UBER SUPPLY - DEMAND GAP ANALYSIS

- There is a clear and consistent supply-demand gap in Uber ride services, during the night (12 AM–6 AM) and early morning hours, with the Airport being the most affected pickup point. This gap is due to driver unavailability and high cancellation rates
- Solving these time and location specific gaps through changes and solutions will lead to better rider experience, improved driver engagement, and a more efficient and profitable system for Uber.

EXCEL ANALYSIS:

In the following figures we can see that

- Requests of Rides per Hour (1.1)
- Status Distribution (1.2)
- Share of Total Requests by Pickup Points (1.3)
- Requests by Hours and Pickup Points (1.4)
- Status of Requests by Pickup Points (1.5)

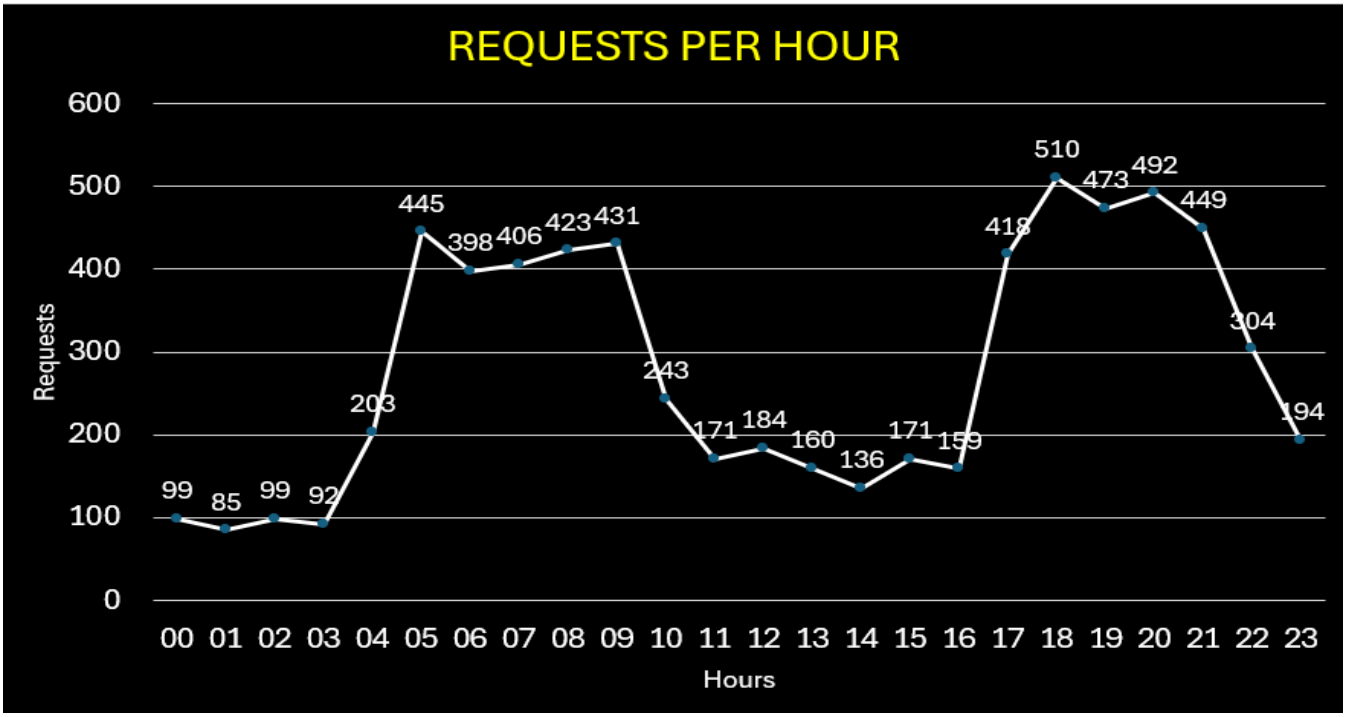


Fig 1.1 – This Figure shows the Number of Requests made per Hour

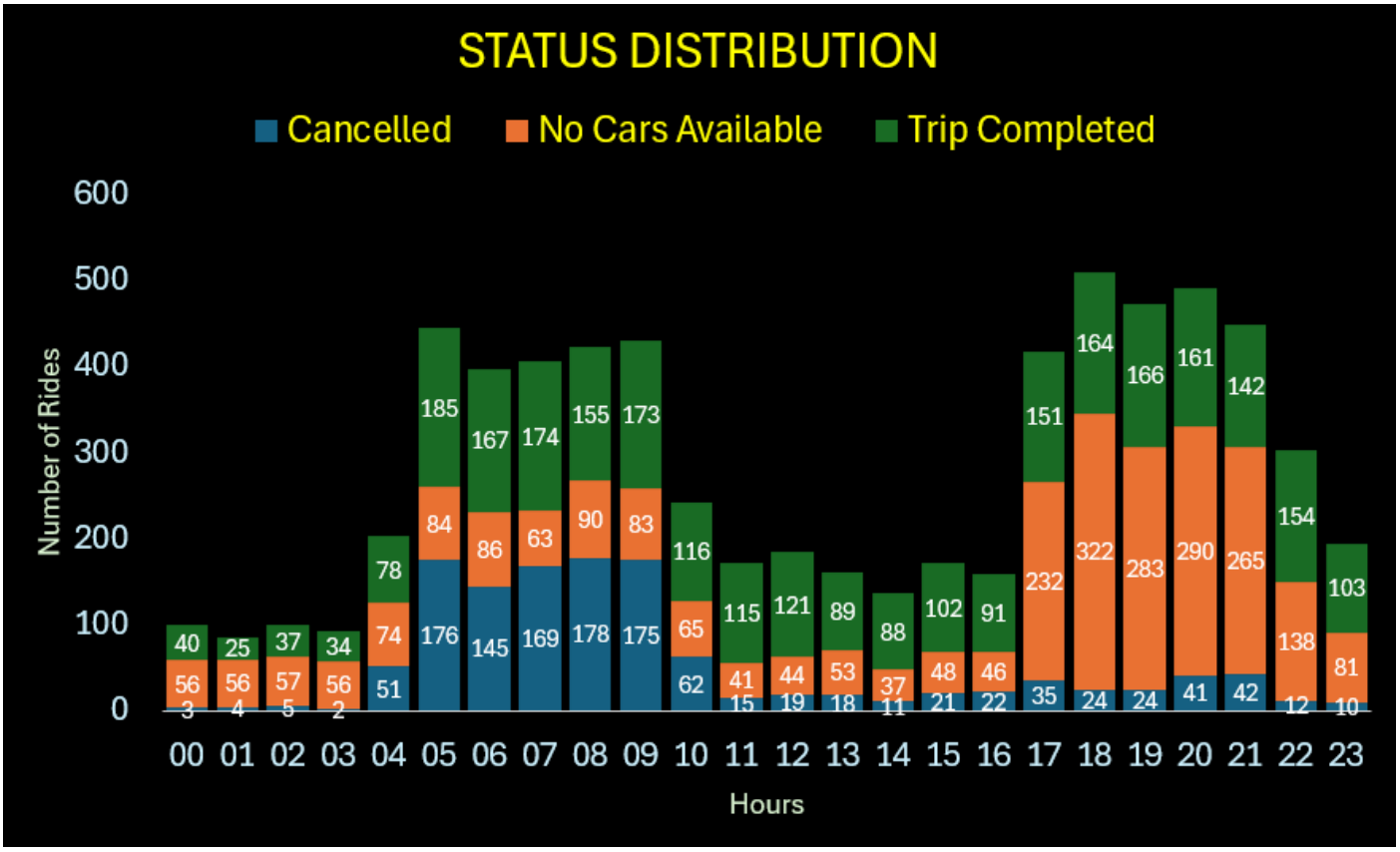


Fig 1.2 – This Figure shows the distribution of Request status per Hour

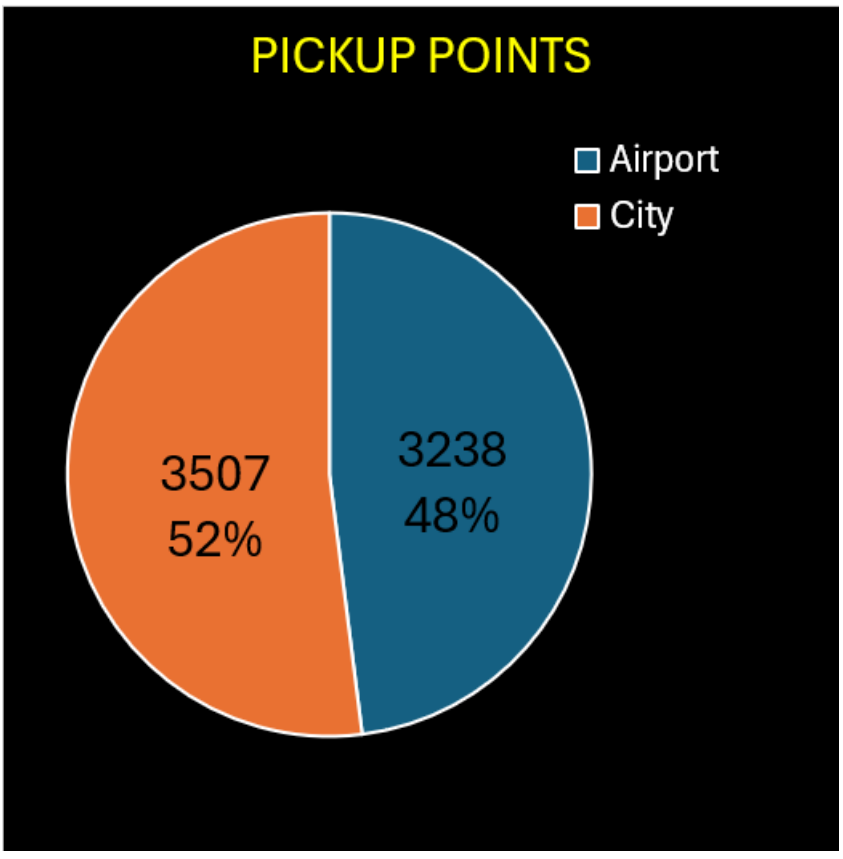


Fig 1.3 – This Figure shows the Requests Share of Pickup Points

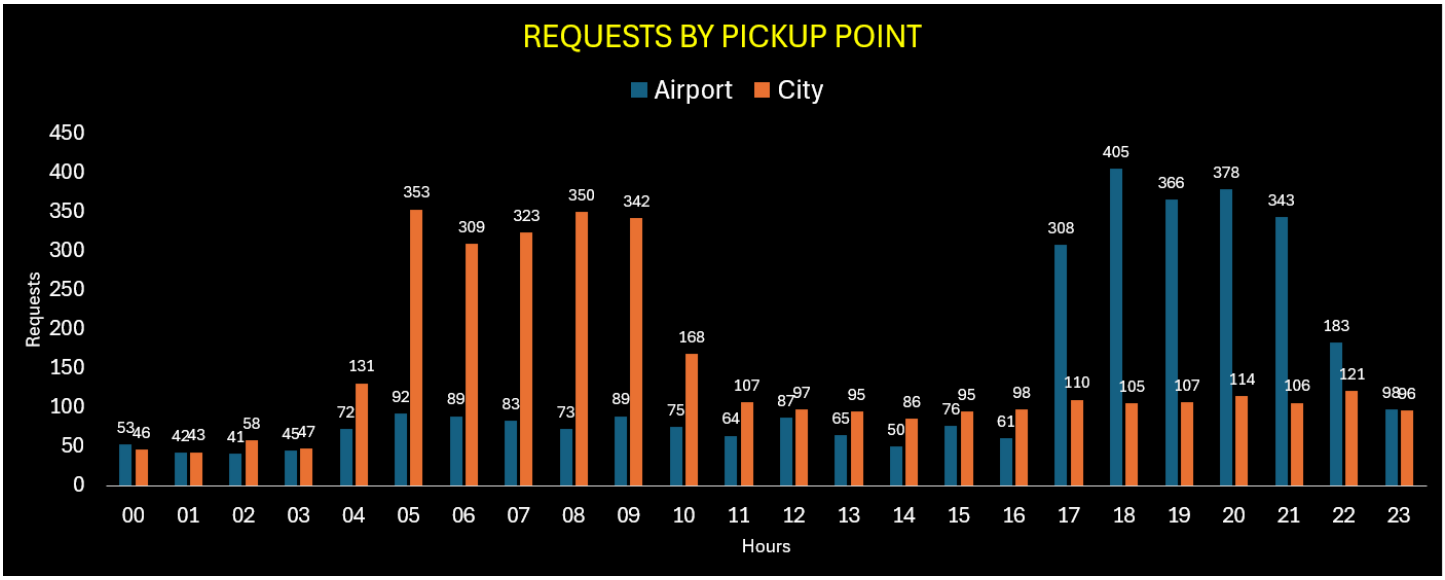


Fig 1.4 – This Figure shows the Requests by Pickup point per Hour

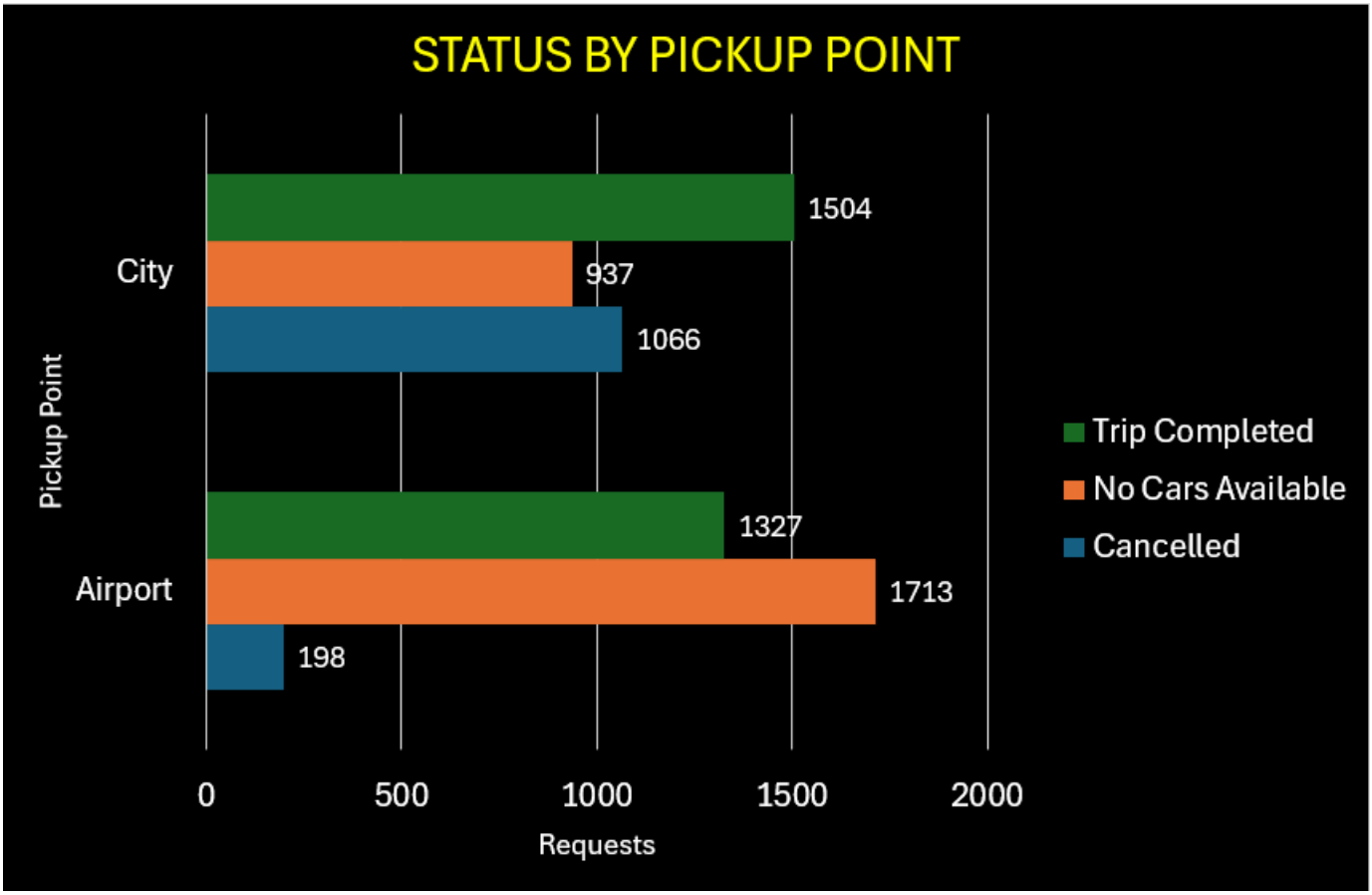


Fig 1.5 – This Figure Shows the Requests Status by Pickup points

SQL ANALYSIS:

