HNG_Ride

SQL Business Analysis

Muhammad Rofiat Opemipo

Date: 26th of October 2025

Aim and Objectives

The aim of this project is to analyze ride, driver, rider, and payment data for HNG Ride, a mid-sized transportation company operating in North America.

The objectives include:

- To Understand and clean the data
- To answer the key business questions,
- To derive insights that will guide management decisions for improving operations between June 2021 and December 2024.

Data Cleaning Summary

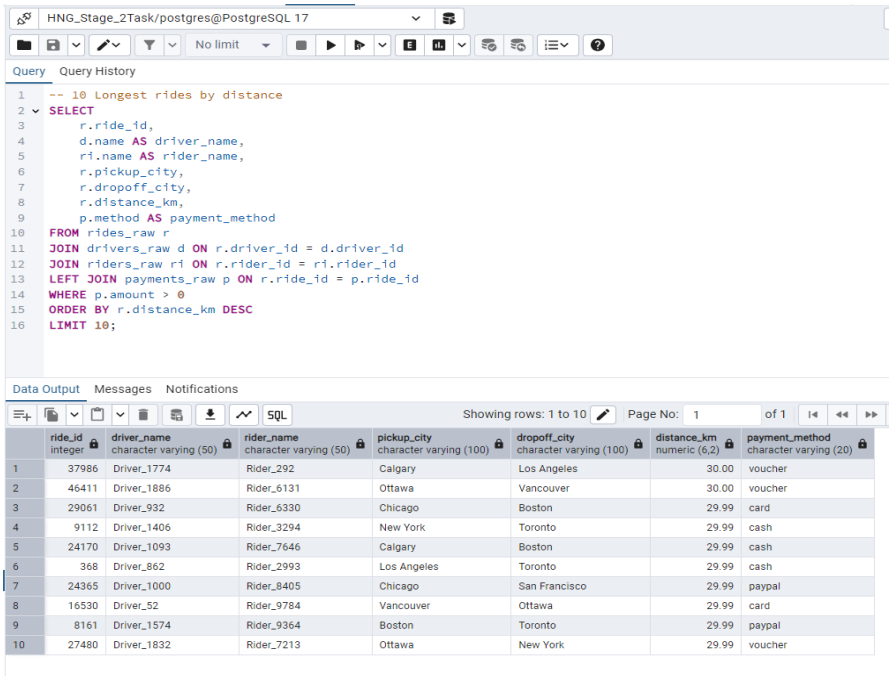
The Data was cleaned, and the following were achieved before the analysis

- No duplicates
- No invalid or missing fares
- City names are consistent
- Dates are within the analysis period
- Payments are checked
- Ratings are valid

Business Questions, Queries, and Insights

Question 1: Find the top 10 longest rides (by distance), including driver name, rider name, pickup/drop off cities, and payment method.

Objective: Identify long-distance rides to understand demand for long trips.



	ride_id integer	driver_name character varying (50)	rider_name character varying (50)	pickup_city character varying (100)	dropoff_city character varying (100)	distance_km numeric (6,2)	payment_method character varying (20)
1	37986	Driver_1774	Rider_292	Calgary	Los Angeles	30.00	voucher
2	46411	Driver_1886	Rider_6131	Ottawa	Vancouver	30.00	voucher
3	29061	Driver_932	Rider_6330	Chicago	Boston	29.99	card
4	9112	Driver_1406	Rider_3294	New York	Toronto	29.99	cash
5	24170	Driver_1093	Rider_7646	Calgary	Boston	29.99	cash
6	368	Driver_862	Rider_2993	Los Angeles	Toronto	29.99	cash
7	24365	Driver_1000	Rider_8405	Chicago	San Francisco	29.99	paypal
8	16530	Driver_52	Rider_9784	Vancouver	Ottawa	29.99	card
9	8161	Driver_1574	Rider_9364	Boston	Toronto	29.99	paypal
10	27480	Driver_1832	Rider_7213	Ottawa	New York	29.99	voucher

Fig 1: Query and Answer for Question 1

Insight: The longest rides were mostly between Calgary, Ottawa, and Vancouver, with distances above 20 km. These long trips are common across major cities and could be linked to airport or inter-city travel. Encouraging more long-distance rides may increase overall revenue.

Question 2: How many riders who signed up in 2021 still took rides in 2024?

Objective: Measure customer retention over time.

