5DMACP04

Sendy Logistics Challenge

Team 5D06

Tejaswini Kale	257	01FE17BCS231
Usman Khan	261	01FE17BCS235
Vibha Hegde	266	01FE17BCS240
K L Vijeth	273	01FE17BCS247

Introduction

- Sendy is an e-commerce platform which offers door to door deliveries of goods in Kenya. It is a delivery solution for businesses.

- The problem statement focuses on estimating the delivery time taken by the riders to deliver the orders in Nairobi Region.

Runner

Convenient and fast delivery with our walking couriers.

BASE PRICE: KSH 90

CBD to CBD delivery

Delivered within 1 hr

Standard Bike

One fixed price anywhere in Nairobi! Great for day-to-day deliveries.

FLAT PRICE: KSH 250

Anywhere in Nairobi

Same-day delivery

Express Bike

Your designated Bike Rider for urgent deliveries.

BASE PRICE: KSH 340

Extra: Ksh 20/KM after 5KM

Delivered within 1 hr

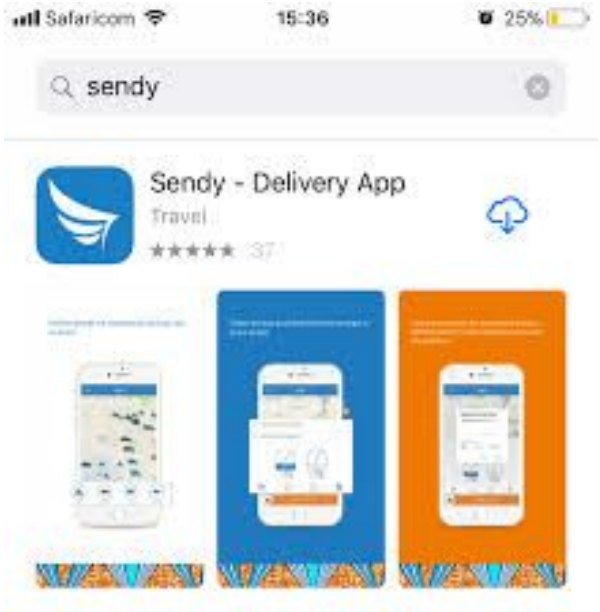
Problem Statement

To Predict the estimated time of arrival (ETA) for motorbike deliveries in Nairobi,Kenya.

(Time taken to deliver order from the point of driver pickup to the point of arrival at final destination.)

Problem Description

- The dataset comes from the sendy API and the sendy web and mobile app which connects customers with transporters.
- When customers place delivery requests through the platform, Sendy dispatches these orders in real time to the closest available Sendy Partner Drivers.
- The orders mentioned in the dataset are sendy express orders which are delivered in and around Nairobi, Kenya.



Data Description

Train.csv (3.9 MB)
Test.csv (1.2 MB)
Riders.csv (29.1 KB)
21201 Records
29 Attributes

Attribute	Description	Example
Order No.	A number which identifies the orders uniquely.	Ex: Order_No_4211
User Id	A unique number for the customers who placed the order	Ex: User_Id_2642
Vehicle Type	Category of vehicle used for delivery.	Value: Bike
Platform Type	Categorical value identifying the booking platform.	Values: 1,2,3,4
Order Type	Defines the type of order.	Values: Business/Personal
Rider Id	A unique number to identify the riders.	Ex: Rider_Id_432

Table 1: Order Details

Attribute	Description	Example
Day of Month	Categorical attribute that describes the date	Values: 1-31
Weekday	Categorical attribute of day of week.	Values: Sunday - Saturday
Time	HH:MM:SS in 12 hour format.	9:40:10 AM

Table 2: Time attribute format

Attribute type
Placement Times
Confirmation Times
Arrival at Pickup Times
Pickup Times

Table 3: Time attribute Types

Attribute	Description	Example
Distance	Distance between pickup and drop location.	Ex: 3 (in kms)
Pickup Latitude	Coordinate of pickup.	Ex: -1.1875
Pickup Longitude	Coordinate of pickup.	Ex: 35.2314
Destination Latitude	Coordinate of drop location	Ex: -1.3425
Destination Longitude	Coordinate of drop location	Ex: 36.9024

Table 4: Distance Attributes

Attribute	Description	Example
Temperature	Temperature at the time of placement of order	Ex: 25 (in degree celsius)
Precipitation	Precipitation at the time of placement of order	Ex: 3 (in mm)

Table 5: Natural Factors affecting delivery times

Attribute	Description	Example
Rider Id	A unique number to identify the riders.	Ex: Rider_Id_432
No of Orders	Number of orders the rider has completed.	Ex: 380
Age	Age of rider as an employee of the company in days.	Ex: 2298
Average Rating	Average rating of the rider	Ex: 13.5
Number of ratings	The rating of rider. Optional to user.	Ex: 519

Table 6: Attributes in Riders.csv

Attribute	Description	Example
Pickup Time	Time at which order is picked by rider	Ex: 10:15:55 AM
Arrival at Destination time	Time at which order reaches the destination	Ex: 10:30:05 AM
Time from pickup to arrival	The time taken in seconds	Ex: 850 seconds

Table 7: Prediction Variables

Attribute	Description	Example
Order Id	.A number which identifies the orders uniquely.	Ex: Order_No_4211
Time from Pickup to Arrival	The time taken in seconds	Ex: 380

Table 8: Submission Format - csv file

Evaluation Metric

Root Mean Squared Error

Exploratory Data Analysis

- Analysis of Attributes
- Outlier Analysis
- Missing Values

Observations in the Data:

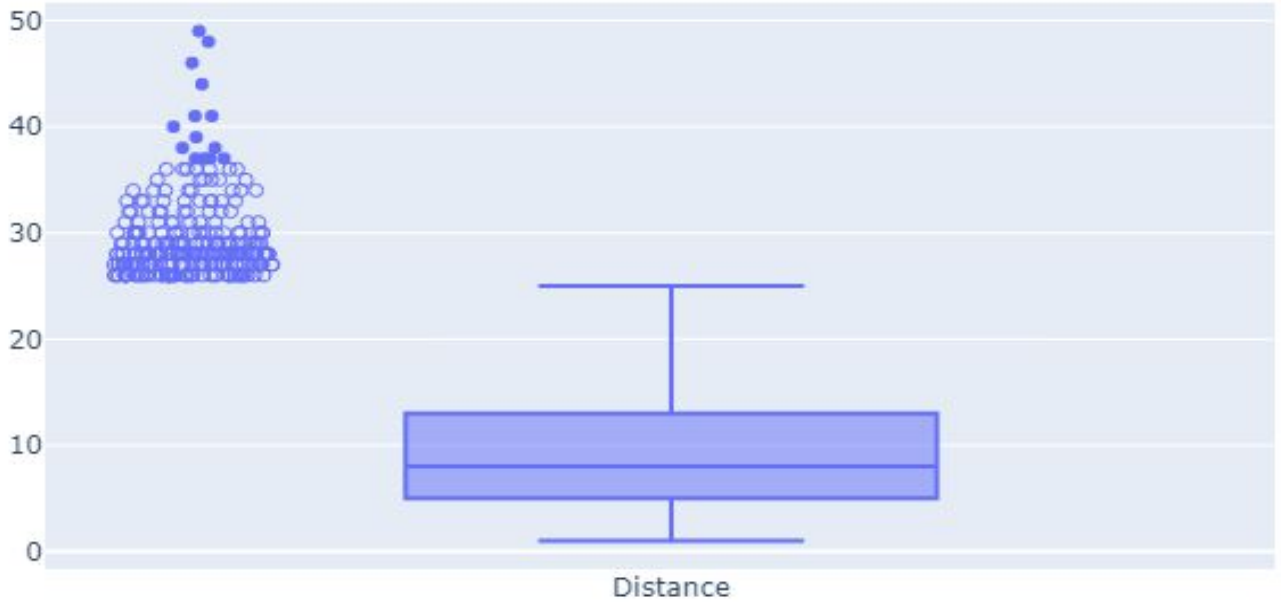
- Percentage of Null Values

Temperature	20.593368%	4091
Precipitation	97.396348%	20649

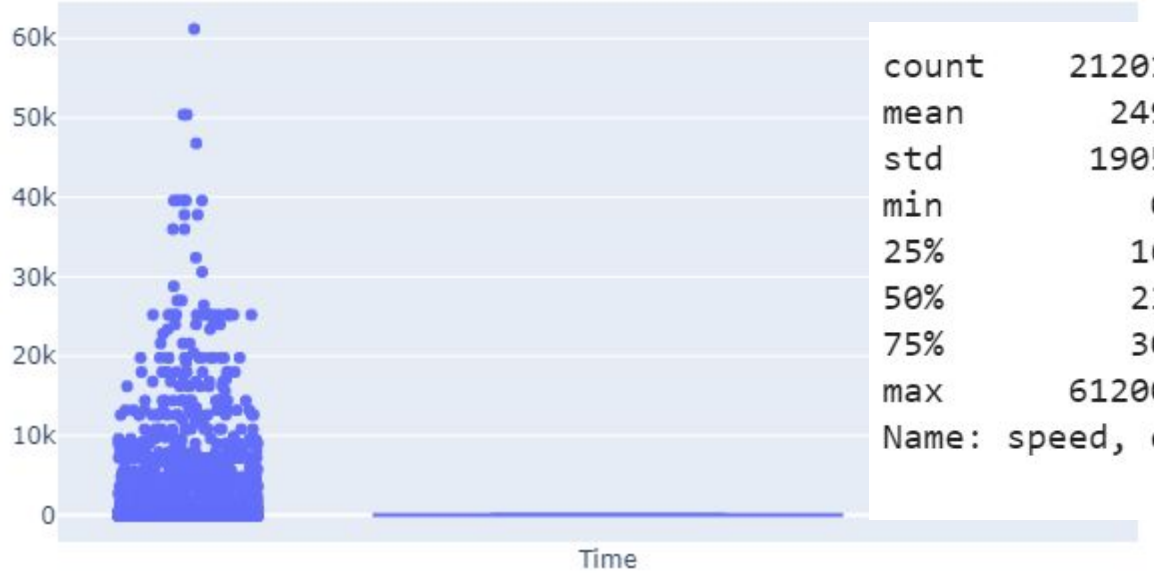
- All the time attributes belong on the same day but at different time stamps.