1. SHOWING ALL DATA

```
select * from uberdata order by Request_timestamp;
```

Request_id	Pickup_point	Driver_id	Status	Request_timestamp	Drop_timestamp
1363	Airport	136	Trip Completed	2016-07-11 00:00:00	2016-07-11 01:11:00
1362	City	NULL	No Cars Available	2016-07-11 00:02:00	NULL
1367	City	155	Trip Completed	2016-07-11 00:05:00	2016-07-11 01:06:00
1364	City	NULL	No Cars Available	2016-07-11 00:06:00	NULL
1365	City	229	Cancelled	2016-07-11 00:09:00	NULL
1366	City	NULL	No Cars Available	2016-07-11 00:09:00	NULL
1	Airport	285	Trip Completed	2016-07-11 00:20:00	2016-07-11 00:51:00
2	Airport	NULL	No Cars Available	2016-07-11 00:23:00	NULL
3	Airport	80	Trip Completed	2016-07-11 00:24:00	2016-07-11 01:31:00
7	Airport	NULL	No Cars Available	2016-07-11 00:30:00	NULL
5	Airport	264	Trip Completed	2016-07-11 00:36:00	2016-07-11 01:35:00
6	City	NULL	No Cars Available	2016-07-11 00:36:00	NULL
4	City	NULL	No Cars Available	2016-07-11 00:37:00	NULL
8	City	NULL	No Cars Available	2016-07-11 00:40:00	NULL
9	Airport	235	Trip Completed	2016-07-11 00:45:00	2016-07-11 02:00:00
10	City	228	Trip Completed	2016-07-11 00:54:00	2016-07-11 01:59:00
11	City	198	Trip Completed	2016-07-11 01:00:00	2016-07-11 01:53:00
12	City	NULL	No Cars Available	2016-07-11 01:08:00	NULL
13	City	119	Trip Completed	2016-07-11 01:08:00	2016-07-11 01:58:00
16	Airport	NULL	No Cars Available	2016-07-11 01:09:00	NULL
14	Airport	NULL	No Cars Available	2016-07-11 01:10:00	NULL
19	City	NULL	No Cars Available	2016-07-11 01:14:00	NULL
15	City	NULL	No Cars Available	2016-07-11 01:15:00	NULL
17	Airport	NULL	No Cars Available	2016-07-11 01:16:00	NULL

2. TOTAL NUMBER OF REQUESTS

```
select count(Request_id) as "Total Number of Cab
Requests" from uberdata;
```

Total Number of Cab Requests
6745

3. REQUESTS BY STATUS

```
select Status,count(Request_id) as "Cab Requests"
from uberdata group by Status;
```

Status	Cab Requests
Trip Completed	2831
Cancelled	1264
No Cars Available	2650

4. PERCENTAGE OF RIDES BY STATUS

```
SELECT Status,COUNT(*) AS
total_requests,ROUND(COUNT(*) * 100.0 / (SELECT
COUNT(*) FROM uberdata), 2) AS percentage
FROM uberdata
GROUP BY Status;
```

Status	total_requests	percentage
Trip Completed	2831	41.97
Cancelled	1264	18.74
No Cars Available	2650	39.29

5. CAB REQUESTS BY PICKUP POINT

```
select Pickup_point, count(Request_id) as "Cab Requests"
from uberdata
group by Pickup_point;
```

Pickup_point	Cab Requests
Airport	3238
City	3507

6. PERCENTAGE OF CAB REQUESTS BY PICKUP POINT

```
SELECT Pickup_point, COUNT(*) AS Total_Requests,
COUNT(*) * 100.0 / (SELECT COUNT(*) FROM uberdata) AS
Percentage
FROM uberdata
GROUP BY Pickup_point;
```

	Pickup_point	Total_Requests	Percentage
▶	Airport	3238	48.00593
	City	3507	51.99407

7. DRIVER PERFORMANCE

```
select Driver_id, count(*) as Number_of_Rides from
uberdata
where Driver_id is not null
group by Driver_id
order by Number_of_Rides desc;
```

	Driver_id	Number_of_Rides
▶	27	22
	22	21
	70	21
	84	21
	176	21
	177	21
	24	20
	69	20
	114	20
	142	20
	197	20
	9	19
	16	19
	54	19
	78	19
	107	19
	212	19
	37	18
	80	18
	98	18
	105	18
	109	18
	125	18

8. REQUESTS BY HOURS

```
select hour(Request_timestamp)as Hours,count(*) as
Requests from uberdata
```

```
group by Hours
order by Hours;
```

Hours	Requests
0	99
1	85
2	99
3	92
4	203
5	445
6	398
7	406
8	423
9	431
10	243
11	171
12	184
13	160
14	136
15	171
16	159
17	418
18	510
19	473
20	492
21	449
22	304
23	194

9. BUSIEST HOURS

```
select hour(Request_timestamp)as Hours,count(*) as
Requests from uberdata
group by Hours
order by Requests desc;
```

	Hours	Requests
▶	18	510
	20	492
	19	473
	21	449
	5	445
	9	431
	8	423
	17	418
	7	406
	6	398
	22	304
	10	243
	4	203
	23	194
	12	184
	11	171
	15	171
	13	160
	16	159
	14	136
	2	99
	0	99
	3	92
	1	85

10. TIME GAP

```
SELECTRequest_timestamp,Drop_timestamp,ROUND(TIMES
TAMPDIFF(SECOND, Request_timestamp, Drop_timestamp)
/ 3600, 2) AS Ride_Time
FROM uberdata
WHERE Drop_timestamp IS NOT NULL AND
Request_timestamp IS NOT NULL
ORDER BY Request_timestamp;
```

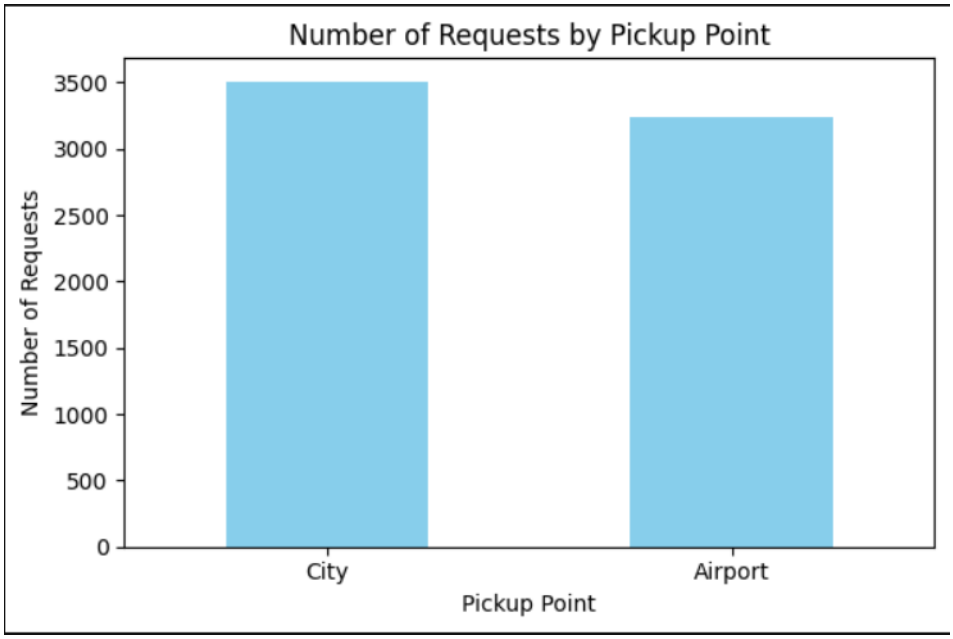
Request_timestamp	Drop_timestamp	Ride_Time
2016-07-11 00:00:00	2016-07-11 01:11:00	1.18
2016-07-11 00:05:00	2016-07-11 01:06:00	1.02
2016-07-11 00:20:00	2016-07-11 00:51:00	0.52
2016-07-11 00:24:00	2016-07-11 01:31:00	1.12
2016-07-11 00:36:00	2016-07-11 01:35:00	0.98
2016-07-11 00:45:00	2016-07-11 02:00:00	1.25
2016-07-11 00:54:00	2016-07-11 01:59:00	1.08
2016-07-11 01:00:00	2016-07-11 01:53:00	0.88
2016-07-11 01:08:00	2016-07-11 01:58:00	0.83
2016-07-11 01:29:00	2016-07-11 02:05:00	0.60
2016-07-11 02:04:00	2016-07-11 02:37:00	0.55
2016-07-11 02:17:00	2016-07-11 03:31:00	1.23
2016-07-11 02:20:00	2016-07-11 03:34:00	1.23
2016-07-11 02:36:00	2016-07-11 03:03:00	0.45
2016-07-11 02:47:00	2016-07-11 03:23:00	0.60
2016-07-11 03:14:00	2016-07-11 04:19:00	1.08
2016-07-11 03:17:00	2016-07-11 04:09:00	0.87
2016-07-11 03:22:00	2016-07-11 04:35:00	1.22
2016-07-11 03:22:00	2016-07-11 03:57:00	0.58
2016-07-11 03:30:00	2016-07-11 04:22:00	0.87

