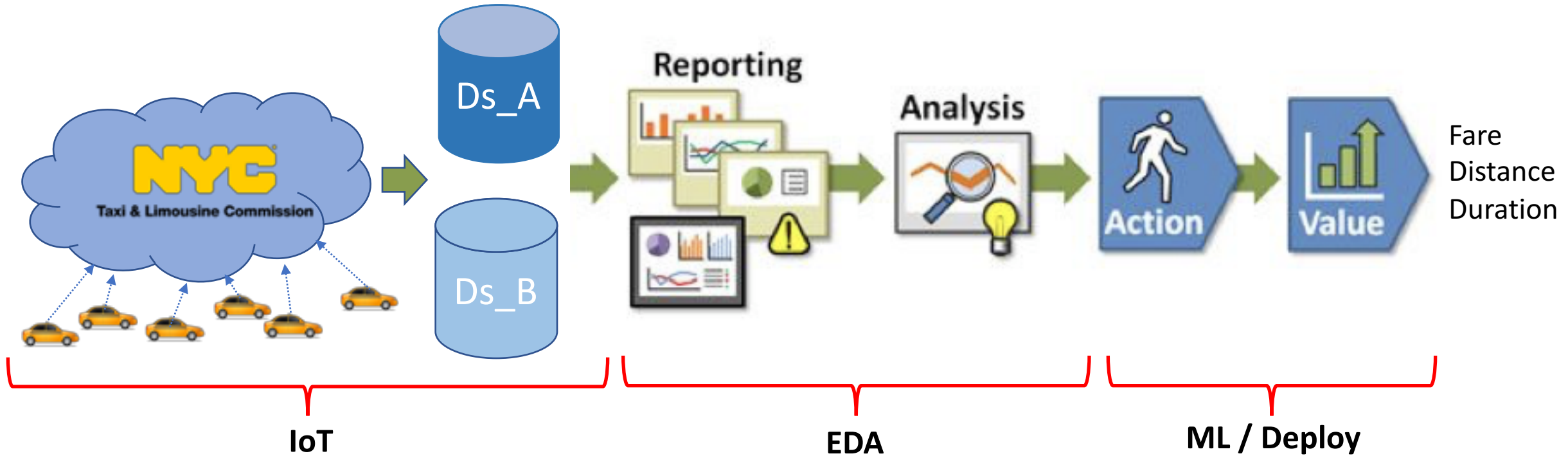




NYC Taxi - Trip Prediction Project

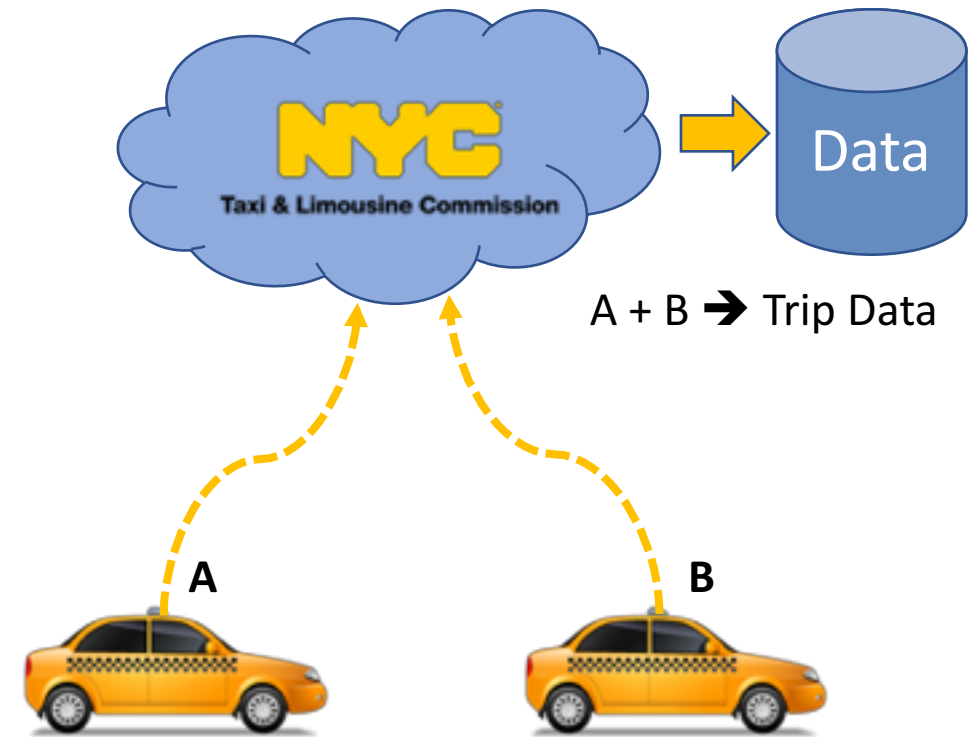
NYC Taxi Trip Prediction Project



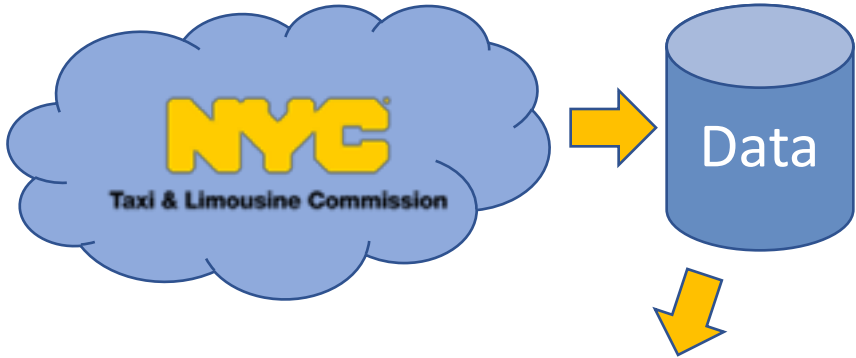
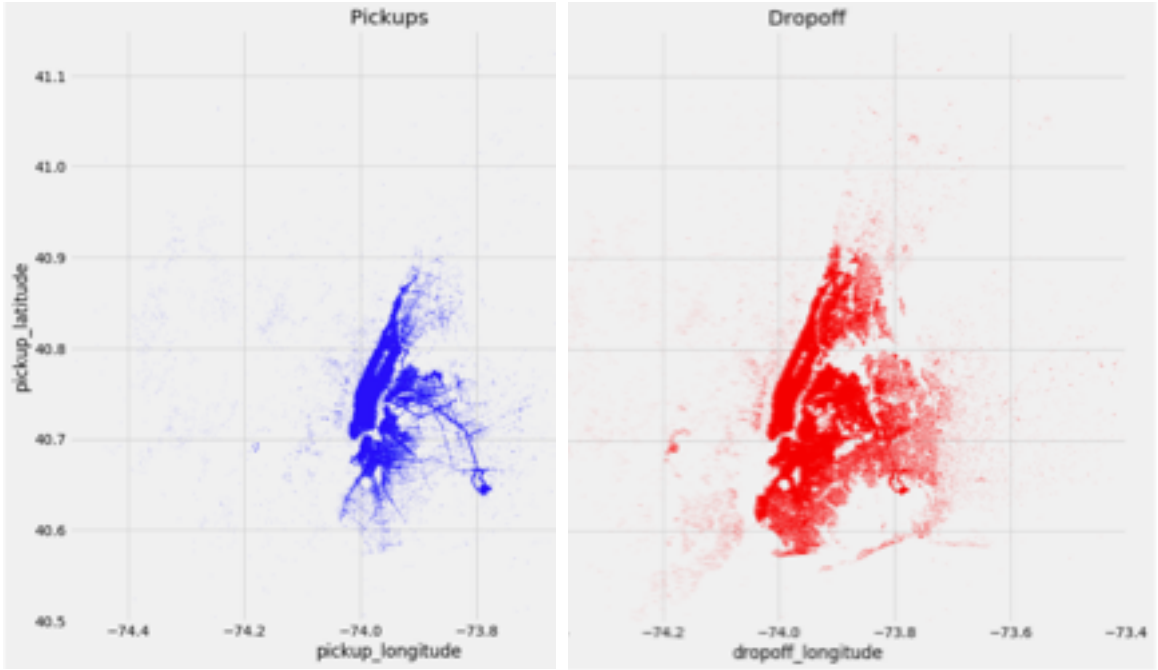
Ds_A → 14M Trip dataset (full June/15) → EDA → Sample of 120K dataset → Machine Learning

Ds_B → 120K Trip dataset (sample all months 2015) → EDA and Machine Learning

NYC Taxi - IoT – Getting data



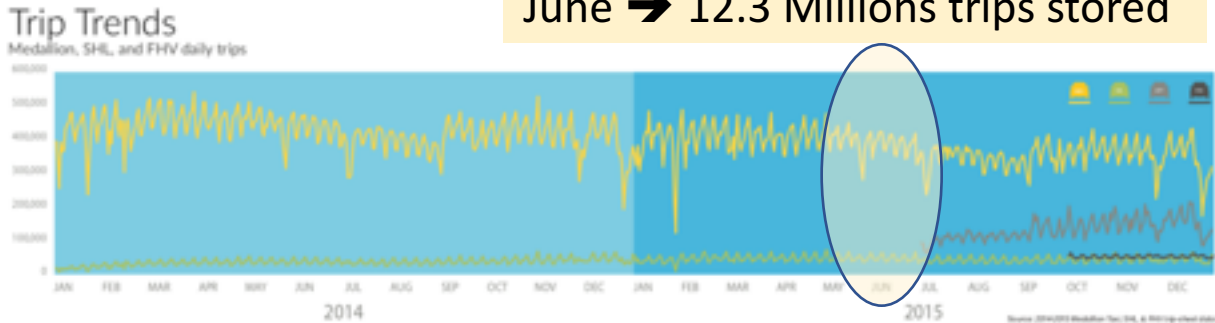
NYC Taxi - Dataset



	VendorID	tpep_pickup_datetime	tpep_dropoff_datetime	passenger_count	trip_distance	pickup_longitude	pickup_latitude	RateCodeID
0	2	2015-06-02 11:19:29	2015-06-02 11:47:52	1	1.63	-73.954430	40.764141	1
1	2	2015-06-02 11:19:30	2015-06-02 11:27:56	1	0.46	-73.971443	40.758942	1
2	2	2015-06-02 11:19:31	2015-06-02 11:30:30	1	0.87	-73.978111	40.738434	1
3	2	2015-06-02 11:19:31	2015-06-02 11:39:02	1	2.13	-73.945892	40.773529	1
4	1	2015-06-02 11:19:32	2015-06-02 11:32:49	1	1.40	-73.979088	40.776772	1

