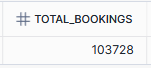
**1. Total Bookings in the given timeline.**

*SELECT COUNT(Trip\_id) AS Total\_bookings*

*FROM uber\_trip\_details;*



**2. Total revenue generated BY Uber in the given timeline.**

SELECT ROUND(SUM(fare\_amnt) + SUM(Surge\_Fee), 0) AS Total\_revenue

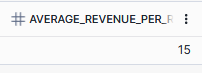
FROM uber\_trip\_details;



**3. Find Average Revenue Value per ride.**

SELECT ROUND((SUM(fare\_amnt) + SUM(Surge\_Fee))/COUNT(Trip\_id), 0) AS Average\_Revenue\_Per\_Ride

FROM uber\_trip\_details;



**4. Find Total Trips Distance**

SELECT CONCAT(ROUND(SUM(TRIPDIST), 0), ' Km') AS Total\_Trip\_Distance

FROM uber\_trip\_details;



**5. Find Average Trips Distance and also add 'Km' in output.**

SELECT

CONCAT(ROUND(SUM(TRIPDIST)/COUNT(Trip\_id), 1), ' Km') AS Average\_Trip\_Distance

FROM uber\_trip\_details;



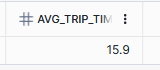
**6. Find Average Trip Time Per Ride in seconds.**

SELECT

ROUND(SUM(TOTIMESTAMPDIFF(SECOND, TO\_TIMESTAMP(pickupdate || ' ' || pickuptime) , TO\_TIMESTAMP(dropday || ' ' || droptime)))/

(SELECT COUNT(\*) FROM uber\_trip\_details)/60, 1) Avg\_Trip\_Time

FROM uber\_trip\_details



**7. Find Total Rev generated BY Transaction Type**

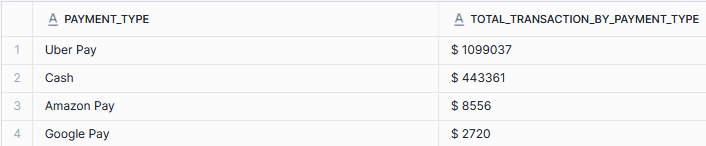
SELECT payment\_type,

CONCAT('$ ', ROUND(SUM(fare\_amnt) + SUM(surge\_fee), 0)) AS Total\_TransactiON\_BY \_Payment\_Type

FROM uber\_trip\_details

GROUP BY payment\_type

ORDER BY ROUND(SUM(fare\_amnt) + SUM(surge\_fee), 0) DESC



**8. Find Total no of bookings, total distance covered, and revenue based ON day and Night.**

SELECT

CASE WHEN pickuptime> = '20:00:00' or pickuptime< = '05:00:00' then 'Night' ELSE 'Day' END AS DayorNight,

COUNT(\*) AS Total\_bookings,

ROUND(SUM(tripdist), 1) Total\_Distance\_in\_Km,

ROUND(SUM(fare\_amnt) + SUM(Surge\_Fee), 0) AS Total\_revenue\_$

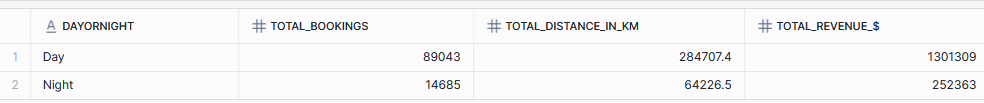
FROM uber\_trip\_details

GROUP BY 1

ORDER BY CASE DayorNight

WHEN ‘Day' then 1

ELSE 2 END

****

**9. Find key performance indicators like Total Bookings, Total Booking Value, Avg Booking Value, Total Trip Distance across different Vehicle Types in Uber trips.**

SELECT

Vehicle,

COUNT(\*) AS Total\_Bookings,

ROUND(SUM(fare\_amnt) + SUM(Surge\_Fee), 0) AS Total\_Booking\_val\_$,

ROUND((SUM(fare\_amnt) + SUM(Surge\_Fee))/COUNT(\*), 1) Avg\_Booking\_Value,

ROUND(SUM(tripdist), 1) Total\_Trip\_Distance\_KM

FROM uber\_trip\_details

GROUP BY 1

****

**10. Find Total No of Bookings BY Day\_of\_Week, BY Total Percent Value and Total Distance coverage.**

SELECT

dayname(pickupdate) Day\_of\_Week,

COUNT(Trip\_id) AS Total\_Bookings,

ROUND(COUNT(Trip\_id)\*100.0/(SELECT COUNT(trip\_id) FROM uber\_trip\_details), 1) Percent\_of\_Total,

ROUND(SUM(tripdist), 1) Total\_Trip\_Distance\_KM

FROM uber\_trip\_details

GROUP BY Day\_of\_Week

ORDER BY CASE Day\_of\_week

WHEN 'Sun' then 1

WHEN 'MON' then 2

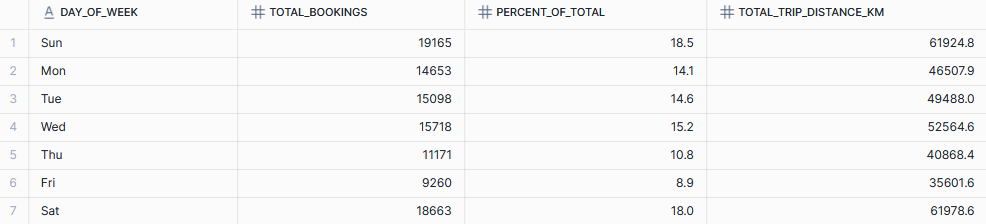
WHEN 'Tue' then 3

WHEN 'Wed' then 4

WHEN 'Thu' then 5

WHEN 'Fri' then 6

ELSE 7 END



**11. Find most Pickup Location and observe locations which contributes 80% of Total COUNT (PARETO ANALYSIS)**

WITH cte AS

(SELECT

location AS Pickup\_Location ,

COUNT(location ) AS Total\_COUNT,

ROUND((COUNT(location )\*100.0)/(SELECT COUNT(\*) FROM uber\_trip\_details), 1) AS Percent\_of\_total,