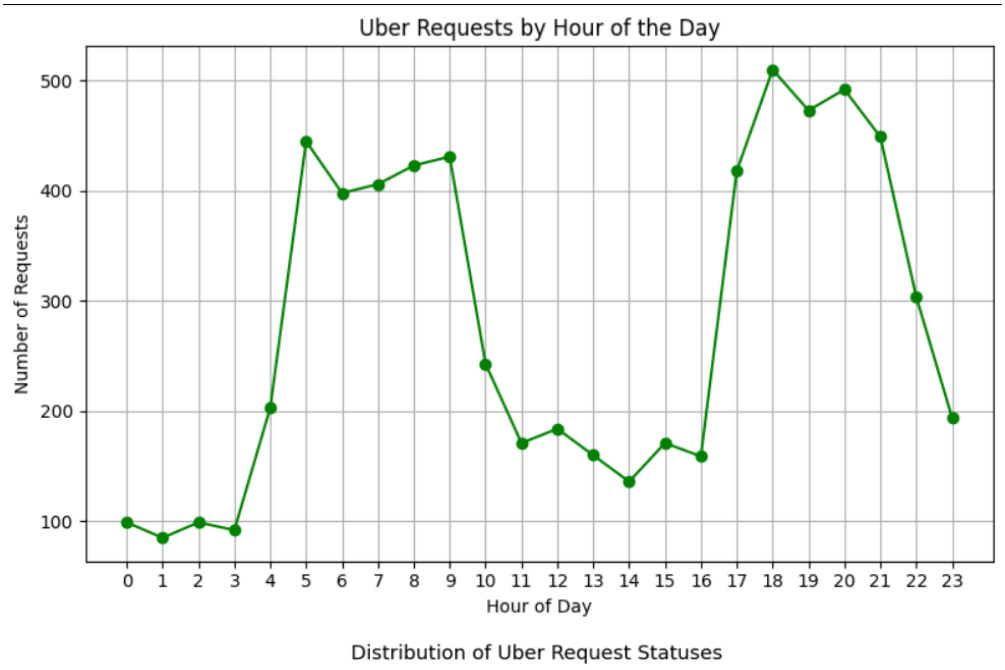
11. AVERAGE RIDE TIME

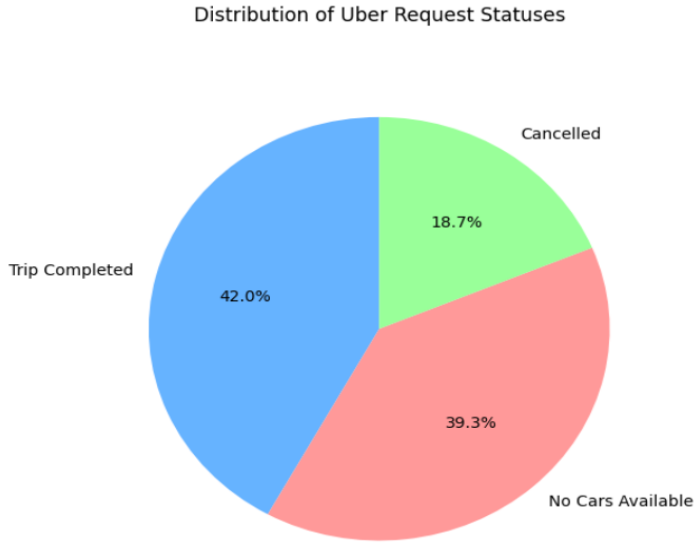
```
SELECT ROUND(AVG(TIMESTAMPDIFF(SECOND,
Request_timestamp, Drop_timestamp) / 3600), 2) AS
Average_Ride_Time
FROM uberdata
WHERE Drop_timestamp IS NOT NULL AND
Request_timestamp IS NOT NULL;
```

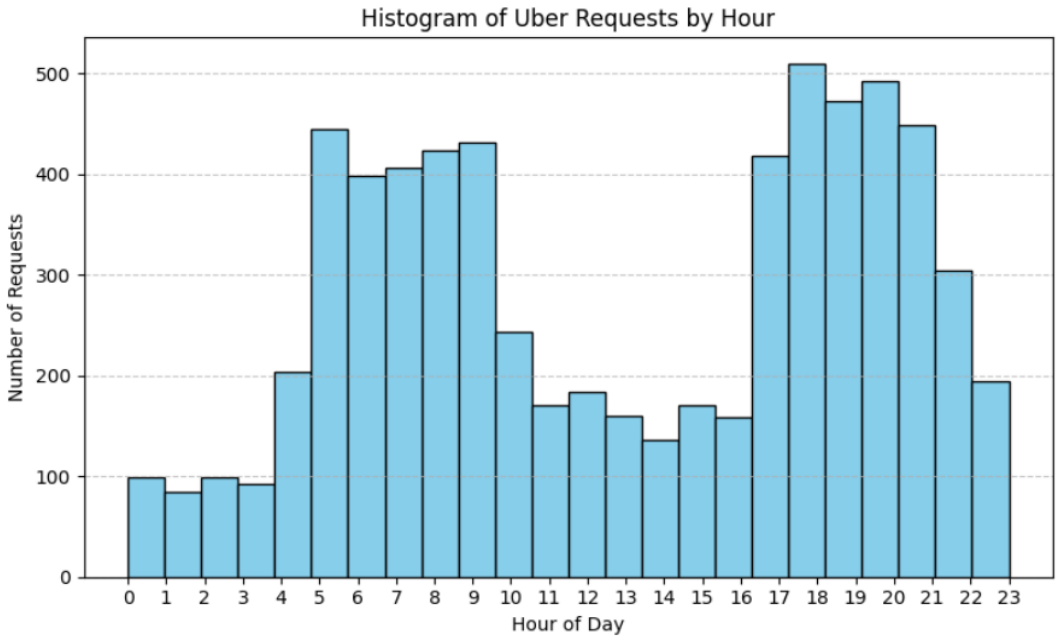
Average_Ride_Time
0.87

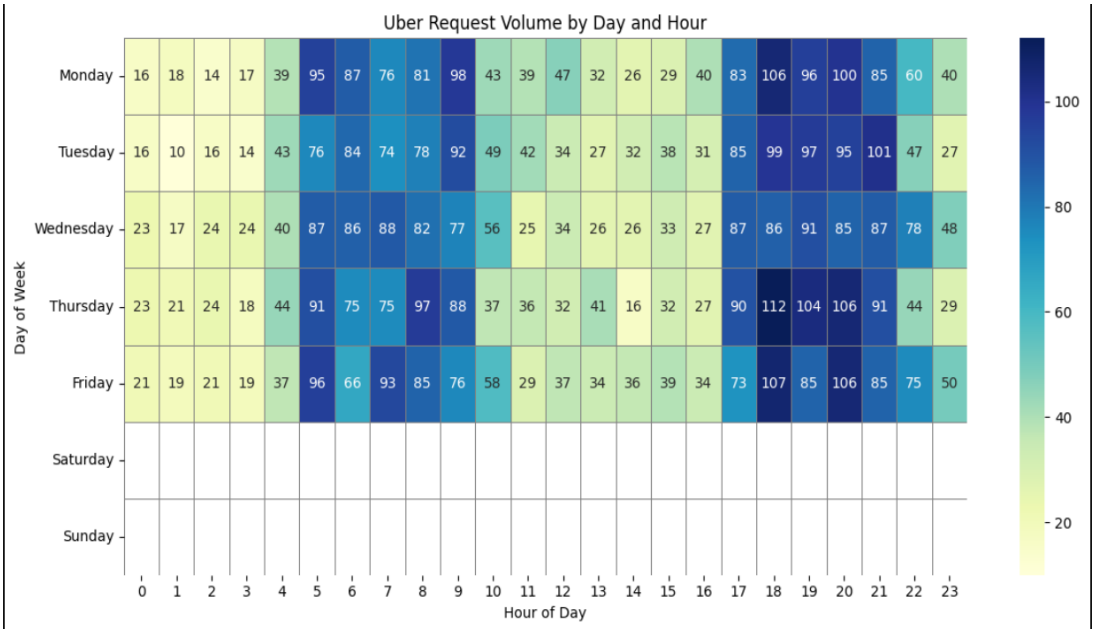
EDA:

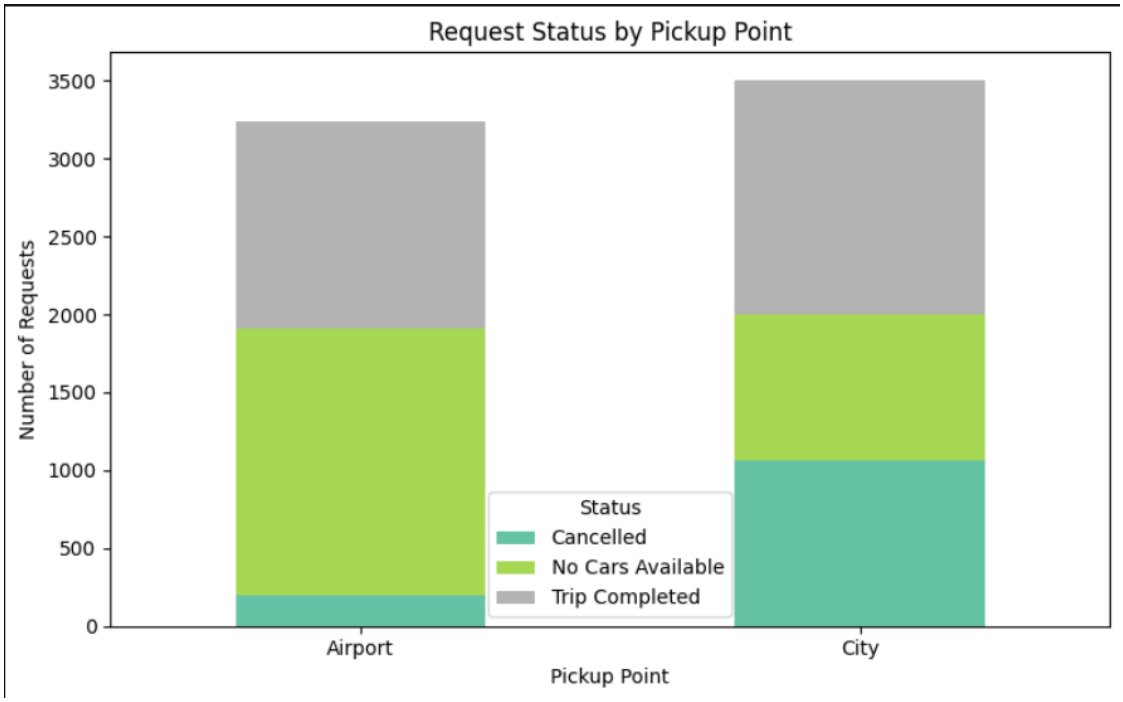

Number of Requests by Pickup Point
Number of Requests
3500
3000
2500
2000
1500
1000
500
0
City
Airport
Pickup Point

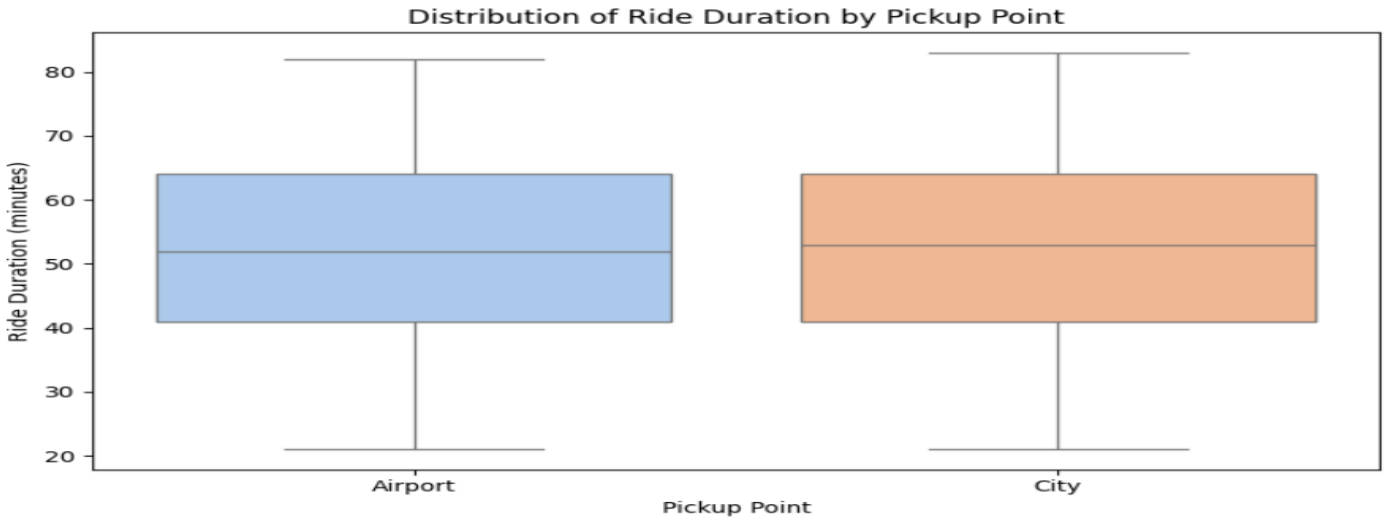

Uber Requests by Hour of the Day
Number of Requests
500
400
300
200
100
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Hour of Day

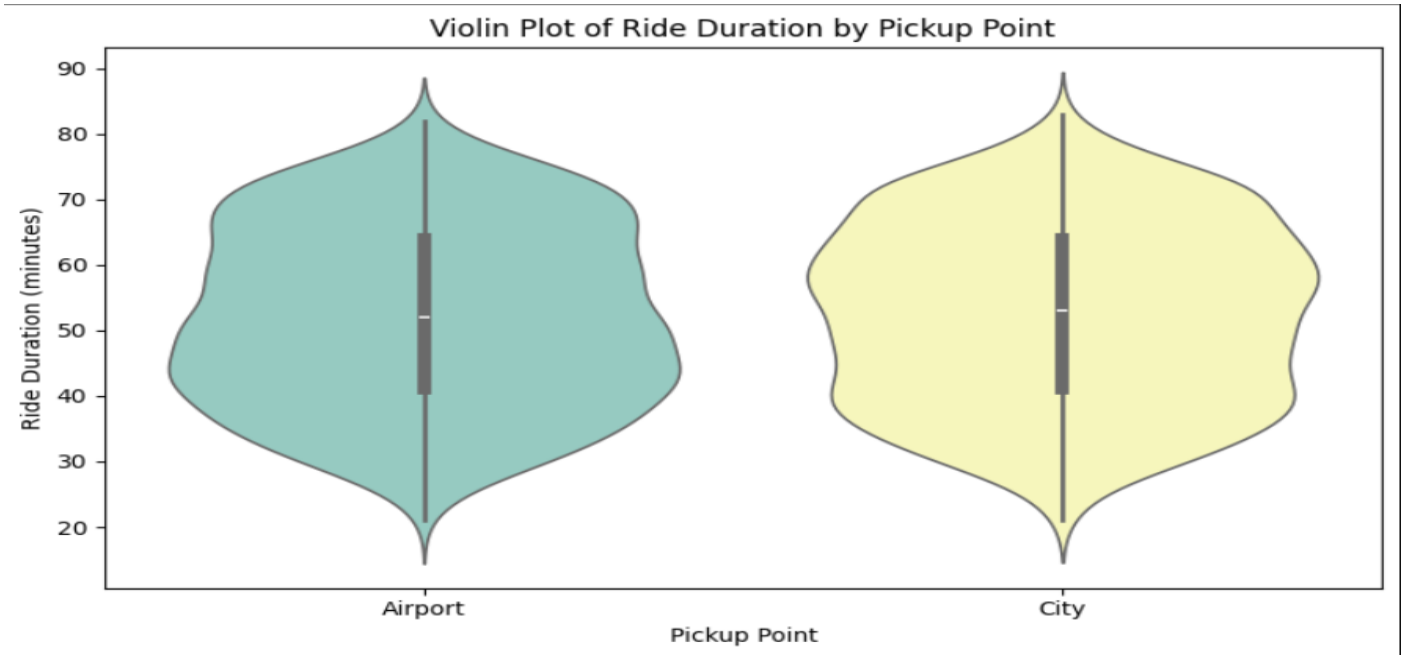

Distribution of Uber Request Statuses
Cancelled
18.7%
Trip Completed
42.0%
39.3%
No Cars Available