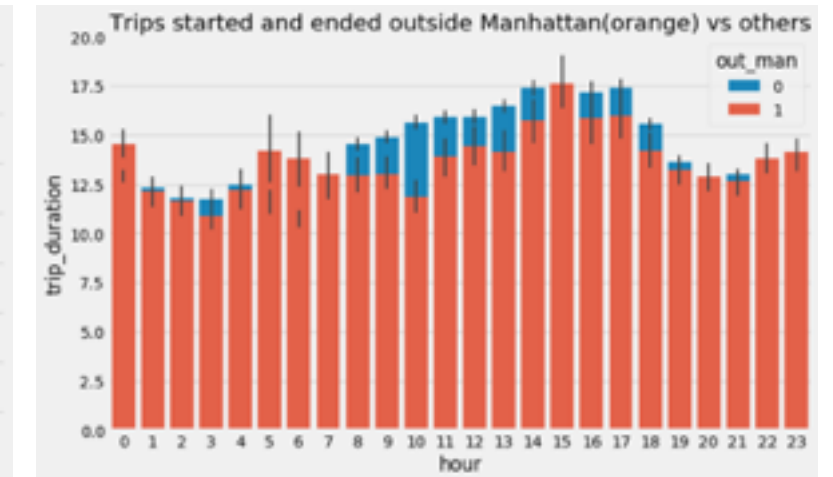
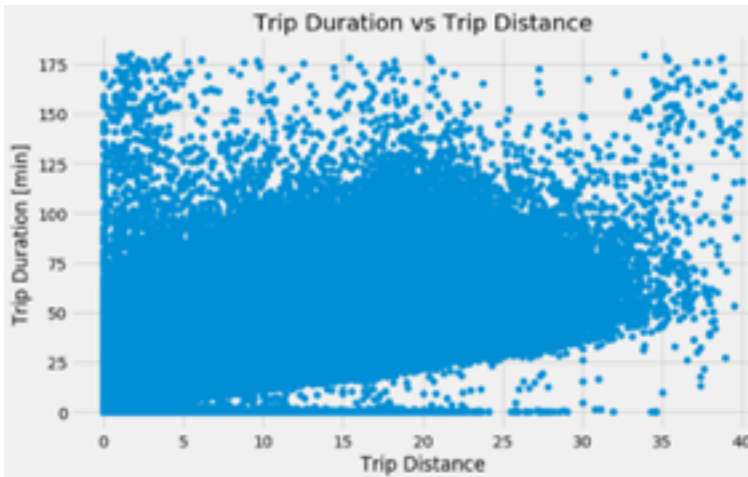
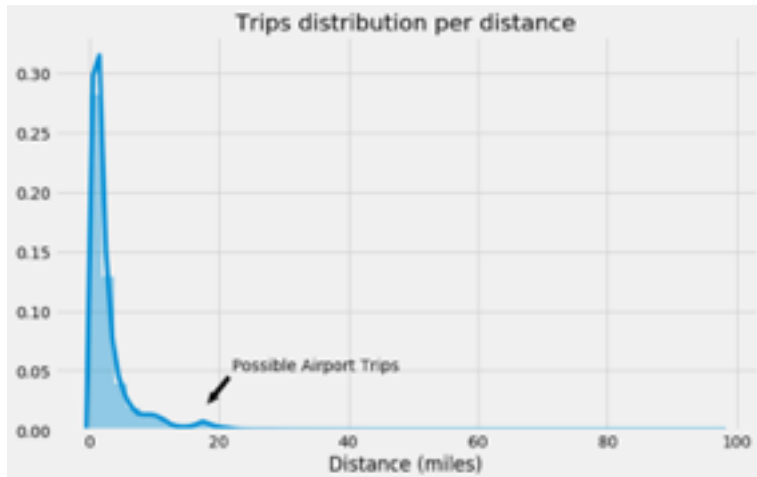
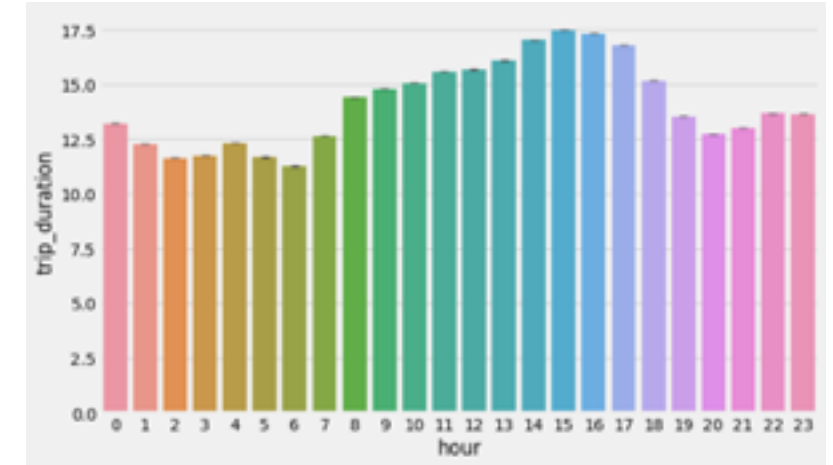
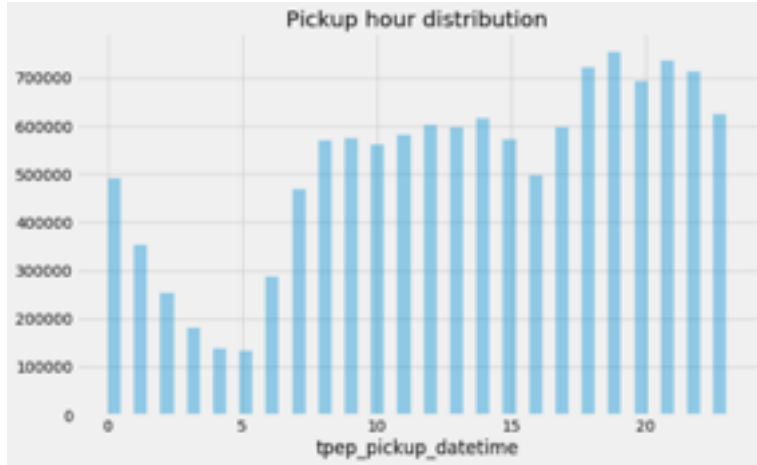
	dropoff_longitude	dropoff_latitude	payment_type	fare_amount	extra	mta_tax	tip_amount	tolls_amount	improvement_surcharge	total_amount
-73.974754	40.754093	2	17.0	0.0	0.5	0.00	0.0	0.3		17.80
-73.978539	40.761909	1	6.5	0.0	0.5	1.00	0.0	0.3		8.30
-73.990273	40.745438	1	8.0	0.0	0.5	2.20	0.0	0.3		11.00
-73.971527	40.760330	1	13.5	0.0	0.5	2.86	0.0	0.3		17.16
-73.982162	40.758999	2	9.5	0.0	0.5	0.00	0.0	0.3		10.30

