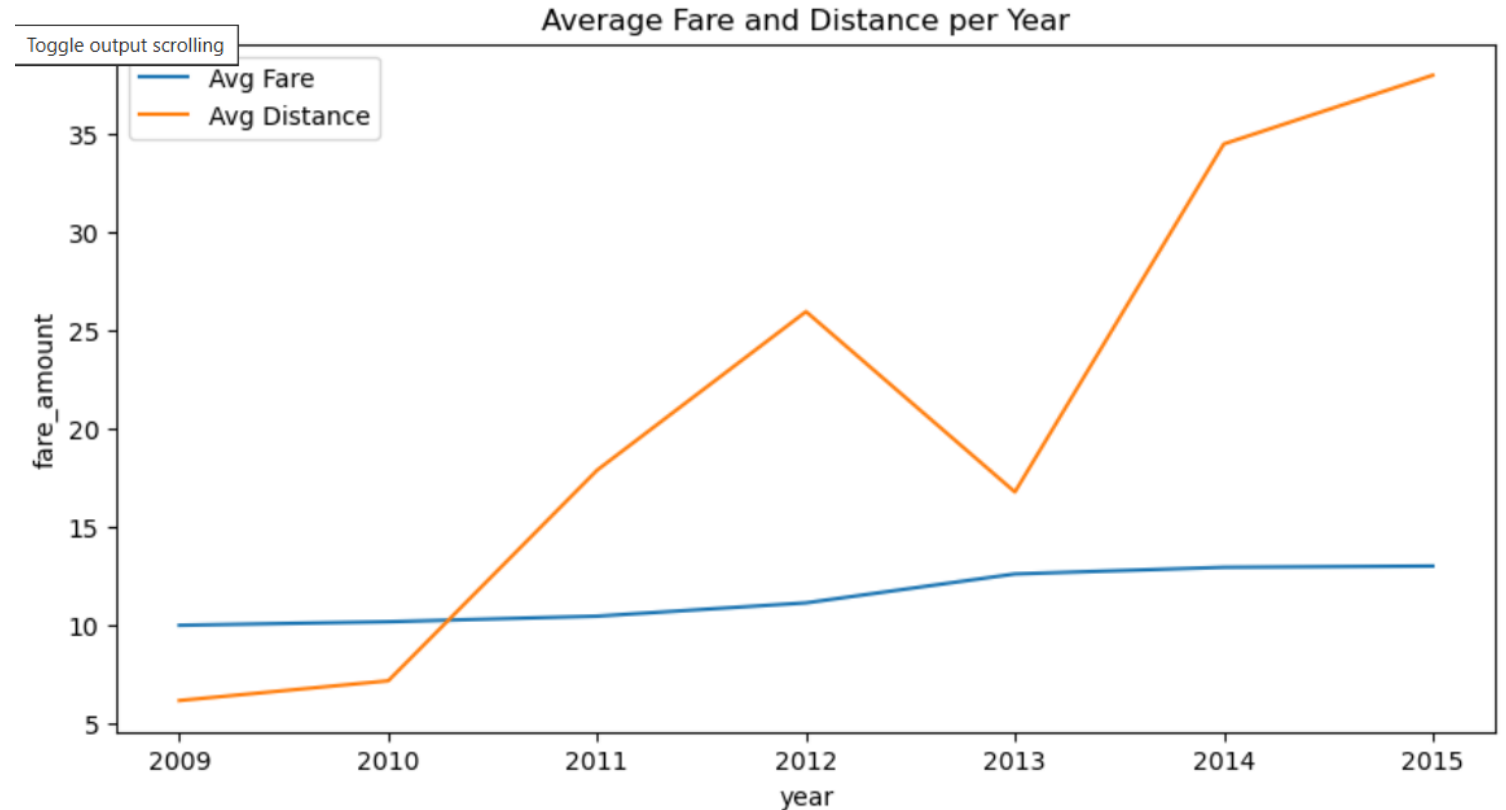


After removing outliers



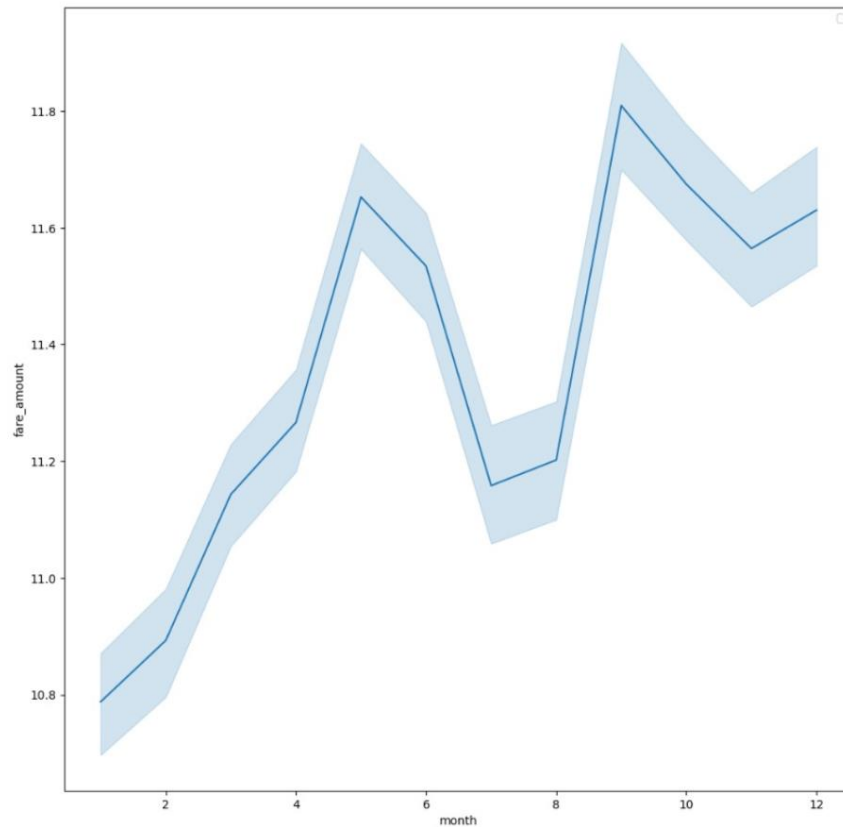
How is the fare affected along with distance among different years?

