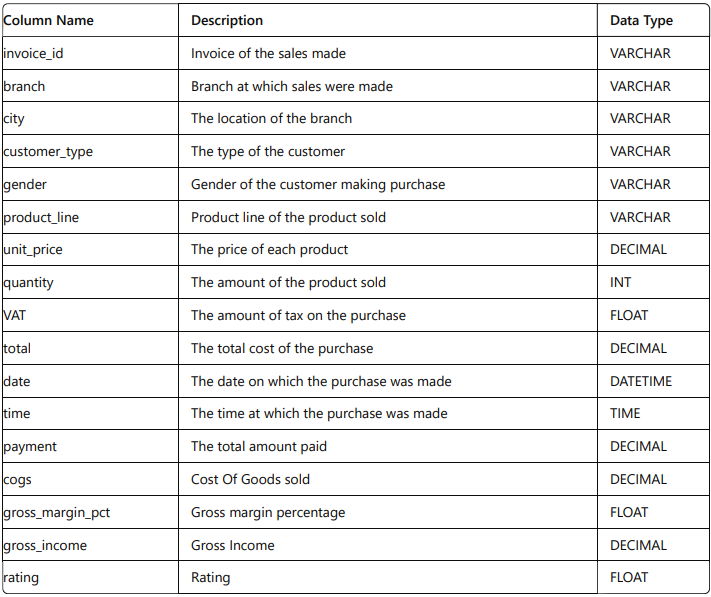
***Superstore Sales Analysis Report***

**Purposes of the Project:**

To analyse Superstore Sales data using SQL to uncover trends, patterns, and correlations that can provide valuable insights and actionable information to improve business performance and decision-making.

The primary purpose of this project is for learning and practicing SQL.

**About Data:**

****

**Step-by-Step Project Approach:**

**1. Build a Database**

* Database Creation: Establish a database structure for storing sales data.
* Table Creation: Define a table schema with fields like invoice\_id, branch, product\_line, etc.
* Data Insertion: Insert the data. Null values are not present in our database because, in creating the tables, NOT NULL was specified for each field, effectively filtering out any null values.

**2. Feature Engineering**

* Time of Day Analysis:
  + New Column: Create time\_of\_day to categorize sales into Morning, Afternoon, Evening.
  + Purpose: Understand peak sales periods during different times of the day.
* Day of Week Analysis:
  + New Column: Introduce day\_name to identify days (Mon, Tue, etc.) of transactions.
  + Purpose: Analyse which days see the highest transaction volumes per branch.
* Month of Year Analysis:
  + New Column: Add month\_name to capture transaction months (Jan, Feb, etc.).
  + Purpose: Determine which months contribute most to sales and profit.

**3. Exploratory Data Analysis (EDA)**

* Conducting exploratory data analysis is essential to address the project's objectives.

-- EXTRACT YEAR FROM DATE

select odate, year(odate) from sales;

-- add year column

alter table sales add column oyear int not null;

update sales set oyear = year(odate);

-- extract months from date

select odate, monthname(odate) from sales;

-- add month column

alter table sales add column omonth varchar(20) not null;

update sales set omonth = monthname(odate) ;

-- extract dayname from date

select otime, dayname(otime) from sales;

-- add day\_name column

alter table sales add column oday\_name varchar(20) not null;

update sales set oday\_name =dayname(odate) ;

-- divide time in three categories

select otime from sales;

select otime,

(case

when otime between "00:00:00" and "12:00:00" then "MORNING"

when otime between "12:00:01" and "16:00:00" then "AFTERNOON"

else "EVENING"

end) AS time\_of\_day

from sales;

-- add time of day column

alter table sales add column time\_of\_day varchar(20) not null;

update sales set time\_of\_day = (case

when otime between "00:00:00" and "12:00:00" then "MORNING"

when otime between "12:00:01" and "16:00:00" then "AFTERNOON"

else "EVENING"

end) ;

-- DISTINCT BREANCH

select distinct(branch) from sales;

-- A,C,b

-- DISTINCT CITY

select distinct(city) from sales;

-- yangon,naypyitaw,mandalay

-- DISTINCT CUSTOMER\_TYPE

select distinct(customer\_type) from sales;

-- normal,member

-- DISTINCT PRODUCT\_LINE

select distinct(product\_line) from sales;

-- food & brverages,health & beauty,sports & travel,fashions accessories,home & lifstyle,electtronic accessories

-- DISTINCT PAYMENT TYPE

select distinct(payment) as payment\_mode from sales;

-- credit card,ewallet,cash

-- What is the most selling product line

select product\_line,sum(quantity) as Qty from sales group by product\_line order by Qty desc; -- Electronic Accessories 971

-- What product line had the largest revenue

select product\_line,sum(total) as Revenue from sales group by product\_line order by Revenue desc;

-- Food and beverages 56144.8440

-- What product line had the largest tax\_pct

select product\_line,sum(tax\_pct) as VAT from sales group by product\_line order by VAT desc; -- Food and beverages 2673.563990712166

-- What product line had the largest cogs

select product\_line,sum(cogs) as COGS from sales group by product\_line order by COGS desc; -- Food and beverages 53471.28

-- What is the average rating of each product line

select product\_line,round(avg(rating),1) as Rating from sales group by product\_line order by Rating desc;

-- Food and beverages 7.1

-- What is the total revenue by month

select omonth as Month,sum(total) as TotalRevenue from sales group by Month order by TotalRevenue desc;

-- January 116291.8680

-- What month had the largest tax pct

select omonth as Month,sum(tax\_pct) as VAT from sales group by Month order by VAT desc; -- January 5537.708010196686