SUM(ROUND((COUNT(location )\*100.0)/(SELECT COUNT(\*) FROM uber\_trip\_details), 1))

OVER(ORDER BY COUNT(location ) DESC) Cum\_SUM\_percentage

FROM uber\_trip\_details u

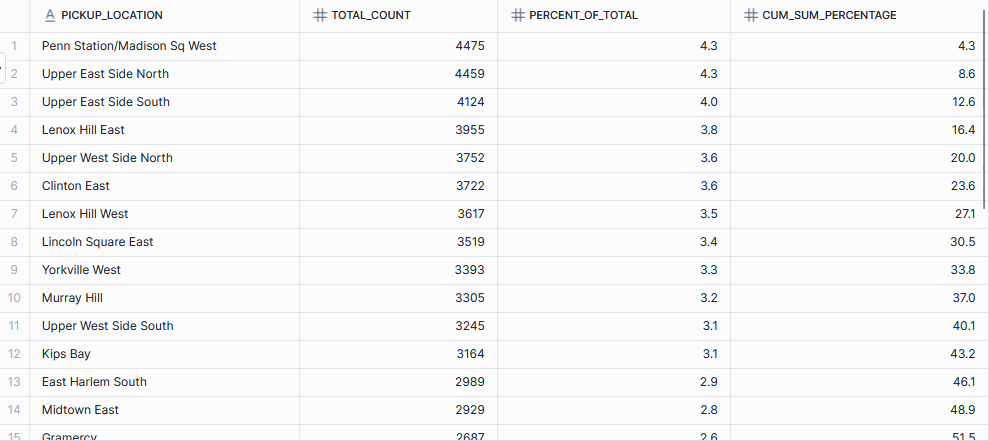
INNER JOIN loc\_info l ON l.location \_id = u.pickup\_locid

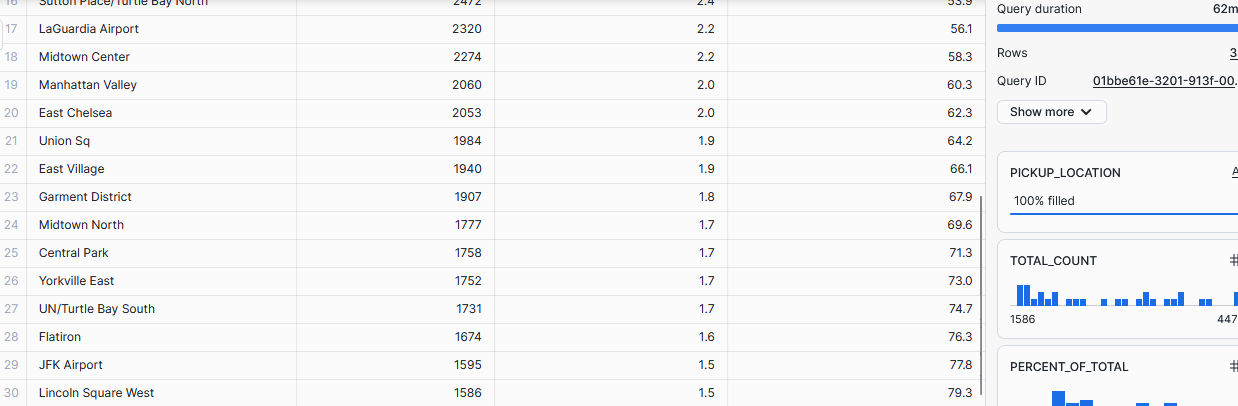
GROUP BY 1

ORDER BY 2 DESC)

SELECT \* FROM cte

WHERE Cum\_SUM\_percentage< = 80





**12. Find most Drop-off Location and observe locations which contributes 80% of Total COUNT (PARETO ANALYSIS)**

WITH cte AS

(SELECT location AS Drop\_Location , COUNT(location ) AS Total\_COUNT,

ROUND((COUNT(location )\*100.0)/(SELECT COUNT(\*) FROM uber\_trip\_details), 1) AS Percent\_of\_total,

SUM(ROUND((COUNT(location )\*100.0)/(SELECT COUNT(\*) FROM uber\_trip\_details), 1)) OVER(ORDER BY COUNT(location ) DESC) Cum\_SUM\_percentage