Insight : Average fare is very stable over time while distance increases or decreases.

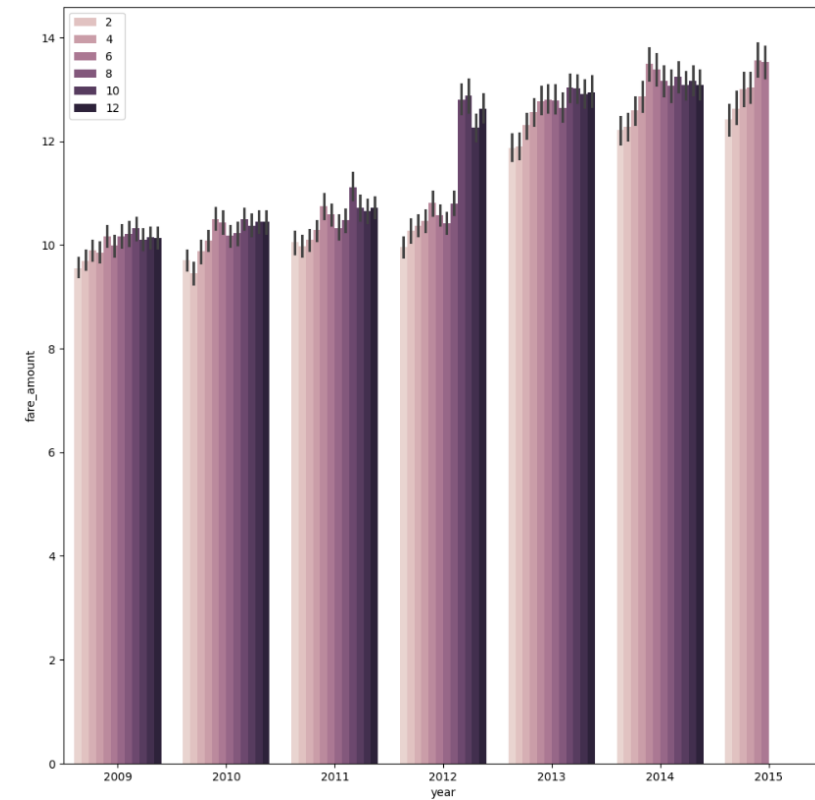


Do month affect Fare price?

