



# Best Time to Use NYC Taxi

---

Group : 23



# Motivation and Objectives

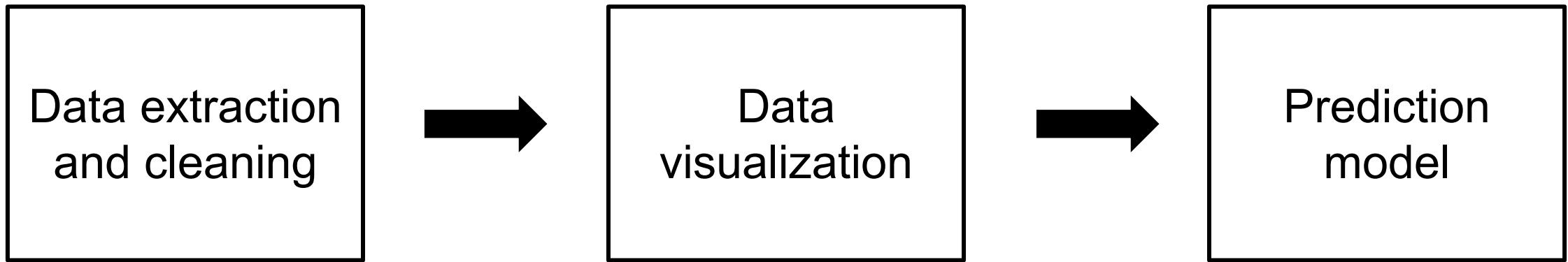
- Finding a taxi at peak hours or traveling in traffic can be very taxing in New York City
  
- Can we predict the best time of day to take a cab?
  1. To reduce the amount of stress
  2. or, Maximize the amount of time/money spent on things we like doing?

Who would want to sit through this?



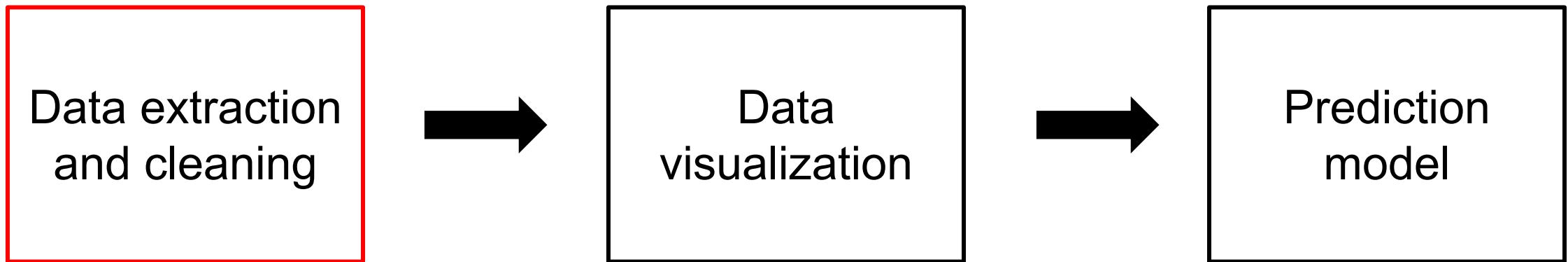


# Methodology





# Methodology





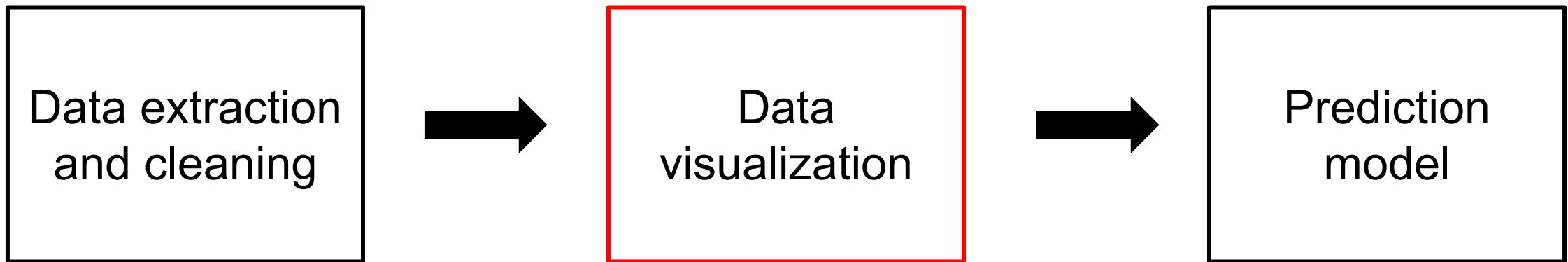
# Dataset

- The dataset is publicly available on the NYC taxi and Limousine commission webpage.
- Using the yellow taxi trip records from 2019.
- There are 12 csv files for 12 months and each csv file includes **~1M data points**.

