

Data Analysis Online Retail Sales

Flipmart

Flip Mart Dataset

SQL Analysis

SQL Question Addressed:

1. Write a SQL query to retrieve all columns for sales made on '2022-11-05'.
2. Write a SQL query to retrieve all transactions where the category is 'Clothing' and the quantity sold is equal or more than 4 in the month of 'Nov-22'.
3. Write a SQL query to calculate the total sales for each category.
4. Write a SQL query to find the average age of customers who purchased items from 'Beauty' category.
5. Write a SQL query to find all transactions where the total sales is greater than 1000.
6. Write a SQL query to find the total number of transactions made by each gender in each category.
7. Write a SQL query to calculate the average sale of each month. Find out best selling month in each year.
8. Write a SQL query to find the top 5 customers based on the highest total sales.
9. Write a SQL query to find the number of unique customers who purchased items from each category.
10. Write a SQL query to create each shift and number of orders (Example Morning ≤ 12 , Afternoon Between 12 and 17, Evening < 17).

1. Write a SQL query to retrieve all columns for sales made on '2022-11-05'.

Input Query

```
1  -- Write sql query to retrieve all columns for sales made on 2022-11-05
2  •  SELECT
3      *
4  FROM
5      retail_sales
6  WHERE
7      sale_date = '2022-11-05'
```

Result

	transactions_id	sale_date	sale_time	customer_id	gender	age	category	quantity	price_per_unit	cogs	total_sale
▶	180	2022-11-05	10:47:00	117	Male	41	Clothing	3	300	129	900
	214	2022-11-05	16:31:00	53	Male	20	Beauty	2	30	8.1	60
	240	2022-11-05	11:49:00	95	Female	23	Beauty	1	300	123	300
	856	2022-11-05	17:43:00	102	Male	54	Electronics	4	30	9.3	120
	943	2022-11-05	19:29:00	90	Female	57	Clothing	4	300	318	1200
	1137	2022-11-05	22:34:00	104	Male	46	Beauty	2	500	145	1000
	1256	2022-11-05	09:58:00	29	Male	23	Clothing	2	500	190	1000
	1265	2022-11-05	14:35:00	86	Male	55	Clothing	3	300	111	900
	1587	2022-11-05	20:06:00	140	Female	40	Beauty	4	300	105	1200
	1819	2022-11-05	20:44:00	83	Female	35	Beauty	2	50	13.5	100
	1896	2022-11-05	20:19:00	87	Female	30	Electronics	2	25	30.75	50

2. Write a SQL query to retrieve all transactions where the category is 'Clothing' and the quantity sold is equal or more than 4 in the month of 'Nov-22'.

Input Query

```
1  -- Write a sql query to retrieve where the category is
2  • SELECT
3      *
4  FROM
5      retail_sales
6  WHERE
7      category = 'Clothing' AND quantity >= 4
8      AND DATE_FORMAT(sale_date, '%b-%y') = 'Nov-22'
```

Result

	transactions_id	sale_date	sale_time	customer_id	gender	age	category	quantity	price_per_unit	cogs	total_sale
▶	64	2022-11-15	06:34:00	7	Male	49	Clothing	4	25	8.5	100
	146	2022-11-10	22:01:00	74	Male	38	Clothing	4	50	49	200
	159	2022-11-10	21:30:00	42	Male	26	Clothing	4	50	23.5	200
	284	2022-11-12	09:17:00	129	Male	43	Clothing	4	50	20.5	200
	547	2022-11-14	07:36:00	3	Male	63	Clothing	4	500	250	2000
	699	2022-11-21	22:21:00	129	Female	37	Clothing	4	30	16.2	120
	735	2022-11-26	21:38:00	153	Female	64	Clothing	4	500	515	2000
	943	2022-11-05	19:29:00	90	Female	57	Clothing	4	300	318	1200
	965	2022-11-27	21:45:00	84	Male	22	Clothing	4	50	13	200
	1259	2022-11-03	17:31:00	105	Female	45	Clothing	4	50	21	200
	1296	2022-11-26	20:42:00	45	Female	22	Clothing	4	300	342	1200
	1476	2022-11-11	22:27:00	130	Female	27	Clothing	4	500	555	2000
	1484	2022-11-23	09:29:00	22	Female	19	Clothing	4	300	147	1200
	1497	2022-11-19	21:44:00	109	Male	41	Clothing	4	30	32.4	120
	1615	2022-11-17	13:43:00	82	Female	61	Clothing	4	25	13.5	100
	1696	2022-11-21	17:59:00	24	Female	50	Clothing	4	50	55	200
	1885	2022-11-09	07:32:00	148	Female	52	Clothing	4	30	10.8	120

