Uber Supply Demand Gap Analysis Insights Using SQL

GitHub Link:

Uber Supply Demand Gap Analysis Insights

Introduction:

With the EDA and data cleaning phases completed, this task is about getting useful business insights using SQL. The cleaned data has been loaded into a MySQL structured table, allowing us to perform precise query-based analysis on fundamental operating metrics such as demand-supply gap, trip completion ratio, driver behavior, and time-based patterns.
Through writing and performing targeted SQL queries, we are trying to find trends and difficulties in the Uber ride request system and ultimately make data-driven suggestions towards better service availability and efficiency.

Data Analysis & Insights:

1. Overall Demand vs Supply

Question: How many ride requests were successfully served during the study period?

Answer:

```
SELECT
    COUNT(*) AS Total_requests,
    SUM(Status = 'Trip Completed') AS completed,
    ROUND(SUM(Status = 'Trip Completed') / COUNT(*), 2) AS completion_rate
FROM uber_requests;
```

	Total_requests	completed	completion_rate
▶	6745	2831	0.42

Insights:

From the analysis, we can see that 42% of requests are completed and the rest failed. This indicates an acute shortage in supply and provides the starting point for every subsequent dive.

2. Completion Rate by Pickup Location

Question: Does the success rate differ between the city and the airport?

Answer:

```sql
SELECT
    Pickup_point,
    SUM(Status = 'Trip Completed') AS Completed,
    ROUND(SUM(Status = 'Trip Completed') / COUNT(*), 2) AS Completion_Rate
FROM uber_requests
GROUP BY Pickup_point;
```

Pickup_point	Completed	Completion_Rate
Airport	1327	0.41
City	1504	0.43

Insights:

Completion rate at both points is low, but airport rides are slightly less successful (41%) than city rides (43%). The airport thus requires proportionally more cars or more intelligent rebalancing.

3. Hour-By-Hour Supply Gap

Question: At what hours does unmet demand spike?

Answer:

```sql
SELECT
    Request_hour,
    ROUND(SUM(Gap_Flag = 'Demand Unfulfilled') / COUNT(*), 2) AS Gap_Ratio
FROM uber_requests
GROUP BY Request_hour
ORDER BY Gap_Ratio DESC;
```

	Request_hour	Gap_Ratio
▶	1	0.71
	18	0.68
	21	0.68
	20	0.67
	19	0.65
	17	0.64
	2	0.63
	3	0.63
	8	0.63
	4	0.62

	0	0.60
	9	0.60
	5	0.58
	6	0.58
	7	0.57
	10	0.52
	22	0.49
	23	0.47
	13	0.44
	16	0.43
	15	0.40
	14	0.35
	12	0.34
	11	0.33

Insights:

The worst times are nighttime departures at 01:00 and return commuter peak 6:00pm - 9:00pm, during which there is zero opportunity for 70% of passengers to get a car. These areas are best to target driver incentives.

4. Time-of-Day Slot Performance

Question: Which time of day slot suffers the most?

Answer:

```
SELECT
    Time_of_day_slot,
    ROUND(SUM(Status = 'Trip Completed') / COUNT(*), 2) AS Completion_Rate
FROM uber_requests
GROUP BY Time_of_day_slot
ORDER BY Completion_Rate DESC;
```

	Time_of_day_slot	Completion_Rate
▶	Late Morning	0.59
	Afternoon	0.59
	Night	0.42
	Early Morning	0.40
	Late Night	0.39
	Evening	0.34

Insights:

Evening peaks are in the lower end (34% completion), then late night. Increasing pricing or guaranteed earnings planning could bring the most gain.

5. Why requests fail (root cause by Location)

Question: When trips fail, is it driver cancellation or no cars?

Answer:

```
SELECT
    Pickup_point,
    SUM(Status = 'Cancelled') AS Cancelled,
    SUM(Status = 'No Cars Available') AS No_Cars,
    ROUND(SUM(Status = 'Cancelled')
          / SUM(Status IN ('Cancelled','No Cars Available')), 2) AS Cancel_Share
FROM uber_requests
WHERE Status IN ('Cancelled','No Cars Available')
GROUP BY Pickup_point;
```

	Pickup_point	Cancelled	No_Cars	Cancel_Share
▶	City	1066	937	0.53
	Airport	198	1713	0.10

Insights:

Airport breakdowns are "no cars available," but in the city more than half are driver cancellations. Airport solutions are thus a matter of just vehicle count, but city solutions must be based on decreasing cancellations.

6. Peak Demand Hour For Each Location

Question: When does each pickup point experience its single busiest hour?

Answer:

```
SELECT
    Pickup_point,
    Request_hour,
    Request_Count
FROM (
    SELECT
        Pickup_point,
        Request_hour,
        COUNT(*) AS Request_Count,
        ROW_NUMBER() OVER (PARTITION BY Pickup_point ORDER BY COUNT(*) DESC) AS rn
    FROM uber_requests
    GROUP BY Pickup_point, Request_hour
) AS ranked
WHERE rn = 1;
```

	Pickup_point	Request_hour	Request_Count
▶	Airport	18	405
	City	5	353

Insights:

Airport demand is high at 18:00 (after-work flights), and city demand gathers at 05:00 to 09:00 (morning rush). Adjusting driver shifts during these peaks can smooth out the gap.

7. Average Trip Duration By Pickup & Slot

Question: Does traffic or anything else make some trips inherently longer?