	Order_No	User_Id	Vehicle_Type	Platform_Type	Type	Placement_Day_of_Month	Placement_Weekday	Placement_Time	Confirmation_Day_of_Month	Confirmation_Weekday	Confirmation_Time
5096	Order_No_3530	User_Id_2457	Bike	3	Business	11	2	12:22:05	13	4	13:27:30
16629	Order_No_2800	User_Id_1402	Bike	3	Business	17	1	14:34:37	18	2	13:16:53

Outlier Analysis of Distance Attribute

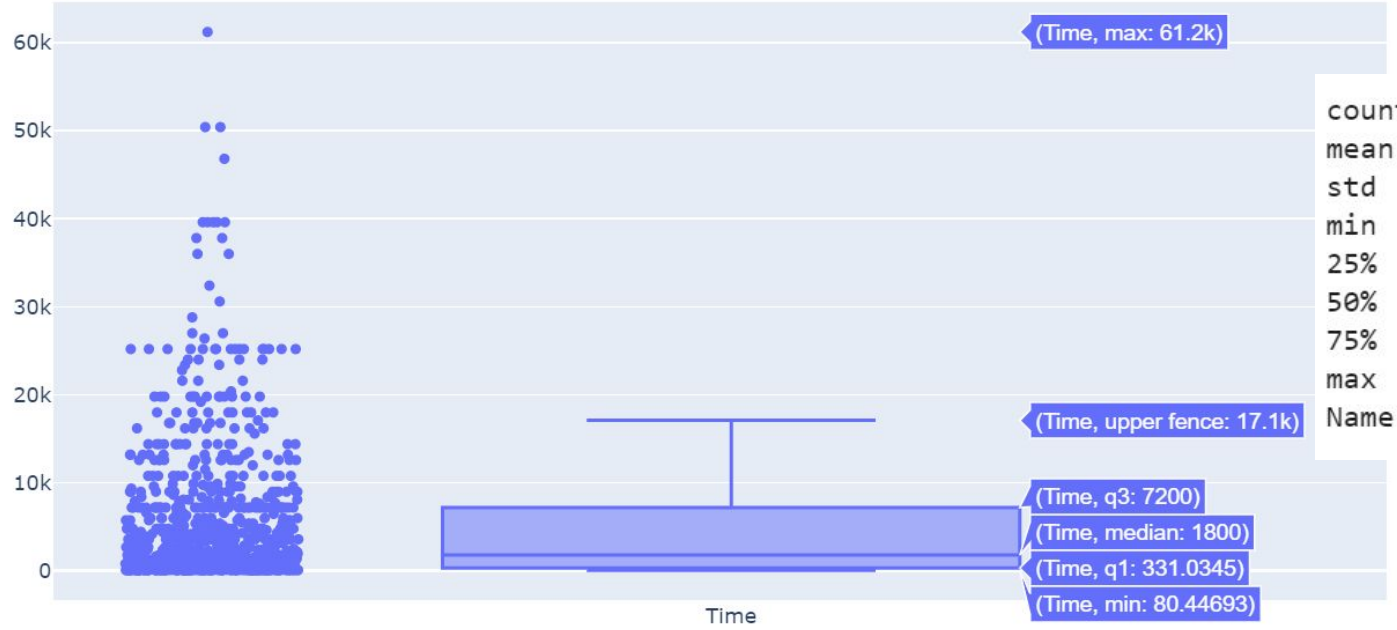


Outlier Analysis of Speed



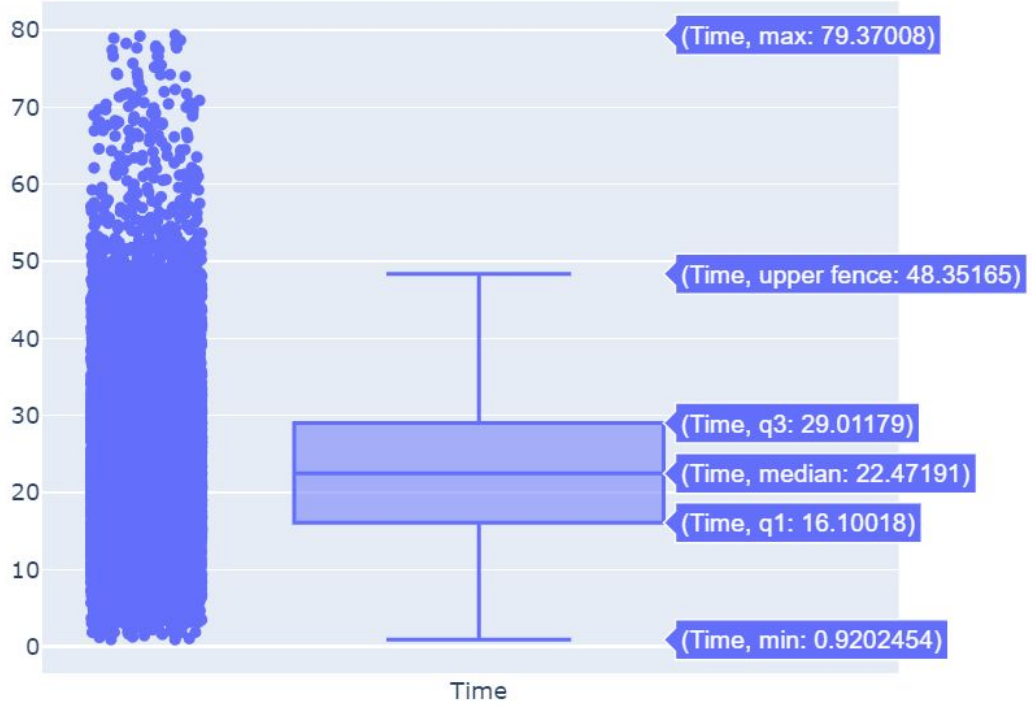

```
count    21201.000000
mean       249.305379
std       1905.297307
min          0.920245
25%         16.463415
50%         23.064602
75%         30.288462
max      61200.000000
Name: speed, dtype: float64
```

Outlier Analysis of Speed (More than 80kmph)



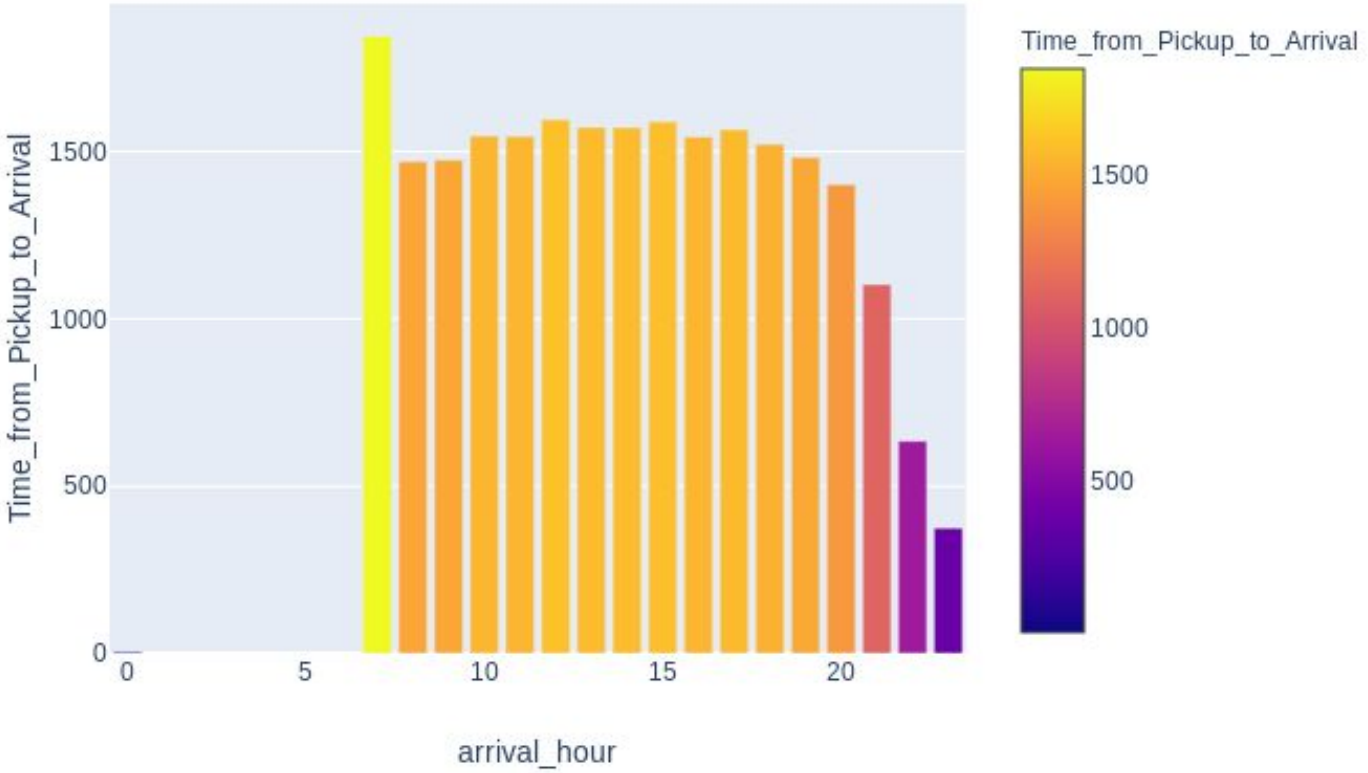

```
count       966.000000
mean       4984.404015
std        7498.835461
min          80.446927
25%         331.609195
50%        1800.000000
75%        7200.000000
max       61200.000000
Name: speed, dtype: float64
```

Outlier Analysis of Speed (Less than 80kmph)

