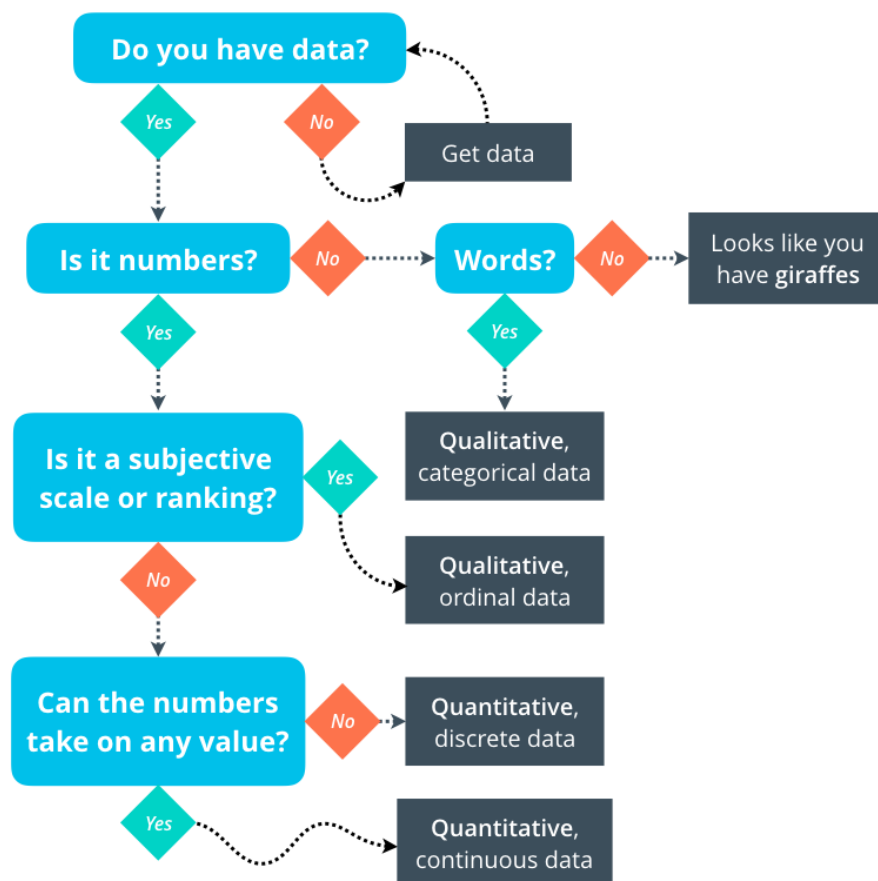


# OSHO AGYEYA

## PUCHO ROUND 2

### New York City Taxi Trip



**Osho Agyeya**  
**oshoagyeya123@gmail.com**  
**+917800191911**

# DATASET USED

Tableau Public - Pucho@osho

File Data Window Help

train

train.csv

Filters  
0 | Add

Connections  
Add  
train  
Text File

Files  
train.csv  
New Union

Sort fields Modified

Show aliases Show hidden fields 1,000 rows

Alt	#	Pickup Dateti...	Dropoff Dateti...	Passenger Co...	Pickup Longit...	Pickup Latitude	Dropoff Longi...	Dropoff Latitu...	Store And Fw...	Trip Duration	Trip Duration ...
id01904...	2	01-01-2016 12 0...	01-01-2016 12 1...	5	-73.981743	40.719158	-73.938828	40.829182	N	849	:
id16655...	1	01-01-2016 12 0...	01-01-2016 12 2...	1	-73.985085	40.747166	-73.958038	40.717491	N	1,294	:
id12103...	2	01-01-2016 12 0...	01-01-2016 12 0...	5	-73.965279	40.801041	-73.947479	40.815170	N	408	:
id38882...	1	01-01-2016 12 0...	01-01-2016 12 0...	1	-73.982292	40.751331	-73.991341	40.750340	N	280	:
id09242...	1	01-01-2016 12 0...	01-01-2016 12 1...	1	-73.970108	40.759800	-73.989357	40.742989	N	736	:
id22943...	2	01-01-2016 12 0...	01-01-2016 12 1...	1	-73.984993	40.773891	-73.936493	40.847771	N	712	:
id10782...	2	01-01-2016 12 0...	01-01-2016 12 0...	1	-73.973335	40.764072	-73.974854	40.761734	N	114	:
id36094...	1	01-01-2016 12 0...	01-01-2016 12 2...	2	-73.993103	40.752632	-73.953903	40.816540	N	1,204	:
id29143...	2	01-01-2016 12 0...	01-01-2016 12 2...	1	-73.985443	40.735710	-73.957489	40.811611	N	1,310	:
id36759...	2	01-01-2016 12 0...	01-01-2016 12 1...	1	-73.993736	40.741760	-74.004669	40.745468	N	445	:
id08461...	2	01-01-2016 12 0...	01-01-2016 12 0...	1	-73.976860	40.683376	-73.983643	40.688866	N	263	:
id05661...	2	01-01-2016 12 0...	01-01-2016 12 2...	1	-74.006439	40.739475	-73.967201	40.769188	N	1,384	:
id03725...	1	01-01-2016 12 0...	01-01-2016 12 1...	1	-73.969788	40.757584	-73.951576	40.759445	N	770	:

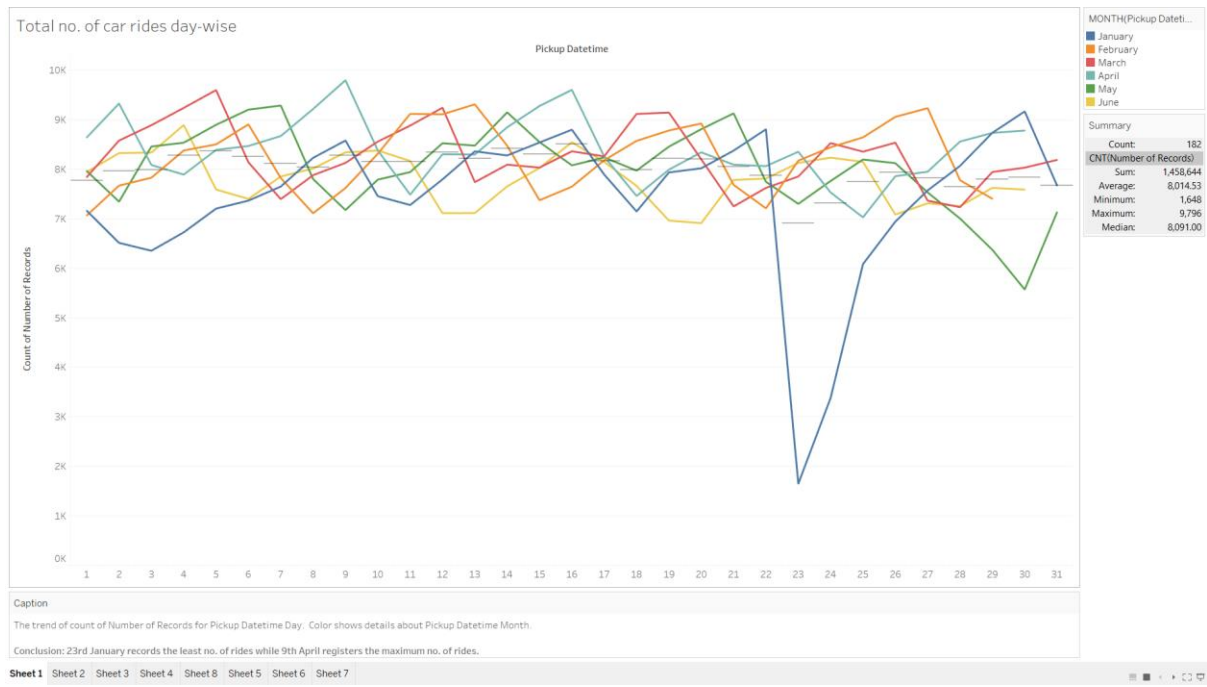
Data Source Sheet 1 Sheet 2 Sheet 3 Sheet 4 Sheet 5 Sheet 6 Sheet 7

