TODAY'S GOOD MOOD IS SPONSORED BY COFFEE

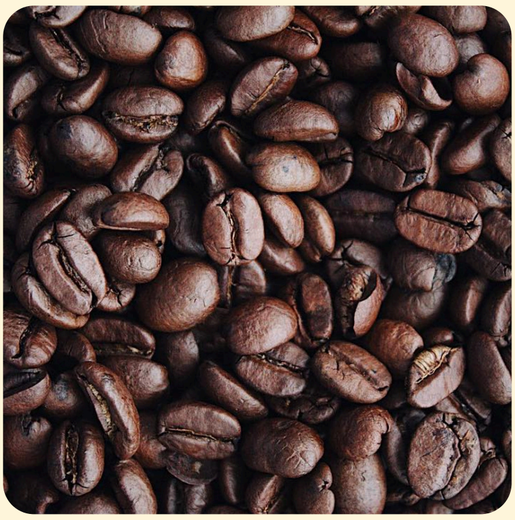
"ESPRESSO INSIGHTS: DATA-DRIVEN COFFEE SALES ANALYSIS WITH SQL"

Presented by : Priti Patil

OBJECTIVE

The business plans to grow by establishing three coffee shops in India’s top three major cities. Since its inception in January 2023, the company has achieved significant success in online product sales, receiving highly positive feedback from various cities. As a data analyst, your responsibility is to evaluate the sales data and deliver insights to recommend the best three cities for this expansion.

DOCUMENTATION



DOCUMENTATION

DATABASE SCEMA

QUESTIONS

- **Coffee Consumers Count**

How many people in each city are estimated to consume coffee, given that 25% of the population does?

```
SELECT city_name,
ROUND(
(population * 0.25)/1000000, 2) AS coffee_consumers_in_millions,
city_rank
FROM city
ORDER BY 2 DESC;
```

- **Sales Count for Each Product**

How many units of each coffee product have been sold?

```
SELECT product_name , sum(s.total) as total_sales
FROM products AS p
JOIN sales AS s ON p.product_id = s.product_id
GROUP BY product_name
ORDER BY total_sales DESC;
```

- **High Rating**

List customers who have given a rating of 5.

```
SELECT customer_name , rating
FROM customers as c
JOIN sales AS s ON c.customer_id= s.customer_id
WHERE rating = 5;
```

- **Total Revenue from Coffee Sales**

What is the total revenue generated from coffee sales across all sales.

```
SELECT sum(total) AS total_revenue
FROM sales;
```

QUESTIONS

Find all sales details for customers who are from cities ranked in the top 2.

```sql
SELECT s.sale_id, s.sale_date, s.product_id, s.customer_id,
c.customer_name, s.total, s.rating , a.city_rank
FROM sales AS s
JOIN customers AS c ON s.customer_id = c.customer_id
JOIN products AS p ON p.product_id = s.product_id
JOIN city AS a ON a.city_id = c.city_id
WHERE a.city_rank <= 2;
```

Identify cities where the average rent is greater than $2,000 and total sales exceed $10,000.

```sql
SELECT city_name, AVG (estimated_rent) AS average_rent,
SUM(s.total) AS total_sales
FROM sales AS s
JOIN customers AS c ON s.customer_id = c.customer_id
JOIN city AS a ON a.city_id = c.city_id
GROUP BY city_name
HAVING  AVG (a.estimated_rent) > 2000  AND SUM(s.total) > 10000;
```

List all products sold on '2023-01-01' along with their total sales and ratings.

```sql
SELECT p.product_name, s.sale_date,
SUM(s.total) AS total_sales,
AVG(s.rating) AS avg_rating
FROM sales AS s
JOIN products AS p ON s.product_id = p.product_id
WHERE s.sale_date = '2023-01-01'
GROUP BY p.product_name;
```

Determine the correlation between city rank and total revenue generated in that city.

```sql
SELECT a.city_rank, SUM(s.total) AS total_revenue
FROM sales AS s
JOIN customers AS c ON s.customer_id = c.customer_id
JOIN city AS a ON a.city_id = c.city_id
GROUP BY a.city_rank
ORDER BY a.city_rank;
```

QUESTIONS

Find Total Revenue Per City

```sql
SELECT a.city_name , sum(s.total) AS total_revenue
FROM city AS a
JOIN customers AS c ON a.city_id =  c.city_id
JOIN sales AS s ON s.customer_id = c.customer_id
GROUP BY a.city_name
ORDER BY total_revenue DESC;
```

Identify customers with high spending but low ratings (dissatisfied)

```sql
SELECT c.customer_name, SUM(s.total) AS total_spent,
AVG(s.rating) AS avg_rating
FROM sales s
JOIN customers AS c ON s.customer_id = c.customer_id
GROUP BY c.customer_id, c.customer_name
HAVING AVG(s.rating) < 3 AND SUM(s.total) > 50;
```

Average Rent Per Customer Per City

```sql
SELECT a.city_name,
ROUND(SUM(a.estimated_rent) /
COUNT(DISTINCT c.customer_id), 2) AS avg_rent_per_customer
FROM city a
JOIN customers c ON a.city_id = c.city_id
JOIN sales s ON c.customer_id = s.customer_id
GROUP BY a.city_name
ORDER BY avg_rent_per_customer ASC;
```

Find cities with more population, low total revenue and less customers.

```sql
SELECT a.city_name, a.population,
    COUNT(DISTINCT c.customer_id) AS total_customers,
    SUM(s.total) AS total_revenue
FROM city a
LEFT JOIN customers c ON a.city_id = c.city_id
LEFT JOIN sales s ON c.customer_id = s.customer_id
GROUP BY a.city_name, a.population
HAVING a.population > 1000000
    AND total_revenue < 500000
    AND total_customers < 50
ORDER BY  a.population DESC, total_revenue ASC;
```

RECOMMENDATIONS & REASONS

- **City 1: Pune**

Average rent per customer is very low.
Highest total revenue.
Average sales per customer is also high.

- **City 2: Delhi**

Highest estimated coffee consumers at 7.7 million
Highest total number of customers, which is 68.
Average rent per customer is 330.

- **City 3: Jaipur**

Highest number of customers, which is 69
Average rent per customer is very low at 156
Average sales per customer is better.

- **Improve Customer Satisfaction**

Identify customers with low ratings but high spending. Address their concerns via personalized services or feedback collection.

THANK YOU