VendorID	tpep_pickup_datetime	tpep_dropoff_datetime	passenger_count	trip_distance	RatecodeID	store_and	PULocationID	DOLocationID	payment_type	fare_amount	extra	mta_tax	tip_amount	tolls_amount	improvement_surcharge	total_amount
1	01-01-19 0:46	01-01-19 0:53	1	1.5	1	N	151	239	1	7	0.5	0.5	1.65	0	0.3	9.95
1	01-01-19 0:59	01-01-19 1:18	1	2.6	1	N	239	246	1	14	0.5	0.5	1	0	0.3	16.3
2	21-12-18 13:48	21-12-18 13:52	3	0	1	N	236	236	1	4.5	0.5	0.5	0	0	0.3	5.8
2	28-11-18 15:52	28-11-18 15:55	5	0	1	N	193	193	2	3.5	0.5	0.5	0	0	0.3	7.55
2	28-11-18 15:56	28-11-18 15:58	5	0	2	N	193	193	2	52	0	0.5	0	0	0.3	55.55
2	28-11-18 16:25	28-11-18 16:28	5	0	1	N	193	193	2	3.5	0.5	0.5	0	5.76	0.3	13.31
2	28-11-18 16:29	28-11-18 16:33	5	0	2	N	193	193	2	52	0	0.5	0	0	0.3	55.55
1	01-01-19 0:21	01-01-19 0:28	1	1.3	1	N	163	229	1	6.5	0.5	0.5	1.25	0	0.3	9.05
1	01-01-19 0:32	01-01-19 0:45	1	3.7	1	N	229	7	1	13.5	0.5	0.5	3.7	0	0.3	18.5
1	01-01-19 0:57	01-01-19 1:09	2	2.1	1	N	141	234	1	10	0.5	0.5	1.7	0	0.3	13
1	01-01-19 0:24	01-01-19 0:47	2	2.8	1	N	246	162	1	15	0.5	0.5	3.25	0	0.3	19.55
1	01-01-19 0:21	01-01-19 0:28	1	0.7	1	N	238	151	1	5.5	0.5	0.5	1.7	0	0.3	8.5
1	01-01-19 0:45	01-01-19 1:31	1	8.7	1	N	163	25	1	34.5	0.5	0.5	7.15	0	0.3	42.95
1	01-01-19 0:43	01-01-19 1:07	1	6.3	1	N	224	25	1	21.5	0.5	0.5	5.7	0	0.3	28.5
1	01-01-19 0:58	01-01-19 1:15	1	2.7	1	N	141	234	1	13	0.5	0.5	1	0	0.3	15.3
2	01-01-19 0:23	01-01-19 0:25	1	0.38	1	N	170	170	2	3.5	0.5	0.5	0	0	0.3	4.8
2	01-01-19 0:39	01-01-19 0:48	1	0.55	1	N	170	170	1	6.5	0.5	0.5	1.95	0	0.3	9.75



# Methodology



**What can we learn from this  
visualizing this dataset?**

# Most popular pick-up & drop-off zones



Most Pick-Up zones



## Top 5 Pickup zones

- 237: Upper East Side South
- 161: Midtown Center
- 236: Upper East Side North
- 162: Midtown East
- 186: Penn Station/Madison Sq West
- 132: JFK airport