Osho Aggeya

06:48 PM  
31-01-2018

Dataset:train.csv

Source: <https://www.kaggle.com/c/nyc-taxi-trip-duration/data>



## Visualisation 1

### Design/Variables

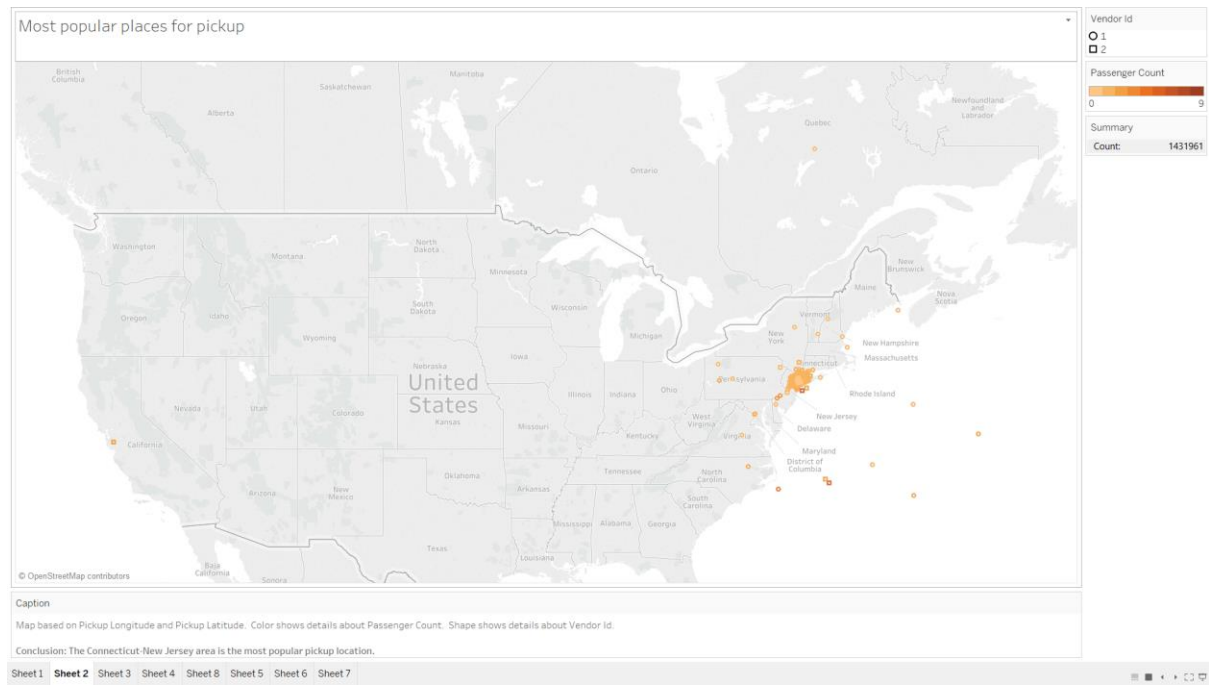
The trend of count of Number of Records for Pickup Datetime Day. Color shows details about Pickup Datetime Month.

### Conclusion

23rd January records the least no. of rides while 9th April registers the maximum no. of rides.

### Questions:

Q) Why is it that 23<sup>rd</sup> January registers a sharp decline in no. of rides?



## Visualisation 2

### Design/Variables

Map based on Pickup Longitude and Pickup Latitude. Color shows details about Passenger Count. Shape shows details about Vendor Id.

