LABMENTIX

UBER REQUEST DATA PROJECT

Summary and Insights

LABMENTIX

Submission By: Muskan Chauhan

Batch: 5th April Batch – 1

Submission Date: 22nd June, 2025

Steps Done

❑Excel

- Data Cleaning
- Pivot Tables
- and Dashboard

❑SQL

- New Insights

❑Python

- EDA

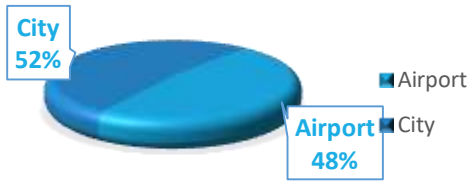
Excel

1. From an un-cleaned data to cleaned data.

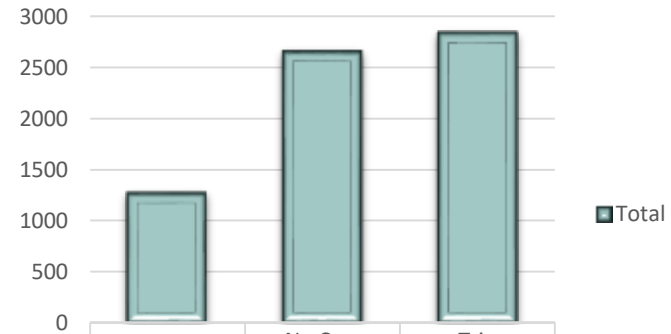
- Ordered the Request Ids in ascending order.
- Convert the timestamp to Date Time Format for further easy working in excel and SQL.
- Created new columns – "*Request Hour*" and "*Day*" by extraction from the "*Request timestamp*" column.
- Removed the duplicates.
- Replacing the NA values in "*Driver id*" and "*Drop timestamp*" column by "*Blank*".

Request id	Pickup point	Driver id	Status	Request timestamp	Drop timestamp	Request Hour	Day
1	Airport	285	Trip Completed	11-07-2016 00:20	11-07-2016 00:51	0	Monday
2	Airport		No Cars Available	11-07-2016 00:23		0	Monday
3	Airport	80	Trip Completed	11-07-2016 00:24	11-07-2016 01:31	0	Monday
4	City		No Cars Available	11-07-2016 00:37		0	Monday
5	Airport	264	Trip Completed	11-07-2016 00:36	11-07-2016 01:35	0	Monday
6	City		No Cars Available	11-07-2016 00:36		0	Monday
7	Airport		No Cars Available	11-07-2016 00:30		0	Monday