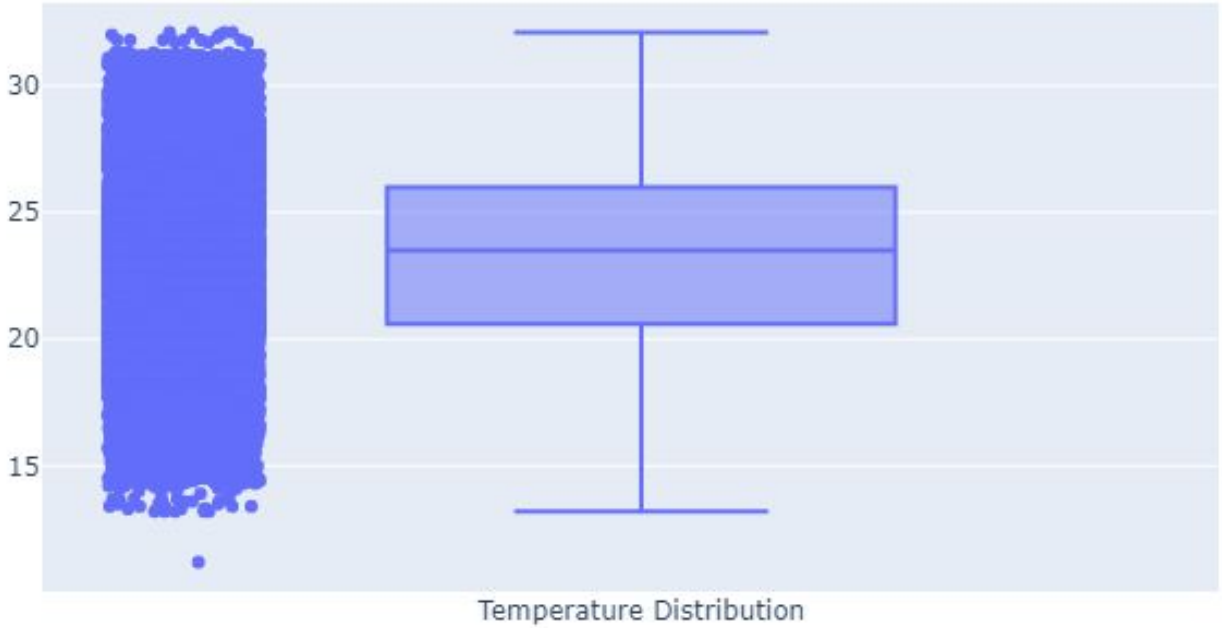



```
count    20235.000000
mean        23.256192
std         10.281375
min          0.920245
25%         16.100179
50%         22.471910
75%         29.010812
max         79.370079
Name: speed, dtype: float64
```

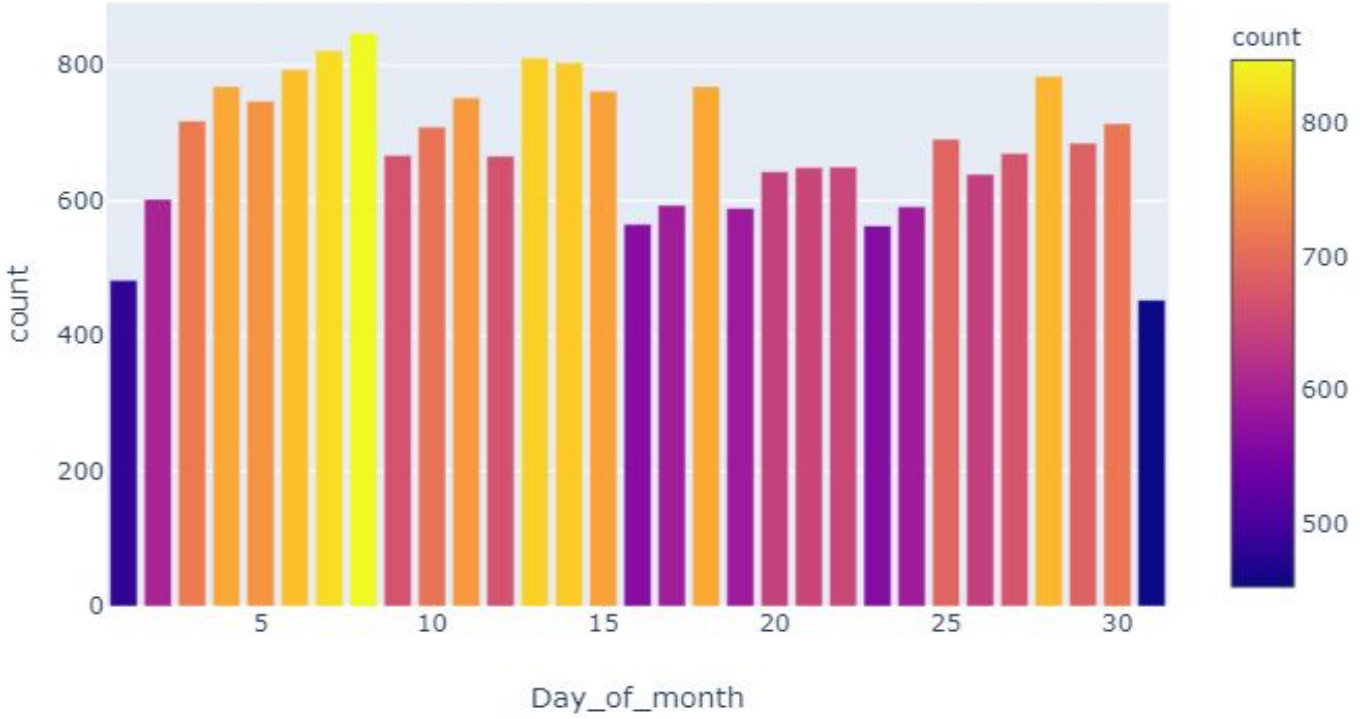
ETA vs arrival hour



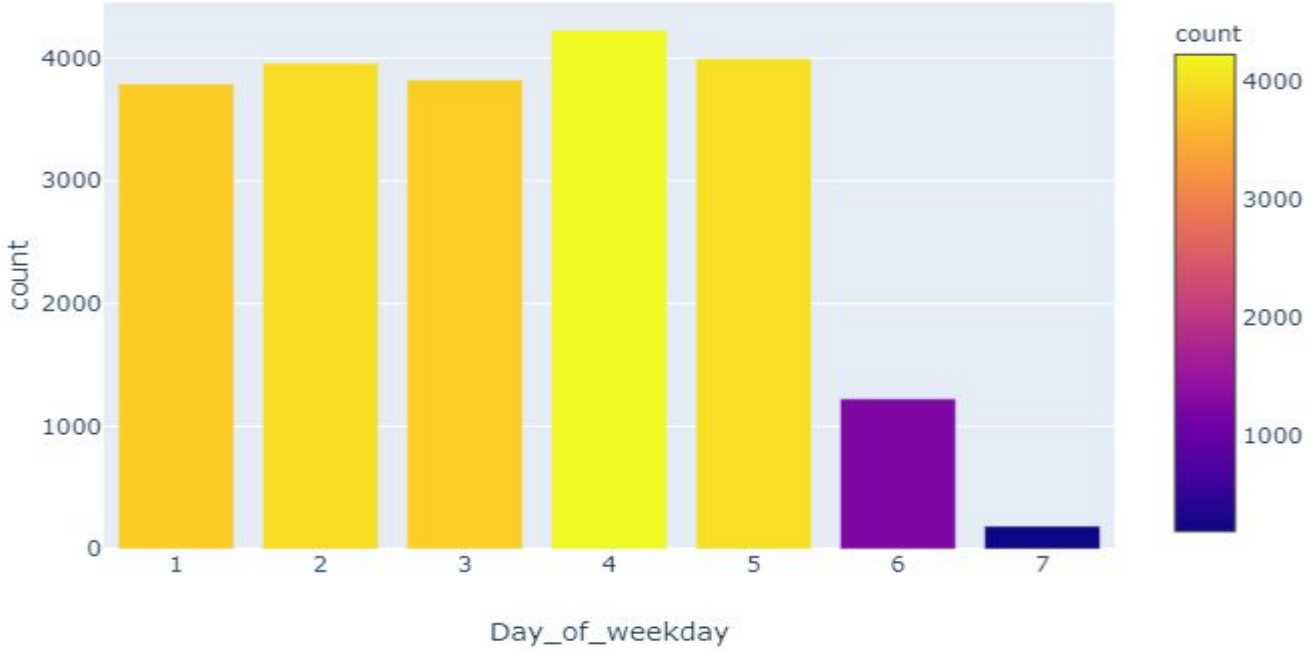
Temperature Distribution



Number of orders on each day of the month



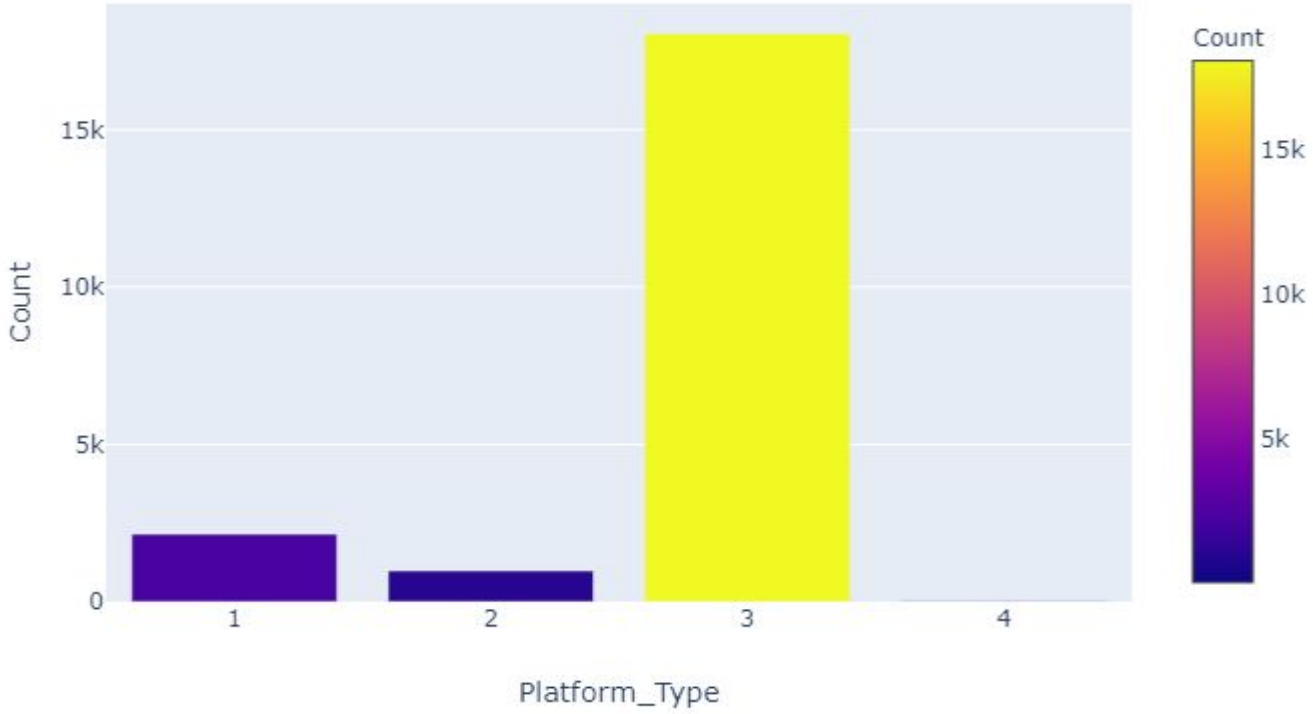
Number of orders on each day of the week