month effect on avg price



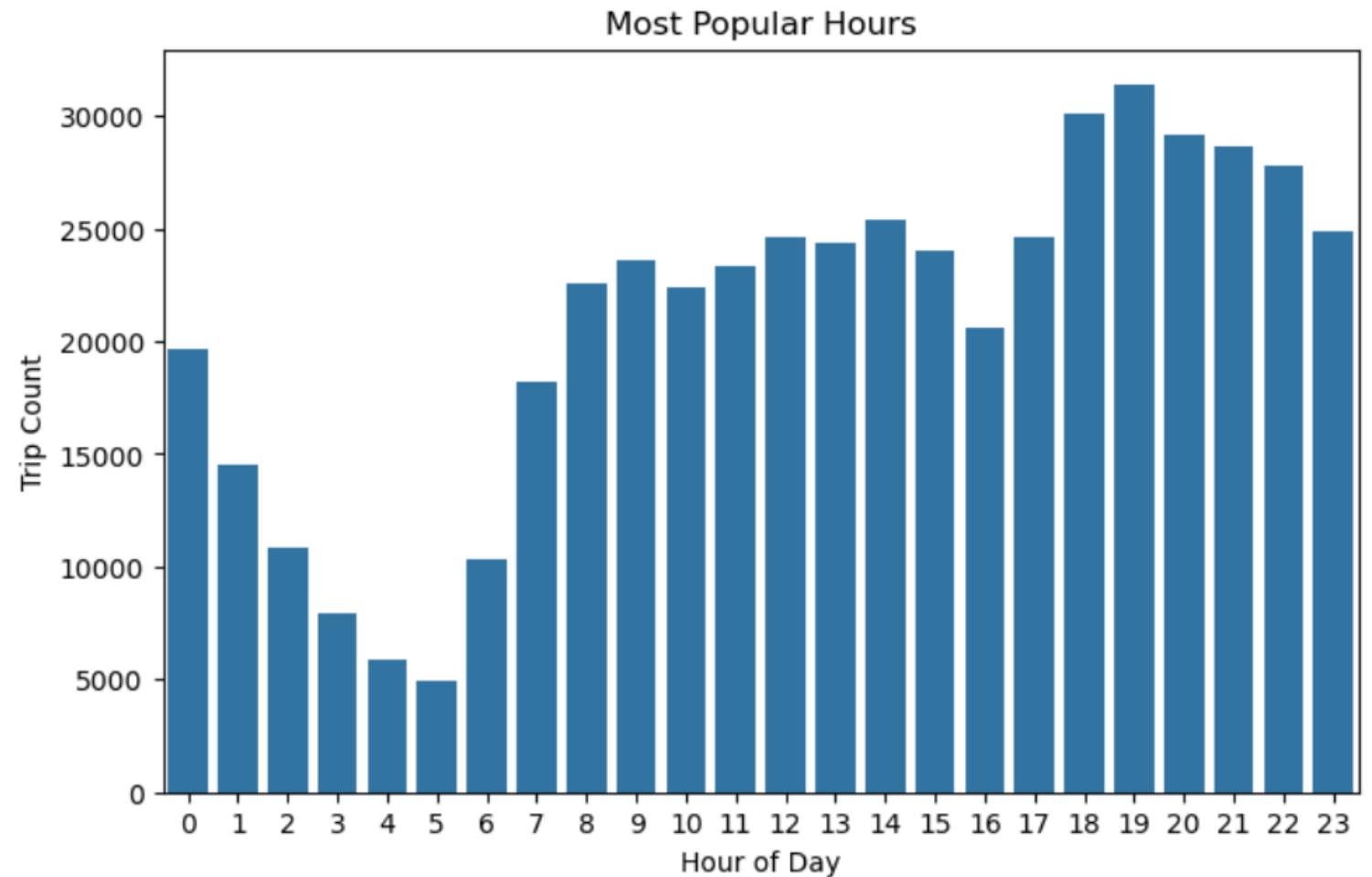
Effect of months over the years



Insight : There is a slight variation in fare prices across different months.

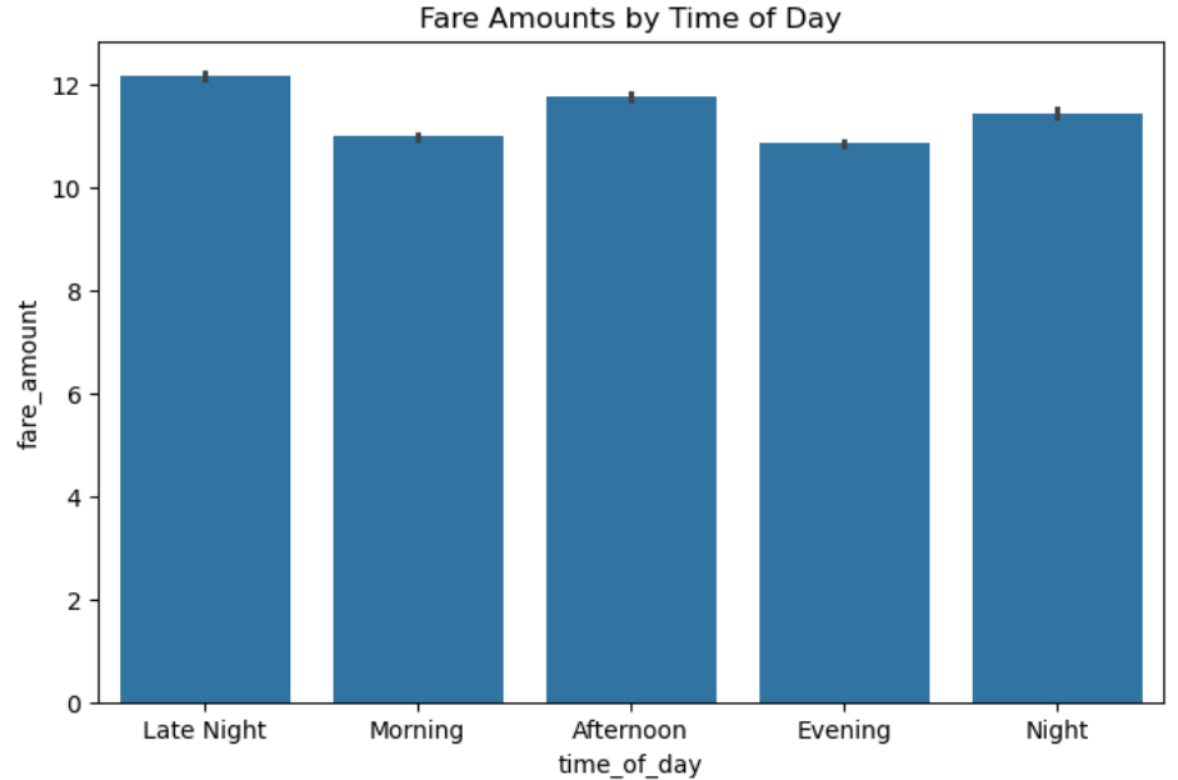
When Do we ride?