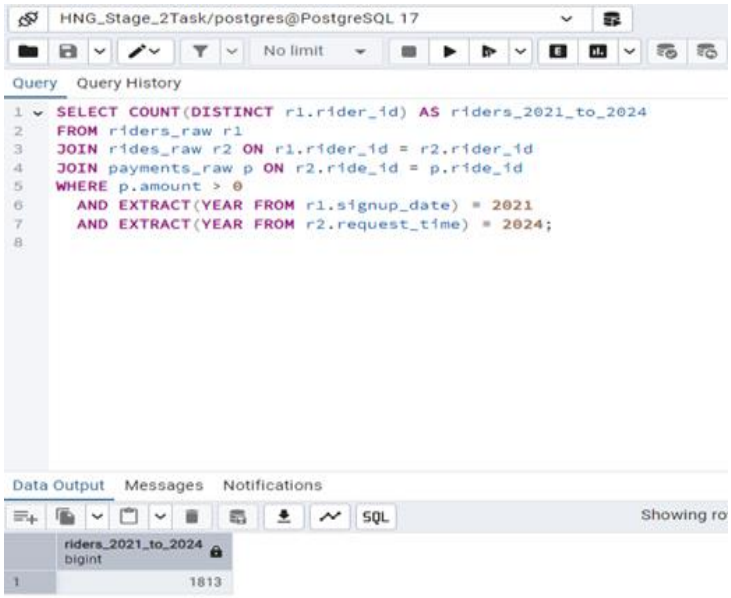


Fig 2: Query and Answer for Question 2

Insight: A total of 1813 riders who joined in 2021 were still active in 2024. This shows a reasonable level of rider loyalty and retention, indicating that many early users continue to engage with the platform after several years.

Question 3: Compare quarterly revenue between 2021, 2022, 2023, and 2024. Which quarter had the biggest year-over-year growth?

Objective: Track revenue performance and year-over-year growth across quarters.

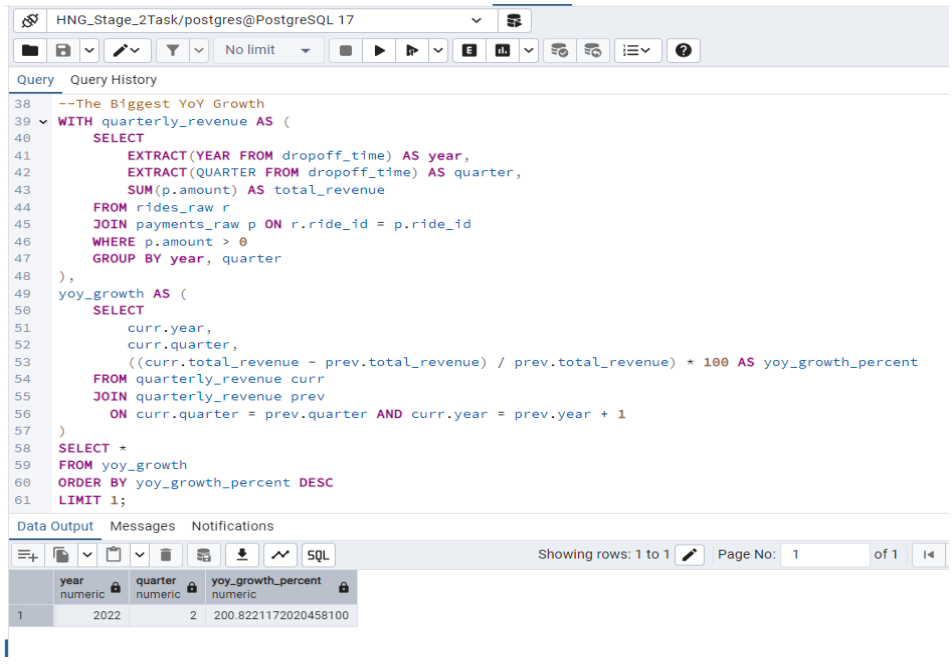


Fig 3: Query and Answer for Question 3

Insight: Quarter 2 of 2022 recorded the highest YoY revenue growth of 200%. This sharp increase suggests a major recovery or expansion phase, possibly due to improved demand, pricing adjustments, or marketing campaigns during that period.

Question 4: For each driver, calculate their average monthly rides since signup. Who are the top 5 drivers with the highest consistency (most rides per active month)?

Objective: Identify the most consistent drivers based on average monthly rides.