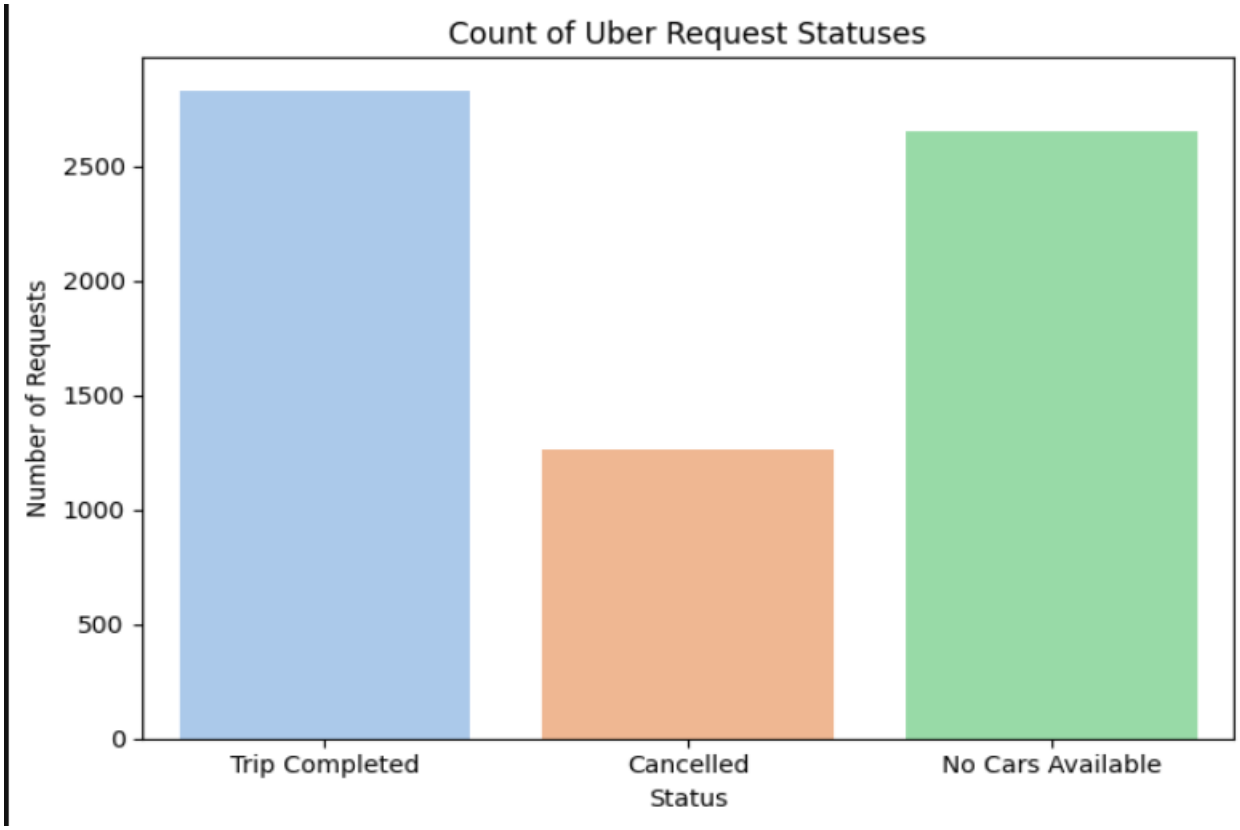

Histogram of Uber Requests by Hour
Number of Requests
Hour of Day
0
100
200
300
400
500
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23

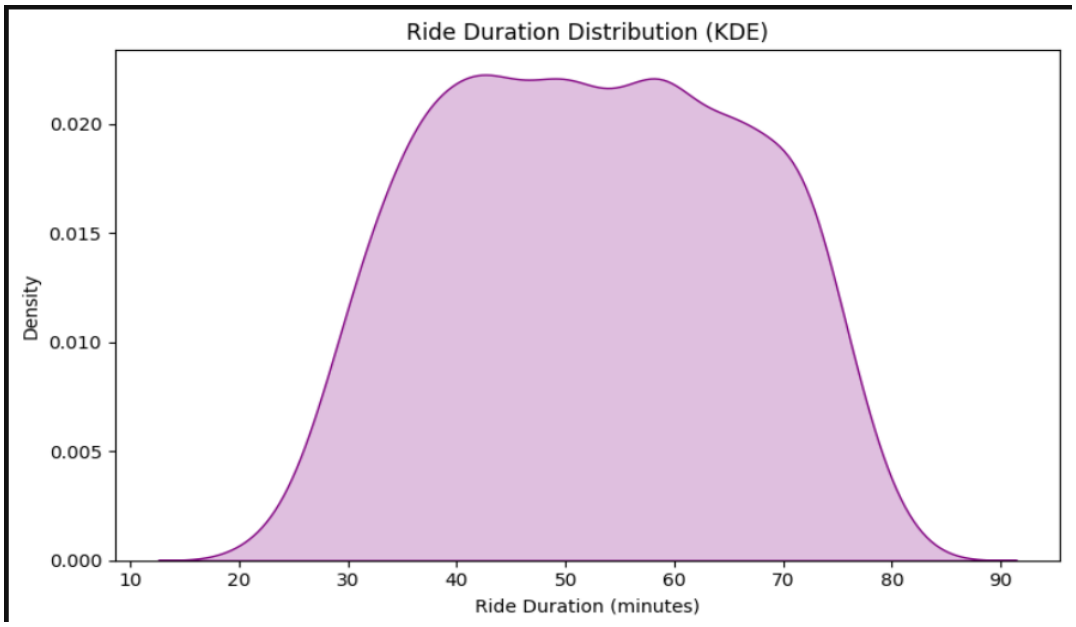

Uber Request Volume by Day and Hour
Day of Week
Hour of Day
Monday 16 18 14 17 39 95 87 76 81 98 43 39 47 32 26 29 40 83 106 96 100 85 60 40
Tuesday 16 10 16 14 43 76 84 74 78 92 49 42 34 27 32 38 31 85 99 97 95 101 47 27
Wednesday 23 17 24 24 40 87 86 88 82 77 56 25 34 26 26 33 27 87 86 91 85 87 78 48
Thursday 23 21 24 18 44 91 75 75 97 88 37 36 32 41 16 32 27 90 112 104 106 91 44 29
Friday 21 19 21 19 37 96 66 93 85 76 58 29 37 34 36 39 34 73 107 85 106 85 75 50
Saturday
Sunday
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
20
40
60
80
100

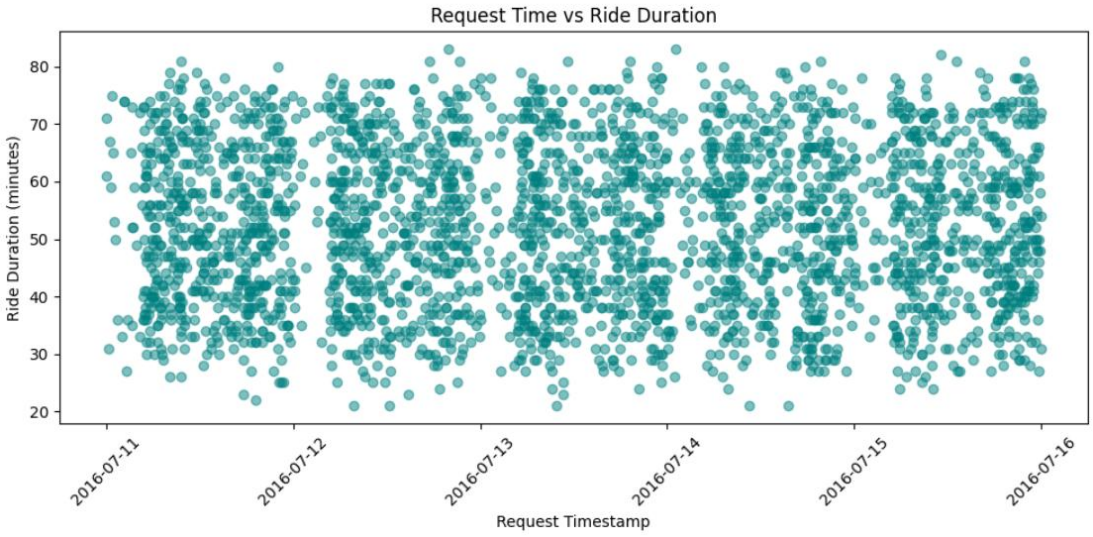

Request Status by Pickup Point
Number of Requests
Pickup Point
Airport
City
Status
Cancelled
No Cars Available
Trip Completed
0
500
1000
1500
2000
2500
3000
3500


Distribution of Ride Duration by Pickup Point
Ride Duration (minutes)
80
70
60
50
40
30
20
Airport
City
Pickup Point