REQUEST DISTRIBUTION BY PICKUP POINT

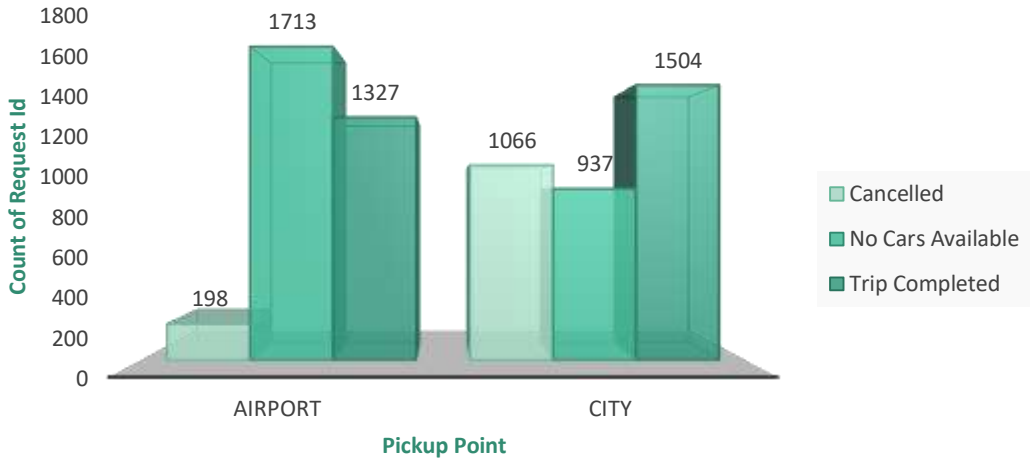


Trip Request Status Summary

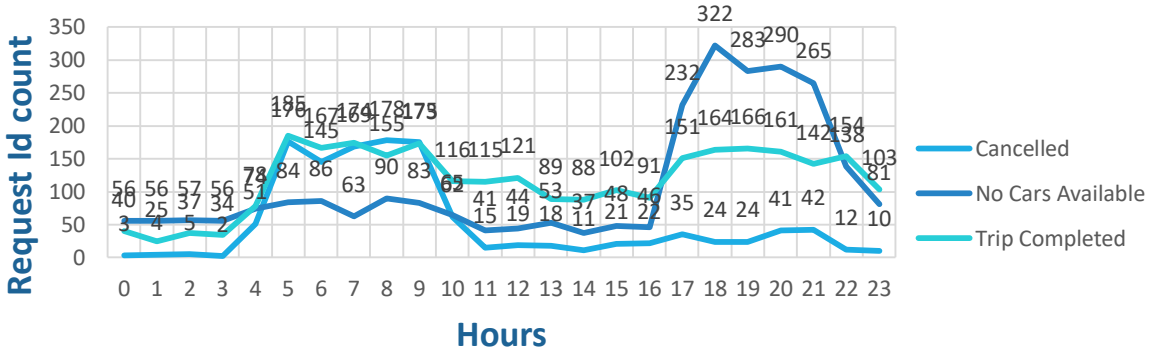


	Cancelled	No Cars Available	Trip Completed
Total	1264	2650	2831

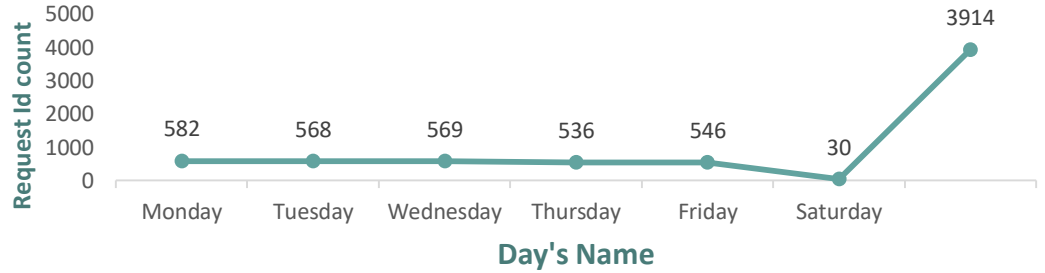
Trip status by Pickup point



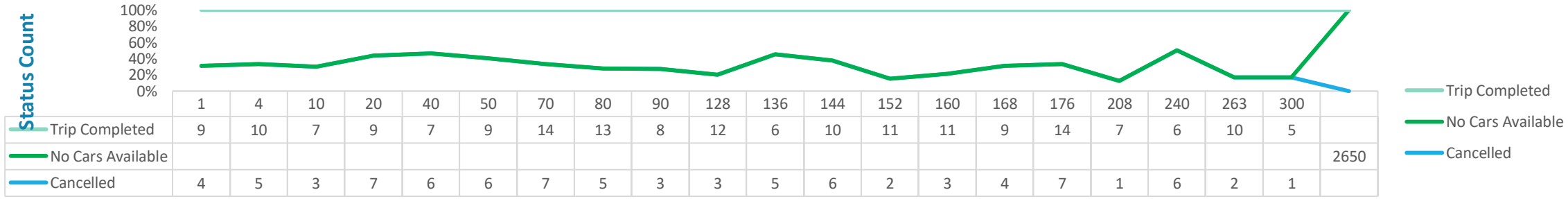
Hourly Trip Status Distribution