Most Drop-Off zones



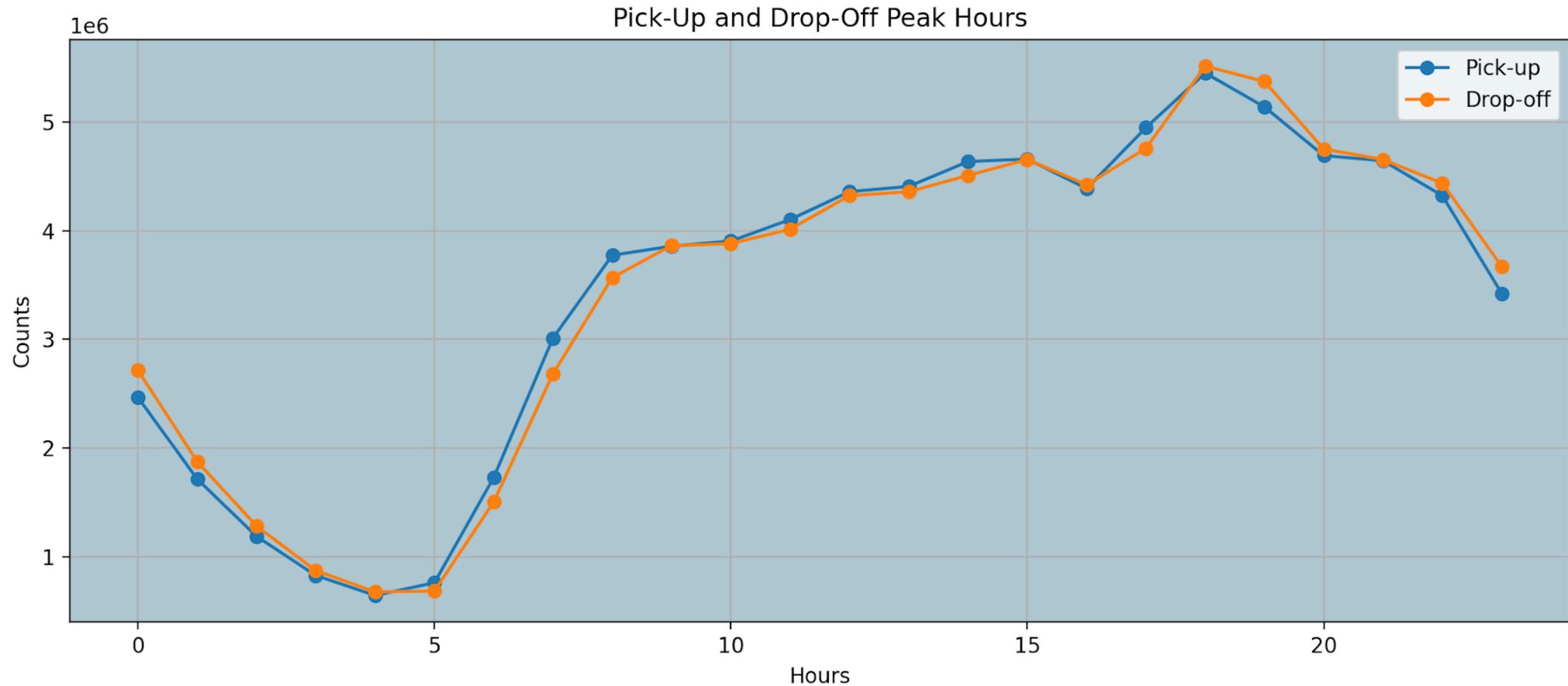
## Top 5 Drop-off zones

- 236: Upper East Side North
- 237: Upper East Side South
- 161: Midtown Center
- 170: Murray Hill
- 230: Times Sq/Theatre District