1	MONDAY
2	TUESDAY
3	WEDNESDAY
4	THURSDAY
5	FRIDAY
6	SATURDAY
7	SUNDAY

Business Vs Personal

Number of orders placed under each platform



	address
AddNum	
Addr_type	POI
Address	Dunga Clos
Block	
City	Nairobi
CountryCode	KEN
District	
LongLabel	Kirloskar Kenya, Dunga Clos, Nairobi, KEN
Match_addr	Kirloskar Kenya
MetroArea	
Neighborhood	Nairobi
PlaceName	Kirloskar Kenya
Postal	
PostalExt	
Region	Nairobi
Sector	
ShortLabel	Kirloskar Kenya
Subregion	Nairobi
Territory	
Type	Business Facility

	location
x	36.8297
y	-1.3005

Data Preprocessing

- Data Cleaning
- Data Integration
- Data Transformation
- Data Reduction

Data Cleaning

Missing Values:

- Temperature attribute null values are replaced with its global mean value.
- Precipitation attribute null values are replaced with with 0.

Noisy Data:

- Speed shows outliers - Tuples with speed greater than 60 kmph are dropped. (1085 tuples)

Data Integration

Riders Data is merged with training file with rider id as joining attribute.

Order_No	Order_No_4211	Order_No_27440	Order_No_14170	Order_No_21727	Order_No_8301	Order_No_10440
User_Id	User_Id_633	User_Id_2642	User_Id_3339	User_Id_3523	User_Id_87	User_Id_2801
Vehicle_Type	Bike	Bike	Bike	Bike	Bike	Bike
Platform_Type	3	3	3	3	2	3
Type	Business	Personal	Business	Personal	Personal	Business
Placement_Day_of_Month	9	18	31	2	22	29
Placement_Weekday	5	5	5	2	2	3
Placement_Time	9:35:46 AM	3:41:17 PM	12:51:41 PM	7:12:10 AM	10:40:58 AM	12:14:43 PM
Confirmation_Day_of_Month	9	18	31	2	22	29
Confirmation_Weekday	5	5	5	2	2	3
Confirmation_Time	9:40:10 AM	3:41:30 PM	1:12:49 PM	7:12:29 AM	10:42:24 AM	12:15:51 PM
Arrival_at_Pickup_Day_of_Month	9	18	31	2	22	29
Arrival_at_Pickup_Weekday	5	5	5	2	2	3
Arrival_at_Pickup_Time	10:04:47 AM	4:07:16 PM	1:20:02 PM	7:35:24 AM	10:56:00 AM	12:21:54 PM
Pickup_Day_of_Month	9	18	31	2	22	29
Pickup_Weekday	5	5	5	2	2	3
Pickup_Time	10:27:30 AM	4:13:37 PM	1:33:08 PM	7:38:46 AM	10:59:40 AM	12:25:10 PM
Arrival_at_Destination_Day_of_Month	9	18	31	2	22	29
Arrival_at_Destination_Weekday	5	5	5	2	2	3
Arrival_at_Destination_Time	10:39:55 AM	5:01:43 PM	2:16:43 PM	8:28:32 AM	11:26:22 AM	1:03:43 PM
Distance	4	20	6	18	7	26
Temperature	20.4	24.5	24.7	15.2	19.2	NaN
Precipitation	NaN	NaN	NaN	NaN	NaN	NaN
Pickup_Lat	-1.31775	-1.32677	-1.25519	-1.29031	-1.27352	-1.26743
Pickup_Long	36.8304	36.7878	36.7822	36.7574	36.7992	36.7871
Destination_Lat	-1.30041	-1.35624	-1.27341	-1.22352	-1.30043	-1.34364
Destination_Long	36.8297	36.9043	36.8182	36.8021	36.7524	36.8925
Rider_Id	Rider_Id_432	Rider_Id_432	Rider_Id_432	Rider_Id_432	Rider_Id_432	Rider_Id_432
Time_from_Pickup_to_Arrival	745	2886	2615	2986	1602	2313
Average_Rating	13.8	13.8	13.8	13.8	13.8	13.8

Data Transformation

- Discretization
  - One hot encoding of platform type (1, 2, 3, 4) and order type (Business,Personal)
- Attribute/Feature Construction
  - Average speed for each rider (Normalization)
  - Time taken for confirmation
  - Time taken to arrive at pickup
  - Time taken to pickup

Attribute Selection for training

Personal	0.000000	1.000000	0.000000	1.000000	1.000000	0.000000	0.000000	1.000000
Business	1.000000	0.000000	1.000000	0.000000	0.000000	1.000000	1.000000	0.000000
confirmation_time_min	4.400000	0.216667	21.133333	0.316667	1.433333	1.133333	0.250000	3.383333
Arrival_time_min	24.616667	25.766667	7.216667	22.916667	13.600000	6.050000	5.466667	14.250000
Pickup_time_min	22.716667	6.350000	13.100000	3.366667	3.666667	3.266667	23.366667	8.783333
Distance	4.000000	20.000000	6.000000	18.000000	7.000000	26.000000	8.000000	7.000000
Temperature	20.400000	24.500000	24.700000	15.200000	19.200000	18.486928	15.200000	24.500000
Precipitation	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Rider_Id	1.899936	1.899936	1.899936	1.899936	1.899936	1.899936	1.899936	1.899936
Average_Rating	13.800000	13.800000	13.800000	13.800000	13.800000	13.800000	13.800000	13.800000

Transformed Attributes

1. Platform types into 4 new attributes (1,2,3,4)
2. Business
3. Personal
4. Average speed for each rider
5. Time taken for confirmation
6. Time taken to arrive at pickup
7. Time taken to pickup

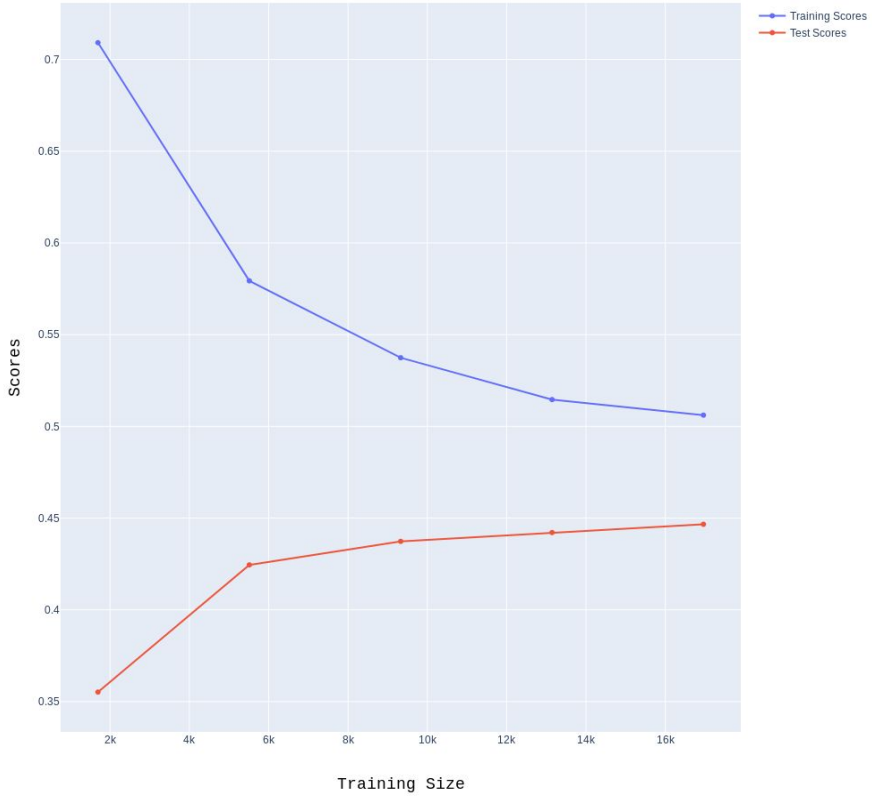
Model Building

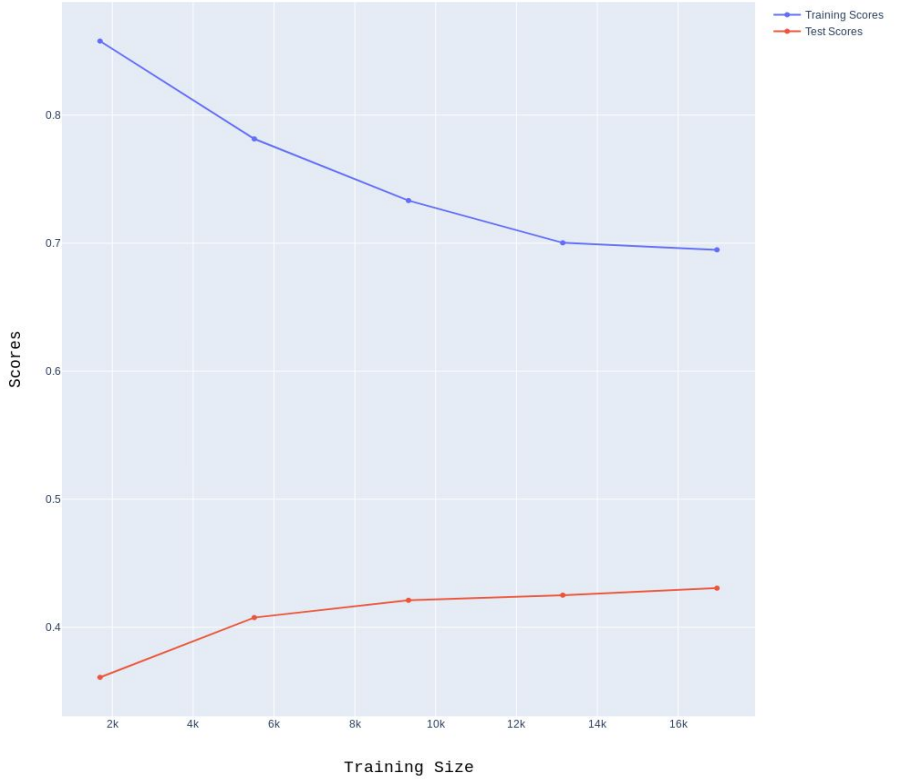
- Linear regression, random forest regressor, XGB regressor, AdaBoost regressor and gradient boosting regressor are used.
- After tuning the hyper parameters and testing against a part of the dataset which wasn't used for model building it was found that XGB regressor, Gradient boosting regression and Ada boost regressor give the most optimum results.
- The average of the results of these three models are taken as the final prediction.