Day-wise Request Distribution



Trip status Distribution by Driver's Id

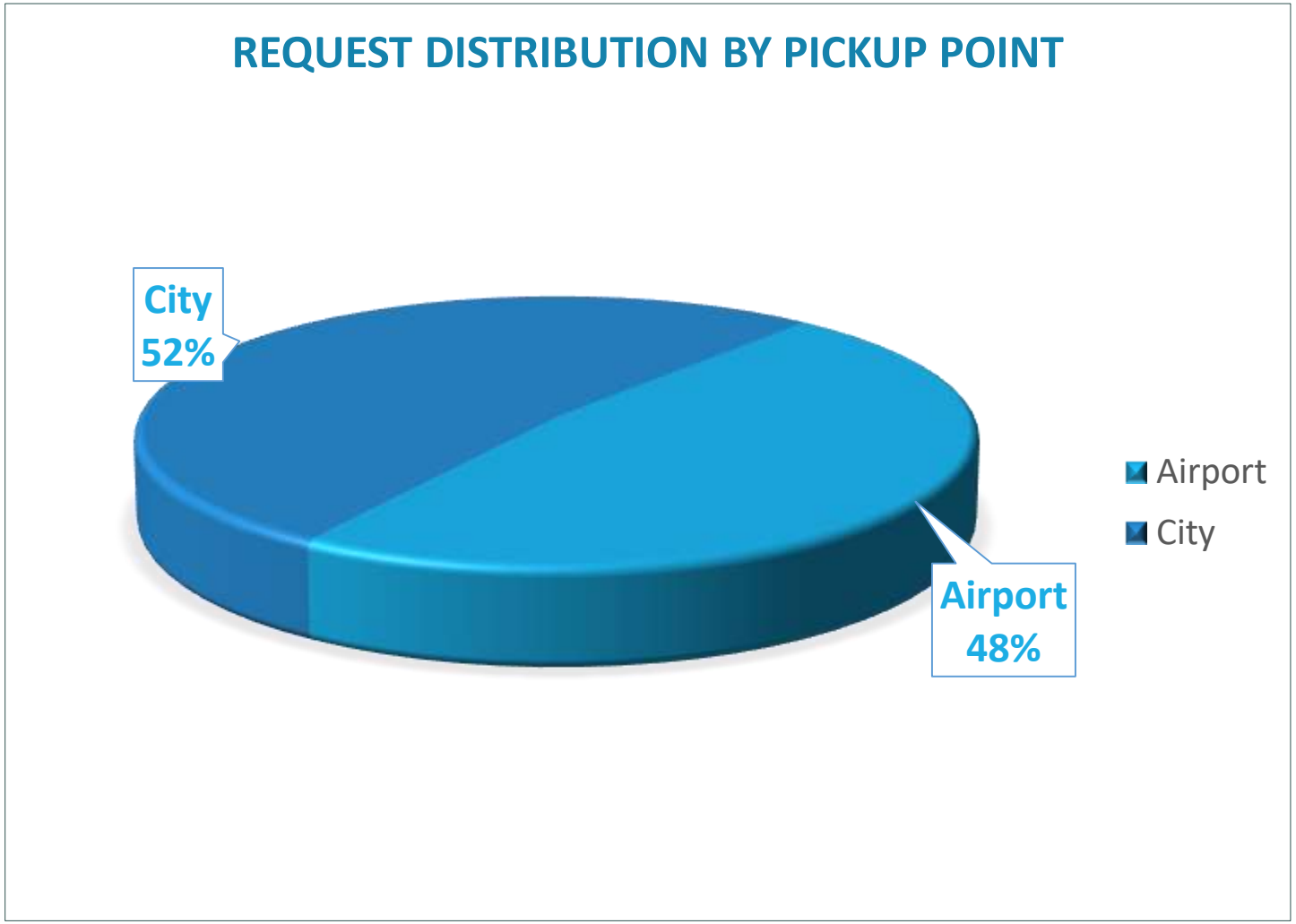


	1	4	10	20	40	50	70	80	90	128	136	144	152	160	168	176	208	240	263	300	
Trip Completed	9	10	7	9	7	9	14	13	8	12	6	10	11	11	9	14	7	6	10	5	
No Cars Available																					2650
Cancelled	4	5	3	7	6	6	7	5	3	3	5	6	2	3	4	7	1	6	2	1	

