PROJECT PROPOSAL

Regression Project Bootcamp T5

Questions

- 1- How do we get the prediction for the value of trips?
- 2- Is a congestion effect in total amount?
- 3- Is a payment type affect in total amount?

Data description

Dataset for yellow taxi trip in NYC in July 2021 This data dictionary describes yellow taxi trip data. For a dictionary describing green taxi data, or a map of the TLC Taxi Zones, please visit http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml . The data provided as CSV file. The data base about 2821515 trips with 18 features. The target in this project is Total amount

Features:

| The feature | Description |
|-----------------------|--|
| VendorID | A code indicating the TPEP provider that provided the record. |
| | 1= Creative Mobile Technologies, LLC; |
| | 2= VeriFone Inc |
| 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. |
| RatecodelD | 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 |

| PULocationID TLC Taxi Zone in which the taximeter was engaged TLC Taxi Zone in which the taximeter was disengaged 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 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. tip_amount Total amount — This field is automatically populated for credit card tips. Cash tips are not included. Total amount of all tolls paid in trip. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | DOLocationID TLC Taxi Zone in which the taximeter was disengaged 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 with a wind overnight charges. 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. tip_amount Tip amount — This field is automatically populated for credit card tips. Cash tips are not included. totls_amount Total amount of all tolls paid in trip. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | | |
|---|---|-----------------------|---|
| 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 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | 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 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | PULocationID | TLC Taxi Zone in which the taximeter was engaged |
| 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | the trip. 1= Credit card 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | DOLocationID | TLC Taxi Zone in which the taximeter was disengaged |
| 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | 2= Cash 3= No charge 4= Dispute 5= Unknown 6= Voided trip 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | payment_type | |
| 3= No charge 4= Dispute 5= Unknown 6= Voided trip 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | 3= No charge 4= Dispute 5= Unknown 6= Voided trip 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | | 1= Credit card |
| 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | 4= Dispute 5= Unknown 6= Voided trip 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | | 2= Cash |
| 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | fare_amount 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | | 3= No charge |
| 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | | 4= Dispute |
| 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | total_amount The time-and-distance fare calculated by the meter Miscellaneous extras and surcharges. Currently, this only includes the \$0.50 and \$1 rush hour and overnight charges. \$0.50 MTA tax that is automatically triggered based on the metered rate in use. Tip amount — This field is automatically populated for credit card tips. Cash tips are not included. Total amount of all tolls paid in trip. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. The total amount charged to passengers. Does not | | 5= Unknown |
| mta_tax Miscellaneous extras and surcharges. Currently, this only includes the \$0.50 and \$1 rush hour and overnight charges. \$0.50 MTA tax that is automatically triggered based on the metered rate in use. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | 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. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | | 6= Voided trip |
| only includes the \$0.50 and \$1 rush hour and overnight charges. \$0.50 MTA tax that is automatically triggered based on the metered rate in use. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | only includes the \$0.50 and \$1 rush hour and overnight charges. \$0.50 MTA tax that is automatically triggered based on the metered rate in use. 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | fare_amount | The time-and-distance fare calculated by the meter |
| tip_amount tolls_amount Total amount of all tolls paid in trip. tolls_amount surcharge \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | 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. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | extra | only includes the \$0.50 and \$1 rush hour and |
| tolls_amount Total amount of all tolls paid in trip. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount Total amount of all tolls paid in trip. \$1.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. The total amount charged to passengers. Does not include cash tips. | tolls_amount Total amount of all tolls paid in trip. \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount Total amount charged to passengers. Does not | mta_tax | |
| \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | improvement_surcharge \$0.30 improvement surcharge assessed trips at the flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | tip_amount | |
| flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not include cash tips. | flag drop. The improvement surcharge began being levied in 2015. total_amount The total amount charged to passengers. Does not | tolls_amount | Total amount of all tolls paid in trip. |
| include cash tips. | | improvement_surcharge | flag drop. The improvement surcharge began being |
| | · | total_amount | |
| congestion_surcharge Additional Congestion Charge | congestion_surcharge Additional Congestion Charge | congestion_surcharge | Additional Congestion Charge |

Tools

In project will use for programming python3 by Jupyter notebook (anaconda33).

Data analysis in pandas.

Python visualization libraries (matplotlib & seaborn).

Module data by linear regression.

The goal of project

Explore and analyze the data.

Prediction the value of a target.