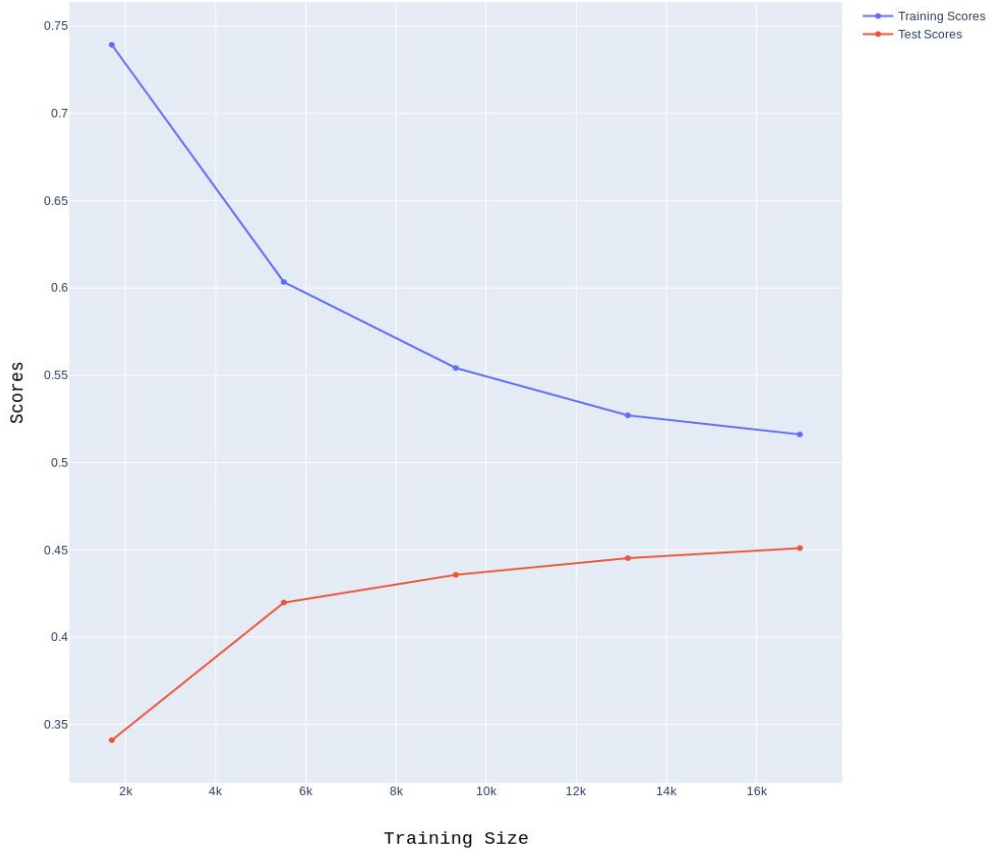
Model Selection

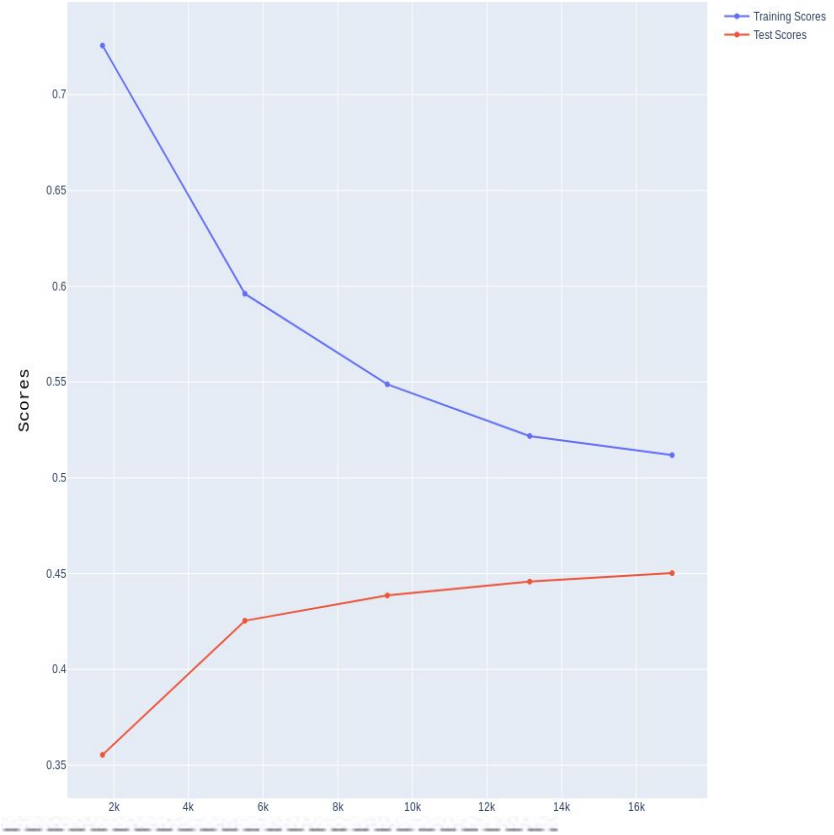
Model	Validation Error
Linear Regressor	163.58
Random Forest Regressor	153.412
XG Boost	148.89
Ada Boost	147.97
Gradient Boost	149.7

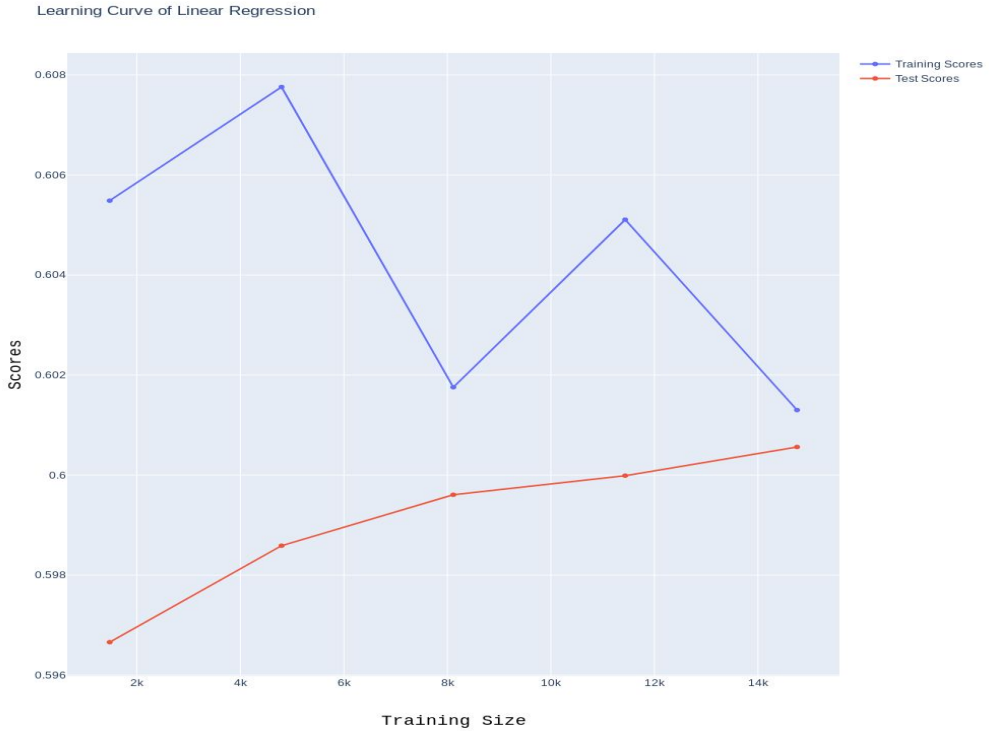
Post Processing Evaluation






Training Scores
Test Scores
Scores
Training Size
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
2k
4k
6k
8k
10k
12k
14k
16k