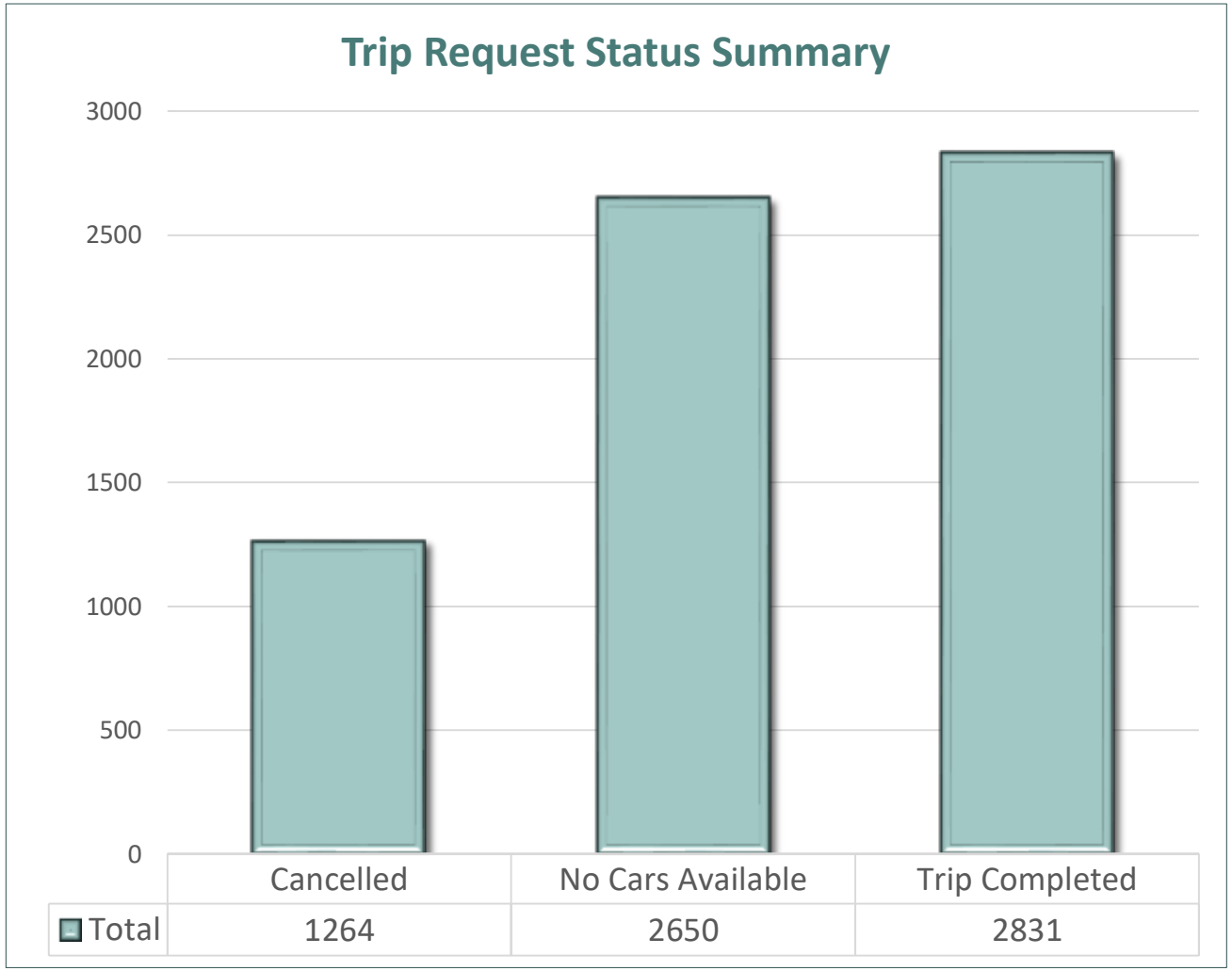
DASHBOARD

Visualization

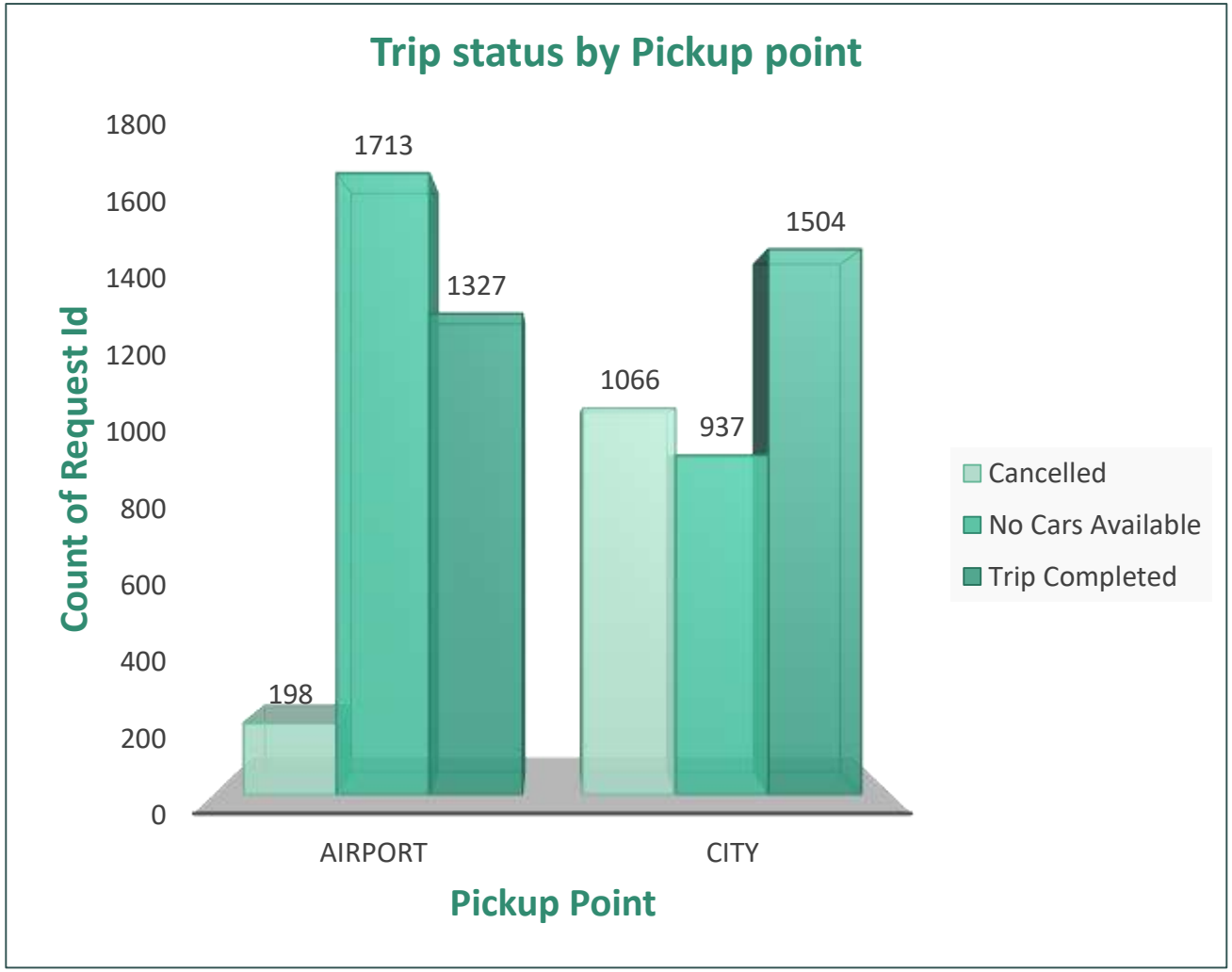
1. The graph concludes that the Uber requests calls have been received from “City” more than that of “Airport” pickup point.
2. This might be due to number of population residing in the “City” area as compared to near by “Airport”.
3. More people commute in the “City”.