FROM uber\_trip\_details u

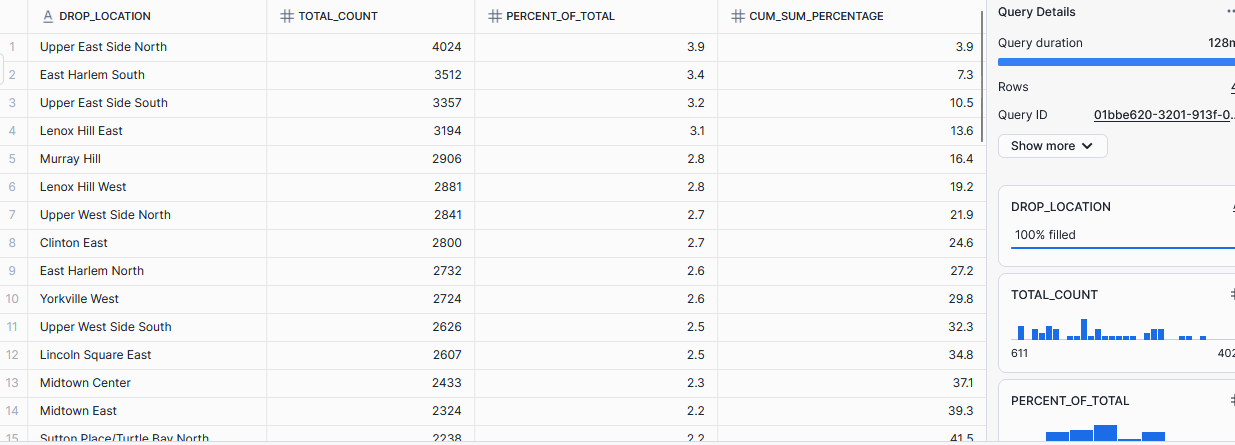
INNER JOIN loc\_info l ON l.location \_id = u.drop\_locid

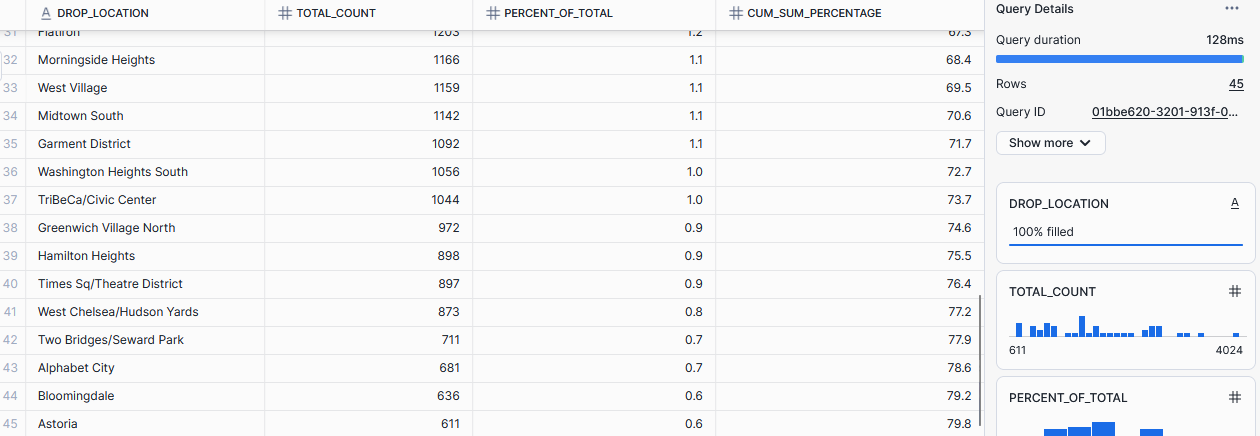
GROUP BY 1)

SELECT \* FROM cte

WHERE Cum\_SUM\_percentage<80

ORDER BY Cum\_SUM\_percentage





**13. We are working ON Vehicle Distribution. Find vehicles Distribution for top 20 Pickup locations**

WITH cte AS

(SELECT

location , vehicle ,

COUNT(vehicle) cnt, dense\_rank() OVER(partition BY location ORDER BY COUNT(vehicle)DESC) rn

FROM uber\_trip\_details u

INNER JOIN loc\_info l ON l.location \_id = u.pickup\_locid

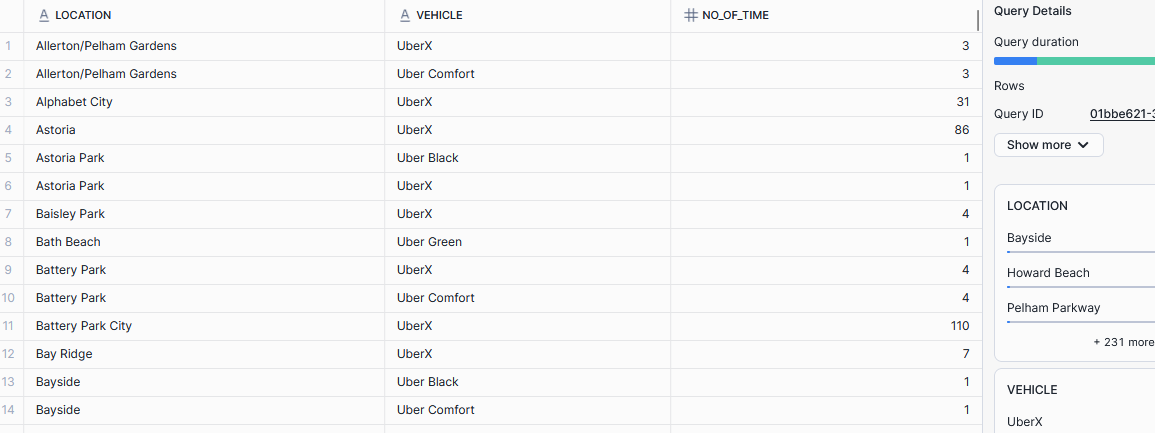
GROUP BY Location , Vehicle

ORDER BY location ASC, cnt DESC)

SELECT location , vehicle, CASE WHEN rn = 1 then cnt END AS No\_of\_time

FROM cte

WHERE rn = 1



**14. Cluster trip times into dayparts (Morning, Afternoon, Evening, Night) and analyze trip frequency & revenue share per part.**

SELECT

CASE

WHEN (pickuptime) > = '04:30:00' AND (pickuptime) < '11:30:00' THEN 'Morning'