Answer:

```
SELECT
    Pickup_point,
    Time_of_day_slot,
    ROUND(AVG(Trip_duration_minutes), 1) AS Avg_Minutes
FROM uber_requests
WHERE Status = 'Trip Completed'
GROUP BY Pickup_point, Time_of_day_slot
ORDER BY Avg_Minutes DESC;
```

	Pickup_point	Time_of_day_slot	Avg_Minutes
▶	City	Late Night	53.5
	City	Early Morning	53.4
	Airport	Late Morning	53.1
	Airport	Night	52.9
	City	Afternoon	52.5
	City	Late Morning	52.4
	City	Night	52.2
	Airport	Evening	52.1
	Airport	Early Morning	52.0
	Airport	Late Night	51.9
	Airport	Afternoon	51.5
	City	Evening	51.3

Insights:

Trip durations remain fairly fixed (52 minutes) between places and times. It's supply imbalance, not traffic, that is the problem.

8. High-performing Drivers

Question: Who completes the most trips?

Answer:

```
SELECT
    Driver_id,
    COUNT(*) AS Trips_Completed
FROM uber_requests
WHERE Status = 'Trip Completed'
GROUP BY Driver_id
ORDER BY Trips_Completed DESC
LIMIT 5;
```

Driver_id	Trips_Completed
22	16
233	15
184	15
126	14
107	14

Insights:

The best driver takes 16 trips over five days, showing that top drivers can perform better when optimally routed. Spreading their schedule could enhance overall vehicle productivity.

9. Driver Cancellation Prevalence

Question: What proportion of active drivers cancel at least once?

Answer:

```
WITH all_drivers AS (
    SELECT DISTINCT Driver_id
    FROM uber_requests
    WHERE Driver_id IS NOT NULL
),
cancellers AS (
    SELECT DISTINCT Driver_id
    FROM uber_requests
    WHERE Status = 'Cancelled'
)
SELECT
    ROUND((SELECT COUNT(*) FROM cancellers) / (SELECT COUNT(*) FROM all_drivers), 2) AS pct_cancelled;
```

pct_cancelled
0.98

Insights:

From analysis, 98% of drivers cancel one or more rides. Eliminating even a small fraction of such cancellations would make a difference to the completion rate, particularly in the city zone.

10. Weekday Performance Swing

Question: Do some weekdays perform better than others?

Answer:

```
SELECT
    Weekday,
    COUNT(*) AS Total_Requests,
    SUM(Status = 'Trip Completed') AS Completed,
    ROUND(SUM(Status = 'Trip Completed') / COUNT(*), 2) AS Completion_Rate
FROM uber_requests
GROUP BY Weekday
ORDER BY FIELD(Weekday,
               'Monday','Tuesday','Wednesday',
               'Thursday','Friday','Saturday','Sunday');
```

Weekday	Total_Requests	Completed	Completion_Rate
Monday	1367	601	0.44
Tuesday	1307	562	0.43
Wednesday	1337	577	0.43
Thursday	1353	530	0.39
Friday	1381	561	0.41

Insights:

Friday experiences the greatest demand and poorest completion (41%), with Monday being somewhat better (44%). Weekends need to be planned around this Friday peak.

11. Driver Assigned Status check

Question: Do “No Cars Available” events ever happen after a driver is assigned?

Answer:

```sql
SELECT
    Driver_status,
    Status,
    COUNT(*) AS Count
FROM uber_requests
GROUP BY Driver_status, Status;
```

Driver_status	Status	Count
Driver Assigned	Trip Completed	2831
No Driver Assigned	No Cars Available	2650
Driver Assigned	Cancelled	1264

Insights:

Each "No Cars Available" occurrence carries a Driver_status = 'No Driver Assigned' indicator. That is, demand failure happens prior to any driver acceptance, so that the gap diagnosis would need to be aimed at vehicle availability rather than driver dropouts.

12. Hourly City-vs-Airport Gap Comparison

Question: During which hours is the ride demand unfulfilled most often, and how does this differ between city and airport pickups?

Answer:

```sql
SELECT
    Request_hour,
    Pickup_point,
    ROUND(SUM(Gap_Flag = 'Demand Unfulfilled') / COUNT(*), 2) AS Gap_Ratio
FROM uber_requests
GROUP BY Request_hour, Pickup_point
ORDER BY Gap_Ratio DESC
LIMIT 10;
```

Request_hour	Pickup_point	Gap_Ratio
21	Airport	0.82
18	Airport	0.80
20	Airport	0.80
19	Airport	0.77
17	Airport	0.76
8	City	0.75
1	City	0.72
6	City	0.72
5	City	0.72
9	City	0.71

Insights:

Late-night airport runs and city evening cancellations make up the top tiers. Such information is helpful in building zone-by-hour driver incentives.

Conclusion:

This project analyzed Uber's demand and supply trends through a cleaned dataset as well as structured SQL queries. Through complete research, we found that the largest gaps are found late at night and during morning rush hours, primarily at the airport, because of very few drivers. The analysis showed that "No Cars Available" was the leading cause for unfulfilled requests, which indicates a supply-side problem and not cancellations by the users. These findings suggest the necessity for improved driver assignment, real-time scheduling, and possible incentives to address the demand–supply gap. Minimization of these gaps can go a long way in significantly enhancing service reliability, customer satisfaction, and operating efficiency for Uber.