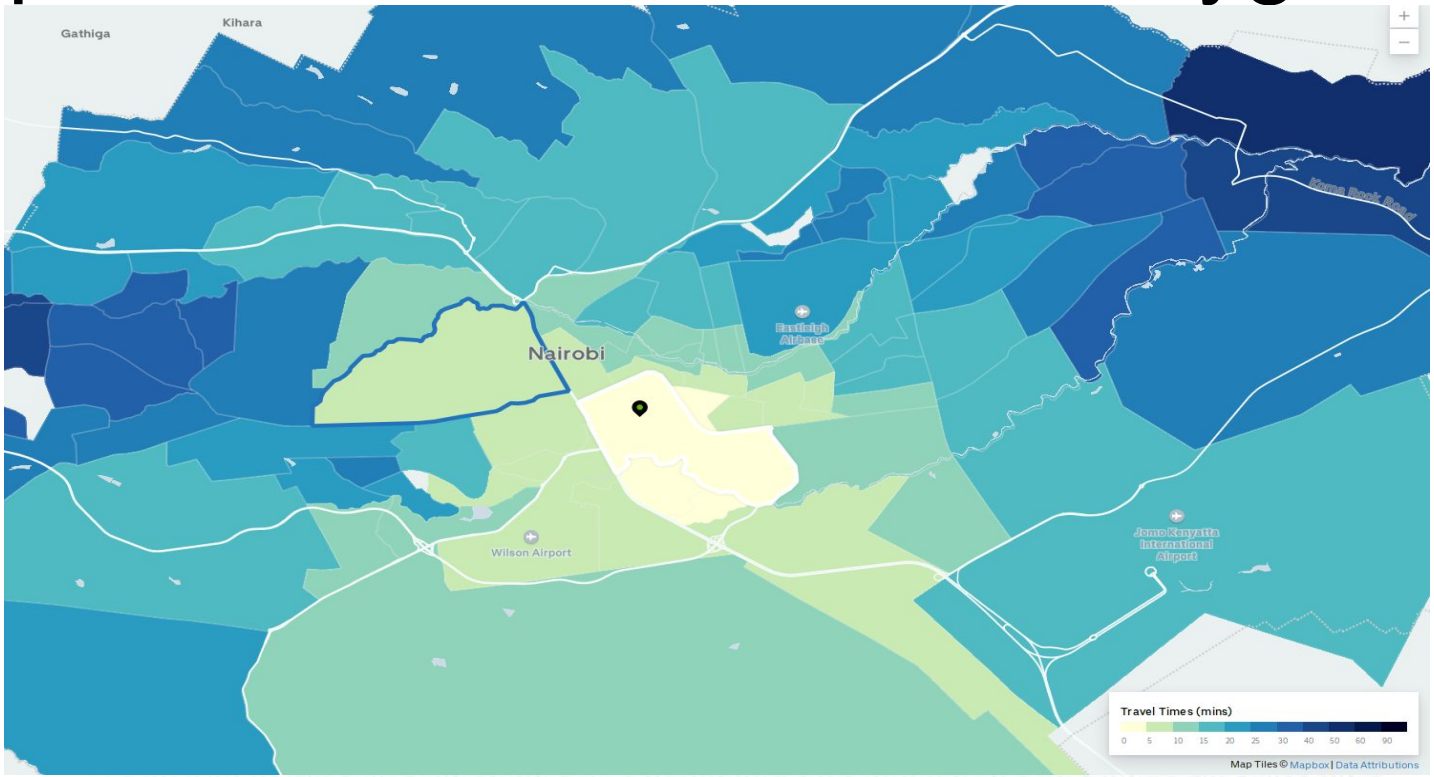

Training Scores
Test Scores
Scores
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
2k
4k
6k
8k
10k
12k
14k
16k


Learning Curve of Linear Regression
Training Scores
Test Scores
Scores
0.608
0.606
0.604
0.602
0.6
0.598
0.596
2k
4k
6k
8k
10k
12k
14k
Training Size

Preprocessing - 2

- Integrating other available datasets. - Uber Dataset, OSRM Dataset
- Rechecking dropped tuples - 8k tuples, lot to lose.
- Using median speed instead of mean speeds to integrate rider data.
- Model building - LGBM regressor, Stacking.

Map Of Nairobi divided into Polygons



Polygons Description

```
['Feature', 'Kiwanja', '1', 'Polygon', '36.8985938', '-1.1610626', '36.8993376', '-1.1613591', '36.8997611', '-1.1615452', '36.9002913', '-1.161406', '36.9007186', '-1.16111
['Feature', 'Kahawa West', '2', 'Polygon', '36.8944471', '-1.169341', '36.8948057', '-1.1696855', '36.8955915', '-1.1705776', '36.8962934', '-1.1714102', '36.8966902', '-1.1
['Feature', 'Kongo Soweto', '3', 'Polygon', '36.8942945', '-1.1820167', '36.9014052', '-1.1833947', '36.902172', '-1.1837041', '36.9095994', '-1.1866997', '36.9099198', '-1.
['Feature', 'Njathaini', '4', 'Polygon', '36.8999442', '-1.1934933', '36.9007987', '-1.1945416', '36.9017944', '-1.1959393', '36.902008', '-1.1964769', '36.9020805', '-1.196
['Feature', 'Karura', '5', 'Polygon', '36.7995202', '-1.1955376', '36.8028963', '-1.1973208', '36.8262389', '-1.2102867', '36.8348069', '-1.2150136', '36.834929', '-1.215243
['Feature', 'Githurai', '6', 'Polygon', '36.8999442', '-1.1934933', '36.9007224', '-1.1937199', '36.902729', '-1.1940585', '36.9041633', '-1.1943616', '36.9048233', '-1.1945
['Feature', 'Roysambu', '7', 'Polygon', '36.8967093', '-1.2025012', '36.8962286', '-1.203322', '36.8957975', '-1.2040838', '36.8946684', '-1.2055889', '36.8943479', '-1.2064
['Feature', "Ng'undu", '8', 'Polygon', '37.1023944', '-1.2621932', '37.1023715', '-1.262723', '37.1027186', '-1.2631853', '37.1032527', '-1.263561', '37.1035998', '-1.263794
['Feature', 'Kasarani', '9', 'Polygon', '36.9132539', '-1.2076712', '36.9139063', '-1.2082124', '36.9144022', '-1.2088888', '36.9146044', '-1.2093172', '36.9146273', '-1.209
['Feature', 'Mwiki', '10', 'Polygon', '37.0070555', '-1.236692', '37.0066587', '-1.2373005', '37.0064337', '-1.2378004', '37.0061933', '-1.2381337', '37.0057852', '-1.238207
['Feature', 'Garden Estate', '11', 'Polygon', '36.8384843', '-1.2247699', '36.845679', '-1.2217581', '36.8459842', '-1.2218647', '36.8475559', '-1.2251785', '36.8482349', '-
['Feature', 'Muthaiga', '12', 'Polygon', '36.8419405', '-1.2309648', '36.8425662', '-1.2315894', '36.8429705', '-1.2321429', '36.8431193', '-1.2325011', '36.843558', '-1.235
['Feature', 'Kitisuru', '13', 'Polygon', '36.7534148', '-1.2263977', '36.7535407', '-1.2264049', '36.7536819', '-1.2265114', '36.7539108', '-1.2266359', '36.7541091', '-1.22
['Feature', 'Ruai', '14', 'Polygon', '37.0197854', '-1.2270544', '37.0363415', '-1.2736742', '37.033984', '-1.2778598', '37.0315921', '-1.2827524', '37.0296733', '-1.2865453
['Feature', 'Babadogo', '15', 'Polygon', '36.9121973', '-1.2379274', '36.9119913', '-1.2380254', '36.9117586', '-1.2381608', '36.9112741', '-1.2382188', '36.9109651', '-1.23
['Feature', "Dandora'b'", '16', 'Polygon', '36.9313169', '-1.2402377', '36.9275479', '-1.2426423', '36.9267392', '-1.2429293', '36.9260296', '-1.2429462', '36.9248623', '-1.
['Feature', 'Saika', '17', 'Polygon', '36.9355933', '-1.2421142', '36.9354331', '-1.2423388', '36.9347884', '-1.2425224', '36.9344031', '-1.2425224', '36.934075', '-1.242228
['Feature', 'Njiru', '18', 'Polygon', '36.9540873', '-1.2443695', '36.9536677', '-1.244404', '36.9534464', '-1.2445902', '36.9532443', '-1.244793', '36.953023', '-1.244776',
['Feature', 'Loresho', '19', 'Polygon', '36.7282869', '-1.2458139', '36.7288553', '-1.2461122', '36.7298395', '-1.2465596', '36.7305987', '-1.2468952', '36.7312281', '-1.246
['Feature', "Dandora 'a'", '20', 'Polygon', '36.9004096', '-1.2420881', '36.9018325', '-1.2467947', '36.9021797', '-1.2473745', '36.9051056', '-1.2515129', '36.9045982', '-1
['Feature', 'Gitathuru/nyayo', '21', 'Polygon', '36.9001426', '-1.242337', '36.9000319', '-1.2425333', '36.8998374', '-1.2428235', '36.8992766', '-1.2432298', '36.8986777',
['Feature', 'Utalii', '22', 'Polygon', '36.8695481', '-1.24671', '36.8693993', '-1.2472062', '36.8691513', '-1.2477174', '36.869075', '-1.2480418', '36.8691361', '-1.248478'
['Feature', 'Korogocho', '23', 'Polygon', '36.8899037', '-1.2463371', '36.8920209', '-1.2505854', '36.8920896', '-1.2511044', '36.8925283', '-1.2519531', '36.8926427', '-1.2
['Feature', 'Mathare North', '24', 'Polygon', '36.8780207', '-1.2491533', '36.877788', '-1.2492049', '36.8773722', '-1.2497497', '36.8770784', '-1.2501926', '36.8770136', '-
['Feature', 'Spring Valley', '25', 'Polygon', '36.7939468', '-1.2498275', '36.7944161', '-1.251359', '36.7947174', '-1.2518254', '36.7951714', '-1.2520312', '36.7960755', '-
```