WHEN (pickuptime) > = '11:30:00' AND (pickuptime) < '16:30:00' THEN 'Afternoon'

WHEN (pickuptime) > = '16:30:00' AND (pickuptime) < '20:30:00' THEN 'Evening'

ELSE 'Night'

END AS time\_of\_day, COUNT(\*) AS No\_of\_bookings,

ROUND(COUNT(CASE

WHEN (pickuptime) > = '04:30:00' AND (pickuptime) < '11:30:00' THEN 'Morning'

WHEN (pickuptime) > = '11:30:00' AND (pickuptime) < '16:30:00' THEN 'Afternoon'

WHEN (pickuptime) > = '16:30:00' AND (pickuptime) < '20:30:00' THEN 'Evening'

ELSE 'Night'

END)\*100.0/(SELECT COUNT(\*) FROM uber\_trip\_details), 1) AS Perc\_of\_Booking,

ROUND((SUM(fare\_amnt) + SUM(Surge\_Fee))/(SELECT SUM(fare\_amnt) + SUM(Surge\_Fee) FROM uber\_trip\_details)\*100.0, 1) AS Revenue\_percent\_of\_Total

FROM uber\_trip\_details GROUP BY 1

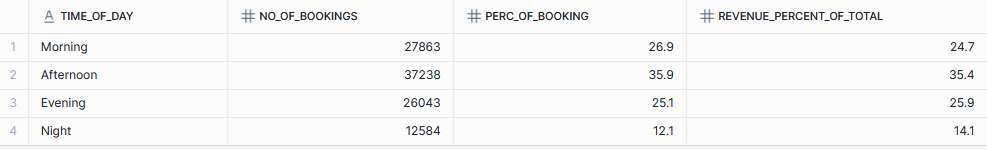
ORDER BY CASE Time\_of\_Day

WHEN 'Morning' then 1

WHEN 'Afternoon' then 2

WHEN 'Evening' then 3

WHEN 'Night' then 4 END



**15. Total Revenue Generated BY Vehicles**

SELECT

vehicle,

ROUND(SUM(fare\_amnt) + SUM(Surge\_Fee), 0) AS Total\_revenue,

ROUND((SUM(fare\_amnt) + SUM(Surge\_Fee))/(SELECT SUM(fare\_amnt) + SUM(Surge\_Fee) FROM uber\_trip\_details)\*100, 1) Perc\_of\_total

FROM uber\_trip\_details

GROUP BY Vehicle

ORDER BY Total\_revenue DESC

****

**16. Which hours of the day generate the most revenue per passengers ON average?**

SELECT

hour(pickuptime) hour\_of\_the\_day,

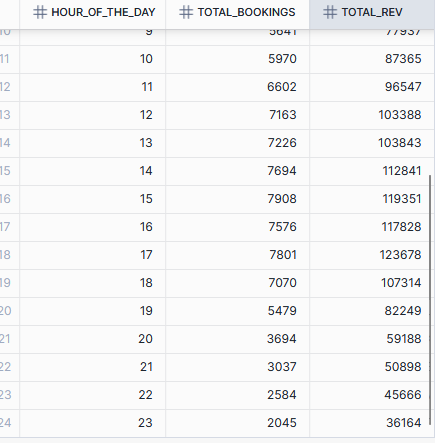
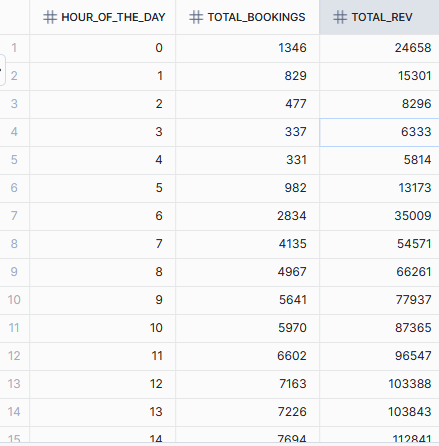
COUNT(\*) AS Total\_Bookings,

ROUND(SUM(fare\_amnt) + SUM(Surge\_Fee)) AS Total\_Rev

FROM uber\_trip\_details

GROUP BY hour\_of\_the\_day

ORDER BY hour\_of\_the\_day

****

**17. What is the average fare per km across different vehicle types?**

SELECT

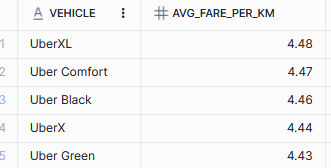
vehicle,

ROUND((SUM(fare\_amnt) + SUM(Surge\_Fee))/SUM(TRIPDIST), 2) AS Avg\_fare\_per\_km

FROM uber\_trip\_details

GROUP BY vehicle

ORDER BY Avg\_fare\_per\_km DESC

****

**18. What pickup location s have the highest surge fee ratios compared to base fare?**

SELECT

location ,

SUM(surge\_fee) Surge\_Fee,

SUM(fare\_amnt) FARE\_AMNT,

ROUND((SUM(surge\_fee)/SUM(fare\_amnt))\*100, 1) Ratio

FROM uber\_trip\_details u

INNER JOIN loc\_info l ON l.location \_id = u.pickup\_locid

GROUP BY location

ORDER BY ratio DESC

limit 10



**19. Do longer trips (distance-wise) always correspond to higher fare? (Correlation analysis between distance and fare)**