All of them are in Manhattan

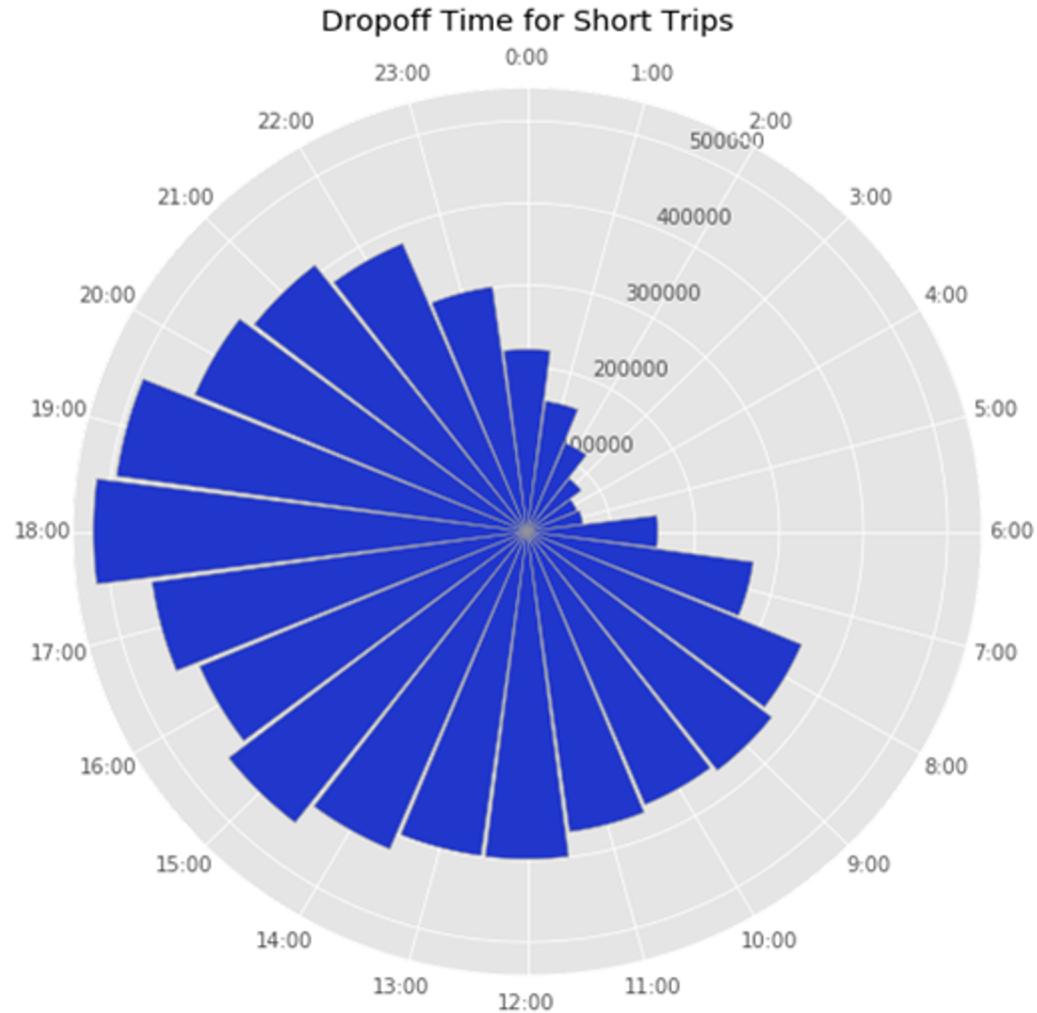