Insight : Riding is more active at night, very low at 5 morning



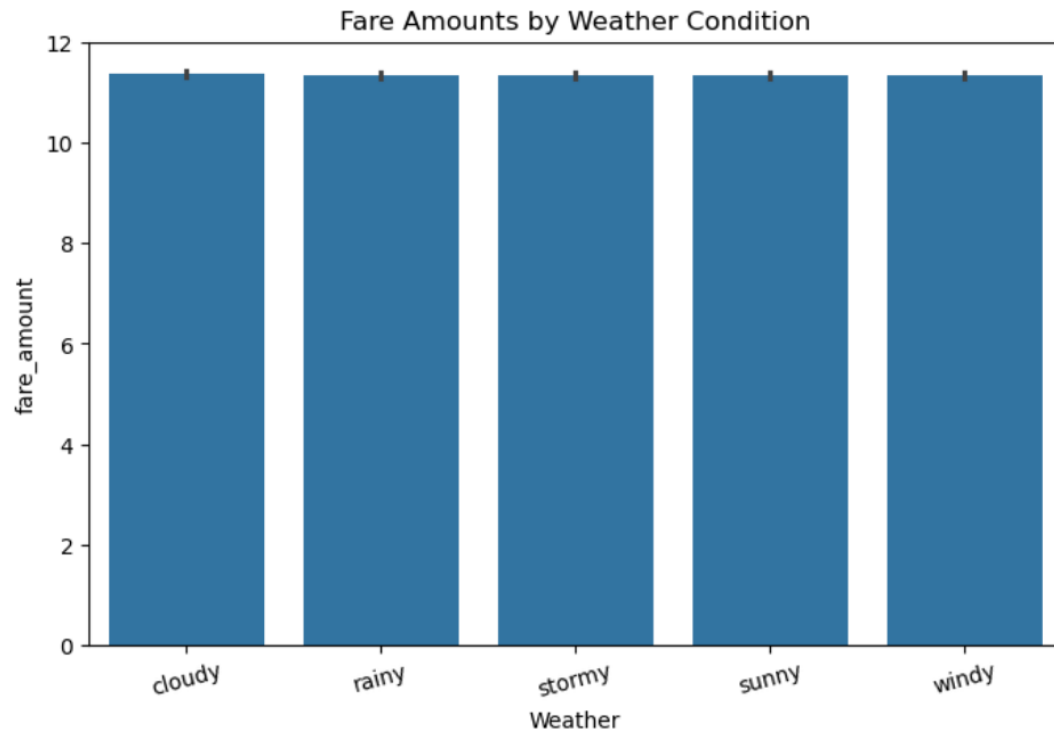
Does this timing affect Fares?

Insight : Late nights are the most paid, maybe as it is more non safe

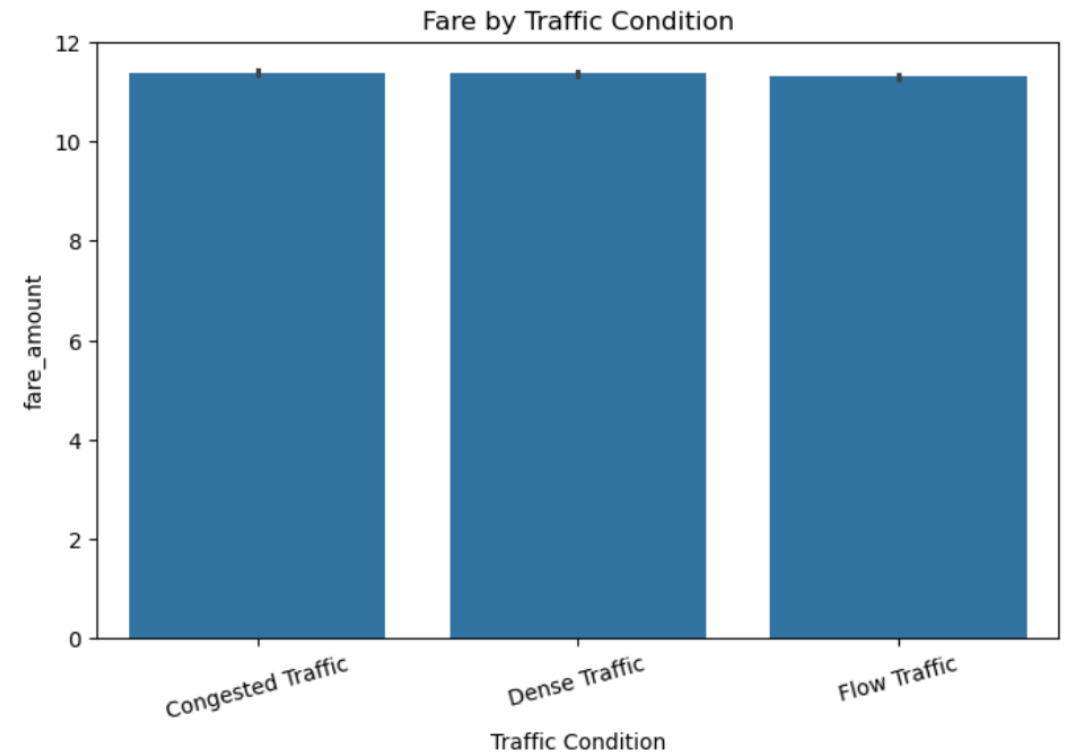


Do Conditions affect Fares?