SELECT

corr(tripdist, (fare\_amnt + surge\_fee)) AS CorrelatiON\_val

FROM uber\_trip\_details

WHERE surge\_fee>0



**20. Create Pivot Table for Weekday vs Hour ON No of Bookings**

SELECT

dayname(PickupDate) AS Weekday,

COUNT(CASE WHEN hour(PickupTime) = 0 THEN 1 END) AS "00",

COUNT(CASE WHEN hour(PickupTime) = 1 THEN 1 END) AS "01",

COUNT(CASE WHEN hour(PickupTime) = 2 THEN 1 END) AS "02",

COUNT(CASE WHEN hour(PickupTime) = 3 THEN 1 END) AS "03",

COUNT(CASE WHEN hour(PickupTime) = 4 THEN 1 END) AS "04",

COUNT(CASE WHEN hour(PickupTime) = 5 THEN 1 END) AS "05",

COUNT(CASE WHEN hour(PickupTime) = 6 THEN 1 END) AS "06",

COUNT(CASE WHEN hour(PickupTime) = 7 THEN 1 END) AS "07",

COUNT(CASE WHEN hour(PickupTime) = 8 THEN 1 END) AS "08",

COUNT(CASE WHEN hour(PickupTime) = 9 THEN 1 END) AS "09",

COUNT(CASE WHEN hour(PickupTime) = 10 THEN 1 END) AS "10",

COUNT(CASE WHEN hour(PickupTime) = 11 THEN 1 END) AS "11",

COUNT(CASE WHEN hour(PickupTime) = 12 THEN 1 END) AS "12",

COUNT(CASE WHEN hour(PickupTime) = 13 THEN 1 END) AS "13",

COUNT(CASE WHEN hour(PickupTime) = 14 THEN 1 END) AS "14",

COUNT(CASE WHEN hour(PickupTime) = 15 THEN 1 END) AS "15",

COUNT(CASE WHEN hour(PickupTime) = 16 THEN 1 END) AS "16",

COUNT(CASE WHEN hour(PickupTime) = 17 THEN 1 END) AS "17",

COUNT(CASE WHEN hour(PickupTime) = 18 THEN 1 END) AS "18",

COUNT(CASE WHEN hour(PickupTime) = 19 THEN 1 END) AS "19",

COUNT(CASE WHEN hour(PickupTime) = 20 THEN 1 END) AS "20",

COUNT(CASE WHEN hour(PickupTime) = 21 THEN 1 END) AS "21",

COUNT(CASE WHEN hour(PickupTime) = 22 THEN 1 END) AS "22",

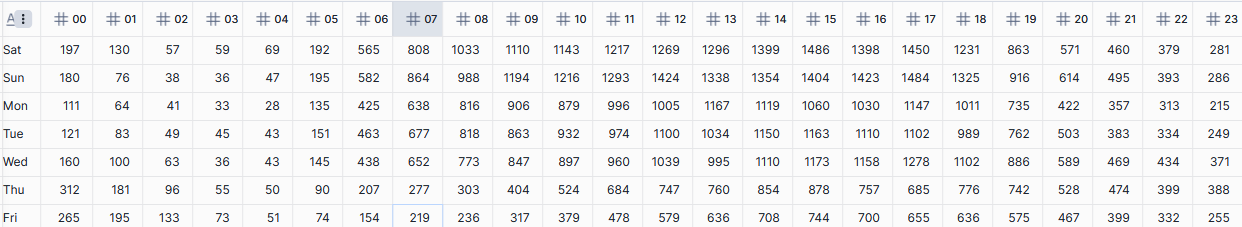
COUNT(CASE WHEN hour(PickupTime) = 23 THEN 1 END) AS "23"

FROM

uber\_trip\_details

GROUP BY

Weekday



**21. What is the average number of passengers per trip across different payment types?**

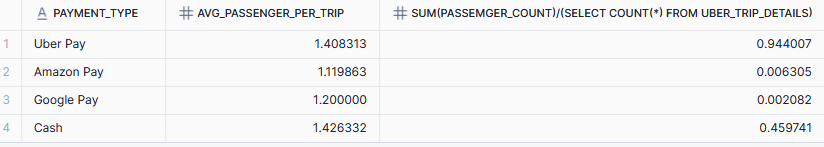
SELECT

payment\_type,

avg(passemger\_COUNT) AS avg\_passenger\_per\_trip , SUM(passemger\_COUNT)/(SELECT COUNT(\*) FROM uber\_trip\_details)

FROM uber\_trip\_details

GROUP BY payment\_type



**22. What is the average number of passengers per trip across different vehicle types?**

SELECT

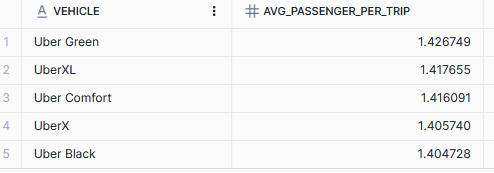
vehicle,

avg(passemger\_COUNT) AS Avg\_passenger\_per\_trip

FROM uber\_trip\_details

GROUP BY vehicle

ORDER BY Avg\_passenger\_per\_trip DESC



**23. Analyze fare Distribution for each vehicle type and see which has the most consistent pricing throughout week.**

SELECT vehicle,

SUM(CASE WHEN dayname(pickupdate) = 'Sun' then ROUND((FARE\_AMNT + surge\_fee), 1) END) AS "Sunday",

SUM(CASE WHEN dayname(pickupdate) = 'MON' then ROUND((FARE\_AMNT + surge\_fee), 1) END) AS "Monday",

SUM(CASE WHEN dayname(pickupdate) = 'Tue' then ROUND((FARE\_AMNT + surge\_fee), 1) END) AS "Tuesday",

SUM(CASE WHEN dayname(pickupdate) = 'Wed' then ROUND((FARE\_AMNT + surge\_fee), 1) END) AS "Wednesday",

