

Question Bank

SQL

Easy

1. Top 100 customers' revenue wise for a particular category.
2. Find out the categories in the month of December' 2019 for which the revenue ≥ 1 million.
3. Eliminate the duplicate rows from a table
4. Top 3 categories revenue wise.
5. After creating a filtered table from a given table, how do we make a permanent table out of that filtered table i.e. so that it is accessible to other users?
6. Different kind of joins
7. What does Union do?
8. What is the order of SQL execution?
9. What does Substring and charindex do? Extract the First_name, middle_name and the last_name.
10. Main difference between WHERE and GROUP BY
11. What does Having clause do?
12. What does partition do?
13. Difference between RANK() and DENSE_RANK()
14. What is primary key and foreign key?
15. Calculate the mean , median , mode using sql queries

Medium

1. Fetch Top 5 master brands in descending order of brand frequency from dim_product table and transpose that table.
2. Given the fact_core_item and dim_product tables, query the mrp and qty_sold of all the brands for calendar year 2018. Round MRP AND QTY SOLD to the nearest integer. Note: fact_core_item.SKU_ID and dim_product.SKU_ID are matching key columns.
3. Fetch the style_IDS of Top 200 Jeans by qty_sold for calendar year 2018 for 'Moda rapido' brand. Use table fact_core_item and dim_product. Note: fact_core_item.SKU_ID and dim_product.SKU_ID are matching key columns.

4. Perform data cleaning for the column "city" from fact_Core_item table for date-2nd June 19

Table 1	Table 2
1	1
2	1
3	2
5	4
Null	Null
Null	

Perform all kinds of joins with the above two tables

6. Find second and third largest salary using :

(a) Rank Function

(b) without using rank function

7. Write a query to extract top 25% of records based on salary.

8.

Table1:

style_id	origin	url_count
11	china	234
22	usa	345
33	germany	456
44	england	677
55	india	567
66	china	234

Table2:

Style_id	origin	url_count
33	germany	234
44	england	5345
55	india	354

Join these two table using left join and exclude the style_ids which are common in both table

9. If there are 11 teams , and each team plays with every other team only once .How many matches will be played? Write an sql query to display all the matches.

10. Write a query to display the revenue from each gender by the date . The output should be such that there should be three columns displaying the date , revenue of male, revenue of female.

11. Write a query to display the cumulative revenue in the transaction table.

12. Two tables T1 and T2 , T1 having 1 columns with 4 rows and T2 having 1 column with 6 rows. Find the maximum and minimum rows , when you do all kinds of joins

13.For a particular order in the fact_core_item, get the brands of the styles and the respective GMV.

Table Name :

Dim_style

Fact_core_item

14.Get the article types generating revenue more than average revenue for an article type.

15.Get the brands generating revenue less than 30lakh in the year 2018.

16. Get the total number of loads and clicks for Android and IOS devices for date range : 19 june 2019 to 21 June 2019

Table name : Widget Entity View

Load event : ProductListLoaded-ViewPort

Click Event : product_click_list_page

17. Get the unique session_id from the given table, having sessions more than 5 minutes.

Table name : session_daily_view

Column to be referred for session time : session_duration(its in milliseconds in table)

18. Get the frequency distribution of Number of items vs No.of orders, like
1 items are in how many orders, 2 items are in how many orders and so on

Hard

1. Cumulative Sum of monthly sales(with and without self-join)

2. Column has value blue and red .update the value blue=red and red=blue.(one query)
3. MPP
4. Meta data
5. External table
6. Dynamic SQL
7. Procedures and how will you create the dynamic table, what situation you created Dynamic table in your project.
8. Hive- partition ,bucket, schema read, schema write
9. Duplicate rows
10. Python lib
11. Lookup, fuzzy lookup, Merge join, incremental load, SCD..Etc. –SSIS
12. DW concept. Explain how did u implement the DW in your project .what are the DW techniques used in your project.
13. Query optimization and performance.
14. Procedure performance suddenly very slow in the same data size, what is the problem and how will you fix?
15. Types of dimensions
16. What is slowly changing dimension
17. Query execution plan
18. Massive parallel processing
19. Trigger dimension
20. Hash distribution
21. Round robin distribution
22. Output of all joins in below tables.

Tabl 1

ID
1
1
1
Null

Tbl2

ID
1
1
2
3

23. Types of fact
24. Degenerate dimension
25. Calculate the total gmv, total quantity sold, number of customers at brand, gender, article type level?
26. Calculate the difference between the first purchase date and second purchase date at:
 - i. Platform level
 - ii. Platform mastercategory level for 'footwear', 'Apparel'
27. Calculate the average difference between the subsequent purchases at:
 - i. Platform level
 - ii. Platform mastercategory level for 'footwear', 'Apparel'
28. Find the gmv, customer count of the 8 personas
29. Calculate the the revenue of article type gender in decreasing order of each brand.
30. Find out the GMV, Quantity, customer count of those customers who purchased only two times in the last two years.

31. Calculate Platform gmv, brand gmv, brand total discount, brand total quantity, Platform frequency, brand frequency, ISP brand.

32. For dresses women find customer count, Platform acquisition customer count, gmv, quantity sold at business unit level.

33. Calculate style_id, article_type, gender, season, article_mrp, base_colour, style_attributes before date 20190331.

34. Find master_category, order_created_date, purchase sequence for all Men customers where mastercategory is 'Apparel', 'Footwear' for dates between 20170401 and 20190331

35. What is the preferred G-AT in terms of GMV, Quantity, and Frequency for all the customers who are:

- i. New to Brand New to Platform?
- ii. New to Brand Repeat to Platform
- iii. Repeat to Brand Repeat to Platform

36. For Private Label Women which persona is producing the highest revenue?