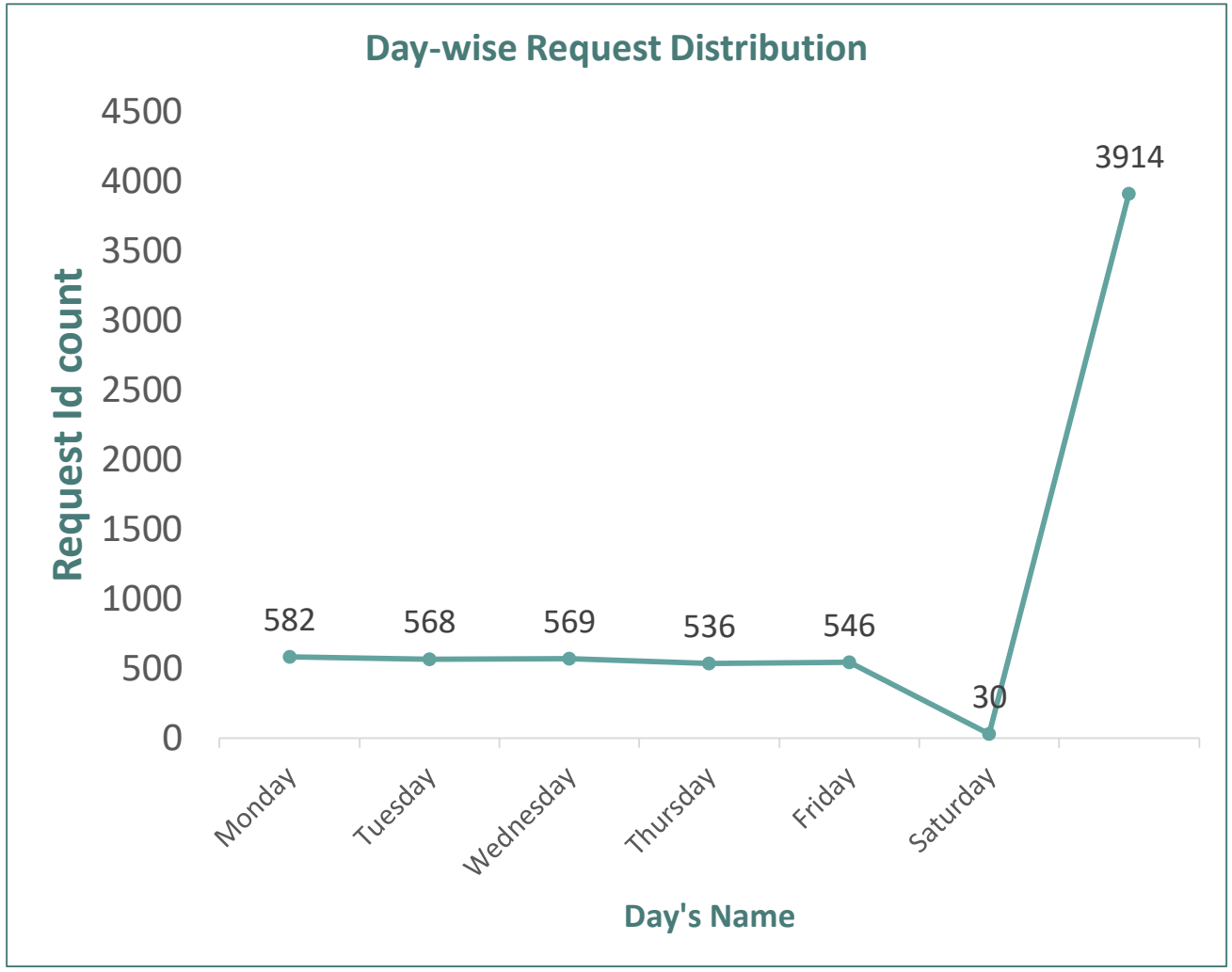
1. There are total of 6,745 requests made from both "*City*" and "Airport" combined.
2. Uber needs to hire more cars or cabs to fulfill its customer's demand as there is quite high number of "*no availability of cars/cabs*".
3. It also positively indicates that number of cabs already hired are in use most probably, hence, people are using the facility to its full efficiency.
4. Cancelled trips might be due to the heavy traffic or inability to find the exact pickup location.



1. The graph distributes the Trip status by their pickup point.
2. Most trips got cancelled in the “City” due to traffic, inability to reach the exact location or cancelled due to short trips.
3. “No cars available” at the “Airport” might be due to high rate of cab bookings to commute to other places from airport.



1. Monday is the most busiest day with 582 Uber requests where as Saturday is the least.
2. Most people travel to Schools, Offices and other workplaces using the cabs after a weekend hence the highest demand on Monday.
3. Low Uber demand on Saturday might be due to day off on weekends and people prefer to stay in.