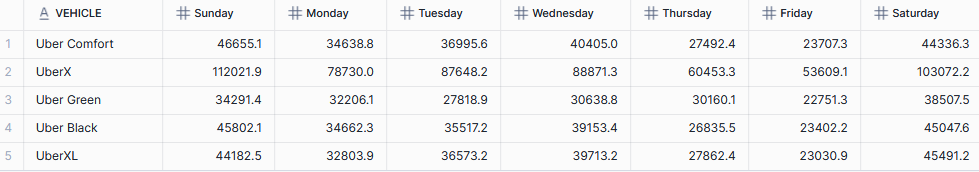
SUM(CASE WHEN dayname(pickupdate) = 'Thu' then ROUND((FARE\_AMNT + surge\_fee), 1) END) AS "Thursday",

SUM(CASE WHEN dayname(pickupdate) = 'Fri' then ROUND((FARE\_AMNT + surge\_fee), 1) END) AS "Friday",

SUM(CASE WHEN dayname(pickupdate) = 'Sat' then ROUND((FARE\_AMNT + surge\_fee), 1) END) AS "Saturday"

FROM uber\_trip\_details

GROUP BY vehicle;



**24. Which pickup locations have the highest average surge fee and how often do they appear?**

SELECT

location ,

ROUND(avg(surge\_fee), 2) surge\_fee,

COUNT(\*) frequency

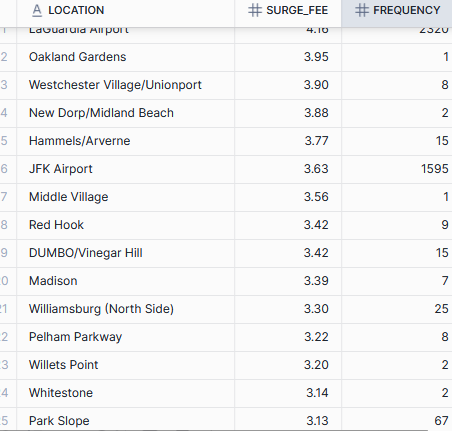
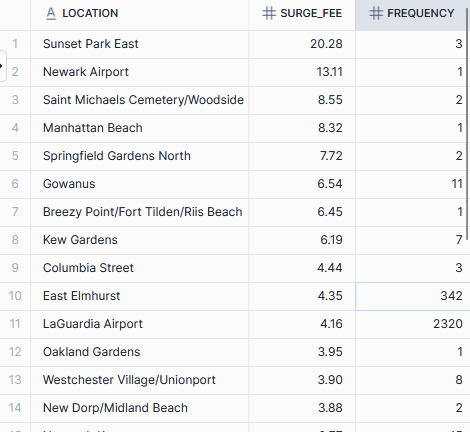
FROM uber\_trip\_details u

INNER JOIN loc\_info l ON l.location \_id = u.pickup\_locid

GROUP BY location

ORDER BY surge\_fee DESC

limit 25;



**25. Create Distance bins and find total bookings and revenue BY percentage**.

SELECT

CASE WHEN ROUND(tripdist, 0)< = 4 then '<4 KM'

WHEN ROUND(tripdist, 0)>4 and ROUND(tripdist, 0)< = 10 then '4-10 KM'

WHEN ROUND(tripdist, 0)>10 and ROUND(tripdist, 0)< = 25 then '10-25 KM'

WHEN ROUND(tripdist, 0)>25 and ROUND(tripdist, 0)< = 60 then '25-60 KM'

WHEN ROUND(tripdist, 0)>60 then '>60 KM'

END AS Distance\_bins

, COUNT(\*) no\_of\_bookings,

ROUND((SUM(fare\_amnt) + SUM(surge\_fee))/(SELECT (SUM(fare\_amnt) + SUM(surge\_fee)) FROM uber\_trip\_details)\*100.0, 1) AS Rev\_Perc\_of\_total

FROM uber\_trip\_details

GROUP BY Distance\_bins

ORDER BY CASE Distance\_bins

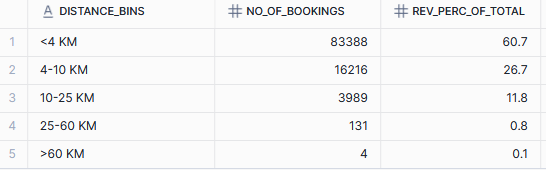
WHEN '<4 KM' THEN 1

WHEN '4-10 KM' THEN 2

WHEN '10-25 KM' THEN 3

WHEN '25-60 KM' THEN 4

ELSE 5 END



**26. Cluster trips BY pickup time(3 hour slab), trip distance, fare, and passengers COUNT**

SELECT

CASE WHEN hour(pickuptime)< = 2 then '0-3'

WHEN hour(pickuptime)>2 and hour(pickuptime)< = 5 then '3-6'

WHEN hour(pickuptime)>5 and hour(pickuptime)< = 8 then '6-9'

WHEN hour(pickuptime)>8 and hour(pickuptime)< = 11 then '9-12'

WHEN hour(pickuptime)>11 and hour(pickuptime)< = 14 then '12-15'

WHEN hour(pickuptime)>14 and hour(pickuptime)< = 17 then '15-18'

WHEN hour(pickuptime)>17 and hour(pickuptime)< = 20 then '18-21'

