**SUPERSTORE SALES ANALYSIS**

**OBJECTIVE**

To analyse sales performance across different branches, products, and customer segments of a retail supermarket chain using Google BigQuery for data processing and Looker Studio for visual dashboard reporting.

**TOOLS USED**

* Google Big Query – for data storage, cleaning, and transformation
* Looker Studio – for dashboard and KPI visualization
* Google Cloud Storage – for dataset upload

**ETL PROCESS**

**EXTRACT**

Get the raw data from your local system into GCP.

Steps:

1. Open GCP TO Navigate to Cloud Storage.
2. Create a new bucket .
3. Upload the file into bucket .

**LOAD**

Load the raw file into BigQuery.

Steps

1. Go to BigQuery Console.
2. Create a new dataset as supermarket\_analytics.
3. Click Create Table:
   * Source: Google Cloud Storage → Select the uploaded file.
   * File format: CSV
   * Schema: Auto-detect
   * Table name: Supermarket\_analytics
   * Click Create Table

**TRANSFORM**

Clean and prepare the data for analysis.

Cleaning process

Data cleaning is the process of:

* Fixing incorrect or inconsistent data
* Converting data into the correct formats
* Renaming columns to follow a consistent naming
* Making the dataset ready for analysis

Create a Cleaned View or Table

CREATE OR REPLACE VIEW fair-bearing-455213-a1.supermarket\_analytics. cleaned\_sales as

SELECT

"Invoice ID" AS Invoice\_ID,

  Branch,

  City,

  `Customer type` AS Customer\_Type,

  Gender,

  `Product line` AS Product\_Line,

  payment,

  SAFE\_CAST(`Unit price` AS FLOAT64) AS Unit\_Price,

  SAFE\_CAST(Quantity AS INT64) AS Quantity,

  SAFE\_CAST(`Tax 5%` AS FLOAT64) AS Tax,

  SAFE\_CAST(Rating AS FLOAT64) AS Rating,

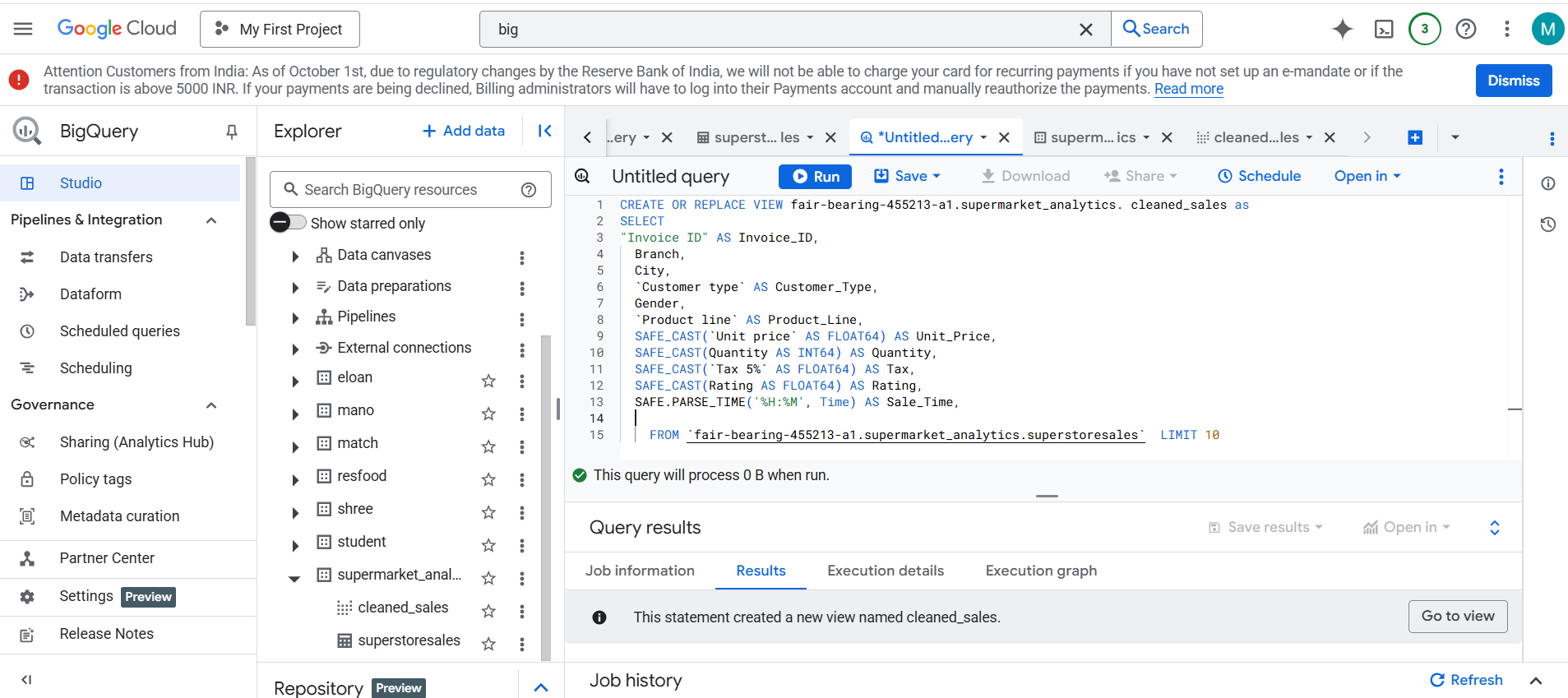
  SAFE.PARSE\_TIME('%H:%M', Time) AS Sale\_Time,