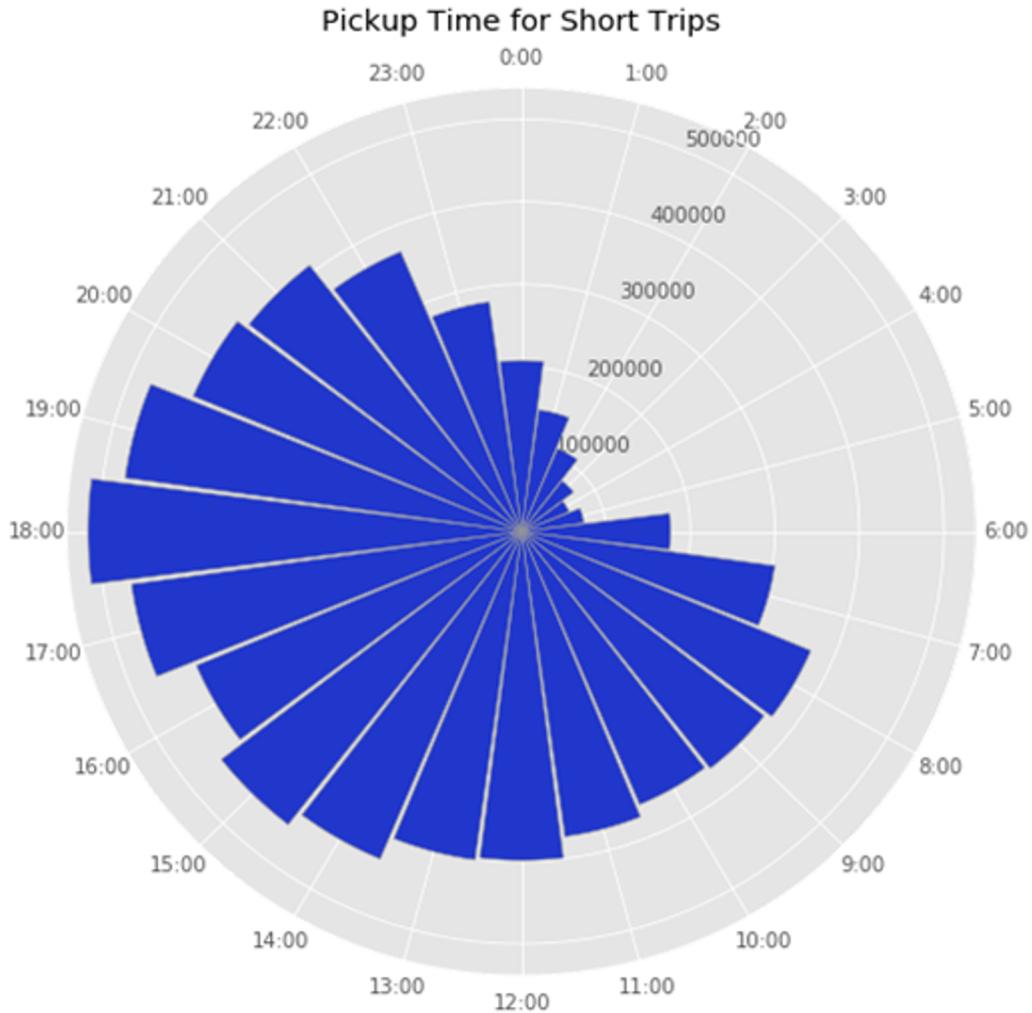