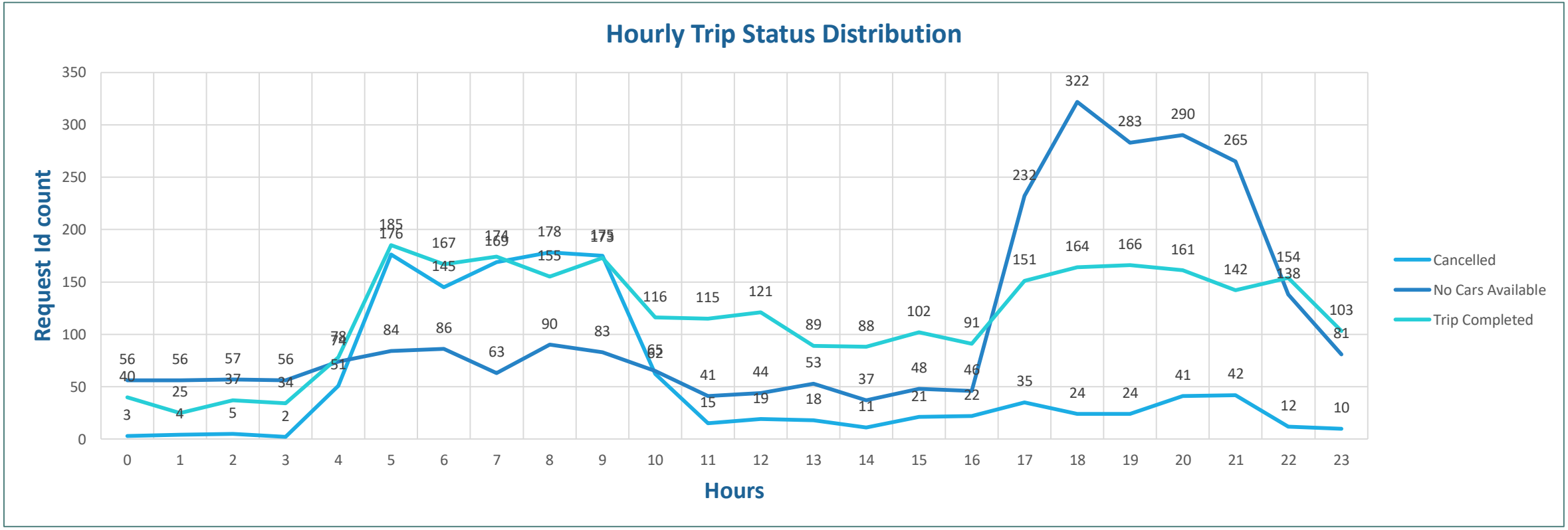


1. Evening Hours (6PM-9PM) are busiest hours as it shows the most cars unavailable along with most of the trips completed. Because people return from work or go out at these hours.

2. The 2nd most busiest hour lie between 4Am to 9Am as people have to go to their work. But it is not as busy as the Evening hour might be because most people do not go out to spend their time for movies, malls etc. in the morning.

3. Less people travel during late night and after noon.

4. From graph we can conclude that, the more the number of the “Trips completed”, more the number of “No cars available”. Uber need to hire more cabs and drivers to bridge this gap.