-- What month had the largest COGS

select omonth as Month,sum(cogs) as COGS from sales group by Month order by COGS desc; -- January 110754.16

-- What is the city with the largest revenue

select city,sum(total) as Revenue from sales group by city order by Revenue desc;

-- Naypyitaw 110568.7065

-- Which city has the largest tax percent

select city,sum(tax\_pct) as tax from sales group by city order by tax desc ;

-- Naypyitaw 5265.176480472088

-- What is the branch with the largest revenue

select branch,sum(total) as Revenue from sales group by branch order by Revenue desc;

-- C 110568.7065

-- Which branch has the largest tax percent

select branch,sum(tax\_pct) as tax from sales group by branch order by tax desc ;

-- C 5265.176480472088

-- Which of the customer types brings the most revenue

select customer\_type,sum(total) as Revenue from sales group by customer\_type order by Revenue;

-- Normal 158743.3050

-- Which of the customer types brings the most tax percent

select customer\_type,sum(tax\_pct) as tax from sales group by customer\_type order by tax desc ;

-- Member 7820.163996100426

-- What is the most common customer type

select customer\_type,count(invoice\_id) as icount from sales group by customer\_type order by icount desc;

-- Member 501

-- Which is the customer type buy most quantity

select customer\_type,sum(quantity) as QTY from sales group by customer\_type order by QTY desc;

-- Member 2785

-- What is the gender of most of the customers

select gender,count(invoice\_id) as icount from sales group by gender order by icount desc;

-- Female 501

-- Which is the gender give most of the revenue

select gender,sum(total) as Revenue from sales group by gender order by Revenue desc;

-- Female 167882.9250

-- Which month customers give most revenue

select omonth,sum(total) as Revenue from sales group by omonth order by Revenue desc;

-- January 116291.8680

-- which month get most tax

select omonth,sum(tax\_pct) as tax from sales group by omonth order by tax desc;

-- January 5537.708010196686

-- which month customer give most rating

select omonth,round(avg(rating),1) as rating from sales group by omonth order by rating desc; -- February 7.1

-- Which time of the day do customers give most ratings

select time\_of\_day,round(avg(rating),2) as rating from sales group by time\_of\_day order by rating desc;

-- AFTERNOON 7.03

-- which day name customer give most rating

select oday\_name as day\_name,round(avg(rating),1) as rating from sales group by day\_name order by rating desc;

-- Monday 7.2

-- which time of day customors shoping most

select time\_of\_day,count(invoice\_id) as icount from sales group by time\_of\_day order by icount desc;

-- EVENING 432

-- Which time of the day do customers give most revenue

select time\_of\_day,sum(total) as Revenue from sales group by time\_of\_day order by Revenue desc;

-- EVENING 138370.9215

-- Which day customers give most revenue

select oday\_name as day\_name,count(invoice\_id) as icount from sales group by day\_name order by icount desc;

-- Saturday 164

-- which is the most used payment type

select payment,count(\*) as payment\_type from sales group by payment order by payment\_type desc;

-- Ewallet 345

**4. Conclusion:**

**1. Product Performance:**

* **Most Selling Product Line**: Electronic Accessories. This indicates a strong demand or potentially effective marketing strategy for these products.
* **Highest Revenue Product Line**: Food and Beverages. Despite not being the most sold in quantity, it generates the highest revenue, suggesting higher price points or premium offerings.
* **Highest Tax Percentage Product Line**: Food and Beverages. This might be due to higher tax rates applied to food items in certain regions or jurisdictions.
* **Highest Cost of Goods Sold (COGS) Product Line:** Food and Beverages. This indicates potentially higher production or acquisition costs for these products.

**2. Revenue Analysis:**

* **City with the Largest Revenue:** Naypyitaw. This could indicate a larger customer base or higher average transaction values in that location**.**
* **Branch with the Largest Revenue:** Branch C. Understanding why this branch outperforms others could reveal insights into local market dynamics or operational effectiveness.

**3. Customer Insights:**

* **Most Revenue-Generating Customer Type:** Members. Despite being fewer in count compared to normal customers, members contribute significantly more revenue, indicating potential loyalty program effectiveness or higher spending per transaction.
* **Gender Insights:** Females contribute more revenue compared to males, suggesting potential targeting or product preference insights that could inform marketing strategies.

**4. Time and Day Analysis:**

* **Most Revenue by Month: January.** This could be due to seasonal factors, promotions, or sales events typically occurring early in the year.
* **Most Revenue by Day:** Saturday. Understanding why Saturdays are peak revenue days could inform staffing and inventory management strategies.
* **Time of Day Insights:** Evenings see the highest revenue and customer activity. This could suggest that evenings are prime shopping times or that more expensive items are purchased during this period.

**5. Payment Preferences:**

* **Most Used Payment Type:** Ewallet. This indicates a trend towards digital payments, potentially influencing future payment processing strategies and partnerships.

**6. Rating and Customer Satisfaction:**

* **Most Common Customer Rating:** 7.1, indicating generally positive customer feedback. However, understanding the factors contributing to lower ratings could provide insights into areas needing improvement.

**5.Strategic Implications:(based on chat-gpt)**

* **Targeting High Revenue Areas:** Allocate resources and marketing efforts towards Naypyitaw and Branch C to capitalize on their revenue potential.
* **Enhancing Product Offerings**: Further develop the Food and Beverages product line due to its high revenue and demand.
* **Customer Engagement:** Further incentivize and engage members to increase their share of wallet and encourage repeat business.
* **Optimizing Sales Timing:** Focus promotional efforts during evenings and weekends to maximize revenue during peak shopping times.