Weather Condition



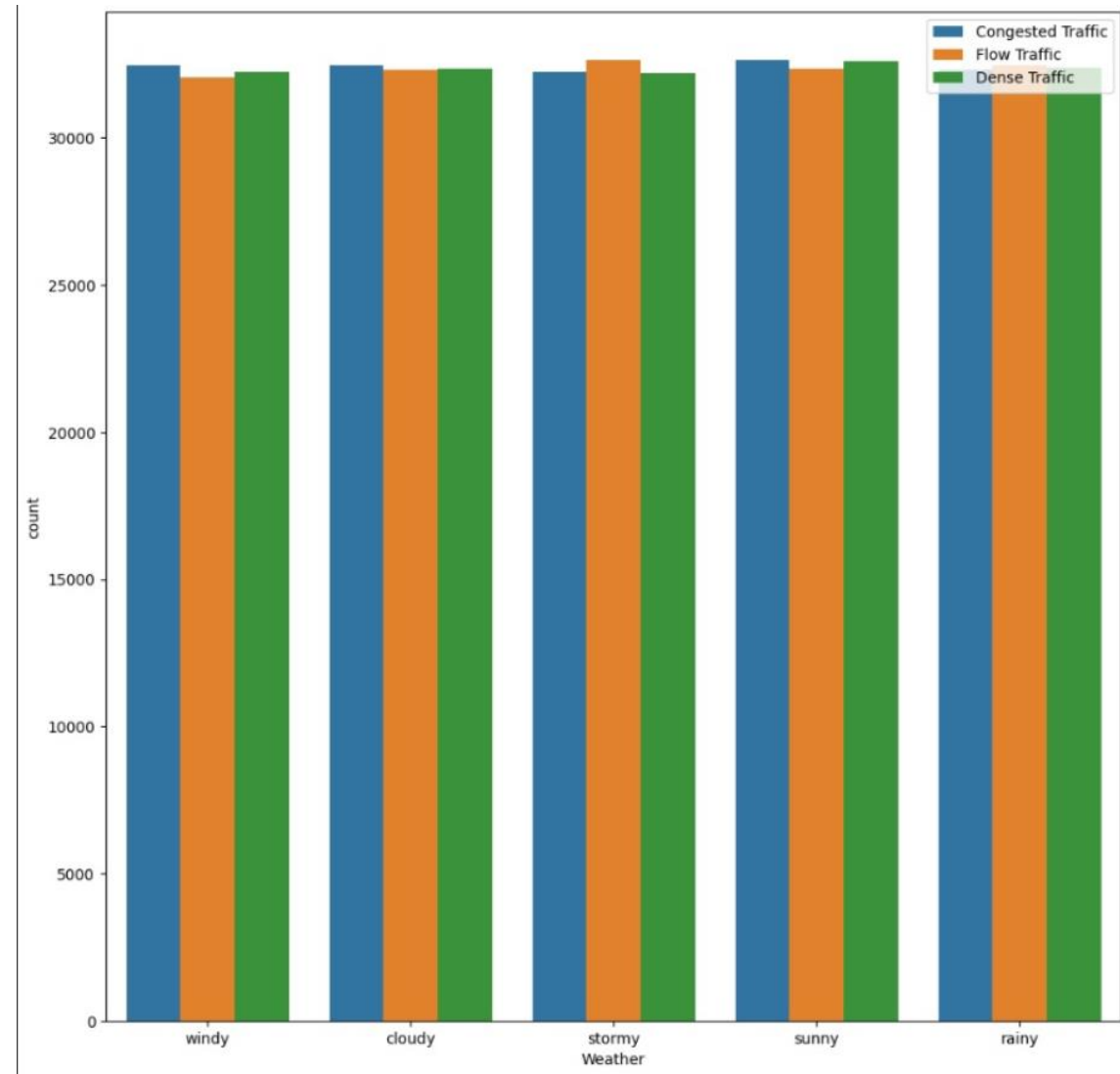
Traffic Conditions



Insight : Fares are almost equal for different weather conditions and traffic conditions