June → 12.3 Millions trips stored

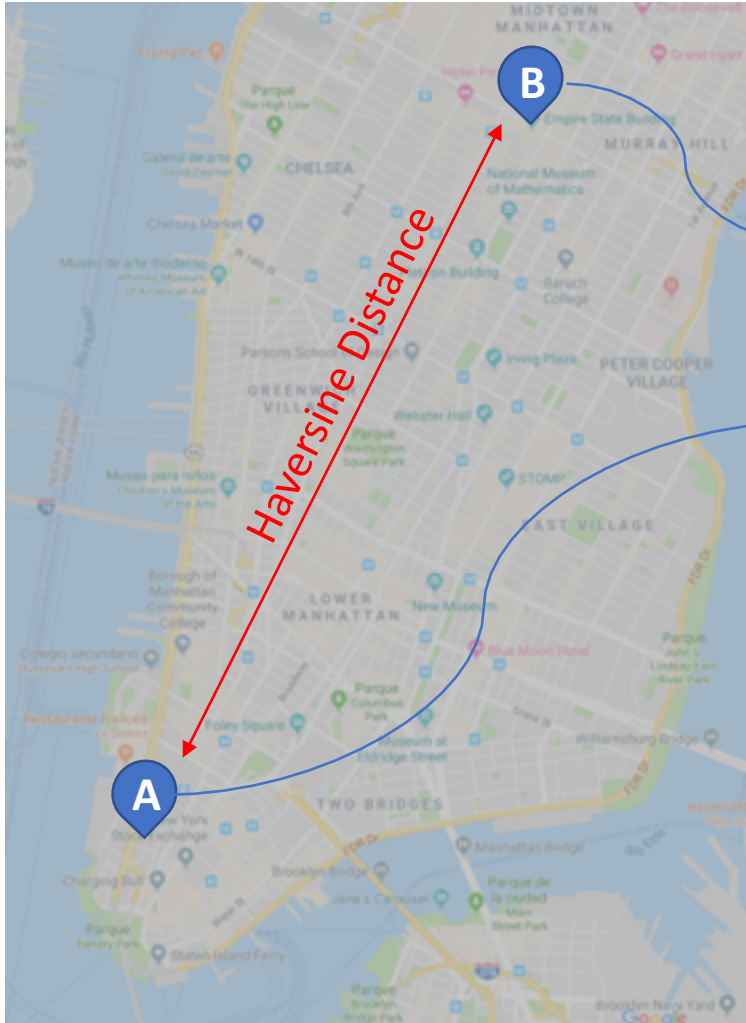


2015: 13,500 Yellow Cabs. → 400,000 trips/day

NYC Taxi Trip Prediction Project - EDA



NYC Taxi Trip Prediction - ML



Dropoff Latitude
Dropoff Longitude

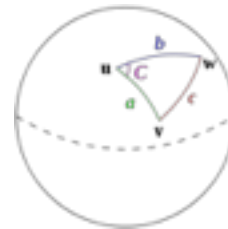
Pickup Latitude
Pickup Longitude
Pickup Time



