GITHUB.COM/IAMTOPU

RETAIL STORE SALES ANALYSIS ON SQL

PRESENTED BY **H.M. KHALED AMBIA TOPU**

INTRODUCTION

Hello, I am H.M. Khaled Ambia Topu. In this project, I have utilized SQL queries to solve questions that are related to Retail Store Sales to analyze their performance.

WRITE A SQL QUERY TO RETRIEVE ALL COLUMNS FOR SALES MADE ON '2022-11-05.

```sql
SELECT * FROM retail_sales
WHERE sale_date = '2022-11-05';
```

	transactions_id [PK] integer	sale_date date	sale_time time without time zone	customer_id integer	gender character varying (10)	age integer
1	180	2022-11-05	10:47:00	117	Male	41
2	240	2022-11-05	11:49:00	95	Female	23
3	1256	2022-11-05	09:58:00	29	Male	23
4	1587	2022-11-05	20:06:00	140	Female	40
5	1819	2022-11-05	20:44:00	83	Female	35
6	943	2022-11-05	19:29:00	90	Female	57
7	1896	2022-11-05	20:19:00	87	Female	30
8	1137	2022-11-05	22:34:00	104	Male	46
9	856	2022-11-05	17:43:00	102	Male	54
10	214	2022-11-05	16:31:00	53	Male	20
11	1265	2022-11-05	14:35:00	86	Male	55

WRITE A SQL QUERY TO RETRIEVE ALL TRANSACTIONS WHERE THE CATEGORY IS 'CLOTHING' AND THE QUANTITY SOLD IS MORE THAN 10 IN THE MONTH OF NOV-2022.

```sql
SELECT * FROM retail_sales
WHERE
category = 'Clothing'
AND
quantiy > 3
AND
TO_CHAR(sale_date, 'YYYY-MM') = '2022-11'
```

	transactions_id [PK] integer	sale_date date	sale_time time without time zone	customer_id integer	gender character varying (10)	age integer
1	1484	2022-11-23	09:29:00	22	Female	19
2	64	2022-11-15	06:34:00	7	Male	49
3	284	2022-11-12	09:17:00	129	Male	43
4	1885	2022-11-09	07:32:00	148	Female	52
5	547	2022-11-14	07:36:00	3	Male	63
6	159	2022-11-10	21:30:00	42	Male	26
7	699	2022-11-21	22:21:00	129	Female	37
8	1259	2022-11-03	17:31:00	105	Female	45
9	146	2022-11-10	22:01:00	74	Male	38
10	1476	2022-11-11	22:27:00	130	Female	27
11	1296	2022-11-26	20:42:00	45	Female	22
12	1696	2022-11-21	17:59:00	24	Female	50
13	1497	2022-11-19	21:44:00	109	Male	41
14	735	2022-11-26	21:38:00	153	Female	64

WRITE A SQL QUERY TO CALCULATE THE TOTAL SALES (TOTAL_SALE) FOR EACH CATEGORY.

```sql
SELECT category, SUM(total_sale) AS total_sales
FROM retail_sales
GROUP BY category
```

	category character varying (20)	total_sales double precision
1	Electronics	311445
2	Clothing	309995
3	Beauty	286790

WRITE A SQL QUERY TO FIND THE AVERAGE AGE OF CUSTOMERS WHO PURCHASED ITEMS FROM THE 'BEAUTY' CATEGORY.

```sql
SELECT ROUND(AVG(age), 1) AS avg_age
FROM retail_sales
WHERE category = 'Beauty'
```

	avg_age numeric
1	40.4

WRITE A SQL QUERY TO FIND ALL TRANSACTIONS WHERE THE TOTAL_SALE IS GREATER THAN 1000.

```sql
SELECT transactions_id, customer_id, gender, age, category, total_sale
FROM retail_sales
WHERE total_sale > 1000
```

	transactions_id [PK] integer	customer_id integer	gender character varying (10)	age integer	category character varying (20)	total_sale double precision
1	522	52	Male	46	Beauty	1500
2	559	5	Female	40	Clothing	1200
3	1522	48	Male	46	Beauty	1500
4	1559	49	Female	40	Clothing	1200
5	421	66	Female	37	Clothing	1500
6	1421	59	Female	37	Clothing	1500
7	484	135	Female	19	Clothing	1200

WRITE A SQL QUERY TO FIND THE TOTAL NUMBER OF TRANSACTIONS (TRANSACTION_ID) MADE BY EACH GENDER IN EACH CATEGORY.

```sql
SELECT category, gender, COUNT(transactions_id) AS total_transactions
FROM retail_sales
GROUP BY 1, 2
ORDER BY 1
```

	category character varying (20)	gender character varying (10)	total_transactions bigint
1	Beauty	Female	330
2	Beauty	Male	281
3	Clothing	Female	347
4	Clothing	Male	351
5	Electronics	Male	343
6	Electronics	Female	335

WRITE A SQL QUERY TO CALCULATE THE AVERAGE SALE FOR EACH MONTH. FIND OUT BEST SELLING MONTH IN EACH YEAR.

```sql
SELECT year, month, avg_sale
FROM
    (SELECT
        EXTRACT (YEAR FROM sale_date) as year,
        EXTRACT (MONTH FROM sale_date) as month,
        AVG(total_sale) AS avg_sale,
        RANK() OVER(PARTITION BY EXTRACT (YEAR FROM sale_date) ORDER BY AVG(total_sale) DESC) AS rank
    FROM retail_sales
    GROUP BY 1, 2)
WHERE rank = '1'
```

	year numeric	month numeric	avg_sale double precision
1	2022	7	541.3414634146342
2	2023	2	535.531914893617

WRITE A SQL QUERY TO FIND THE TOP 5 CUSTOMERS BASED ON THE HIGHEST TOTAL SALES.

```sql
SELECT customer_id, SUM(total_sale) AS highest_sales
FROM retail_sales
GROUP BY customer_id
ORDER BY highest_sales DESC
LIMIT 5
```

	customer_id integer	highest_sales double precision
1	3	38440
2	1	30750
3	5	30405
4	2	25295
5	4	23580

WRITE A SQL QUERY TO FIND THE NUMBER OF UNIQUE CUSTOMERS WHO PURCHASED ITEMS FROM EACH CATEGORY.

```sql
SELECT
      category,
      COUNT(DISTINCT customer_id) AS unique_customer_count
FROM retail_sales
GROUP BY category
```

	category character varying (20)	unique_customer_count bigint
1	Beauty	141
2	Clothing	149
3	Electronics	144

WRITE A SQL QUERY TO CREATE EACH SHIFT AND NUMBER OF ORDERS (EXAMPLE MORNING <=12, AFTERNOON BETWEEN 12 & 17, EVENING >17).

```sql
WITH hourly_sale
AS
(
SELECT *,
    CASE
    WHEN EXTRACT(HOUR FROM sale_time) < 12 THEN 'Morning'
    WHEN EXTRACT(HOUR FROM sale_time) BETWEEN 12 AND 17 THEN 'Afternoon'
    ELSE 'Evening'
    END as shift
FROM retail_sales
)
SELECT
    shift,
    COUNT(*) AS  total_orders
FROM hourly_sale
GROUP BY shift
```

	shift text	total_orders bigint
1	Afternoon	377
2	Evening	1062
3	Morning	548

GITHUB.COM/IAMTOPU

THANK YOU