INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

Grocery Market Analysis Using MySQL

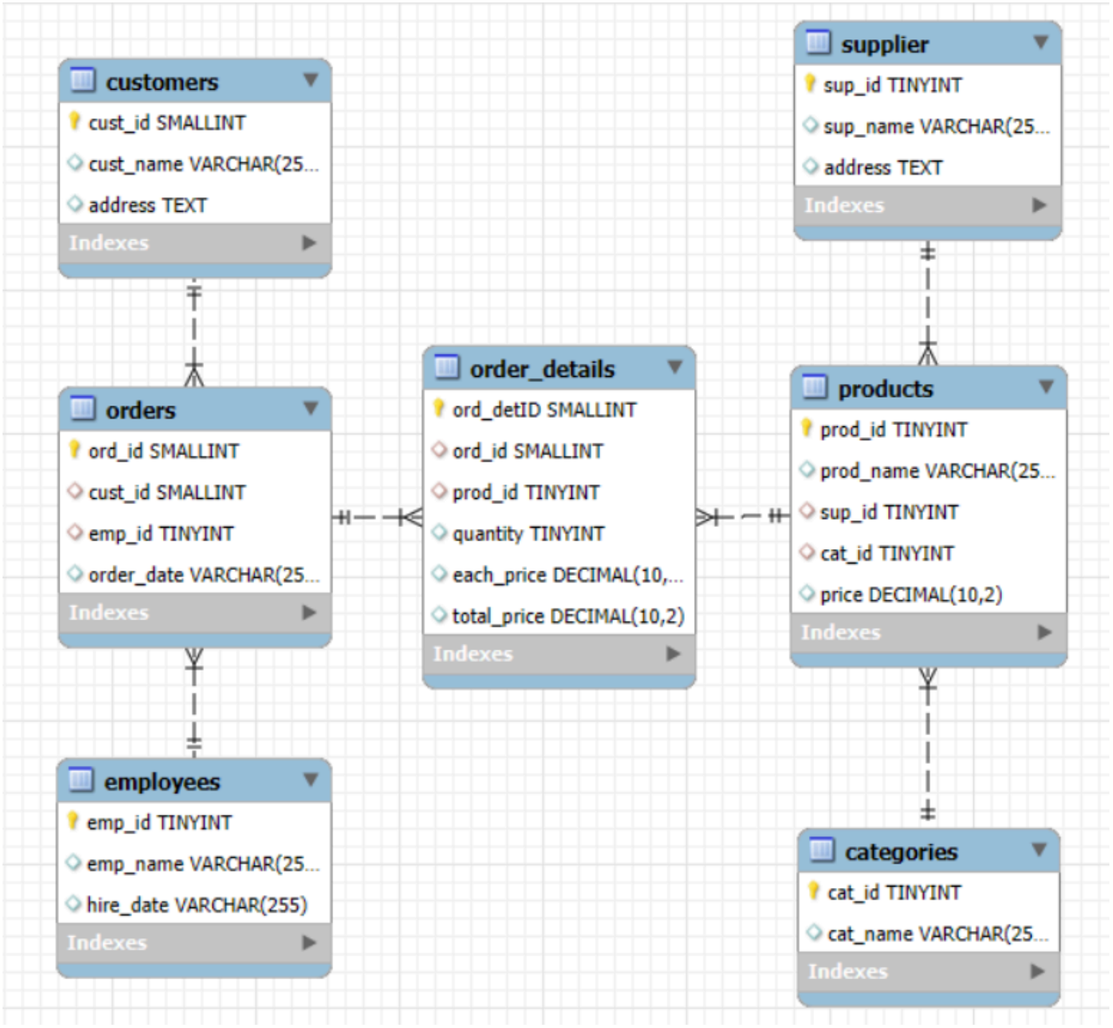
Presented By
PAVAN AHIRE

Objective of the Project

- **Design & Implement Database:** Build a relational database for managing suppliers, customers, products, employees, and orders.
- **Data Retrieval & Manipulation:** Use SQL queries (joins, subqueries, aggregations) to extract meaningful information.
- **Business Insights:** Analyse customer behaviour, sales trends, product performance, and supplier contributions.
- **Decision Support:** Provide insights to improve revenue, inventory management, and employee efficiency.
- **Hands-on SQL Practice:** Strengthen skills in schema design, query optimization, and data-driven problem solving.

INNOMATICS RESEARCH LABS

Entity Relational Diagram (ERD)



- **Customers –** Store customer details; one customer can place many orders.
- **Employees –** Manage orders; one employee can handle many orders.
- **Orders –** Record customer orders, linked to customers and employees.
- **Order_Details –** Capture product quantity and pricing for each order.
- **Products –** Store product info, linked to suppliers and categories.
- **Suppliers –** Provide products; one supplier can supply many items.
- **Categories –** Define product groups; one category can contain many products.

INNOMATICS RESEARCH LABS

Key analysis questions (use cases)

1. Who are the top 5 customers by total purchase amount?

```sql
SELECT c.cust_id, c.cust_name,
       SUM(od.total_price) AS total_purchase
FROM customers c
JOIN orders o ON c.cust_id = o.cust_id
JOIN order_details od ON o.ord_id = od.ord_id
GROUP BY c.cust_id, c.cust_name
ORDER BY total_purchase DESC
LIMIT 5;
```

OUTPUT:

cust_id	cust_name	total_purchase
19	Chetan Naidu	11256.82
166	Kapila	11099.51
67	Eshwar Rao	10819.96
61	Aditi Rao	10230.64
7	Eshwar Iyer	9188.45

Observation

The analysis shows that *Chetan Naidu* is the top customer, followed by Kapila and Eshwar Rao, each with purchases above 10K.
This shows that a few loyal customers make a large share of revenue.