Fare

Distance

Duration

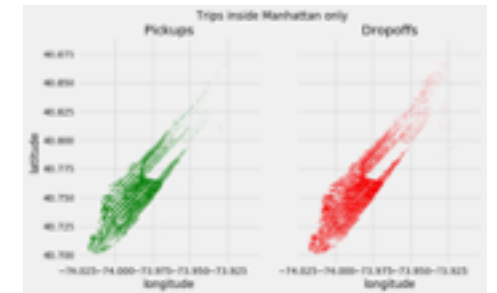


$$2\arcsin\left(\sqrt{\sin^2\left(\frac{\varphi_2 - \varphi_1}{2}\right) + \cos(\varphi_1)\cos(\varphi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right)$$

Hav.
Dist.

Rate
Code

In/Out
Manh.



NYC Taxi Trip Prediction - ML

	Fare DS_A	Fare DS_B		Fare DS_A	Fare DS_B
	RMSE	RMSE	Comparison A-B	Base line comparison	Base line comparison
Base Line	10,25	9,88	3,6%	0,0%	0,0%
Decisión Tree Regresor	3,25	3,21	1,2%	68,3%	67,5%
Random Forest Regresor	3,01	2,92	3,0%	70,6%	70,4%
Random Forest Regresor (Reducing features)	3,00	2,92	2,7%	70,7%	70,4%
Random Forest Regresor (Changing Hyper_param)	2,77	2,72	1,8%	73,0%	72,5%

	Time DS_A	Time DS_B		Time DS_A	Time DS_B
	RMSE	RMSE	Comparison A-B	Base line comparison	Base line comparison
Base Line	14,52	13,94	4,0%	0,0%	0,0%
Decisión Tree Regresor	11,47	10,75	6,3%	21,0%	22,9%
Random Forest Regresor	6,03	6,26	-3,8%	58,5%	55,1%
Random Forest Regresor (Reducing features)	5,60	5,73	-2,3%	61,4%	58,9%
Random Forest Regresor (Changing Hyper_param)	5,26	5,38	-2,3%	63,8%	61,4%

	Dist DS_A	Dist DS_B		Dist DS_A	Dist DS_B
	RMSE	RMSE	Comparison A-B	Base line comparison	Base line comparison
Base Line	3,61	3,46	4,2%	0,0%	0,0%
Decisión Tree Regresor	0,90	0,82	8,9%	75,1%	76,3%
Random Forest Regresor	0,83	0,75	9,6%	77,0%	78,3%
Random Forest Regresor (Changing Hyper_param)	0,78	0,71	9,0%	78,4%	79,5%