### Conclusion

The Connecticut-New Jersey area is the most popular pickup location.

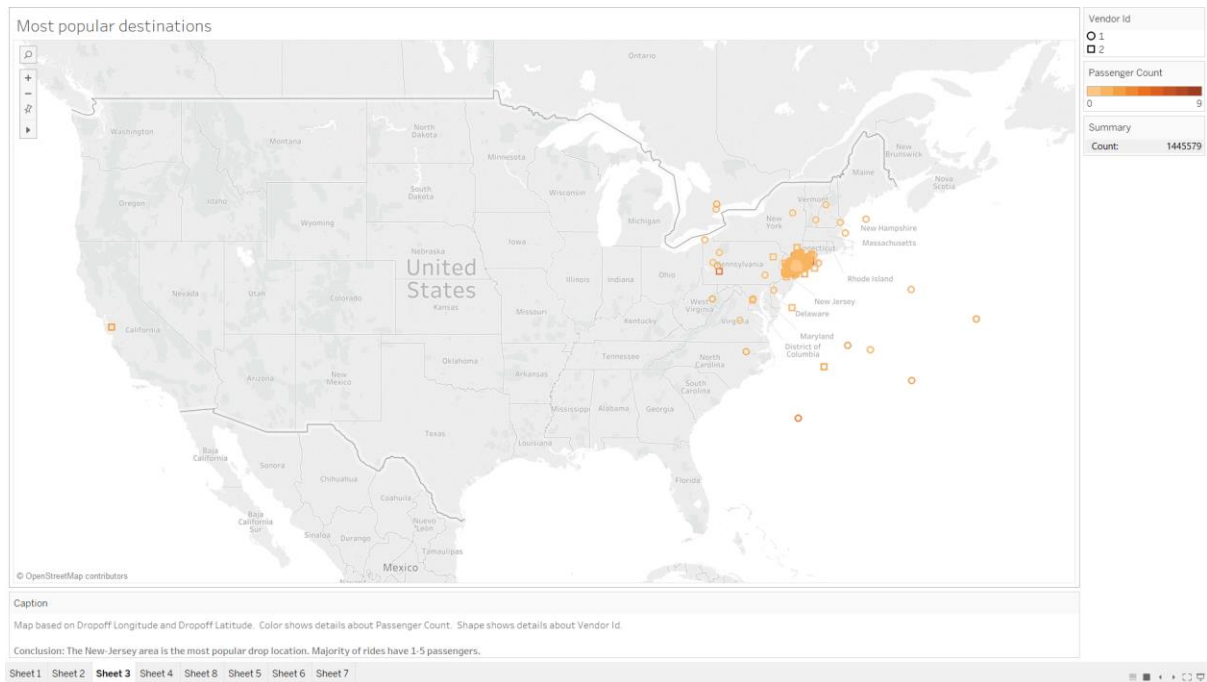
### Questions:

Q)Why is it that Connecticut, Rhode Island and New Jersey are the places from where maximum rides take off?

Q)Why is it that some rides have zero passengers?

Q)Have the drivers wrongly reported to the vendors regarding certain rides so that they can ride their own journey?

Q)Why is it that some rides have pickup location as sea?



## Visualisation 3

### Design/Variables

Map based on Dropoff Longitude and Dropoff Latitude. Color shows details about Passenger Count. Shape shows details about Vendor Id.