# Peak hours

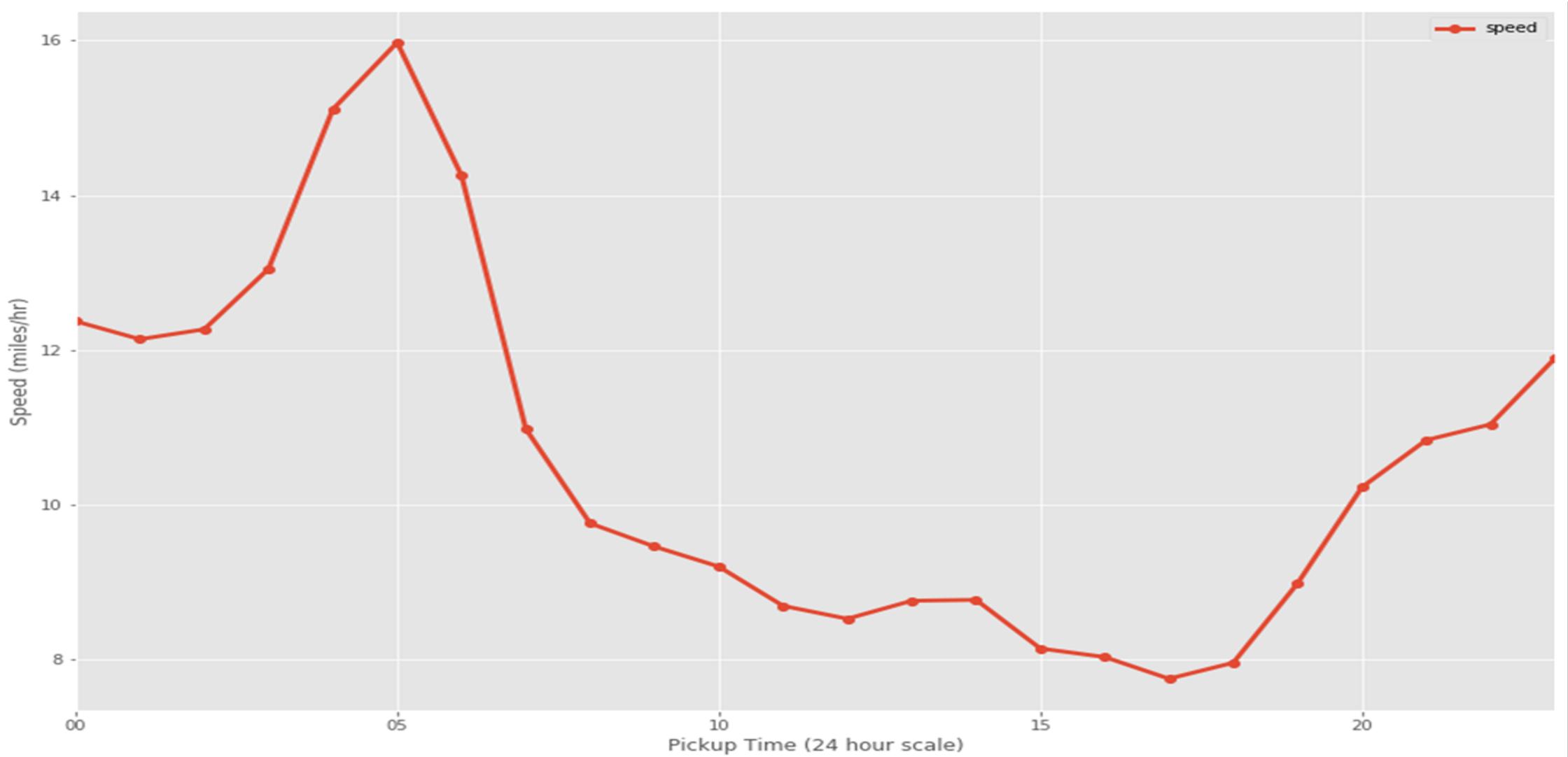




# Number of trips vs the hour of travel



# Average Speed of the taxi



# Number of trips vs location

Short trips < 30 miles

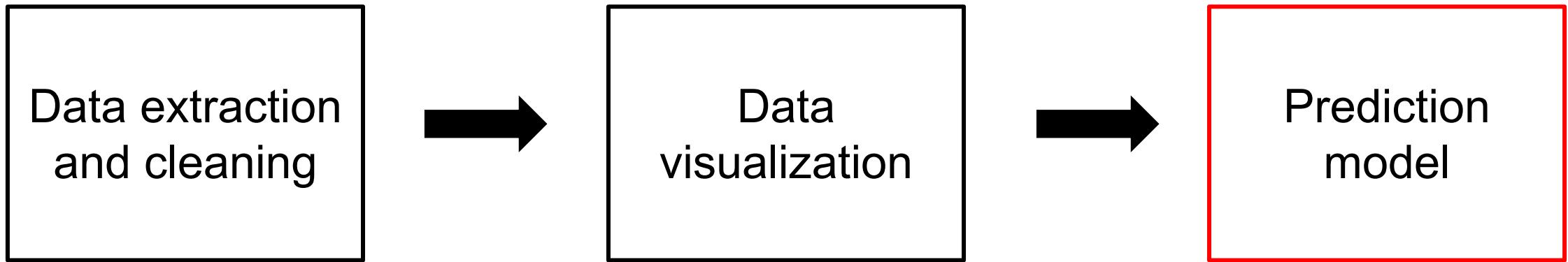
Long trips > 30 miles

pickup zone	dropoff zone	short trips
Upper East Side North	Upper East Side North	47692
Upper West Side South	Upper West Side South	19447
Penn Station/Madison Sq West	Midtown Center	18734
Upper East Side North	Yorkville West	16440
Upper East Side North	Yorkville East	15284

