



GROUP 7

# PROJECT 1:

## DATA WAREHOUSING FOR BUSINESS INTELLIGENCE

Present to: Asst. Prof. Dr. Songsri Tangsripairoj



# Agenda

1

Project Overview

---

2

ER Diagrams and Data Dictionary

---

3

Data Warehouse Design - OLAP

---

4

ETL Process

---

5

Analysis and Visualization Reports

---

6

Discussion and Conclusion

---

# PROJECT OVERVIEW



# SUMMARY

---

Adidas is a global sportswear brand, generates revenue through footwear, apparel, and accessories, operating via an extensive retail and e-commerce network. The company incurs costs in manufacturing, marketing, and R&D while managing key processes in product innovation, supply chain and sales.



# OBJECTIVE

---

This project focuses on enhancing U.S. sales by developing a centralized data warehouse and leveraging BI tools for real-time sales insights, inventory optimization, and improved marketing strategies.

---



# DATA PREPROCESSING

---

## Primary Data Source

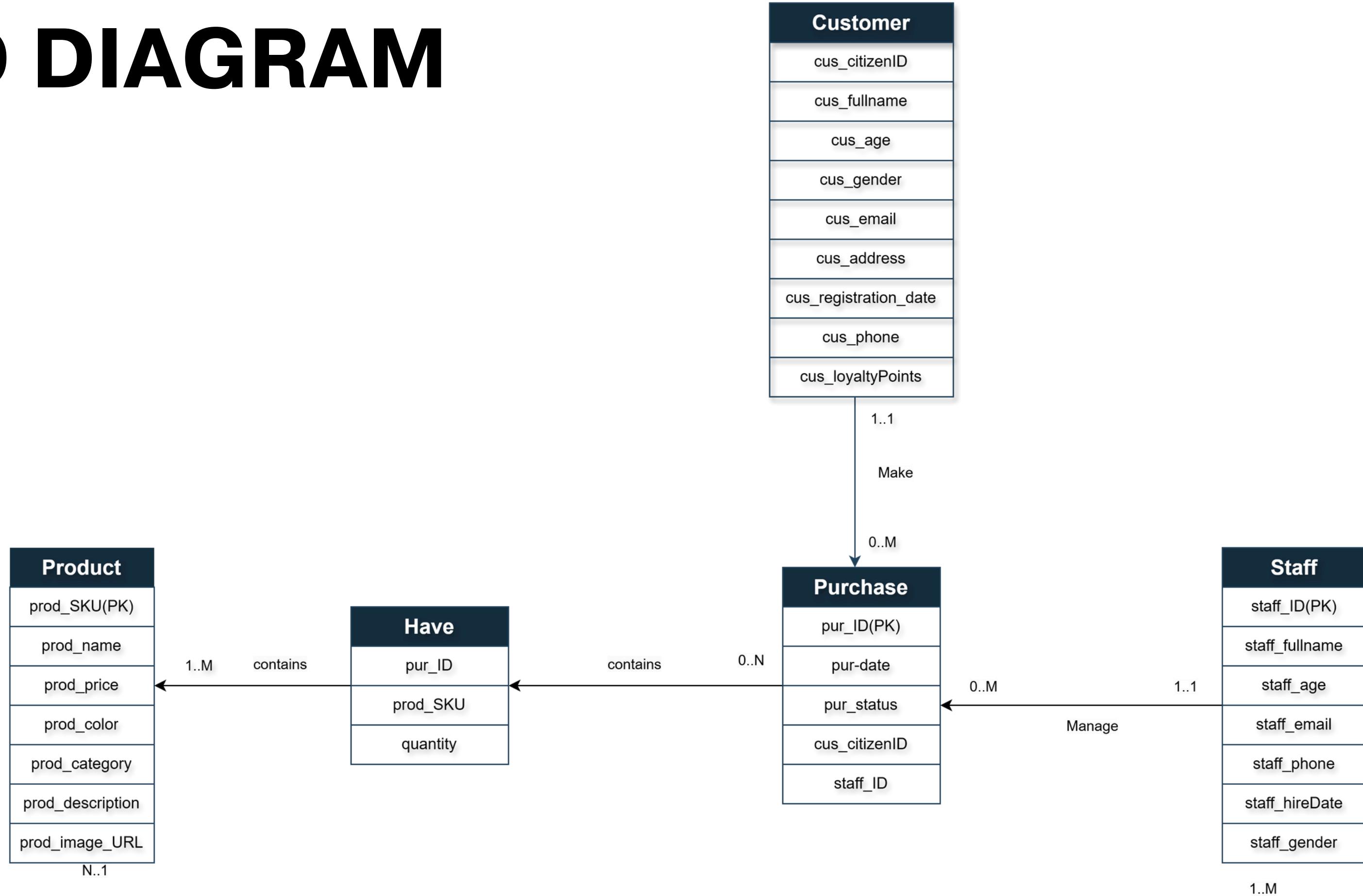
- Adidas webstore shoe data (cleaned to include SKU, name, price, category, description, and image URL) from kaggle dataset.
  - Preprocessed retail data containing retailer details (ID, name, region, state, city, and sales method) from LinkedIn Sale Analysis.
- 