### Conclusion

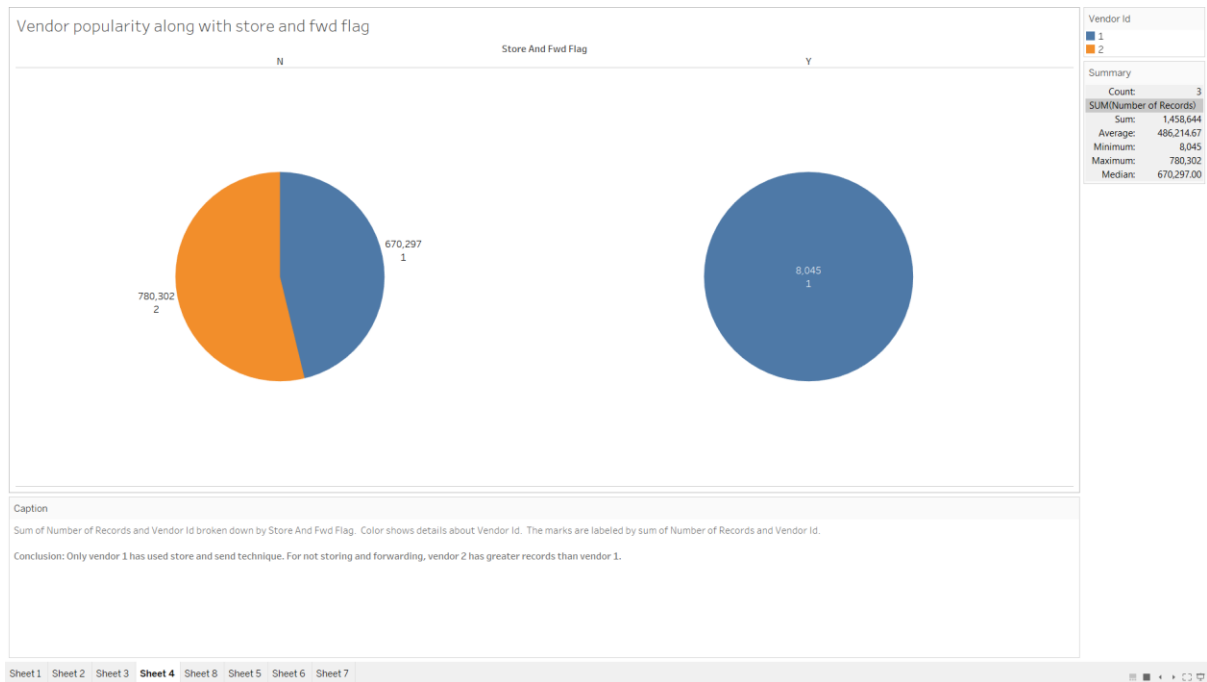
The New-Jersey area is the most popular drop location. Majority of rides have 1-5 passengers.

### Questions:

Q)Why is it that Connecticut, Rhode Island and New Jersey are the places where maximum end?

Q)Why is it that some rides have drop location as sea?

---



## Visualisation 4

### Design/Variables

Sum of Number of Records and Vendor Id broken down by Store And Fwd Flag. Color shows details about Vendor Id. The marks are labeled by sum of Number of Records and Vendor Id.