INNOMATICS RESEARCH LABS

2. How many products exist in each category?

```sql
SELECT c.cat_id, c.cat_name, COUNT(p.prod_id) AS total_products
FROM categories c
LEFT JOIN products p ON c.cat_id = p.cat_id
GROUP BY c.cat_id, c.cat_name;
```

OUTPUT:

cat_id	cat_name	total_products
1	Grains & Cereals	18
2	Dairy Products	6
3	Snacks & Confectioneries	17
4	Personal Care	6
5	Household	3

Observation

The results show that *Grains & Cereals* and *Snacks & Confectioneries* dominate with the highest number of products. Meanwhile, Household has the least with only 3 products. This means the store is more focused on food items compared to household or personal care.

3. What are the monthly trends in order volume and revenue?

```
SELECT DATE_FORMAT(STR_TO_DATE(order_date, '%m/%d/%Y'), '%m-%Y') AS month,
       COUNT(DISTINCT o.ord_id) AS order_count,
       SUM(od.total_price) AS total_revenue
FROM orders o
JOIN order_details od ON o.ord_id = od.ord_id
GROUP BY month
ORDER BY month;
```

OUTPUT:

month	order_count	total_revenue
01-2022	30	70312.45
02-2022	28	66929.42
03-2022	27	45977.16
04-2022	11	29118.54
05-2022	19	41305.62

Observation

January had the highest sales, followed by February. Sales dropped in March and reached the lowest in April, but slightly improved in May. Overall, there is a downward trend in orders and revenue during this period.

4. Which suppliers contribute the most to total product sales (by revenue)?

```
SELECT s.sup_id, s.sup_name, SUM(od.total_price) AS total_revenue
FROM supplier s
JOIN products p ON s.sup_id = p.sup_id
JOIN order_details od ON p.prod_id = od.prod_id
GROUP BY s.sup_id, s.sup_name
ORDER BY total_revenue DESC;
```

OUTPUT:

sup_id	sup_name	total_revenue
3	Aarya	221137.83
2	Sai	113588.51
4	Suresh	101688.78
5	Karthik	81861.96
1	Aarav Sharma	33052.85

Observation

The results show that *Aarya* is the top supplier with the highest revenue of over 2 lakh, which is much higher than the others. This means the business is heavily dependent on one main supplier, which can be risky.

5. What is the total sales value processed by each employee?

```
SELECT e.emp_id, e.emp_name, SUM(od.total_price) AS total_sales
FROM employees e
JOIN orders o ON e.emp_id = o.emp_id
JOIN order_details od ON o.ord_id = od.ord_id
GROUP BY e.emp_id, e.emp_name
ORDER BY total_sales DESC;
```

OUTPUT:

emp_id	emp_name	total_sales
2	Aditya Singh 1	79252.29
6	Zara Verma 1	71562.76
8	Diya Sharma 1	67241.85
3	Pari Kumar 1	66818.39
9	Arjun Kumar 1	54018.31

Observation

I also checked the total sales value handled by each employee.
Aditya Singh leads with the highest sales, followed by Zara Verma and Diya Sharma. This shows that a few employees are driving most of the sales performance.

6. What is the average quantity ordered per product?

```
SELECT p.prod_id, p.prod_name, AVG(od.quantity) AS avg_quantity
FROM products p
JOIN order_details od ON p.prod_id = od.prod_id
GROUP BY p.prod_id, p.prod_name
ORDER BY avg_quantity DESC;
```

OUTPUT:

prod_id	prod_name	avg_quantity
40	Butter	4.5556
31	Toothpaste	3.6667
46	Potato Chips	3.6000
42	Tomato Ketchup	3.5000
22	Mustard Seeds	3.4615

Observation

The results show that Butter has the highest average quantity per order, followed by Toothpaste and Potato Chips. This means these items are purchased more frequently in larger quantities than others.

7. What is the average price of products by category?

```
SELECT c.cat_id, c.cat_name, AVG(p.price) AS avg_price
FROM categories c
JOIN products p ON c.cat_id = p.cat_id
GROUP BY c.cat_id, c.cat_name;
```

OUTPUT:

cat_id	cat_name	avg_price
1	Grains & Cereals	287.673333
2	Dairy Products	366.943333
3	Snacks & Confectioneries	278.892353
4	Personal Care	364.991667
5	Household	363.336667

Observation

The results show that Dairy and Personal Care items are more expensive, while Grains and Snacks are cheaper. This is expected because essential food items usually cost less, while specialty and personal care products are priced higher.

Final Business Insights and Key Recommendations

- Final Business Insights:

- A small group of top customers contributes the majority of revenue.
- Grains & Snacks categories dominate the product mix.
- Business is highly dependent on a few key suppliers.
- Certain products and employees consistently drive higher sales.

- Key Recommendations:

- Launch loyalty programs to retain high-value customers.
- Expand weaker categories like Household and Personal Care.
- Onboard more suppliers to reduce dependency risks.
- Reward top-performing employees and ensure stock of high demand products.

INNOMATICS RESEARCH LABS

Conclusion

- In conclusion, this project helped me design and analyze a Grocery Store Management System using SQL.
- I learned how to create tables, write queries, and generate insights.
- The analysis highlighted customer behavior, product performance, supplier dependency, and employee contributions.
- Overall, it showed how SQL can help in making better business decisions.

THANK YOU

connect with me on

INNOMATICS RESEARCH LABS