## Additional Generated Data

- Due to privacy constraints, our team generated five datasets: store, customer, staff, purchase, and data tables, ensuring comprehensive analysis while protecting sensitive information.



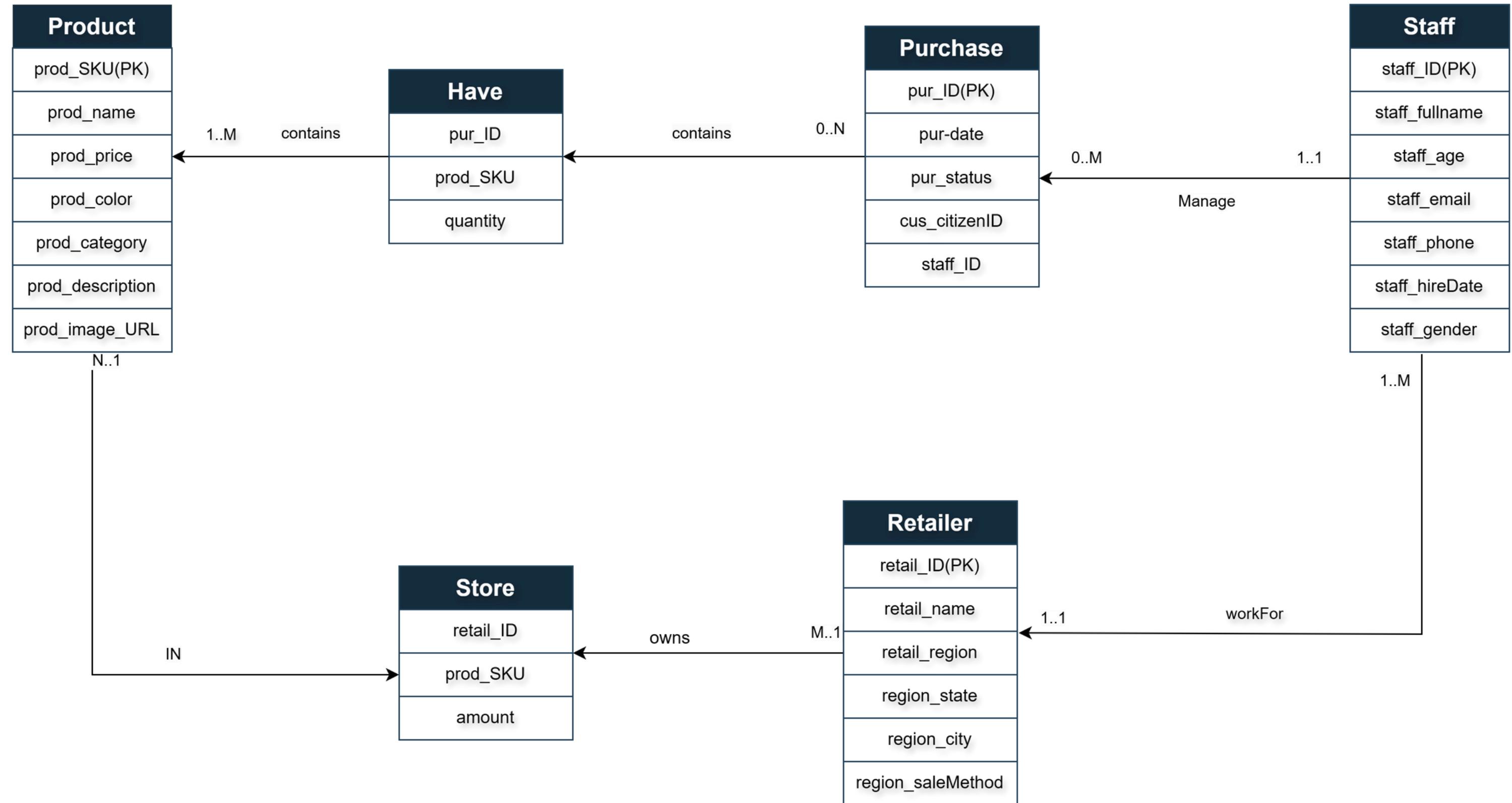
# ERD DIAGRAM



[Back to Agenda](#)



# ERD DIAGRAM



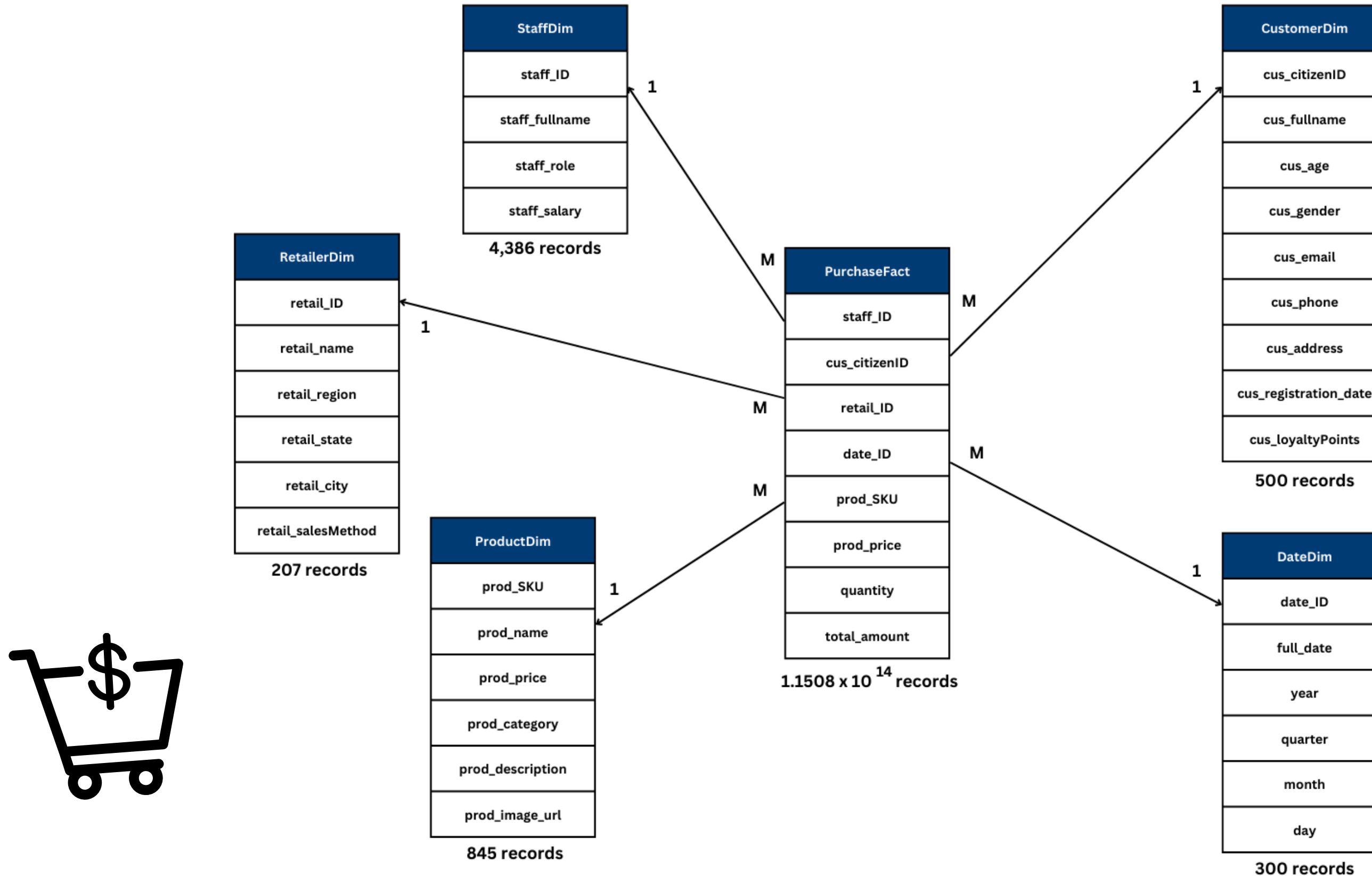
[Back to Agenda](#)