ML Model: Random Forest Regressor

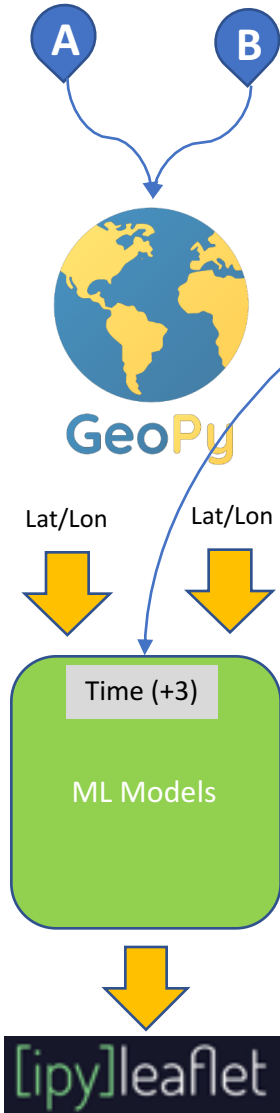
→ Fare: < \$3.00

→ Time: > 5 min

→ Dist: > 0.7 mile

Errors

NYC Taxi Trip Prediction - Deploy



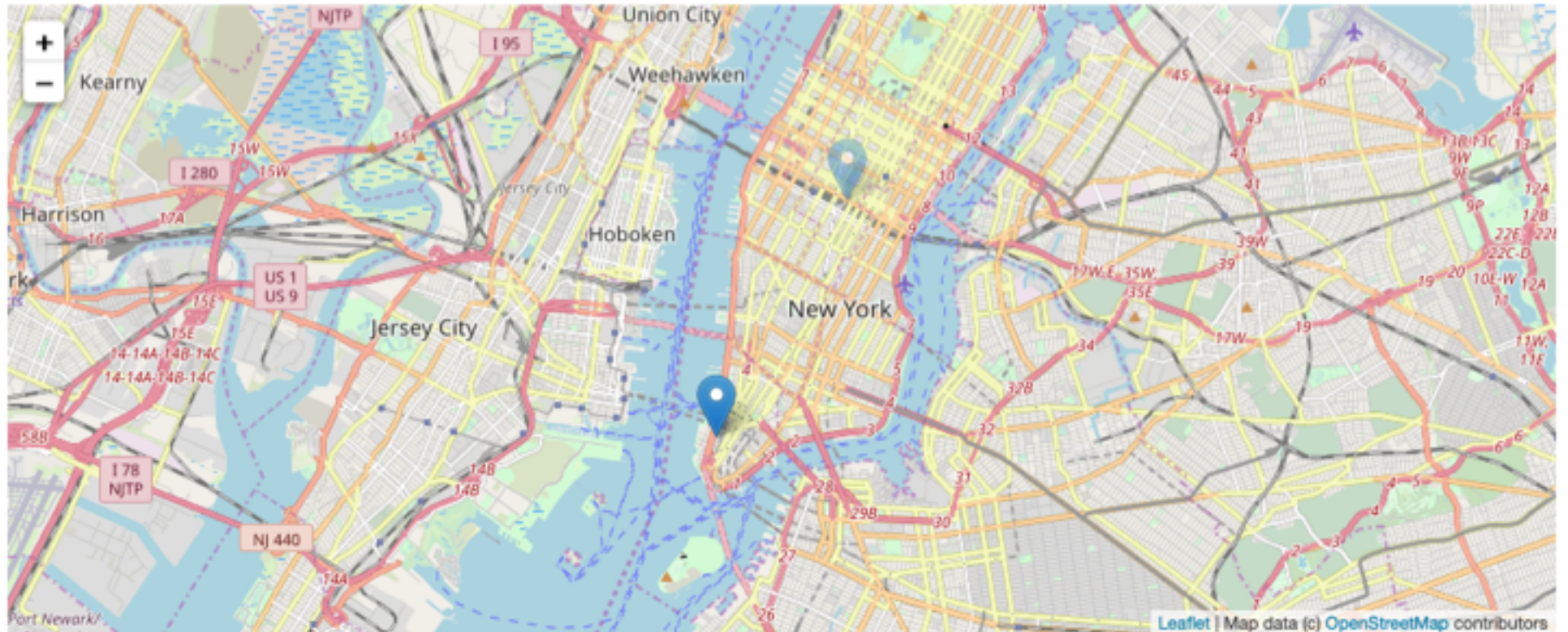
```

1 start_loc = "New York Marriott Downtown"
2 end_loc = "Empire State Building"
3
4 predict_trip(start_loc, end_loc)

```

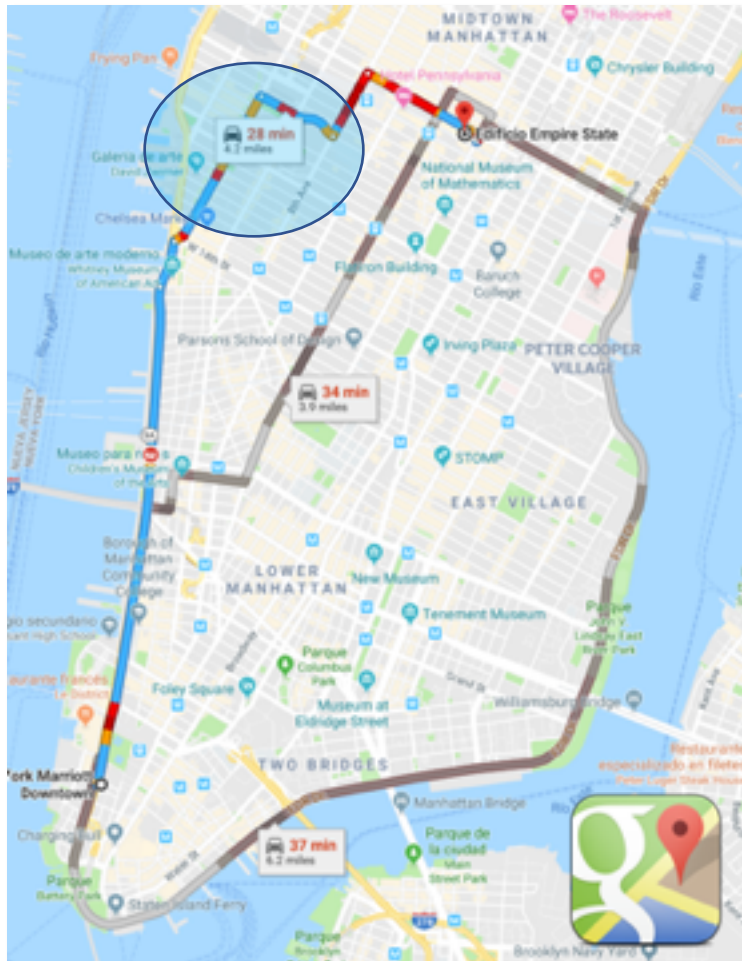
executed in 4.16s, finished 12:03:08 2019-02-08