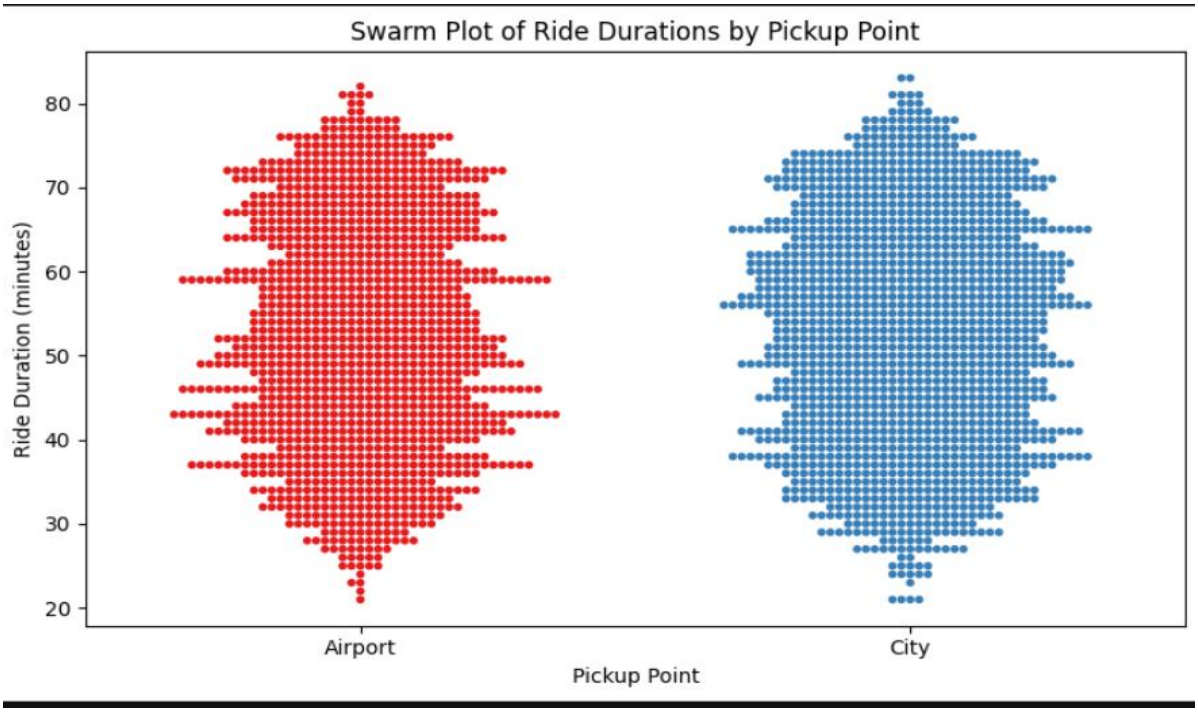

Violin Plot of Ride Duration by Pickup Point
Ride Duration (minutes)
90
80
70
60
50
40
30
20
Airport
City
Pickup Point

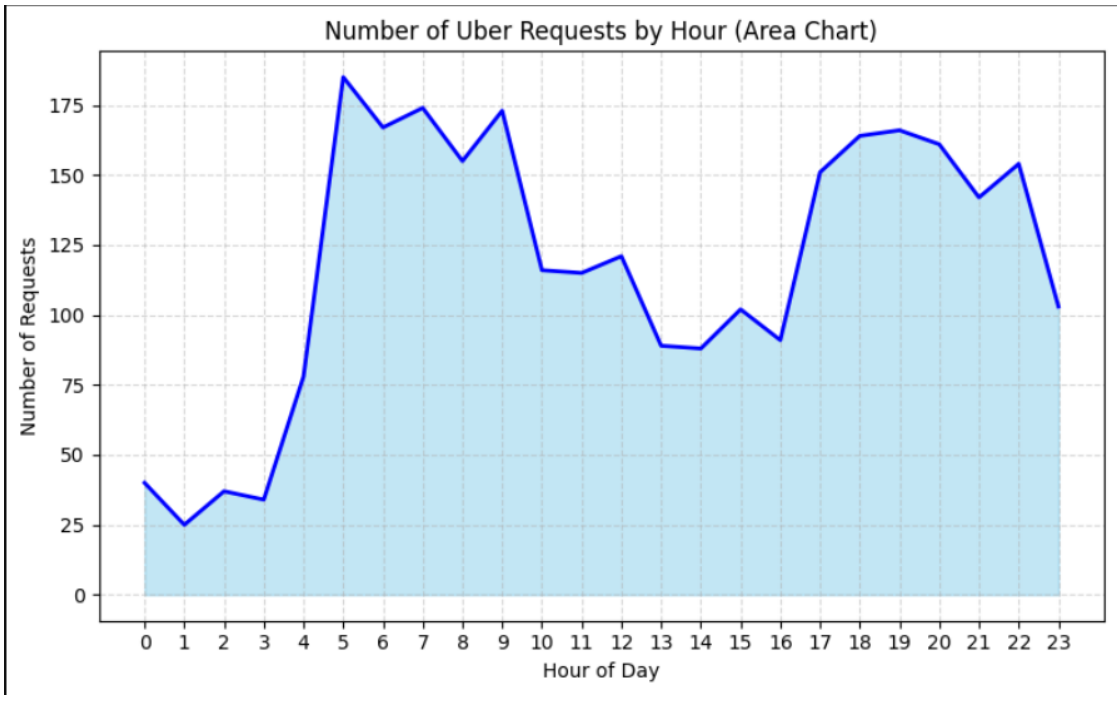

Count of Uber Request Statuses
Number of Requests
2500
2000
1500
1000
500
0
Trip Completed
Cancelled
No Cars Available
Status

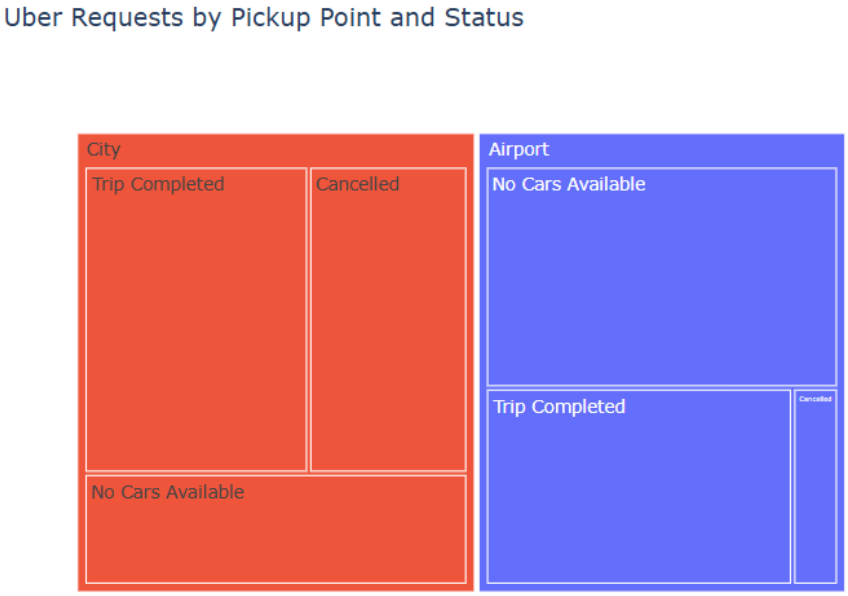

Ride Duration Distribution (KDE)
Density
0.020
0.015
0.010
0.005
0.000
10
20
30
40
50
60
70
80
90
Ride Duration (minutes)

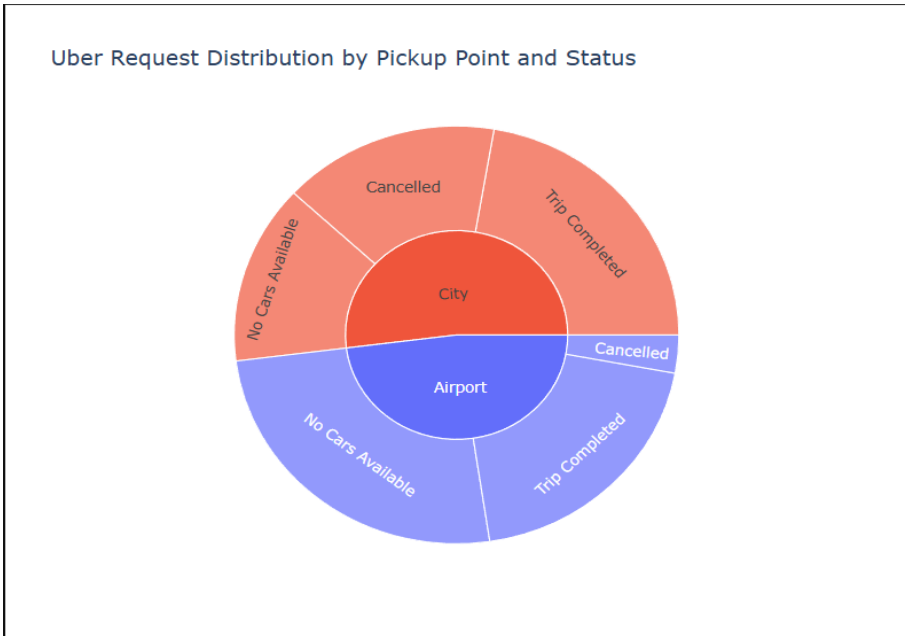

Request Time vs Ride Duration
Ride Duration (minutes)
80
70
60
50
40
30
20
2016-07-11
2016-07-12
2016-07-13
2016-07-14
2016-07-15
2016-07-16
Request Timestamp