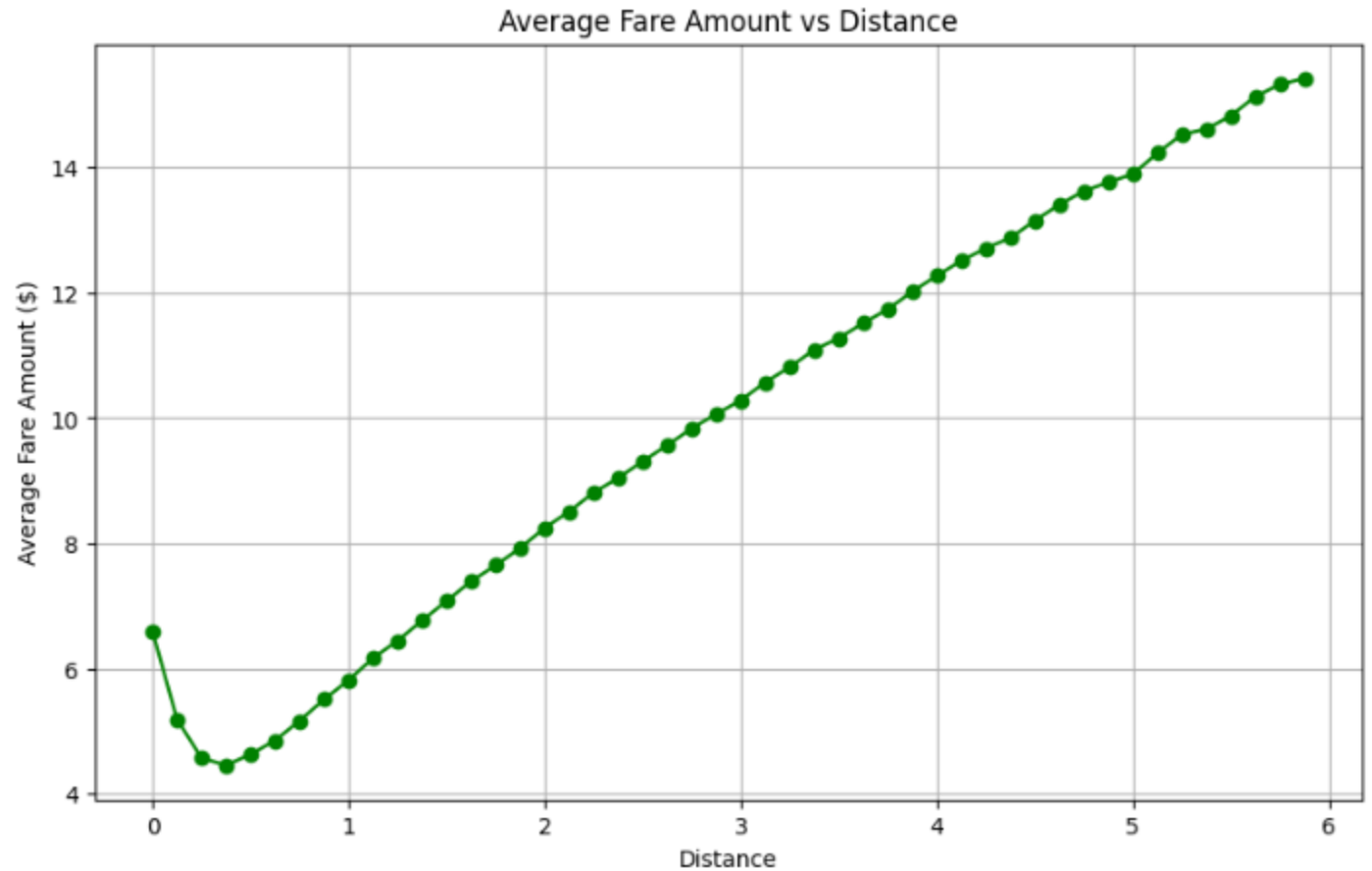
How do weather conditions influence traffic flow??

Insight : it has no significant effect.



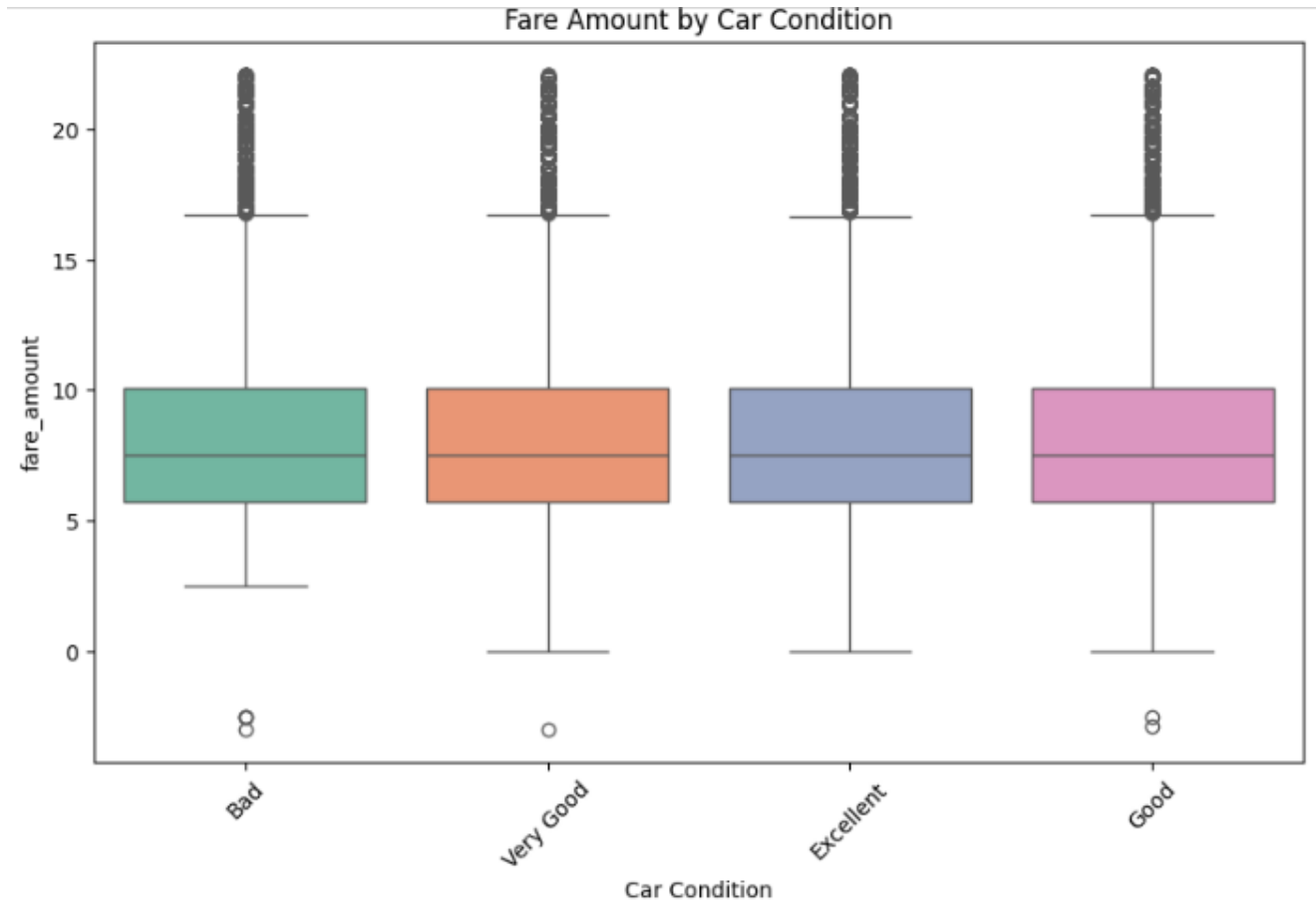
How does distance affect avg fare price?