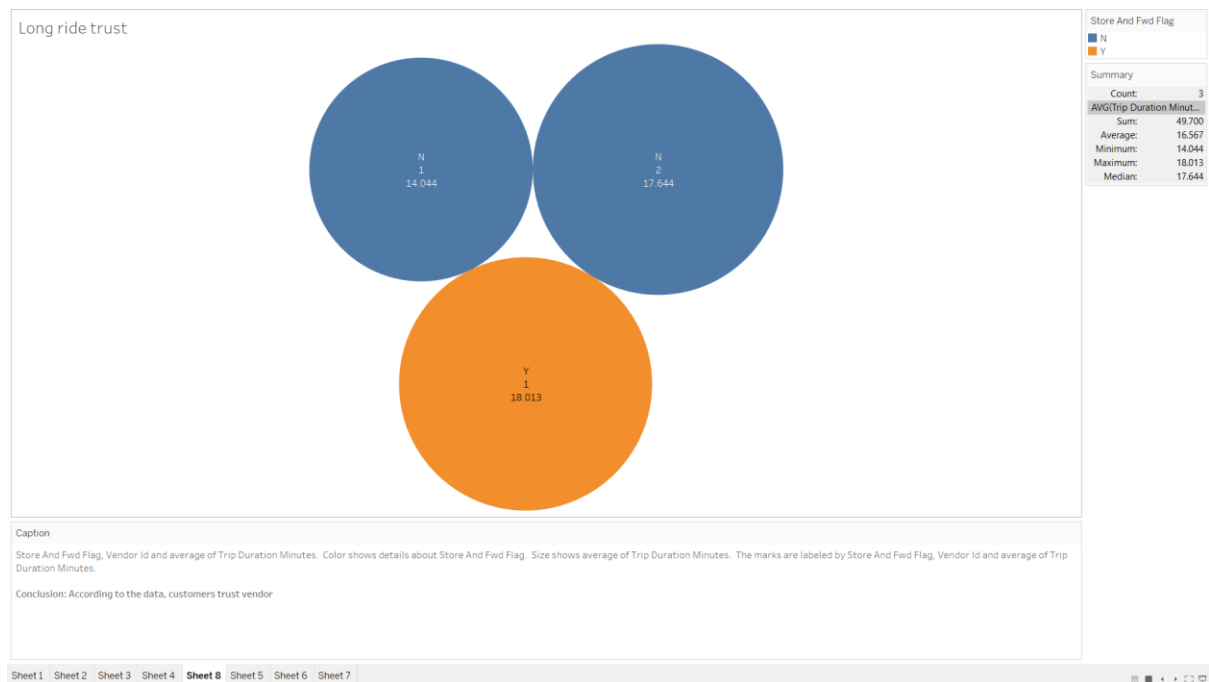
### Conclusion

Only vendor 1 has used store and send technique. For not storing and forwarding, vendor 2 has greater records than vendor 1.

### Questions:

Q) Why is it that vendor 2 does not have any have records where value of store and fwd flag is 1?

---



## Visualisation 5

### Design/Variables

Store And Fwd Flag, Vendor Id and average of Trip Duration Minutes. Color shows details about Store And Fwd Flag. Size shows average of Trip Duration Minutes. The marks are labeled by Store And Fwd Flag, Vendor Id and average of Trip Duration Minutes.