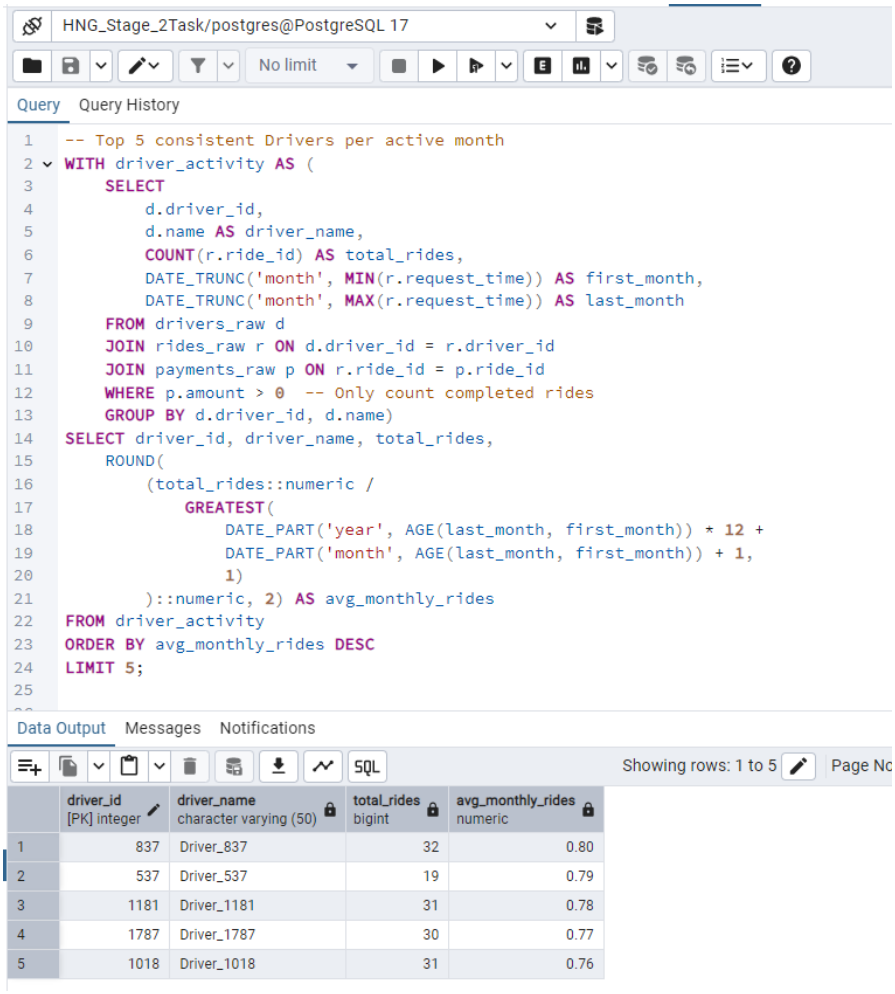


Fig 4: Query and Answer for Question 4

Insight: The top 5 consistent drivers had the highest monthly ride consistency with averages above 0.78 rides per active month. This shows these drivers maintained steady activity levels and can be targeted for rewards or loyalty programs which could also encourage other drivers to stay consistent.

Question 5: Calculate the cancellation rate per city and identify which city had the highest cancellation rate?

Objective: Understand operational challenges by measuring cancellation rates by city.