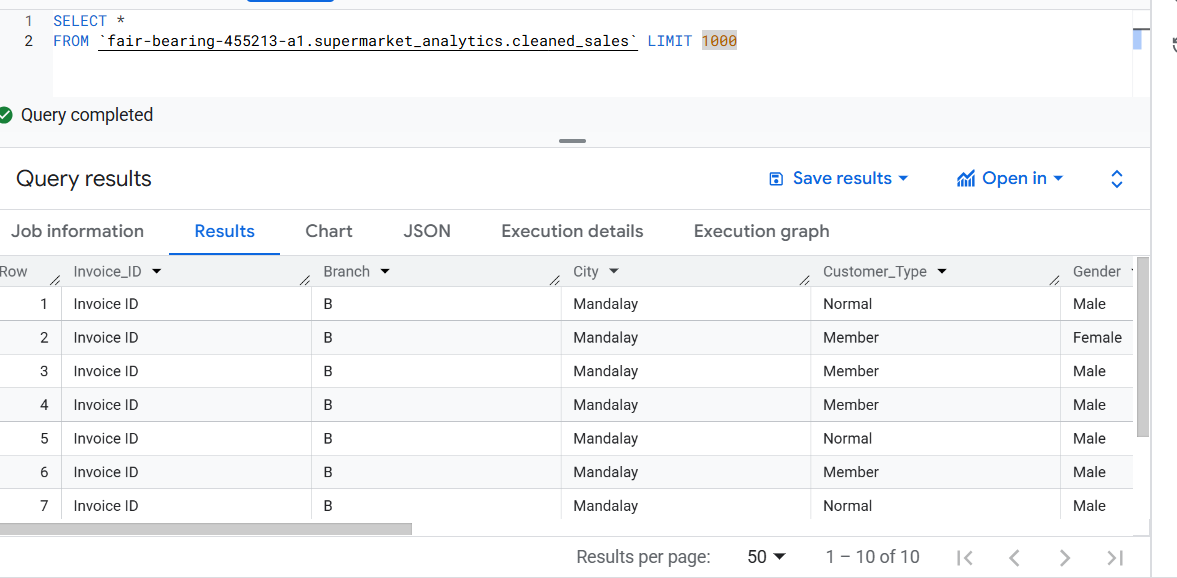
  Date,

  cogs,

  Total,

  FROM `fair-bearing-455213-a1.supermarket\_analytics.superstoresales` where Total is not null LIMIT 10





**Remove Duplicates**

CREATE OR REPLACE TABLE ‘fair-bearing-455213-a1.supermarket\_analytics. ‘ cleaned\_sales as

SELECT \*

FROM (

SELECT \*,

ROW\_NUMBER() OVER (PARTITION BY Invoice\_ID ORDER BY Sale\_DateTime DESC) AS row\_num

