Project proposal – Yellow Taxi Trip

The purpose of this project is to look at the yellow taxi trip records, in order to investigate passenger's behavior.

This project will focus on four things. First, at what time of the day passenger will give more tips. Second, look at the relation between number of passenger and tip amount. Next at what time of the day passengers usually dispute with taxi drivers. Last, know when the demand for taxi increase.

This analysis will help the taxi drivers as well as the company to know when the best time to serve and when to reduce or avoid serving.

Data Obtained

The data gained from NTC taxi and limousine commission https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page downloaded as .csv then comforted to .dp using SQLite.

Characteristics/Features

Data Provider

Field Name	Description
VendorID	A code indicating the TPEP provider that provided the record. 1= Creative Mobile Technologies, LLC; 2= VeriFone Inc.

Trip Information

Field Name	Description
tpep_pickup_datetime	The date and time when the meter was engaged.
tpep_dropoff_datetime	The date and time when the meter was disengaged.
Passenger_count	The number of passengers in the vehicle. This is a driver-entered value.
Trip_distance	The elapsed trip distance in miles reported by the taximeter.
PULocationID	TLC Taxi Zone in which the taximeter was engaged
DOLocationID	TLC Taxi Zone in which the taximeter was disengaged

RateCodeID	The final rate code in effect at the end of the trip. 1= Standard rate 2=JFK 3=Newark 4=Nassau or Westchester 5=Negotiated fare 6=Group ride
Store_and_fwd_flag	This flag indicates whether the trip record was held in vehicle memory before sending to the vendor, aka "store and forward," because the vehicle did not have a connection to the server. Y= store and forward trip N= not a store and forward trip

Payments

Field Name	Description
Payment_type	A numeric code signifying how the passenger paid for the trip. 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip
Fare_amount	The time-and-distance fare calculated by the meter
Extra	Miscellaneous extras and surcharges. Currently, this only includes the $0.50 and $1 rush hour and overnight charges.
MTA_tax	$0.50 MTA tax that is automatically triggered based on the metered rate in use.
Improvement_surcharge	$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015.
Tip_amount	Tip amount – This field is automatically populated for credit card tips. Cash tips are not included.
Tolls_amount	Total amount of all tolls paid in trip.
Total_amount	The total amount charged to passengers. Does not include cash tips.

Tools

- Jupiter: creating various graphical representations to understand the data better using NumPy, Pandas, Matplotlib, and Seaborn library.
- SQLite: managing the tables