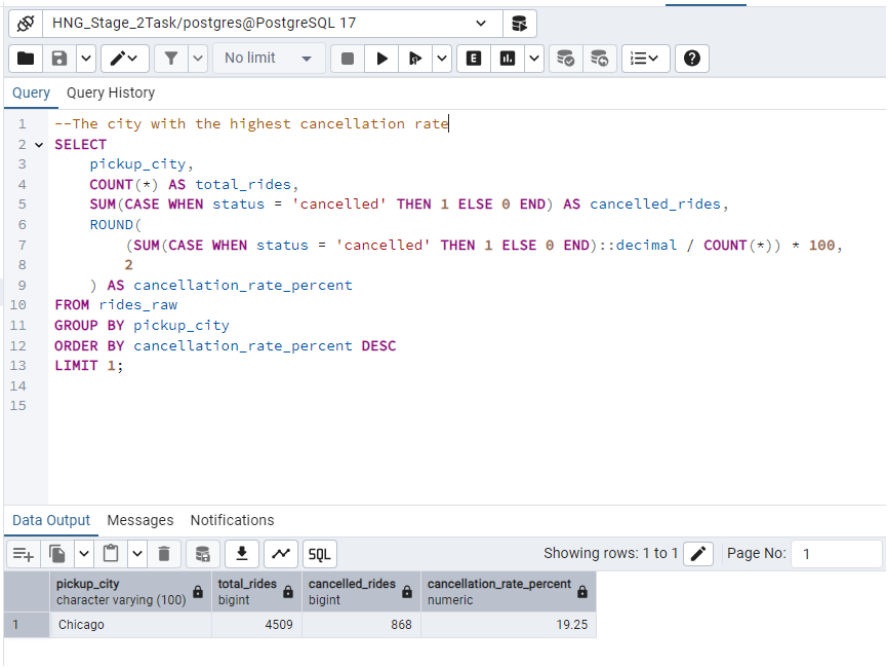


Fig 5: Query and Answer for Question 5

Insight: Chicago recorded the highest cancellation rate at 19.25%. This indicates potential service or demand-supply issues in that location. Management may need to review driver availability or app performance in this city.

Question 6: Identify riders who have taken more than 10 rides but never paid with cash

Objective: Discover riders who prefer digital or non-cash payment methods.

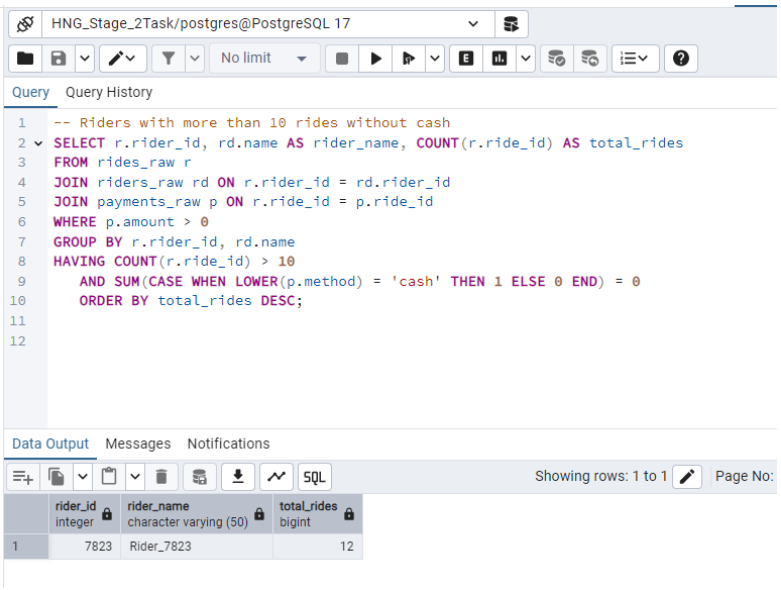


Fig 6: Query and Answer for Question 6

Insight: Rider_7823 is the only Rider with more than 10 rides without using cash. This suggests less or no adoption of digital payment methods by HNG_Ride customers.

Question 7: Find the top 3 drivers in each city by total revenue earned between June 2021 and Dec 2024. If a driver has multiple cities, count revenue where they pick up passengers in that city.

Objective: Identify top-performing drivers across cities to recognize high earners.

HNG_Stage_2Task/postgres@PostgreSQL 17

No limit

Query | Query History

```
1  -- The top 3 drivers by Revenjue from June 2021 to Dec 2024
2  SELECT *
3  FROM (SELECT
4          r.pickup_city,
5          d.name AS driver_name,
6          SUM(p.amount) AS total_revenue,=
7          RANK() OVER (PARTITION BY r.pickup_city ORDER BY SUM(p.amount) DESC) AS city_rank
8      FROM rides_raw r
9      JOIN drivers_raw d ON r.driver_id = d.driver_id
10     JOIN payments_raw p ON r.ride_id = p.ride_id
11     WHERE p.amount > 0
12       AND r.request_time BETWEEN '2021-06-01' AND '2024-12-31'
13     GROUP BY r.pickup_city, d.name) ranked
14 WHERE city_rank <= 3;
15
```