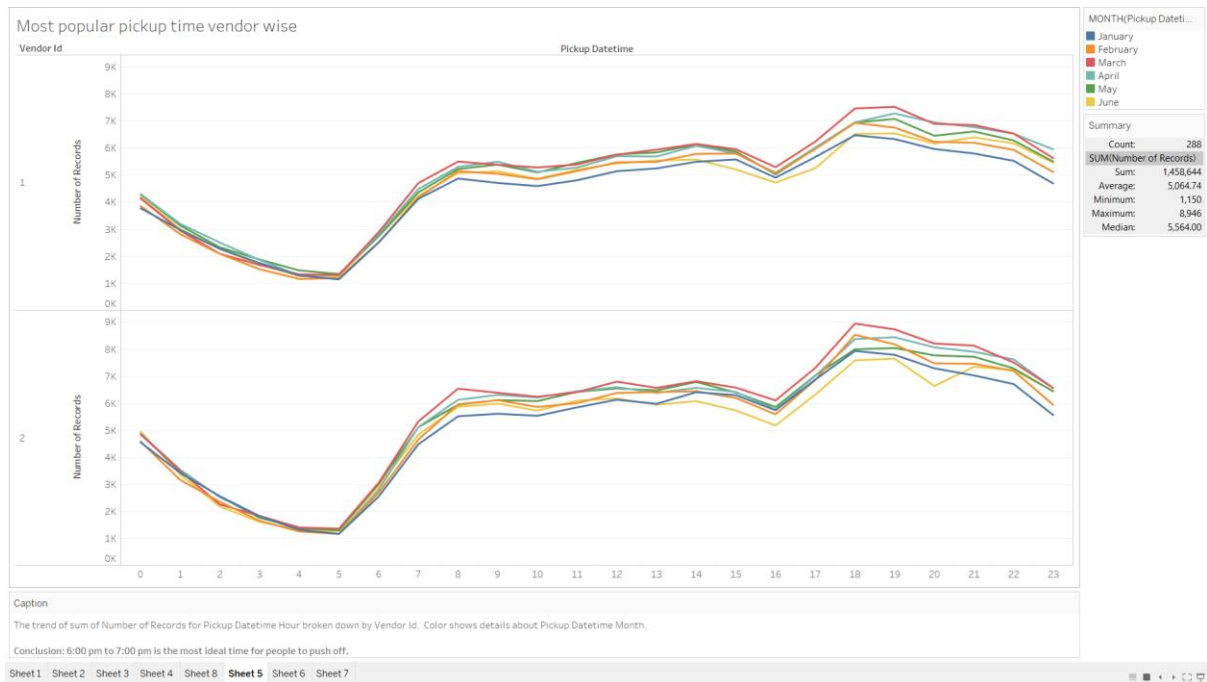
### Conclusion

According to the data, customers trust vendor 2 more than vendor 1 in case where Store and Fwd flag id N, while they prefer vendor 1 in the other case.

### Questions

Q) Why is it that passengers trust vendor 1 more than vendor 2?

---



## Visualisation 6

### Design/Variables

The trend of sum of Number of Records for Pickup Datetime Hour broken down by Vendor Id. Color shows details about Pickup Datetime Month.

### Conclusion

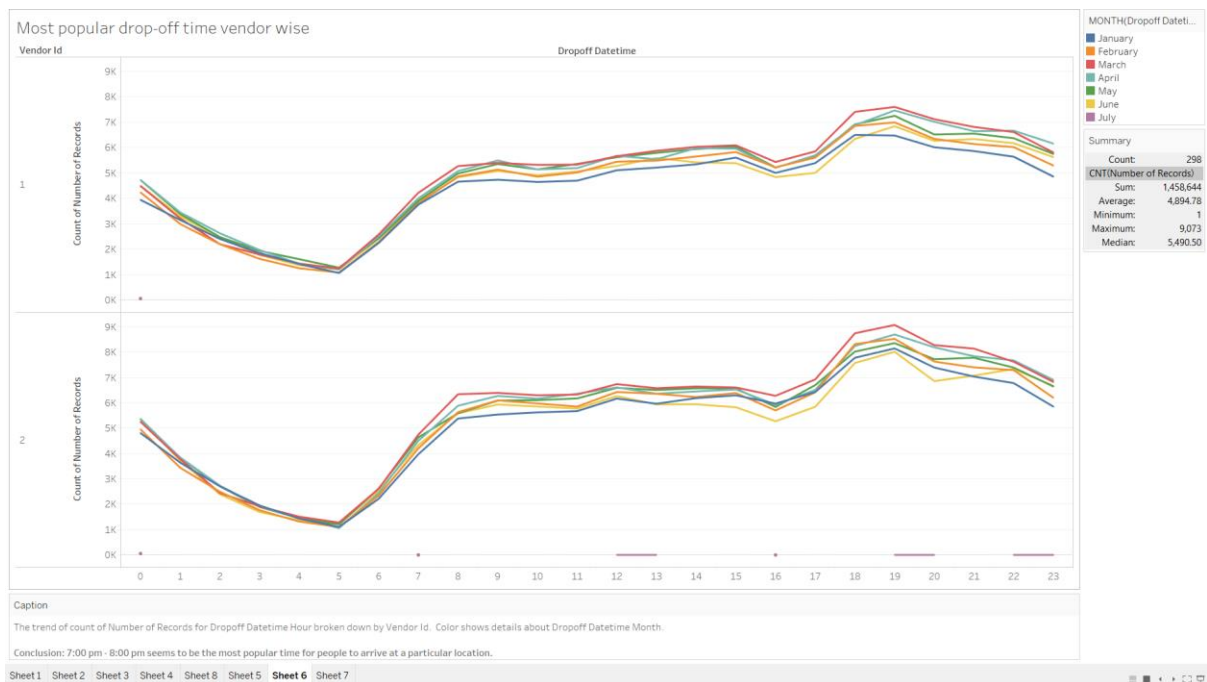
6:00 pm to 7:00 pm is the most ideal time for people to push off.