FROM `fair-bearing-455213-a1.supermarket\_analytics’)

WHERE row\_num = 1;

**Load into Looker Studio**

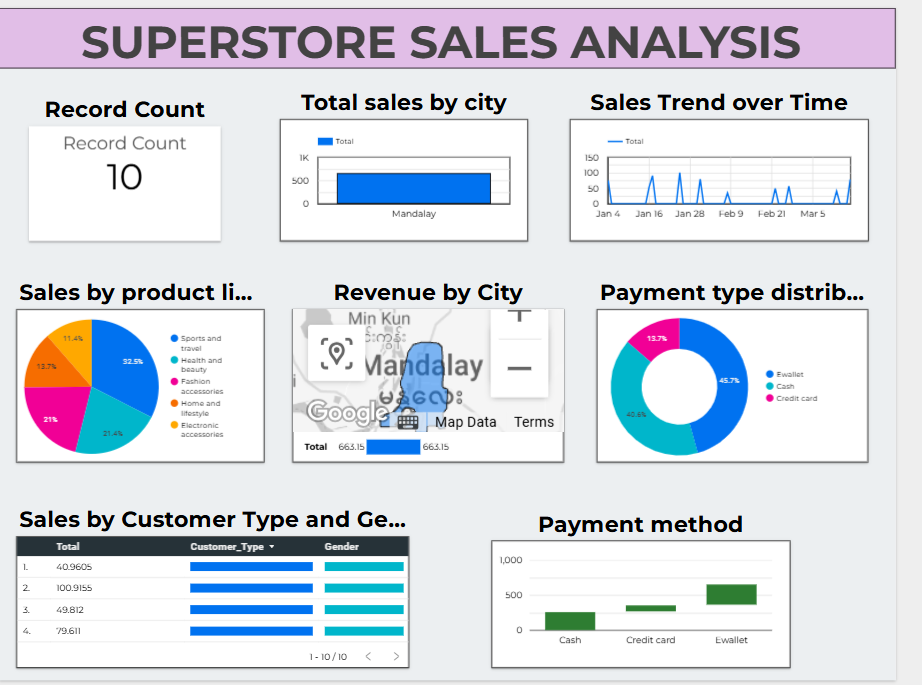
Visualize and analyze the data.

**Steps**

1. Open Looker studio.
2. Create the visualization using the cleaned dataset.

**QUESTIONS**

1. Which city has the highest total sales?
2. What are the trends in total sales over time?
3. Which product line generates the highest sales?
4. What is the distribution of payment methods among customers?
5. Which city contributes the most to overall revenue?
6. How do customer types and genders differ in purchasing behavior?
7. Which payment method yields the highest total revenue?



**CONCLUSION**

The analysis of the Superstore Sales dataset using BigQuery and Looker Studio has provided meaningful insights into sales performance, customer behavior, and revenue trends. Here's a summary of key findings:

1. **Top Performing City**:  
   The city of **Mandalay** shows the highest total sales and revenue contribution, indicating a strong market presence and customer base in that region.
2. **Sales Over Time**:  
   The **sales trend over time** shows fluctuating demand, with certain days experiencing noticeable spikes. This suggests opportunities to investigate promotions, holidays, or seasonal patterns influencing sales.
3. **Product Line Performance**:  
   **Sports and travel** products are the leading revenue drivers, followed by **health and beauty** items. These segments can be targeted for expanded marketing or inventory prioritization.
4. **Payment Preferences**:  
   The majority of transactions are made using **eWallets (45.7%)**, followed closely by **Cash (40.6%)**. Credit card usage is less common. This indicates a preference for digital payments among customers and highlights the importance of optimizing online payment systems.
5. **Geographical Revenue Distribution**:  
   Revenue mapping clearly indicates **Mandalay** as the key contributor. This can guide location-based campaigns and resource allocation.
6. **Customer Demographics**:  
   Analysis by **customer type and gender** reveals diverse spending behavior, with both genders and customer types (e.g., Members vs. Normal) actively contributing to sales. This suggests inclusive marketing strategies would be effective.
7. **Payment Method Impact**:  
   **eWallet** transactions not only dominate in volume but also in value, implying a stronger association with higher-ticket purchases.