Data cleaning :

Converting necessary columns to their correct data types and joining both the given tables to answer upcoming questions:

select Restaurants.Restaurant\_Name,Restaurants.Cuisine,Restaurants.zone,Restaurants.Category,Orders.Order\_ID,

orders.Customer\_Name,cast(Orders.order\_date as date) as Order\_date,

cast(orders.restaurant\_id as int) as restaurant\_id,

cast(Orders.quantity\_of\_items as int) as quantity\_of\_items,cast(Orders.order\_amount as int) as order\_amount,orders.payment\_mode,

cast(Orders.delivery\_time\_taken\_mins as int) as delivery\_time\_taken\_mins,cast(Orders.customer\_rating\_food as int) as customer\_rating\_food,

cast(Orders.customer\_rating\_delivery as int) as customer\_rating\_delivery

from orders

left join restaurants

on orders.restaurant\_id=restaurants.restaurant\_id;

Analysis:

Question-1: Perform summary analysis using aggregate functions

Total sales for all the restaurants

with tb1 as (select Restaurants.Restaurant\_Name,Restaurants.Cuisine,Restaurants.zone,Restaurants.Category,Orders.Order\_ID,

orders.Customer\_Name,cast(Orders.order\_date as date) as Order\_date,

cast(orders.restaurant\_id as int) as restaurant\_id,

cast(Orders.quantity\_of\_items as int) as quantity\_of\_items,cast(Orders.order\_amount as int) as order\_amount,orders.payment\_mode,

cast(Orders.delivery\_time\_taken\_mins as int) as delivery\_time\_taken\_mins,cast(Orders.customer\_rating\_food as int) as customer\_rating\_food,

cast(Orders.customer\_rating\_delivery as int) as customer\_rating\_delivery

from orders

left join restaurants

on orders.restaurant\_id=restaurants.restaurant\_id)

select sum(order\_amount) as total\_sales

from tb1;

Average sales for each restaurant

select Restaurant\_Name,avg(order\_amount) as avg\_sales

from tb1

group by Restaurant\_Name;

Number of restaurants in each zone

select zone,count(restaurant\_name) as num\_restaurants

from tb1

group by zone;

Total sales per zone

select zone,sum(order\_amount) as total\_sales

from tb1

group by zone;

Average sales per zone

select zone,avg(order\_amount) as avg\_sales\_per\_zone

from tb1

group by zone;

Average delivery time per zone

select zone,avg(delivery\_time\_taken\_mins) as avg\_time

from tb1

group by zone;

Average customer rating for food

select avg(customer\_rating\_food) as avg\_rating\_food

from tb1;

Average customer rating for delivery

select avg(customer\_rating\_delivery) as avg\_rating\_delivery

from tb1;

Question-2: Which restaurant received the most orders?

select restaurant\_name,count(order\_id) as total\_orders

from tb1

group by restaurant\_name

order by total\_orders desc;

Question-3: Which restaurant saw most sales?

select Restaurant\_Name,sum(order\_amount) as total\_sales

from tb1

group by Restaurant\_Name

order by total\_sales desc

Question-4:Which customer ordered the most?

select Customer\_Name,count(order\_id)as total\_orders

from tb1

group by Customer\_Name

order by total\_orders desc

Question-5: When do customers order the most in a day?

A day consists of Morning,Afternoon,Evening and Night.

Since the time data in the given dataset is in 24 hour clock format hence assuming ‘00:00:00-11:59:59’ as ‘Morning’, ‘12:00:00-15:59:59’ as ‘Afternoon’, ‘16:00:00-18:59:59’ as evening and the rest as night.

with tb1 as (select Restaurants.Restaurant\_Name,Restaurants.Cuisine,Restaurants.zone,Restaurants.Category,Orders.Order\_ID,

orders.Customer\_Name,cast(Orders.order\_date as datetime) as Order\_date,

cast(orders.restaurant\_id as int) as restaurant\_id,

cast(Orders.quantity\_of\_items as int) as quantity\_of\_items,cast(Orders.order\_amount as int) as order\_amount,orders.payment\_mode,

cast(Orders.delivery\_time\_taken\_mins as int) as delivery\_time\_taken\_mins,cast(Orders.customer\_rating\_food as int) as customer\_rating\_food,

cast(Orders.customer\_rating\_delivery as int) as customer\_rating\_delivery

from orders

left join restaurants

on orders.restaurant\_id=restaurants.restaurant\_id),

--extracting time data from 'order\_date' column

tb2 as (SELECT \*,format(order\_date,'HH:mm:ss') as time\_data

from tb1),

--categorizing time data to morning,afternoon,evening and night

tb3 as (select order\_id,case when time\_data between '00:00:00' and '11:59:59' then 'morning'

when time\_data between '12:00:00' and '15:59:59' then 'afternoon'

when time\_data between '16:00:00' and '18:59:59' then 'evening'

when time\_data between '19:00:00' and '23:59:59' then 'night' end as time\_of\_day

from tb2)

--counting orders for time of the day

select time\_of\_day,count(order\_id)as order\_count

from tb3

group by time\_of\_day

Question-6: Which is the most liked cuisine in each zone?