### Questions:

Q) Why is it that least no. of people book a ride at 5.00 am?

---





## Visualisation 7

### Design/Variables

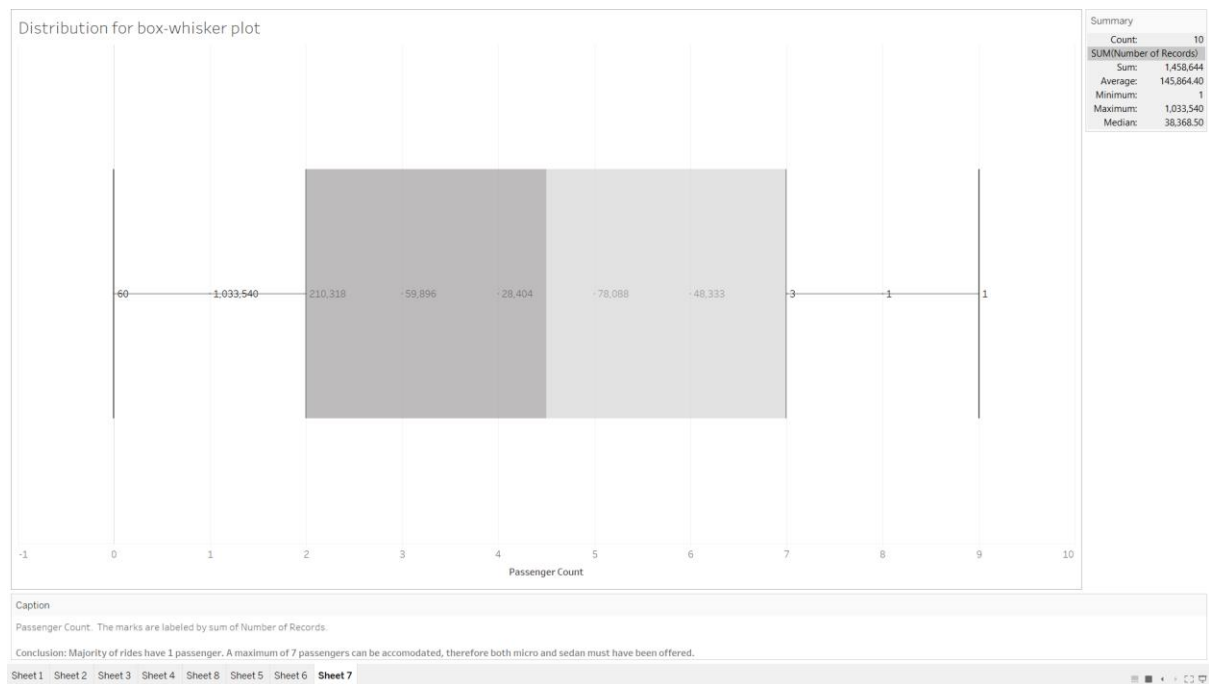
The trend of count of Number of Records for Dropoff Datetime Hour broken down by Vendor Id. Color shows details about Dropoff Datetime Month.

### Conclusion

7:00 pm - 8:00 pm seems to be the most popular time for people to arrive at a particular location.

### Questions:

Q) Why is it that for the dropoff time graph, July has the least no. of records and that to sparsely distributed?



## Visualisation 8

### Design/Variables

Passenger Count. The marks are labeled by sum of Number of Records.

### Conclusion

Majority of rides have 1 passenger. A maximum of 7 passengers can be accommodated, therefore both micro and sedan must have been offered.

### Questions:

Q) How have 9 people been accommodated in a taxi car?

---