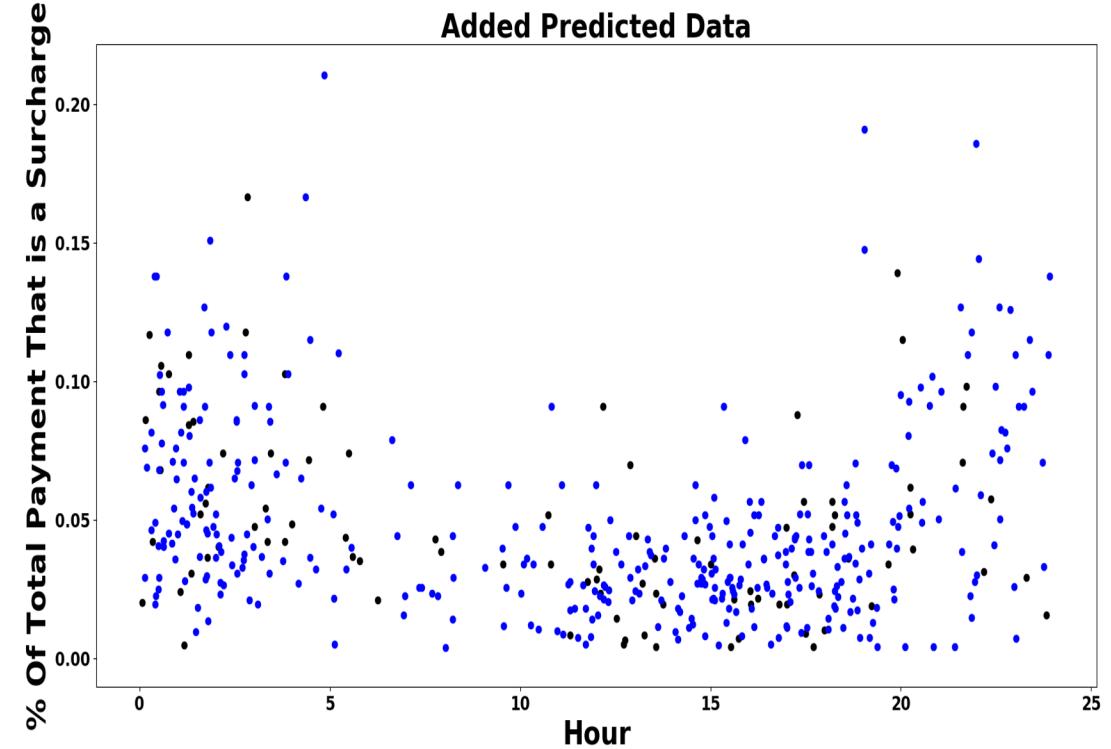
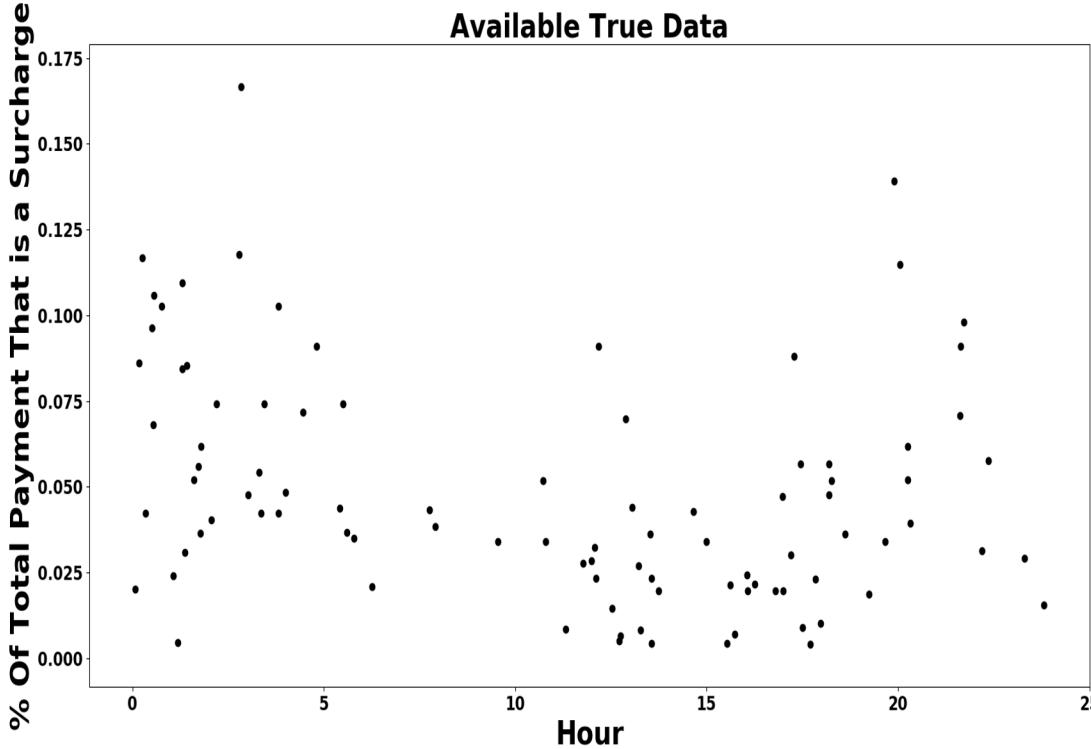
pickup zone	dropoff zone	long trips
JFK Airport	Newark Airport	272
JFK Airport	JFK Airport	67
Meatpacking/West Village West	JFK Airport	31
Midtown North	Newark Airport	9
Battery Park City	JFK Airport	9



# Methodology



# Machine Learning Results



- Some time slots in the data are not complete
- We use simple extrapolation techniques to guess what the missing data may be
- We can also use ML to get even better results
  - In this case, because we want to predict real values we will use regression



# Summary

- We extracted the publicly available data from the NYC TLC website and cleaned it for our purposes.
- We checked for interesting trends in the data using data visualization techniques.
- We devised a prediction model to predict the % surcharge fee.



Have a nice trip!



# Thank you for listening!