Insight : distance have a linear relation with the avg price



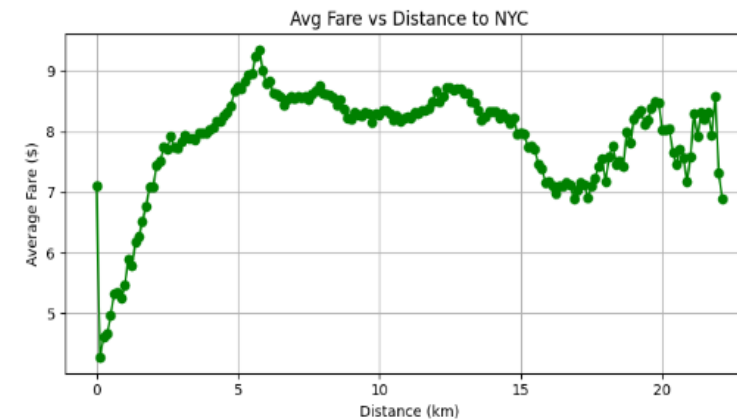
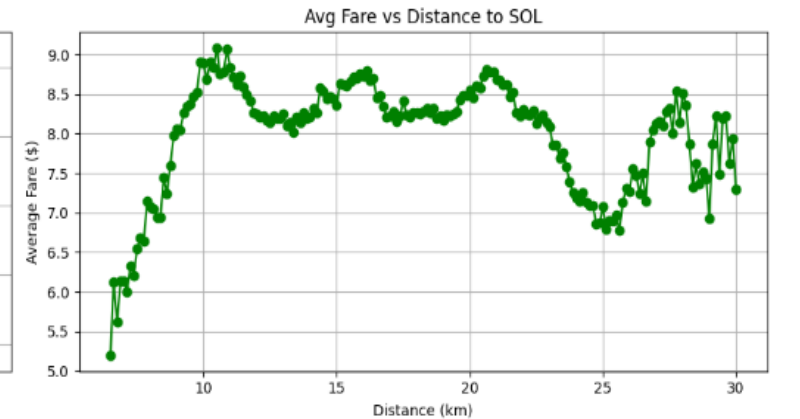
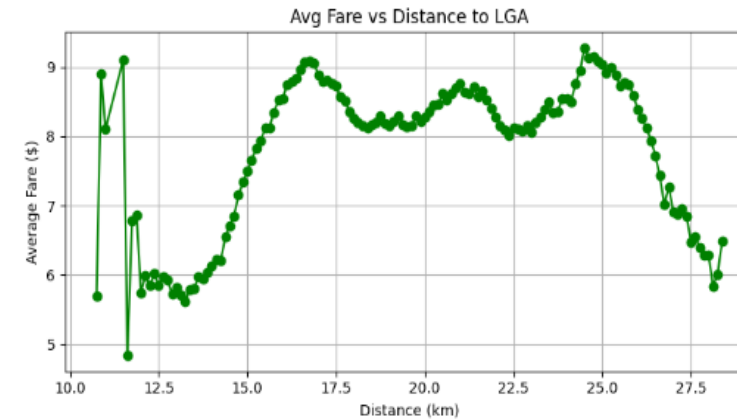
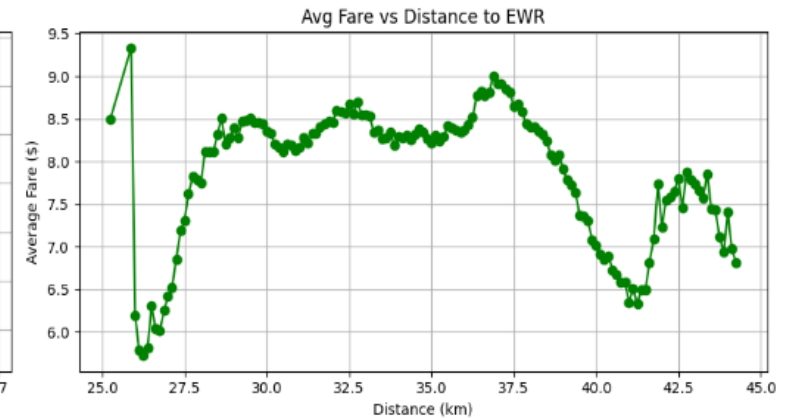
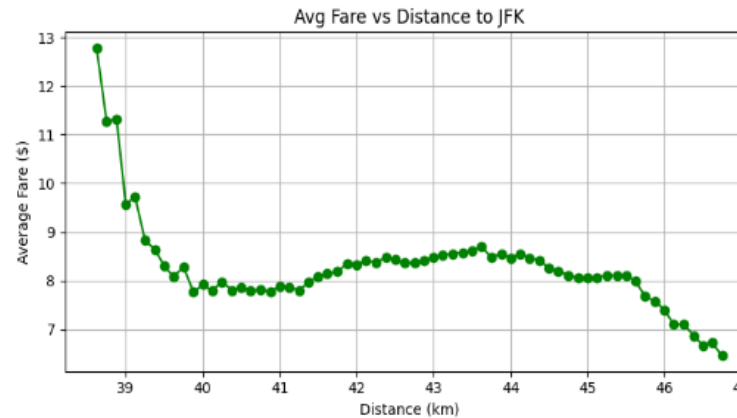
How does car condition affect fare price?

