SQL CASE STUDY

Customer Insights (Challenge-3)

Presented By- Binoy Patra

DATACOACH

CONTENTS

01

Ovewrview

Problem Statements

Basically I am a Customer Insights Analyst for 'The General Store'

I am asked to analyse the following tables to find out crucial information

about our customers to provide to my marketing team.

02
Tables

There are mainly 5 tables:

customers

customer_id	first_shop	age	rewards	can_email
1	2022-03-20	23	yes	no
2	2022-03-25	26	no	no
3	2022-04-06	32	no	no
4	2022-04-13	25	yes	yes
5	2022-04-22	49	yes	yes
6	2022-06-18	28	yes	no
7	2022-06-30	36	no	no
8	2022-07-04	37	yes	yes

orders

order_id	customer_id	date_shop	sales_channel	country_id
1	1	2023-01-16	retail	1
2	4	2023-01-20	retail	1
3	2	2023-01-25	retail	2
4	3	2023-01-25	online	1
5	1	2023-01-28	retail	3
6	5	2023-02-02	online	1
7	6	2023-02-05	retail	1
8	3	2023-02-11	online	3

baskets

order_id	product_id
1	1
1	2
1	5
2	4
3	3
4	2
4	1
5	3
5	5
6	4
6	3
6	1
7	2
7	1
8	3
8	3

products

product_id	category	price
1	food	5.99
2	sports	12.49
3	vitamins	6.99
4	food	0.89
5	vitamins	15.99

country

country_id	country_name	head_office
1	UK	London
2	USA	New York
3	China	Beijing

1. What are the names of all the countries in the country table?

SQL Statements:

```
SELECT C.country_name
FROM datacoach.country C;
```

Output:

	country_name
▶	UK
	USA
	China

2. What is the total number of customers in the customers table?

SQL Statements:

```
SELECT COUNT(*) AS Total_Number_of_Customer
FROM datacoach.customers ;
```

Output:

	Total_Number_of_Customer
▶	8

3. What is the average age of customers who can receive marketing emails (can_email is set to 'yes')?

SQL Statements:

```
SELECT round((avg(c.age)),2) as Average_age
FROM datacoach.customers C
WHERE C.can_email= 'yes';
```

Output:

	Average_age
▶	37.00

4. How many orders were made by customers aged 30 or older?

SQL Statements:

```
SELECT count(o.order_id) as NumberOrders
FROM datacoach.customers c
join datacoach.orders o on c.customer_id=o.customer_id
WHERE c.age>=30;
```

Output:

	NumberOrders
▶	3

5. What is the total revenue generated by each product category?

SQL Statements:

```
SELECT P.category, SUM(P.price) as Revenue
FROM datacoach.products P
JOIN datacoach.baskets B ON P.product_id = B.product_id
JOIN datacoach.orders O ON B.order_id=O.order_id
GROUP BY P.category;
```

Output:

	category	Revenue
▶	food	25.74
	sports	37.47
	vitamins	66.93

6. What is the average price of products in the 'food' category?

SQL Statements:

```
SELECT round((avg(p.price)),3) as Average_Price
FROM datacoach.products P
WHERE P.category='Food'
group by p.category;
```

Output:

	Average_Price
▶	3.440

7. How many orders were made in each sales channel (sales_channel column) in the orders table?

SQL Statements:

```
SELECT sales_channel, COUNT(order_id) AS Numbers_order
FROM datacoach.orders
GROUP BY sales_channel;
```

Output:

	sales_channel	Numbers_order
▶	retail	5
	online	3

8.What is the date of the latest order made by a customer who can receive marketing emails?

SQL Statements:

```
SELECT max(o.date_shop) as LatestOrderDate
FROM datacoach.orders O
JOIN datacoach.customers C ON O.customer_id=C.customer_id
WHERE C.can_email='yes';
```

Output:

	LatestOrderDate
▶	2023-02-02

9. What is the name of the country with the highest number of orders?

SQL Statements:

```
SELECT C.country_id, C.country_name, COUNT(O.order_id) AS NumberOrder
FROM datacoach.country C
JOIN datacoach.orders O ON C.country_id=O.country_id
GROUP BY C.country_id, C.country_name
order by NumberOrder DESC
LIMIT 1;
```

Output:

	country_id	country_name	NumberOrder
▶	1	UK	5

10. What is the average age of customers who made orders in the 'vitamins' product category?

SQL Statements:

```
SELECT round((avg(C.age)),2) AS Average_Age
FROM datacoach.customers C
JOIN datacoach.orders O ON C.customer_id=O.customer_id
JOIN datacoach.baskets B ON O.order_id=B.order_id
JOIN datacoach.products P ON P.product_id=B.product_id
WHERE P.category= 'Vitamins';
```

Output:

	Average_Age
▶	29.71


THANK YOU