Trip has 4.8 miles and will last 28.0 minutes, with a basic cost of \$21.57
 [Trip Info]: Trip inside Manhattan; Rate Code: Standard



NYC Taxi Trip Prediction - Benchmark

Trip has 4.8 miles and will last 28.0 minutes with a basic cost of \$21.57
 [Trip Info]: Trip inside Manhattan; Rate Code: Standard



Sign up for our newsletter!  Like 1056  Tweet New York, NY English

1 | Taxis To and From Airport Cheap and reliable Transfer to/from the Airport. Book your Taxi Today! Taxi2Airport.com
2 | Private Airport Transfer Best price. 24/7. Everywhere. Discounts & special offers. Book online! gettransfer.com

TAXI FARE FINDER

NEW YORK, NY TAXI FARE CALCULATOR FARE COMPARISON RATE CHART NEWSROOM LOCAL TAXI INFO

Please take our quick (2 question) survey on tipping your ride-share drivers and automatically be entered to win one of \$5 RideGuru Swag Bags!

Estimated Fare
\$24.51
 For New York, NY rates \$29.45 incl. 20% tip

See Uber, Lyft, and more...

Other possible fares
 8:00PM to 6:00AM \$30.81
 Cost with Traffic \$19.35 \$24.51 \$47.32

Fare Information
 Trip Information
 Trip is 6.1 mi, 17 mins
 Notes

- Add \$1.00 Peak Hour surcharge (4:00PM to 8:00PM, M-F)
- \$52+toils fairsate JFK -- Manhattan Add \$17.50 to Newark Airport (EWR)
- Trips beyond city limits may be under flat rate or have the portion of the trip beyond the city limit doubled (Rate 4).
- Click here for the NY City Taxi rates information.
- NYC Yellow Taxis can only be street-hailed. They cannot be pre-booked over the phone.
- Tolls & surcharges may apply

How much does a taxi cab cost from Marriott Downtown Grand Ballroom B, New York, New York, United States, 10006 to Empire State Building, 350 5th Ave, New York, New York, United States, 10001 in New York, NY?

Taxi from Airport To Airport Car Service Find Me a Taxi Service

Suggested routes: 1 of 1 (\$24.51) Fastest Cheapest Shortest

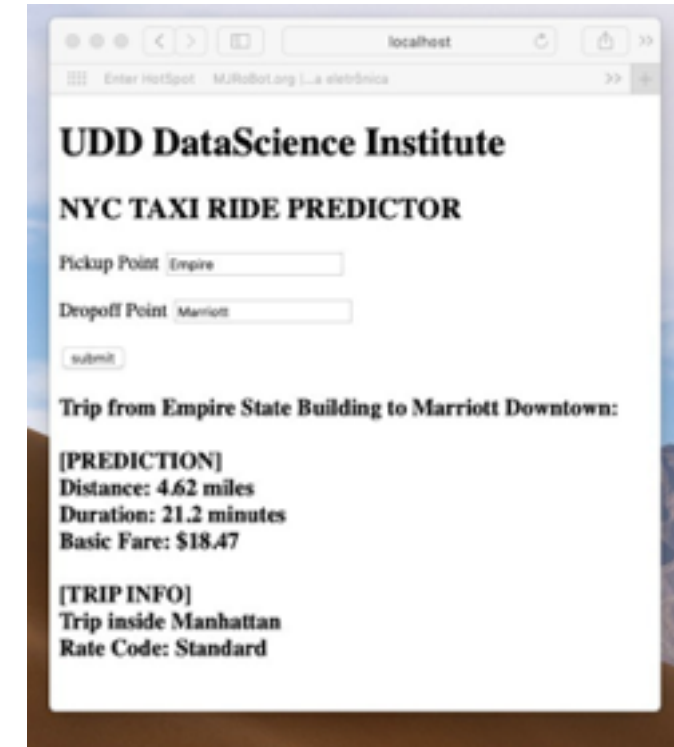
Marriott Downtown Grand Ballroom B, New York, New Y or select a point of interest:
 Empire State Building, 350 5th Ave, New York, New York or select a point of interest:

Get Fare!