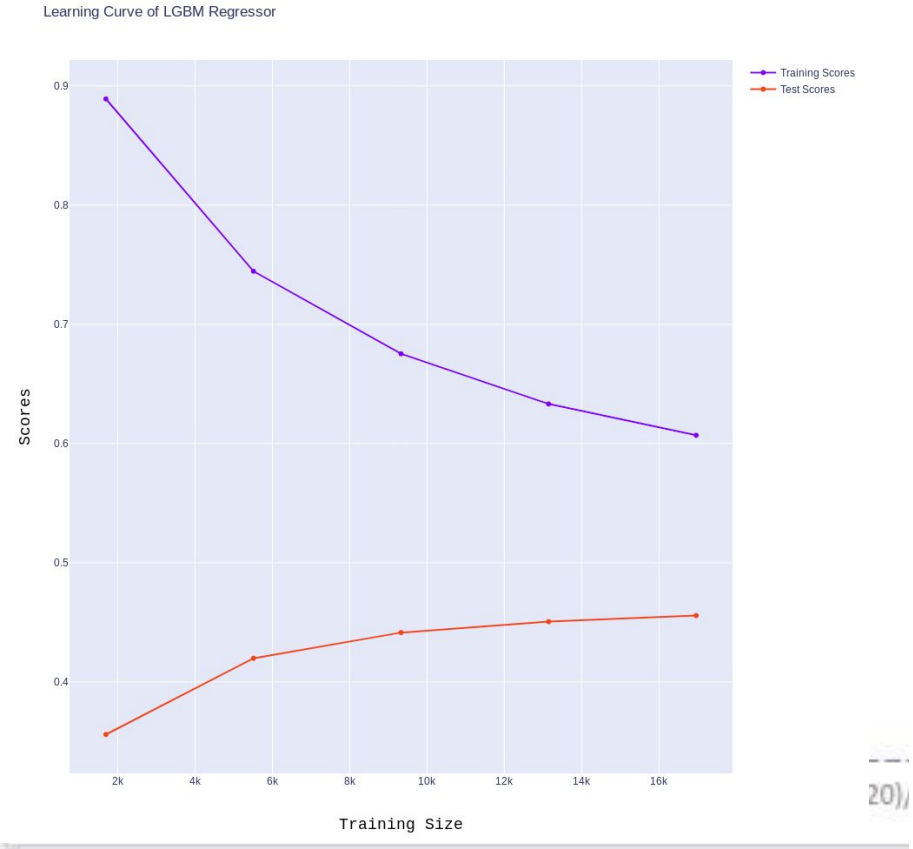
OSRM Data

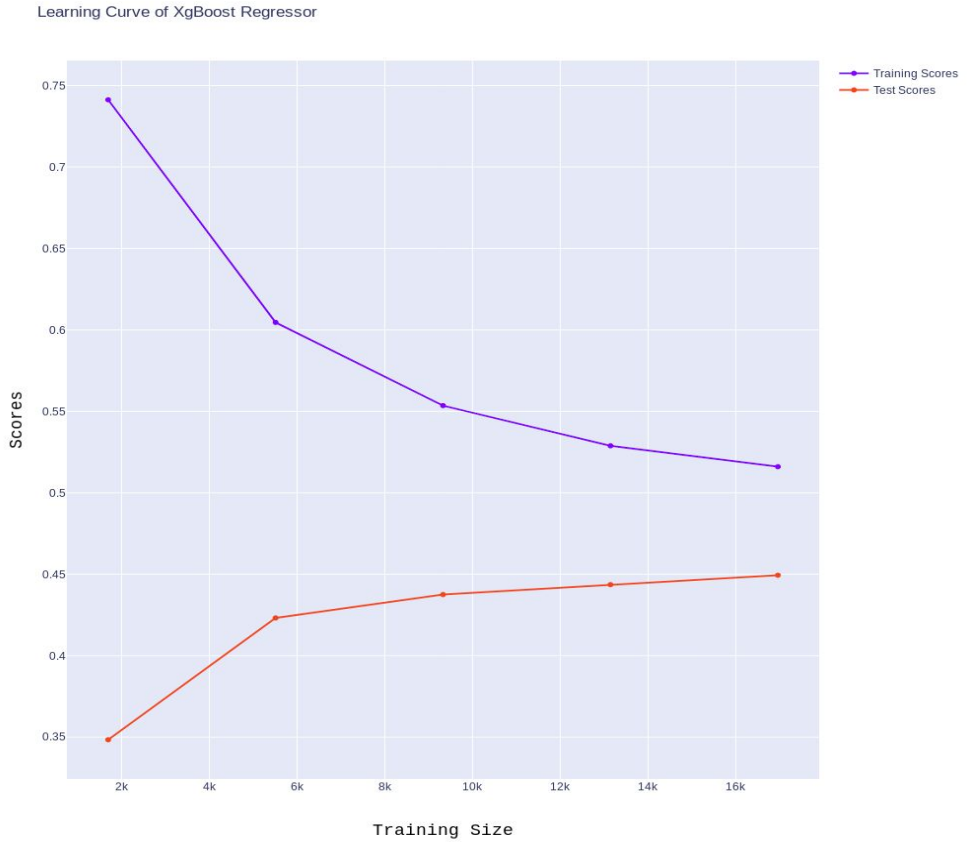
OrderId	routes
Order_No_4211	[{'summary': {'distance': 2528.7, 'duration': 295.5}, 'segments': [{'distance': 2528.7, 'duration': 295.5, 'steps': [{'distance': 389.4, 'duration': 93.5, 'type': 11, 'instruction': 'Head north on Muhc
Order_No_25375	[{'summary': {'distance': 14869.7, 'duration': 991.0}, 'segments': [{'distance': 14869.7, 'duration': 991.0, 'steps': [{'distance': 715.0, 'duration': 128.7, 'type': 11, 'instruction': 'Head southeast',
Order_No_1899	[{'summary': {'distance': 3275.3, 'duration': 245.3}, 'segments': [{'distance': 3275.3, 'duration': 245.3, 'steps': [{'distance': 43.6, 'duration': 10.5, 'type': 11, 'instruction': 'Head north', 'name': '-
Order_No_9336	[{'summary': {'distance': 7912.0, 'duration': 491.1}, 'segments': [{'distance': 7912.0, 'duration': 491.1, 'steps': [{'distance': 27.8, 'duration': 1.8, 'type': 11, 'instruction': 'Head northwest on Ring
Order_No_27883	[{'summary': {'distance': 7947.2, 'duration': 612.4}, 'segments': [{'distance': 7947.2, 'duration': 612.4, 'steps': [{'distance': 113.5, 'duration': 20.4, 'type': 11, 'instruction': 'Head north', 'name':
Order_No_7408	[{'summary': {'distance': 10365.1, 'duration': 730.3}, 'segments': [{'distance': 10365.1, 'duration': 730.3, 'steps': [{'distance': 321.9, 'duration': 38.6, 'type': 11, 'instruction': 'Head northwest o
Order_No_22680	[{'summary': {'distance': 3822.7, 'duration': 383.1}, 'segments': [{'distance': 3822.7, 'duration': 383.1, 'steps': [{'distance': 305.5, 'duration': 73.3, 'type': 11, 'instruction': 'Head east', 'name': '-
Order_No_21578	[{'summary': {'distance': 2805.2, 'duration': 388.9}, 'segments': [{'distance': 2805.2, 'duration': 388.9, 'steps': [{'distance': 69.6, 'duration': 16.7, 'type': 11, 'instruction': 'Head west', 'name': '-',
Order_No_5234	[{'summary': {'distance': 8219.4, 'duration': 590.1}, 'segments': [{'distance': 8219.4, 'duration': 590.1, 'steps': [{'distance': 49.3, 'duration': 11.8, 'type': 11, 'instruction': 'Head south', 'name': '-
Order_No_1768	[{'summary': {'distance': 12917.5, 'duration': 1307.0}, 'segments': [{'distance': 12917.5, 'duration': 1307.0, 'steps': [{'distance': 55.7, 'duration': 40.1, 'type': 11, 'instruction': 'Head west', 'nam
Order_No_823	[{'summary': {'distance': 9081.9, 'duration': 759.9}, 'segments': [{'distance': 9081.9, 'duration': 759.9, 'steps': [{'distance': 221.4, 'duration': 17.7, 'type': 11, 'instruction': 'Head southwest on L
Order_No_2533	[{'summary': {'distance': 13355.3, 'duration': 1024.7}, 'segments': [{'distance': 13355.3, 'duration': 1024.7, 'steps': [{'distance': 15.7, 'duration': 3.8, 'type': 11, 'instruction': 'Head southwest',
Order_No_7842	[{'summary': {'distance': 10463.1, 'duration': 709.3}, 'segments': [{'distance': 10463.1, 'duration': 709.3, 'steps': [{'distance': 40.2, 'duration': 4.8, 'type': 11, 'instruction': 'Head north', 'name':
Order_No_6452	[{'summary': {'distance': 20805.8, 'duration': 1599.9}, 'segments': [{'distance': 20805.8, 'duration': 1599.9, 'steps': [{'distance': 117.8, 'duration': 28.3, 'type': 11, 'instruction': 'Head southwes
Order_No_9463	[{'summary': {'distance': 2535.9, 'duration': 218.0}, 'segments': [{'distance': 2535.9, 'duration': 218.0, 'steps': [{'distance': 602.7, 'duration': 43.4, 'type': 11, 'instruction': 'Head east on Kibera I
Order_No_21335	[{'summary': {'distance': 12091.9, 'duration': 777.9}, 'segments': [{'distance': 12091.9, 'duration': 777.9, 'steps': [{'distance': 2899.0, 'duration': 122.8, 'type': 11, 'instruction': 'Head west on M
Order_No_21307	[{'summary': {'distance': 15198.1, 'duration': 954.7}, 'segments': [{'distance': 15198.1, 'duration': 954.7, 'steps': [{'distance': 152.6, 'duration': 9.2, 'type': 11, 'instruction': 'Head southeast on
Order_No_8201	[{'summary': {'distance': 11266.9, 'duration': 840.3}, 'segments': [{'distance': 11266.9, 'duration': 840.3, 'steps': [{'distance': 68.8, 'duration': 8.3, 'type': 11, 'instruction': 'Head west on School
Order_No_11630	[{'summary': {'distance': 5420.0, 'duration': 416.7}, 'segments': [{'distance': 5420.0, 'duration': 416.7, 'steps': [{'distance': 12.8, 'duration': 3.1, 'type': 11, 'instruction': 'Head southwest on Anf
Order_No_19866	[{'summary': {'distance': 5162.5, 'duration': 393.1}, 'segments': [{'distance': 5162.5, 'duration': 393.1, 'steps': [{'distance': 145.4, 'duration': 11.6, 'type': 11, 'instruction': 'Head northeast on Ci
Order_No_23338	[{'summary': {'distance': 8845.2, 'duration': 692.6}, 'segments': [{'distance': 8845.2, 'duration': 692.6, 'steps': [{'distance': 248.0, 'duration': 59.5, 'type': 11, 'instruction': 'Head east on Kibera I
Order_No_14810	[{'summary': {'distance': 23243.7, 'duration': 1588.6}, 'segments': [{'distance': 23243.7, 'duration': 1588.6, 'steps': [{'distance': 451.2, 'duration': 67.1, 'type': 11, 'instruction': 'Head northeast
Order_No_12412	[{'summary': {'distance': 15897.6, 'duration': 1258.0}, 'segments': [{'distance': 15897.6, 'duration': 1258.0, 'steps': [{'distance': 23.2, 'duration': 2.8, 'type': 11, 'instruction': 'Head west', 'name
Order_No_663	[{'summary': {'distance': 9140.6, 'duration': 849.9}, 'segments': [{'distance': 9140.6, 'duration': 849.9, 'steps': [{'distance': 189.1, 'duration': 45.4, 'type': 11, 'instruction': 'Head southeast on Sa