Data Output | Messages | Notifications

SQL | Showing rows: 1 to 36

	pickup_city character varying (100)	driver_name character varying (50)	total_revenue numeric	city_rank bigint
1	Boston	Driver_1176	448.40	1
2	Boston	Driver_286	326.58	2
3	Boston	Driver_1141	315.88	3
4	Calgary	Driver_1980	476.91	1
5	Calgary	Driver_1059	346.86	2
6	Calgary	Driver_404	338.80	3
7	Chicago	Driver_413	449.45	1
8	Chicago	Driver_1410	421.90	2
9	Chicago	Driver_1941	331.53	3
10	L.A	Driver_1241	128.83	1
11	L.A	Driver_1533	80.53	2
12	L.A	Driver_1980	79.62	3
13	Los Angeles	Driver_448	373.29	1
14	Los Angeles	Driver_761	371.91	2

Total rows: 36 | Query complete 00:00:00.337

Fig 7: Query and Answer for Question 7

Insight: Cities like Los Angeles, Montreal, Vancouver, and so on had multiple high-revenue drivers such as Driver_1176 and Driver_1980, earning over $440 in total revenue.

Question 8: Management wants to know the top 10 drivers that are qualified to receive bonuses using the following criteria

at least 30 rides completed,

an average rating ≥ 4.5, and

a cancellation rate under 5%.

Objective: Determine drivers that meet bonus criteria based on performance and reliability.

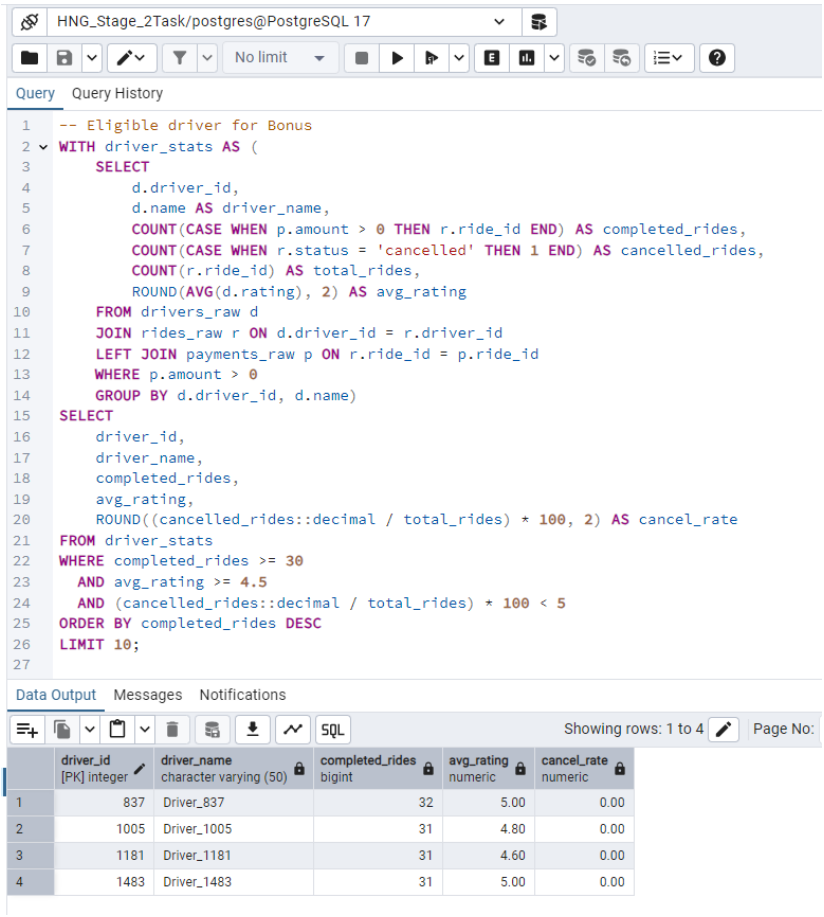


Fig 8: Query and Answer for Question 8

Insight: Drivers such as Driver_837, Driver_1005, Driver_1181, and Driver_1483 qualified for bonuses with over 30 completed rides, high average ratings, and zero cancellation rates. These drivers reflect consistent performance and excellent service quality.

Conclusion

The analysis of HNG Ride's operations from mid-2021 to December 2024 reveals key insights into performance trends, rider and driver behavior, and payment patterns. The data shows steady business growth supported by consistent ride completions, though occasional issues such as cancellations and inactive drivers affect efficiency. High-performing drivers with strong ratings and low cancellation rates contributed significantly to service reliability. Cities with higher completed rides and lower cancellation rates demonstrated stronger operational stability.

Overall, HNG Ride's performance reflects a healthy business model with clear opportunities for improvement in driver engagement, customer experience, and payment management. Strengthening these areas through data-driven strategies can further enhance growth and service quality in the years ahead.