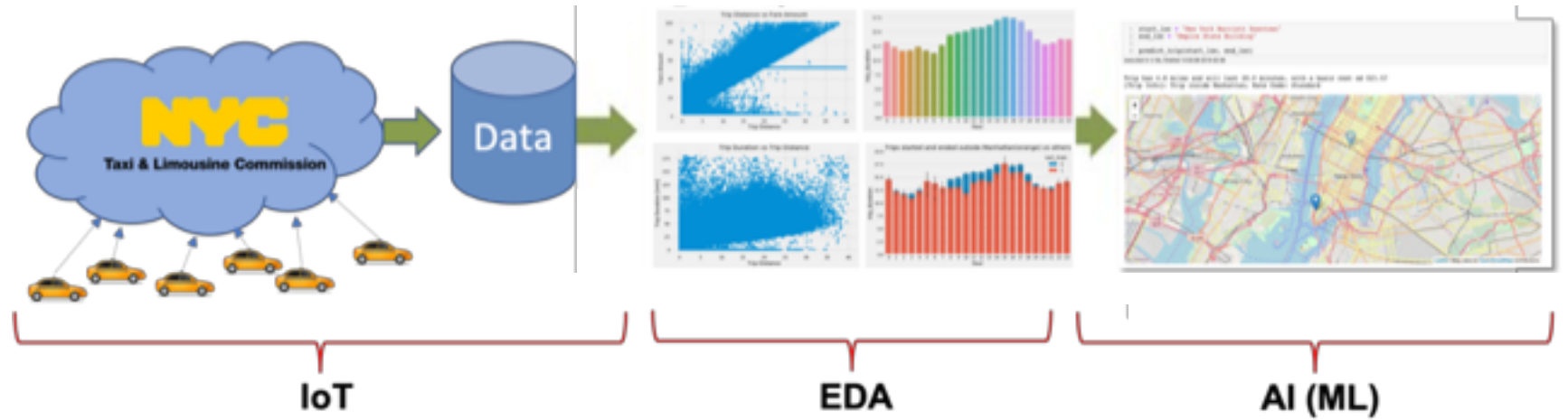
Ad closed by Google
 Report this ad Why this ad?

- New Datasets
 - New Historical Random Samples for each months
- City time Travel at Rush Hour
 - Average traffic speed per hour
 - Number of cabs per hour
- Pickup Demand
 - Specific demand in small zones
- Other ML modes
 - ARIMA
 - Nested Model

- An App
 - Work on a App to be used on mobiles and dektop



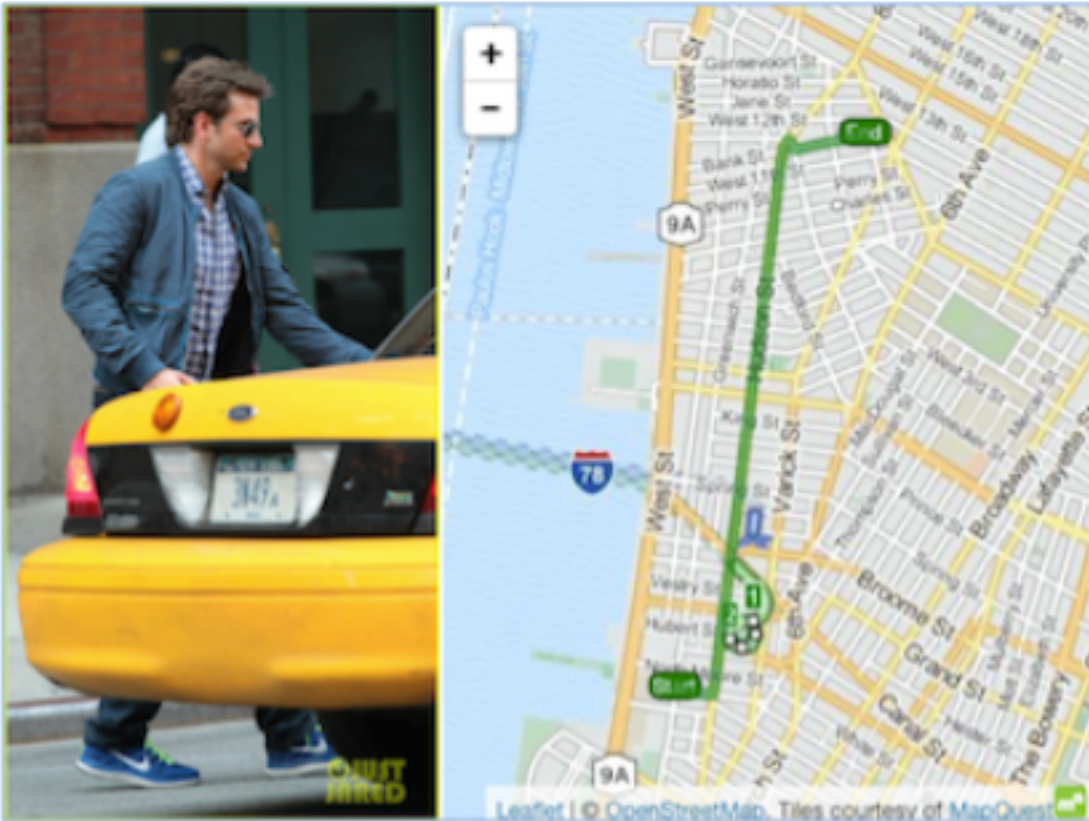
Data Science Environment



https://github.com/Mjrovai/UDD_Master_Data_Science/tree/master/AML-NYC_TAXI_TRIP_PREDICTION

Issue - Violating privacy

Stalking celebrities



A few innocent nights at the gentlemen's club



The End

Briceño, Heriberto
Rovai, Marcelo
Sacasa, Manuel