WHEN hour(pickuptime)>20 then '21-24' END AS Time\_slab

, SUM(tripdist) Trip\_Dist,

ROUND(SUM(fare\_amnt) + SUM(surge\_fee), 1) Total\_Fare

, COUNT(\*) No\_of\_booikngs,

SUM(passenger\_COUNT) Total\_Passengers

FROM uber\_trip\_details

GROUP BY Time\_Slab

ORDER BY CASE Time\_Slab

WHEN '0-3' then 1

WHEN '3-6' then 2

WHEN '6-9' then 3

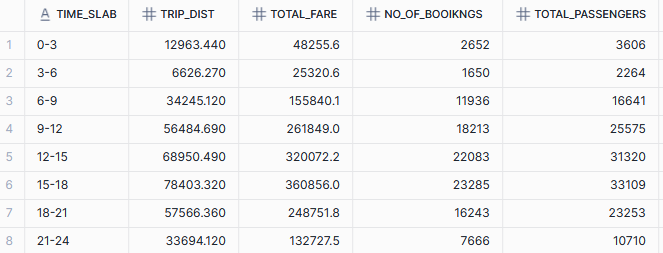
WHEN '9-12' then 4

WHEN '12-15' then 5

WHEN '15-18' then 6

WHEN '18-21' then 7

WHEN '21-24' then 8 END



**27. Cluster trips BY pickup time (3 hour slab), trip distance, fare, and passengers COUNT but in %\_of total**

SELECT

CASE WHEN hour(pickuptime)< = 2 then '0-3'

WHEN hour(pickuptime)>2 and hour(pickuptime)< = 5 then '3-6'

WHEN hour(pickuptime)>5 and hour(pickuptime)< = 8 then '6-9'

WHEN hour(pickuptime)>8 and hour(pickuptime)< = 11 then '9-12'

WHEN hour(pickuptime)>11 and hour(pickuptime)< = 14 then '12-15'

WHEN hour(pickuptime)>14 and hour(pickuptime)< = 17 then '15-18'

WHEN hour(pickuptime)>17 and hour(pickuptime)< = 20 then '18-21'

WHEN hour(pickuptime)>20 then '21-24' END AS Time\_slab

, ROUND(SUM(tripdist)/(SELECT SUM(tripdist) FROM uber\_trip\_details)\*100.0, 1) AS Trip\_Dist\_of\_perc\_in\_total

, ROUND((SUM(fare\_amnt) + SUM(surge\_fee))/(SELECT SUM(fare\_amnt) + SUM(surge\_fee) FROM uber\_trip\_details)\*100.0, 1) AS Total\_Fare\_perc\_of\_total

, ROUND(COUNT(\*)/(SELECT COUNT(\*) FROM uber\_trip\_details)\*100.0, 1) No\_of\_booikngs\_perc\_of\_total

, ROUND(SUM(passemger\_COUNT)/(SELECT SUM(passemger\_COUNT) FROM uber\_trip\_details)\*100.0, 1) Total\_Passengers\_perc\_of\_total

FROM uber\_trip\_details

GROUP BY Time\_Slab

ORDER BY CASE Time\_Slab

WHEN '0-3' then 1

WHEN '3-6' then 2

WHEN '6-9' then 3

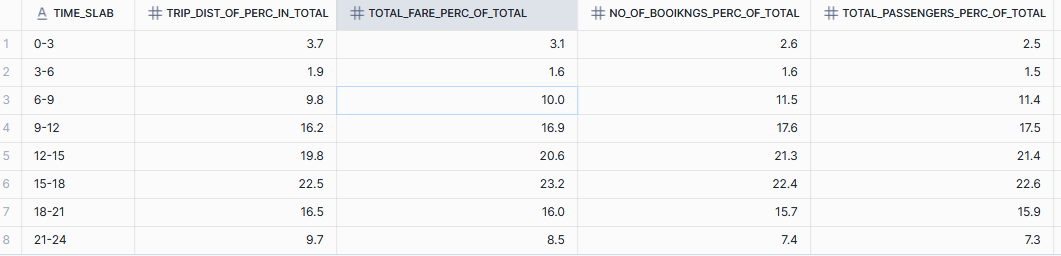
WHEN '9-12' then 4

WHEN '12-15' then 5

WHEN '15-18' then 6

WHEN '18-21' then 7

WHEN '21-24' then 8 END



**28. which Drop-offs locations contributes most to the Surge Fee**

SELECT

location AS DropOFF\_Loc,

SUM(surge\_fee) AS Surge\_Amount,

ROUND((SUM(surge\_fee)/(SUM(fare\_amnt) + SUM(surge\_fee)))\*100.0, 1) location \_wise\_perc\_of\_total\_rev

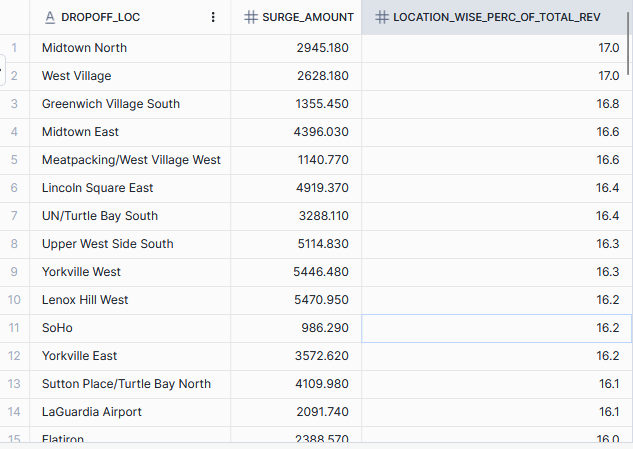
FROM uber\_trip\_details u

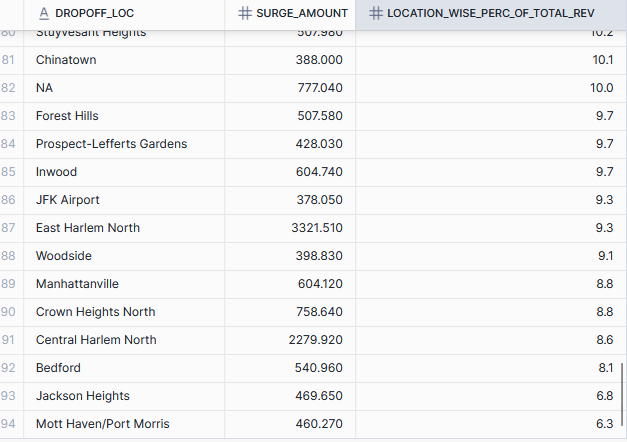
INNER JOIN loc\_info l ON l.location \_id = u.drop\_locid