3. Write a SQL query to calculate the total sales for each category.

Input Query

```
1  -- Write a sql query to calculate
2  • SELECT
3      category,
4      SUM(total_sale) AS Net_Sales,
5      COUNT(*) AS Total_Orders
6  FROM
7      retail_sales
8  GROUP BY category
```

Result

Result Grid			
Filter Rows:			
	category	Net_Sales	Total_Orders
▶	Beauty	286790	611
	Clothing	309995	698
	Electronics	311445	678

4. Write a SQL query to find the average age of customers who purchased items from 'Beauty' category.

Input Query

```
1  -- Write a SQL query to find the averag
2  •  SELECT
3      category,
4      ROUND(AVG(age), 2) AS Average_Age,
5      COUNT(*) AS Total_customers
6  FROM
7      retail_sales
8  WHERE
9      category = 'Beauty'
```

Result

Result Grid			
Filter Rows:			
	category	Average_Age	Total_customers
▶	Beauty	40.42	611

5. Write a SQL query to find all transactions where the total sales is greater than 1000.

Input Query

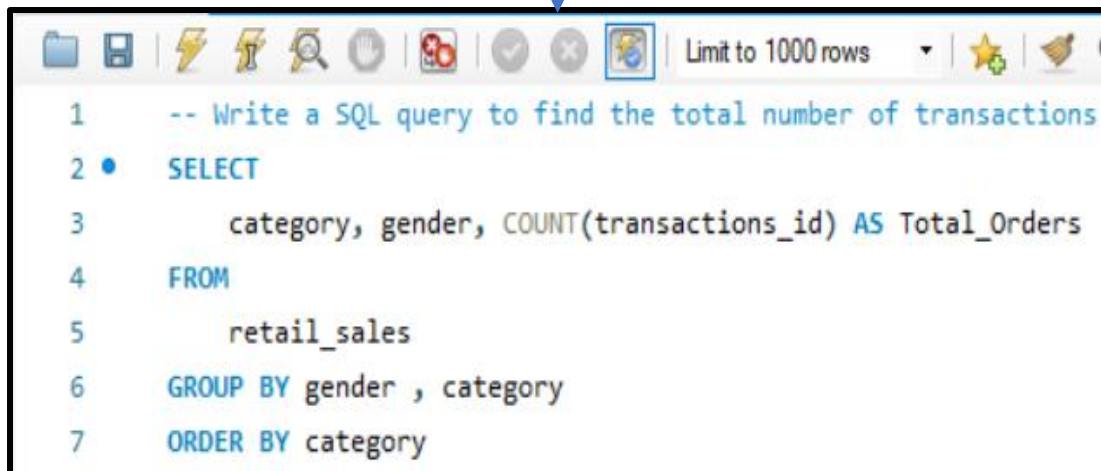
```
1  -- Write a SQL query to find all t
2  •  SELECT
3      total_sale, COUNT(total_sale)
4  FROM
5      retail_sales
6  WHERE
7      total_sale >= 1000
8  GROUP BY total_sale
```

Result

Result Grid			Filter Rows:
	total_sale	COUNT(total_sale)	
▶	1000	96	
	1500	100	
	2000	98	
	1200	108	