OSRM Data

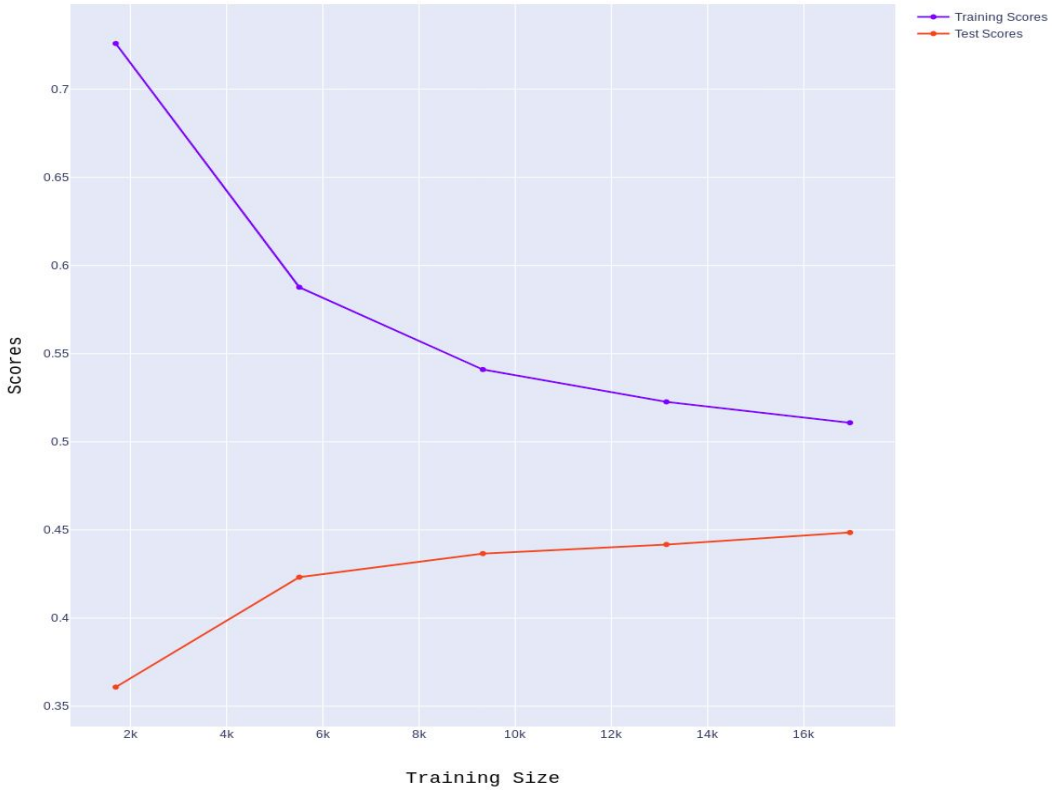
[{'summary': {'distance': 2528.7, 'duration': 295.5}, 'segments': [{'distance': 2528.7, 'duration': 295.5, 'steps': [{'distance': 389.4, 'duration': 93.5, 'type': 11, 'instruction': 'Head north on Muhoho Avenue', 'name': 'Muhoho Avenue', 'way_points': [0, 10]}, {'distance': 42.2, 'duration': 10.1, 'type': 1, 'instruction': 'Turn right', 'name': '-', 'way_points': [10, 13]}, {'distance': 1099.7, 'duration': 88.0, 'type': 2, 'instruction': 'Turn sharp left onto Mombasa Road, A104', 'name': 'Mombasa Road, A104', 'way_points': [13, 29]}, {'distance': 666.9, 'duration': 45.9, 'type': 7, 'instruction': 'Enter the roundabout and take the 3rd exit onto Lusaka Road', 'name': 'Lusaka Road', 'exit_number': 3, 'way_points': [29, 53]}, {'distance': 140.1, 'duration': 12.4, 'type': 7, 'instruction': 'Enter the roundabout and take the 1st exit onto Dunga Road', 'name': 'Dunga Road', 'exit_number': 1, 'way_points': [53, 56]}, {'distance': 190.4, 'duration': 45.7, 'type': 0, 'instruction': 'Turn left onto Dunga Close', 'name': 'Dunga Close', 'way_points': [56, 62]}, {'distance': 0.0, 'duration': 0.0, 'type': 10, 'instruction': 'Arrive at Dunga Close, on the right', 'name': '-', 'way_points': [62, 62]}]}], 'bbox': [36.826235, -1.317632, 36.831209, -1.300476], 'geometry': 'dj`Gyhx_FkD_@e@EqAMs@GmASUAyC[s@K[@I@EOKg@COKFyB~@_A^i@Rg@P_AZuG|BwAf@gFjBuJlDwCdAwC|@kBp@WLONMNCJGJIHMFO?OAKGIIGKAK?K@IA[Ka@_@aAQk@g@qAc@iAw@mBa@cAqA}CiBaES]W_@GCCCsEhBf@nABJAJEFMJoDxA', 'way_points': [0, 62]}]}

Model building





Learning Curve of Gradient Boost Regressor



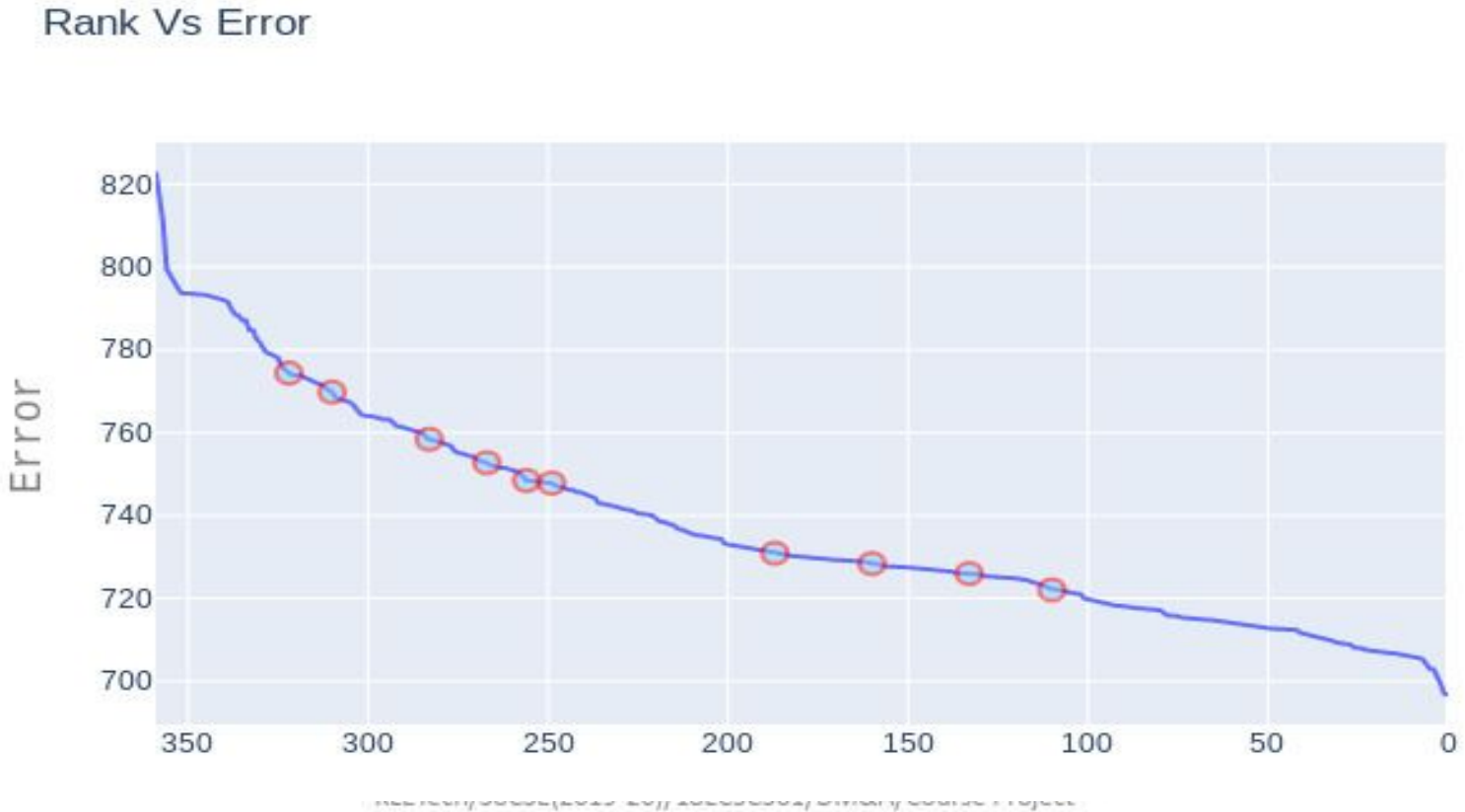
Post Processing Evaluation

	Validation score	Testing Score
Preprocessing 1	147.54	719.73
Preprocessing 2 (UBER)	145.65	724.23
Preprocessing 3 (OSRM)	144.52	725.855

Conclusion

- LGBM regressor gave best results.
- Also learnt to implement stacking and apply clustering techniques.
- Broadened perspective over using other datasets other than ones given by the competition.

Contest Analysis



Milestone	Score	Rank
Linear regression	774.763168625723	322
Use of rider average speeds	769.683184920894	310
Use of polynomial features	758.552123546533	283
Use of latitude longitude and gradient boosting	752.555009486575	267
Use of xg boost	748.818792086984	256
Use of ada boost	747.793532394547	249
After restoring the tuples dropped earlier	728.449405006038	160
Use of lgbm and median rider speeds	**722.072360777788**	**110**
Use of Uber data	730.884446410988	187
Use of OSRM data	725.963964952624	133

Post Contest Revelations

- Pickup time updation error resulted in the anomalies in the time attribute.

References

https://zindi.africa/competitions/sendy-logistics-challenge

https://techtrendske.co.ke/sendy-tech-platform-to-include-freight-services/

https://in.linkedin.com/company/sendy-limited

Thank you