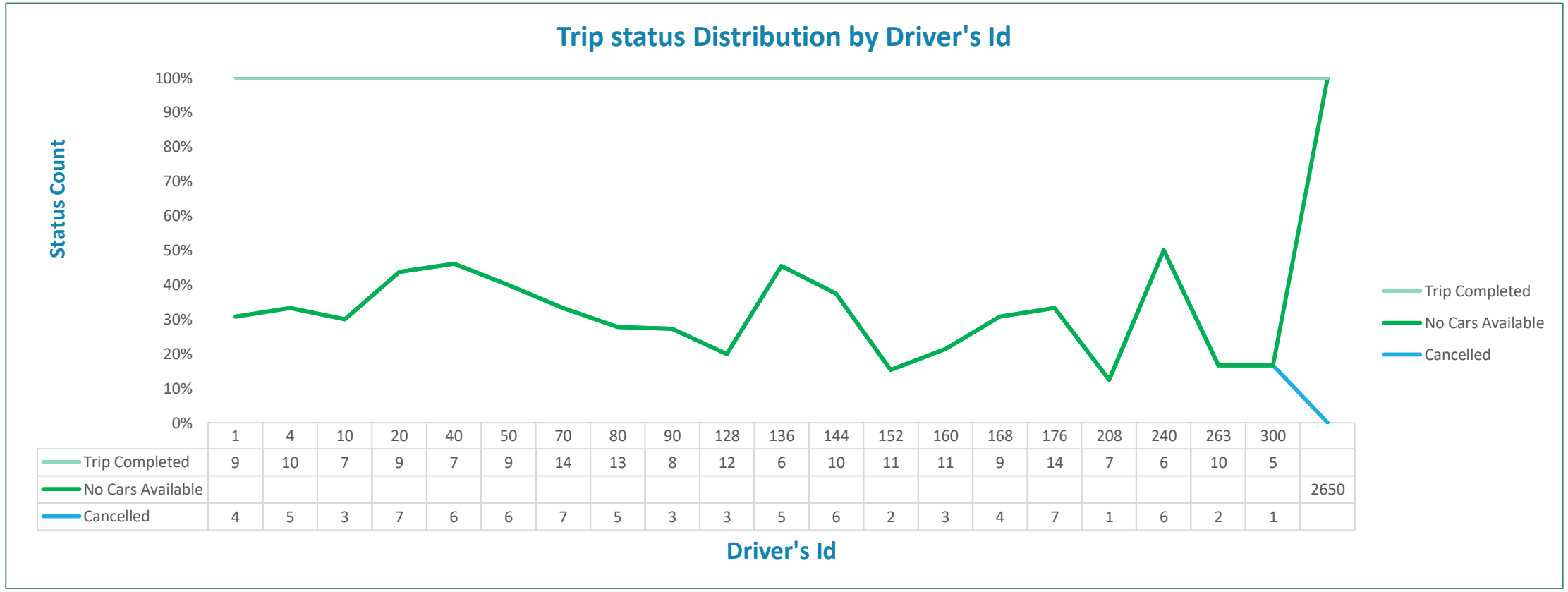


1. There are total of 300 ids of drivers. Among which I have taken random ids.

2. Among the driver's id, the most number of "Trips Completed" by the id no. 22 is 16 and the trips cancelled by this id are 5.

3. The lowest number of trips are done by the id no. 117 which is only 4 completed trips.

4. Most trips are cancelled by the Id no. 84 which is 12 requests.



	1	4	10	20	40	50	70	80	90	128	136	144	152	160	168	176	208	240	263	300	
Trip Completed	9	10	7	9	7	9	14	13	8	12	6	10	11	11	9	14	7	6	10	5	
No Cars Available																					2650
Cancelled	4	5	3	7	6	6	7	5	3	3	5	6	2	3	4	7	1	6	2	1	

MYSQL WORKBENCH

- Created the Database – ‘uber_supply_demand_gap’.
- Imported the dataset through Import wizard in CSV (Comma Separated Values) format.
- Altered the name of the table from `uber request cleaned data - copy` to ` uber_cleaned_data`.

1. Found ‘Distribution of request by pickup points’:

- City – 3507
- Airport – 3238

2. Highest Request Hour:

- 18th Hour is highest hour with most requests followed by 20th and 19th hour.

3. Driver id with most requests:

- Driver's id with most requests is 27 followed by 84, 177, and 176.

4. Request Hour vs. Unfulfilled requests by Status (cancelled/no cars available):

- Hour with most unfulfilled requests turned out to be 8th hour followed by 5,9, and 7 e.g. The morning hours of 8 AM, 5 AM, 9 AM and 7 AM.

5. Distribution of cancelled requests by pickup point:

- Cancelled trips at City – 1066
- Cancelled trips at Airport – 198

6. Busiest Week days:

- Busiest days with most requests are Friday, Monday and Thursday.

GOOGLE COLAB (PYTHON) Exploratory Data Analysis

1. Imported cleaned dataset using pandas and checked rows, columns, and datatypes.
2. Dropped irrelevant columns like *'Drop timestamp*' and *'Driver id*' to focus only on request details.
3. Displayed dataset size:
   - Rows: 6745
   - Columns: 6
4. Checked descriptive statistics: mean, std, min, and max for numeric columns.
5. Plotted visualizations using matplotlib and seaborn:
   - Trips by Day of the Week – to identify peak days
   - Request Status Distribution – to see how many were completed, cancelled, or failed.
   - Pickup Point vs Status – to compare supply issues at City vs Airport.

Thank You