6. Write a SQL query to find the total number of transactions made by each gender in each category.

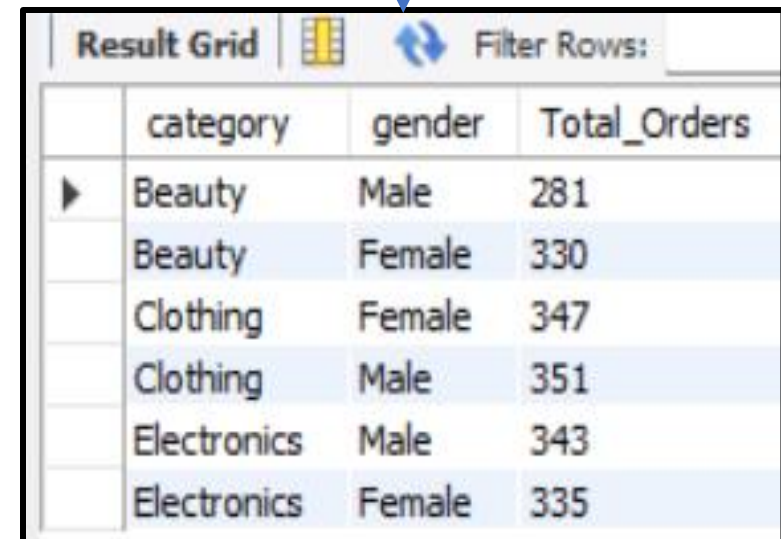
Input Query



The screenshot shows a SQL query editor with a toolbar at the top containing icons for file operations, execution, and search. The query text is as follows:

```
1  -- Write a SQL query to find the total number of transactions
2  •  SELECT
3      category, gender, COUNT(transactions_id) AS Total_Orders
4  FROM
5      retail_sales
6  GROUP BY gender , category
7  ORDER BY category
```

Result



The screenshot shows a result grid with a toolbar at the top containing icons for grid operations and a filter. The data is presented in a table with the following structure:

	category	gender	Total_Orders
▶	Beauty	Male	281
	Beauty	Female	330
	Clothing	Female	347
	Clothing	Male	351
	Electronics	Male	343
	Electronics	Female	335

7. Write a SQL query to calculate the average sale of each month. Find out best selling month in each year.

Input Query

```
1  -- Write a SQL query to calculate the average sale of each month. Find out best selling month in each year
2  •  select year, month, Average_Sale from
3  (select year(sale_date) as year, month(sale_date) as month, avg(total_sale) as Average_Sale, rank() over(partition by year(sale_date) order by avg(total_sale)desc) as rn
4  from retail_sales
5  group by year(sale_date), month(sale_date) ) as t1
6  where rn = 1
```

Result

Result Grid			
	year	month	Average_Sale
▶	2022	7	541.3414634146342
	2023	2	535.531914893617

8. Write a SQL query to find the top 5 customers based on the highest total sales.

Input Query

```
1  -- Write a SQL query to find the top 5 customers
2  •  SELECT
3      customer_id, SUM(total_sale) AS Total_Sales
4  FROM
5      retail_sales
6  GROUP BY customer_id
7  ORDER BY Total_Sales DESC
8  LIMIT 5;
```

Result

	customer_id	Total_Sales
▶	3	38440
	1	30750
	5	30405
	2	25295
	4	23580

9. Write a SQL query to find the number of unique customers who purchased items from each category.

Input Query

```
-- Write a SQL query to find the number of unique customers who purchased items from each category.
SELECT
    category, COUNT(DISTINCT customer_id) AS customers
FROM
    retail_sales
GROUP BY category
ORDER BY customers DESC;
```

Result

	category	customers
►	Clothing	149
	Electronics	144
	Beauty	141

10. Write a SQL query to create each shift and number of orders (Example Morning ≤ 12 , Afternoon Between 12 and 17, Evening < 17).

Input Query

```
-- Write a SQL query to create each shift and number of orders
with hourly_sale as
(select *,
 case when hour(sale_time) < 12 then 'Morning'
 when hour(sale_time) between 12 and 17 then 'Afternoon'
 else 'Evening'
 end as shift
 from retail_sales)
select shift, count(transactions_id) AS Total_Orders
from hourly_sale
group by shift
```

Result

	shift	Total_Orders
►	Evening	1062
	Morning	548
	Afternoon	377