# 🧊 Retail Sales Analytics Warehouse using Snowflake

## 📌 Project Summary

This project demonstrates how to design and implement a mini data warehouse using \*\*Snowflake\*\* to analyze retail sales across products, customers, and time periods. Built as a student portfolio project, it showcases dimensional modeling, SQL querying, and cloud data warehousing best practices.

## 🧠 Objective

Build a data warehouse to answer key business questions such as:

- Which products generate the most revenue?

- What is the monthly trend in sales?

- How do different customer age groups perform in terms of spending?

---

## 🗃️ Schema Design

This project uses a \*\*star schema\*\* with:

- \*\*FACT\_SALES\*\* (fact table)

- \*\*DIM\_PRODUCT\*\*, \*\*DIM\_CUSTOMER\*\*, \*\*DIM\_TIME\*\* (dimension tables)

![Schema Diagram](snowflake\_star\_schema.png)

---

## 📂 Data Files

| File Name | Table Name | Description |

|-------------------|----------------|----------------------------------|

| `dim\_product.csv` | DIM\_PRODUCT | Product details |

| `dim\_customer.csv` | DIM\_CUSTOMER | Customer info with age/location |

| `dim\_time.csv` | DIM\_TIME | Date breakdowns |

| `fact\_sales.csv` | FACT\_SALES | Transaction-level sales data |

---

## 💻 Technologies Used

- ❄️ \*\*Snowflake\*\* (on Azure, Pune)

- 🐍 Python (for data generation)

- 📄 CSV files (local upload)

- 🧾 SQL (Joins, Group By, Aggregates)

---

## 📊 Key SQL Queries & Insights

### 🔹 Total Sales per Product

```sql

SELECT

p.name AS product\_name,

SUM(f.quantity) AS total\_units\_sold,

SUM(f.total\_amount) AS total\_revenue

FROM FACT\_SALES f

JOIN DIM\_PRODUCT p ON f.product\_id = p.product\_id

GROUP BY p.name

ORDER BY total\_revenue DESC;

```

### 🔹 Monthly Sales Trend

```sql

SELECT

t.month,

SUM(f.total\_amount) AS monthly\_revenue

FROM FACT\_SALES f

JOIN DIM\_TIME t ON f.time\_id = t.time\_id

GROUP BY t.month;

```

### 🔹 Sales by Customer Age Group

```sql

SELECT

c.age\_group,

SUM(f.total\_amount) AS total\_sales

FROM FACT\_SALES f

JOIN DIM\_CUSTOMER c ON f.customer\_id = c.customer\_id

GROUP BY c.age\_group

ORDER BY total\_sales DESC;

```

---

## 📈 Future Improvements

- Add new dimensions like Store or Region

- Automate data ingestion using Snowpipe

- Build dashboards with Power BI or Streamlit

---

## 📁 Folder Structure

```

snowflake-retail-sales-warehouse/

├── data/

│ ├── dim\_product.csv

│ ├── dim\_customer.csv

│ ├── dim\_time.csv

│ └── fact\_sales.csv

├── sql/

│ └── analysis\_queries.sql

├── screenshots/

│ └── [add your screenshots here]

├── snowflake\_star\_schema.png

└── README.md

```