Insight : There is no notable difference based on the car's condition.



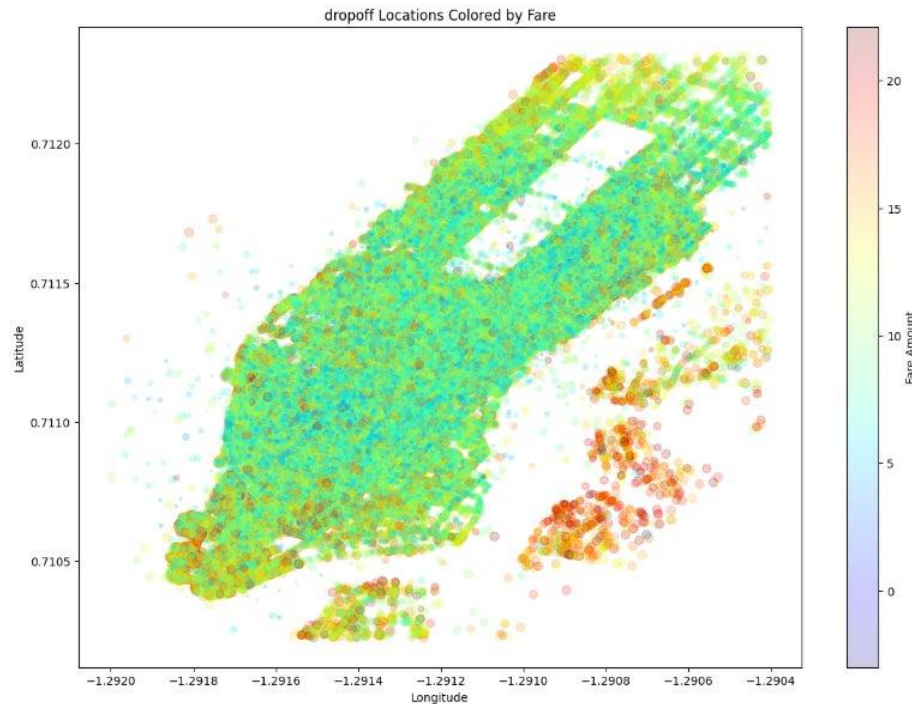
Does proximity to airports influence fare prices? ?

Insight :Average fares vary with the distance to each airport. As distance to JFK increases, fares tend to decrease, while fares to LGA, EWR, and SOL fluctuate with distance

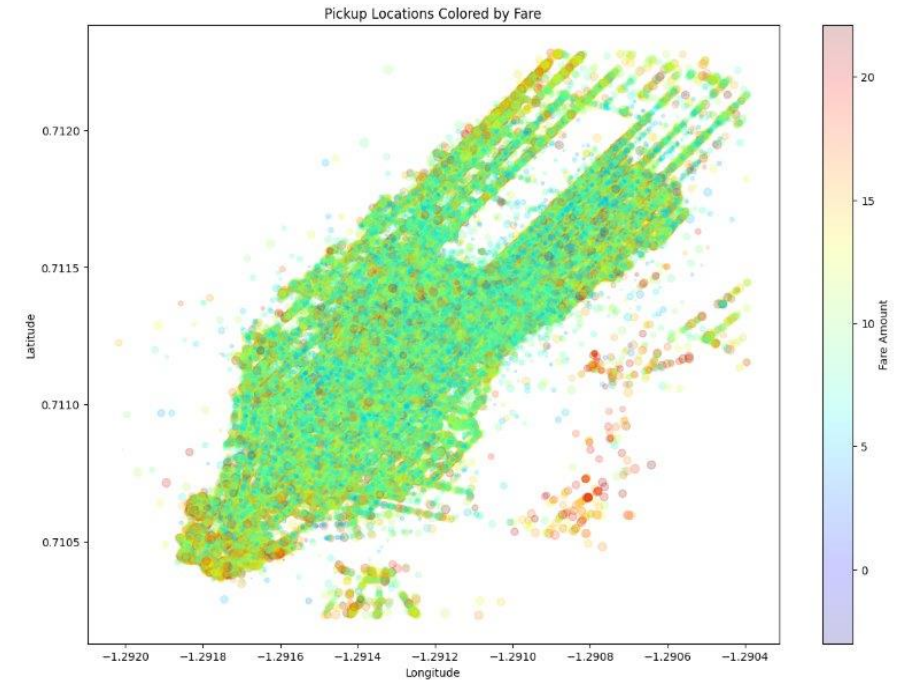


Which drop-off locations tend to have higher fares and longer trip durations?

Drop-off locations



Pickup locations



Insight : Passengers typically incur higher fares and longer trip durations when the drop-off or pickup location is on the outskirts of the city.