with tb1 as (select Restaurants.Restaurant\_Name,Restaurants.Cuisine,Restaurants.zone,Restaurants.Category,Orders.Order\_ID,

orders.Customer\_Name,cast(Orders.order\_date as datetime) as Order\_date,

cast(orders.restaurant\_id as int) as restaurant\_id,

cast(Orders.quantity\_of\_items as int) as quantity\_of\_items,cast(Orders.order\_amount as int) as order\_amount,orders.payment\_mode,

cast(Orders.delivery\_time\_taken\_mins as int) as delivery\_time\_taken\_mins,cast(Orders.customer\_rating\_food as int) as customer\_rating\_food,

cast(Orders.customer\_rating\_delivery as int) as customer\_rating\_delivery

from orders

left join restaurants

on orders.restaurant\_id=restaurants.restaurant\_id),

tb2 as (select zone,Cuisine,count(cuisine) as cuisine\_count

from tb1

group by zone,Cuisine),

tb3 as (select \*,row\_number()over(partition by zone order by cuisine\_count desc) as rank\_count

from tb2)

select \*

from tb3

where rank\_count=1

Question-7:which zone has the most sales?

with tb1 as (select Restaurants.Restaurant\_Name,Restaurants.Cuisine,Restaurants.zone,Restaurants.Category,Orders.Order\_ID,

orders.Customer\_Name,cast(Orders.order\_date as datetime) as Order\_date,

cast(orders.restaurant\_id as int) as restaurant\_id,

cast(Orders.quantity\_of\_items as int) as quantity\_of\_items,cast(Orders.order\_amount as int) as order\_amount,orders.payment\_mode,

cast(Orders.delivery\_time\_taken\_mins as int) as delivery\_time\_taken\_mins,cast(Orders.customer\_rating\_food as int) as customer\_rating\_food,

cast(Orders.customer\_rating\_delivery as int) as customer\_rating\_delivery

from orders

left join restaurants

on orders.restaurant\_id=restaurants.restaurant\_id)

select zone,sum(order\_amount) as total\_sales\_each\_zone

from tb1

group by zone

order by total\_sales\_each\_zone desc

Question-8: Which restaurant is the top rated in each zone in terms of food quality?

Assuming the quality of food to be good when customer rating for the food is 4 and above.

with tb1 as (select Restaurants.Restaurant\_Name,Restaurants.Cuisine,Restaurants.zone,Restaurants.Category,Orders.Order\_ID,

orders.Customer\_Name,cast(Orders.order\_date as datetime) as Order\_date,

cast(orders.restaurant\_id as int) as restaurant\_id,

cast(Orders.quantity\_of\_items as int) as quantity\_of\_items,cast(Orders.order\_amount as int) as order\_amount,orders.payment\_mode,

cast(Orders.delivery\_time\_taken\_mins as int) as delivery\_time\_taken\_mins,cast(Orders.customer\_rating\_food as int) as customer\_rating\_food,

cast(Orders.customer\_rating\_delivery as int) as customer\_rating\_delivery

from orders

left join restaurants

on orders.restaurant\_id=restaurants.restaurant\_id),

--filtering data having customer\_rating\_food 4 and above:

tb2 as (select Restaurant\_Name,zone,count(customer\_rating\_food) as good\_rating

from tb1

where customer\_rating\_food>=4

group by Restaurant\_Name,Zone),

--segrigating data

tb3 as (select \*,row\_number() over (partition by zone order by good\_rating desc) as rank\_count

from tb2)

select \*

from tb3

where rank\_count=1

Question-9:Which restaurant is top rated in each zone in terms of food delivery?

Assuming the delivery of food is top rated when the customer rating for delivery is 4 and above.

with tb1 as (select Restaurants.Restaurant\_Name,Restaurants.Cuisine,Restaurants.zone,Restaurants.Category,Orders.Order\_ID,

orders.Customer\_Name,cast(Orders.order\_date as datetime) as Order\_date,

cast(orders.restaurant\_id as int) as restaurant\_id,

cast(Orders.quantity\_of\_items as int) as quantity\_of\_items,cast(Orders.order\_amount as int) as order\_amount,orders.payment\_mode,

cast(Orders.delivery\_time\_taken\_mins as int) as delivery\_time\_taken\_mins,cast(Orders.customer\_rating\_food as int) as customer\_rating\_food,

cast(Orders.customer\_rating\_delivery as int) as customer\_rating\_delivery

from orders

left join restaurants

on orders.restaurant\_id=restaurants.restaurant\_id),

tb2 as (select Restaurant\_Name,zone,count(customer\_rating\_delivery) as good\_delivery

from tb1

where customer\_rating\_delivery>=4

group by Restaurant\_Name,Zone),

tb3 as (select \*,row\_number() over (partition by zone order by good\_delivery desc) as rank\_count

from tb2)

select \*

from tb3

where rank\_count=1

Question-10: Which is the most preferred payment method?

with tb1 as (select Restaurants.Restaurant\_Name,Restaurants.Cuisine,Restaurants.zone,Restaurants.Category,Orders.Order\_ID,

orders.Customer\_Name,cast(Orders.order\_date as datetime) as Order\_date,

cast(orders.restaurant\_id as int) as restaurant\_id,

cast(Orders.quantity\_of\_items as int) as quantity\_of\_items,cast(Orders.order\_amount as int) as order\_amount,orders.payment\_mode,

cast(Orders.delivery\_time\_taken\_mins as int) as delivery\_time\_taken\_mins,cast(Orders.customer\_rating\_food as int) as customer\_rating\_food,

cast(Orders.customer\_rating\_delivery as int) as customer\_rating\_delivery

from orders

left join restaurants

on orders.restaurant\_id=restaurants.restaurant\_id)

select payment\_mode,count(payment\_mode) as count\_of\_payment\_method

from tb1

group by Payment\_Mode

order by count\_of\_payment\_method desc