GROUP BY DropOFF\_Loc

having Surge\_Amount >350

ORDER BY location \_wise\_perc\_of\_total\_rev DESC





**29. Find Total Bookings ON the basis of Vehicles and Daytime**

SELECT vehicle,

SUM(CASE WHEN hour(pickuptime) > = 5 and hour(pickuptime) < 12 THEN 1 END) AS Morning,

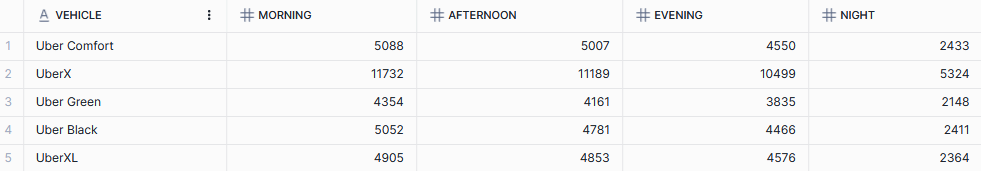
SUM(CASE WHEN hour(pickuptime) > = 12 AND hour(pickuptime) < 16 THEN 1 END) AS Afternoon,

SUM(CASE WHEN hour(pickuptime) > = 16 AND hour(pickuptime) < 20 THEN 1 END) AS Evening,

SUM(CASE WHEN hour(pickuptime) > = 20 or hour(pickuptime) < 5 THEN 1 END) AS Night

FROM uber\_trip\_details

GROUP BY vehicle



**30. Find Total Surge Amount ON the basis of Vehicles and Daytime**

SELECT vehicle,

SUM(CASE WHEN hour(pickuptime) > = 5 and hour(pickuptime) < 12 THEN surge\_fee END) AS Morning,

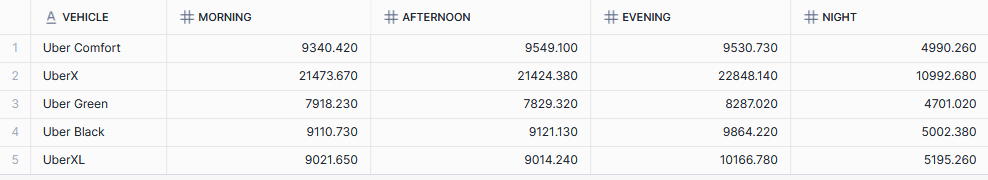
SUM(CASE WHEN hour(pickuptime) > = 12 AND hour(pickuptime) < 16 THEN surge\_fee END) AS Afternoon,

SUM(CASE WHEN hour(pickuptime) > = 16 AND hour(pickuptime) < 20 THEN surge\_fee END) AS Evening,

SUM(CASE WHEN hour(pickuptime) > = 20 or hour(pickuptime) < 5 THEN surge\_fee END) AS Night

FROM uber\_trip\_details

GROUP BY vehicle



**31.Find No of Trips ON the basis of time of this format : 0-5 minutes, 5-15 minutes, 15 minutes to 1 hour, 1 hour to 3 hour, 3-8 hour, 8-24 hour, 24 hour +**

WITH cte AS(

SELECT

trip\_id,

TOTIMESTAMPDIFF(minute, TO\_TIMESTAMP(pickupdate || ' ' || pickuptime), TO\_TIMESTAMP(dropday || ' ' || droptime)) time\_intrvl

FROM uber\_trip\_details),

cte2 AS(

SELECT

CASE WHEN time\_intrvl<5 then 'Very\_Short 0-5min'

WHEN time\_intrvl> = 5 and time\_intrvl<16 then 'Short\_Ride 5-15min'

WHEN time\_intrvl> = 16 and time\_intrvl<61 then 'Medium\_Ride 15min - 1Hour'

WHEN time\_intrvl> = 61 and time\_intrvl<181 then 'Long\_Ride 1Hour-3Hour'

WHEN time\_intrvl> = 181 and time\_intrvl<481 then 'Extra\_long\_ride 3-8Hour'

WHEN time\_intrvl> = 481 and time\_intrvl<1441 then 'OVER\_Haul\_Ride 8-24Hour'

ELSE 'Outlier' END AS Time\_period

FROM cte)

SELECT Time\_period, COUNT(Time\_period) No\_of\_Rides

FROM cte2

GROUP BY 1

ORDER BY CASE Time\_period

WHEN 'Very\_Short 0-5min' then 1

WHEN 'Short\_Ride 5-15min' then 2

WHEN 'Medium\_Ride 15min - 1Hour' then 3

WHEN 'Long\_Ride 1Hour-3Hour' then 4

WHEN 'Extra\_long\_ride 3-8Hour' then 5

WHEN 'OVER\_Haul\_Ride 8-24Hour' then 6

WHEN 'Outlier' then 7 END

