Zuber SQL Analysis Report

Description

You're working as an analyst for Zuber, a new ride-sharing company that's launching in Chicago. Your task is to find patterns in the available information. You want to understand passenger preferences and the impact of external factors on rides.

You'll study a database, analyze data from competitors, and investigate the impact of weather on ride frequency.

A database with info on taxi rides in Chicago:

neighborhoods table: data on city neighborhoods

- *name:* name of the neighborhood
- *neighborhood_id*: neighborhood code

cabs table: data on taxis

- *cab_id:* vehicle code
- *vehicle_id:* the vehicle's technical ID
- *company_name*: the company that owns the vehicle

trips table: data on rides

- *trip_id:* ride code
- *cab_id:* code of the vehicle operating the ride
- *start_ts:* date and time of the beginning of the ride (time rounded to the hour)
- *end_ts:* date and time of the end of the ride (time rounded to the hour)
- *duration_seconds:* ride duration in seconds
- *distance_miles:* ride distance in miles
- *pickup_location_id:* pickup neighborhood code
- *dropoff_location_id:* dropoff neighborhood code

weather_records table: data on weather

- *record_id:* weather record code
- *ts:* record date and time (time rounded to the hour)
- *temperature:* temperature when the record was taken

description: brief description of weather conditions, e.g. "light rain" or "scattered clouds"

Tasks

1. For November 1-7, 2017, the most popular taxi companies were Flash Cab and Taxi Affiliation Services. Find the number of rides for these two companies and name the resulting variable *trips_amount.* Join the rides for all other companies in the group "Other." Group the data by taxi company names. Name the field with taxi company names *company*. Sort the result in descending order by *trips_amount*.
2. Retrieve the identifiers of the O'Hare and Loop neighborhoods from the *neighborhoods* table.

```sql
SELECT
    CASE
        WHEN c.company_name LIKE '%Flash%' THEN 'Flash Cab'
        WHEN c.company_name LIKE '%Taxi Affiliation Services%' THEN 'Taxi Affiliation Services'
        ELSE 'Other'
    END AS company,
    COUNT(t.trip_id) AS trips_amount
FROM
    trips t
INNER JOIN
    cabs c ON t.cab_id = c.cab_id
WHERE
    t.start_ts::date BETWEEN '2017-11-01' AND '2017-11-07'
GROUP BY
    company
ORDER BY
    trips_amount DESC;
```

Result

company	trips_amount
Other	335771
Flash Cab	64084
Taxi Affiliation Services	37583

```sql
SELECT
    neighborhood_id,
    name
FROM
    neighborhoods
WHERE
    name LIKE '%Hare%' OR name = 'Loop'
```

Result

neighborhood_id	name
50	Loop
63	O'Hare

Conclusions

- Dominant Taxi Companies: Flash Cab and Taxi Affiliation Services had the highest number of rides between November 1–7, 2017, suggesting a strong market presence during the sample period.

- Location Reference Established: Successfully identifying the neighborhood IDs for O'Hare and the Loop provides a foundation for geographic ride analysis, especially in terms of airport traffic and downtown activity.

Recommendations

- Target Competitor Strategies: Zuber should study the operational strategies of Flash Cab and Taxi Affiliation Services to understand what drives their high ride volume—considering pricing, availability, and geographic focus.

- Focus Launch Zones: Zuber may benefit from emphasizing service coverage in the Loop and O'Hare neighborhoods where ride demand is likely high.

- Weather-Based Ride Modeling: Integrate weather data further to build predictive models for ride demand. This would allow Zuber to adjust driver availability based on weather forecasts.