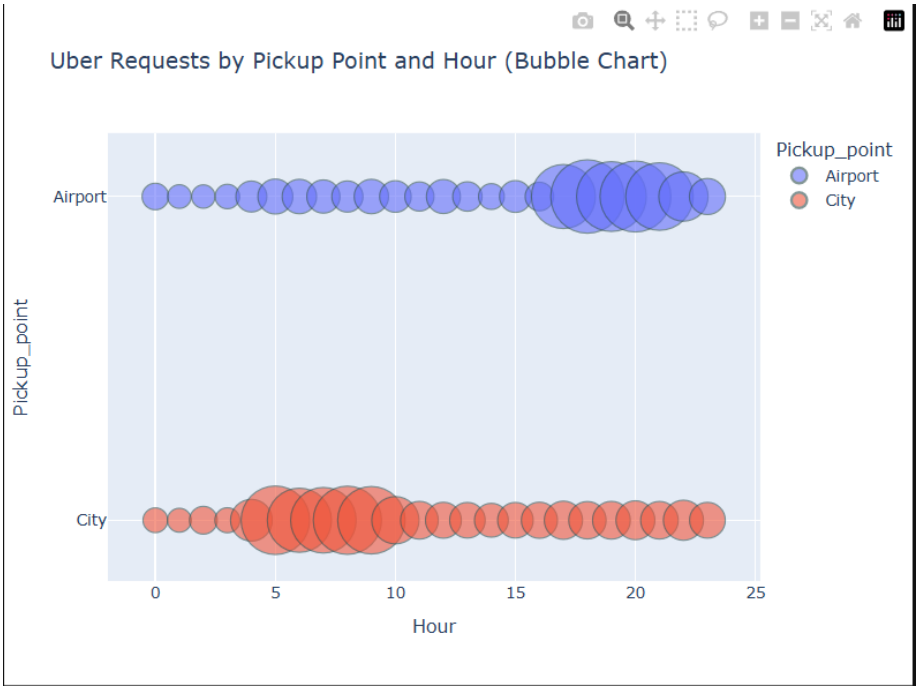

Strip Plot of Ride Durations by Pickup Point
Ride Duration (minutes)
80
70
60
50
40
30
20
Airport
City
Pickup Point

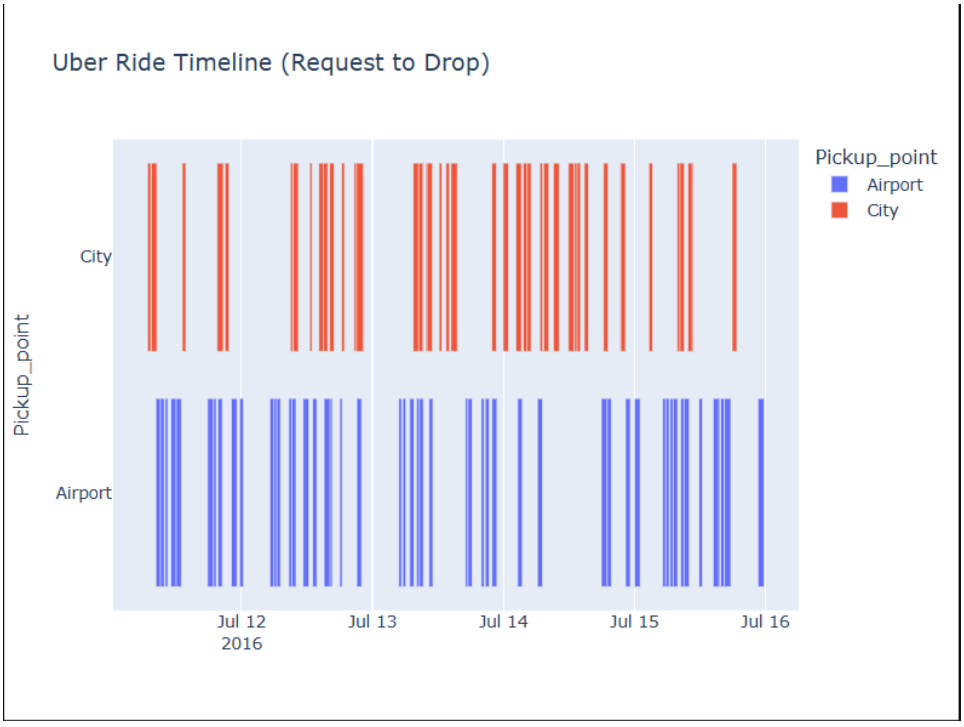

Swarm Plot of Ride Durations by Pickup Point
Ride Duration (minutes)
80
70
60
50
40
30
20
Airport
City
Pickup Point

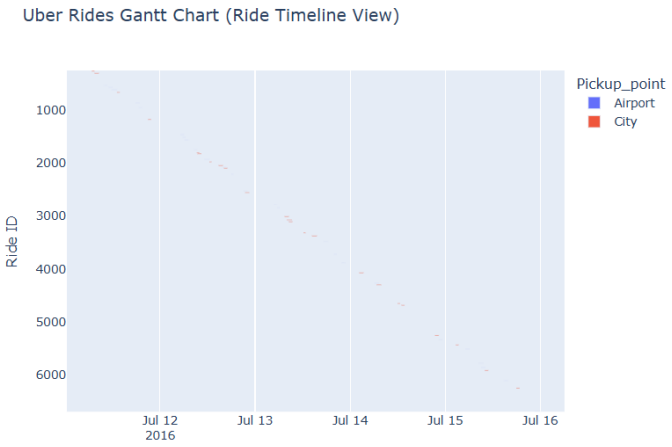

Number of Uber Requests by Hour (Area Chart)
Number of Requests
175
150
125
100
75
50
25
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Hour of Day

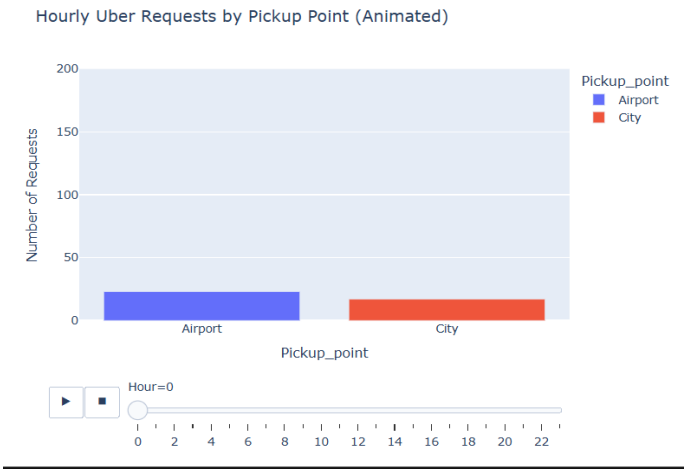

Uber Requests by Pickup Point and Status
City
Trip Completed
Cancelled
No Cars Available
Airport
No Cars Available
Trip Completed
Cancelled


Uber Request Distribution by Pickup Point and Status
Cancelled
Trip Completed
No Cars Available
City
Cancelled
Airport
No Cars Available
Trip Completed


Uber Requests by Pickup Point and Hour (Bubble Chart)
Pickup_point
Airport
City
Airport
City
Pickup_point
0
5
10
15
20
25
Hour


Uber Ride Timeline (Request to Drop)
Pickup_point
Airport
City
City
Airport
Pickup_point
Jul 12
2016
Jul 13
Jul 14
Jul 15
Jul 16





SOLUTIONS:

Therefore by Analyzing everything here are some of the solutions to improve the services, quality, efficiency and profit of the Uber company

1. **Incentivize Drivers During Critical Time Slots**

-> Offer better pricing or bonuses for trips between 12 AM and 6 AM

-> Analysis shows low driver availability and high cancellation rates in this time window

-> Add "Night Shift Incentive" for active drivers between midnight and 6 AM

2. **Rebalance Supply from City to Airport at Night**

-> Encourage city-based drivers to move toward the airport at night

-> Identified a lack of available cars at the airport at night

-> Notify idle city drivers when airport demand is high

-> Offer guaranteed ride minimums or bonus payouts for night airport pickups

3. **Implement Smart Shift**

-> Introduce time-slot-based driver scheduling, focusing on early morning (4–8 AM) and night shifts

-> Cancellations are highest during early morning, possibly due to driver fatigue or low incentives

-> Use driver data to assign smarter shifts with rest gaps

4. **App-Level Improvements (Driver Side)**

-> Detect and discourage frequent cancellation behavior

-> Frequent cancellations hurt customer experience and lead to demand drop

-> Track driver cancellation patterns

-> Add a cancellation penalty or nudge popup ("Are you sure you want to cancel this ride?")