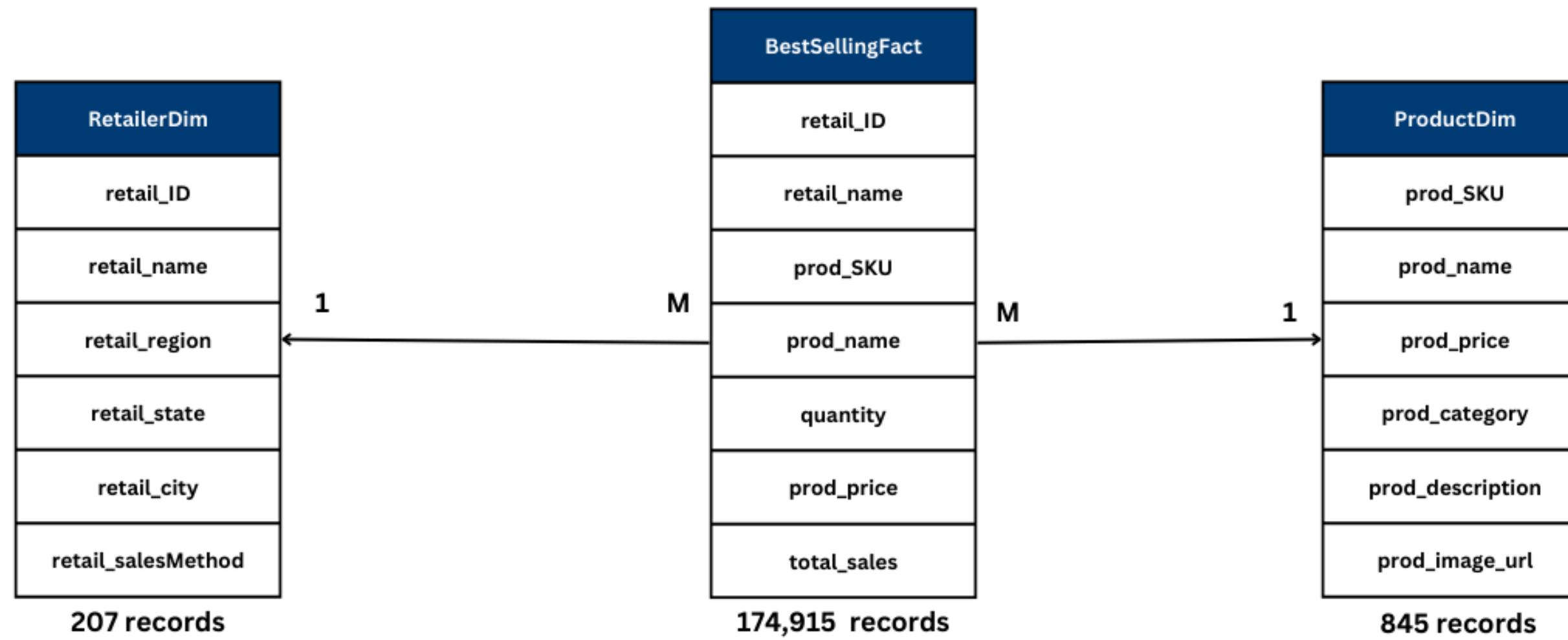


# DATA WAREHOUSE DESIGN - OLAP

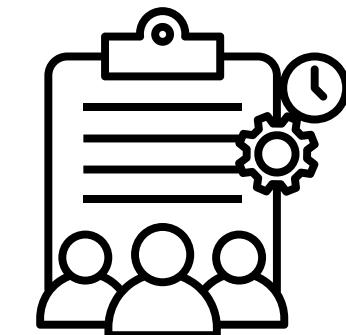
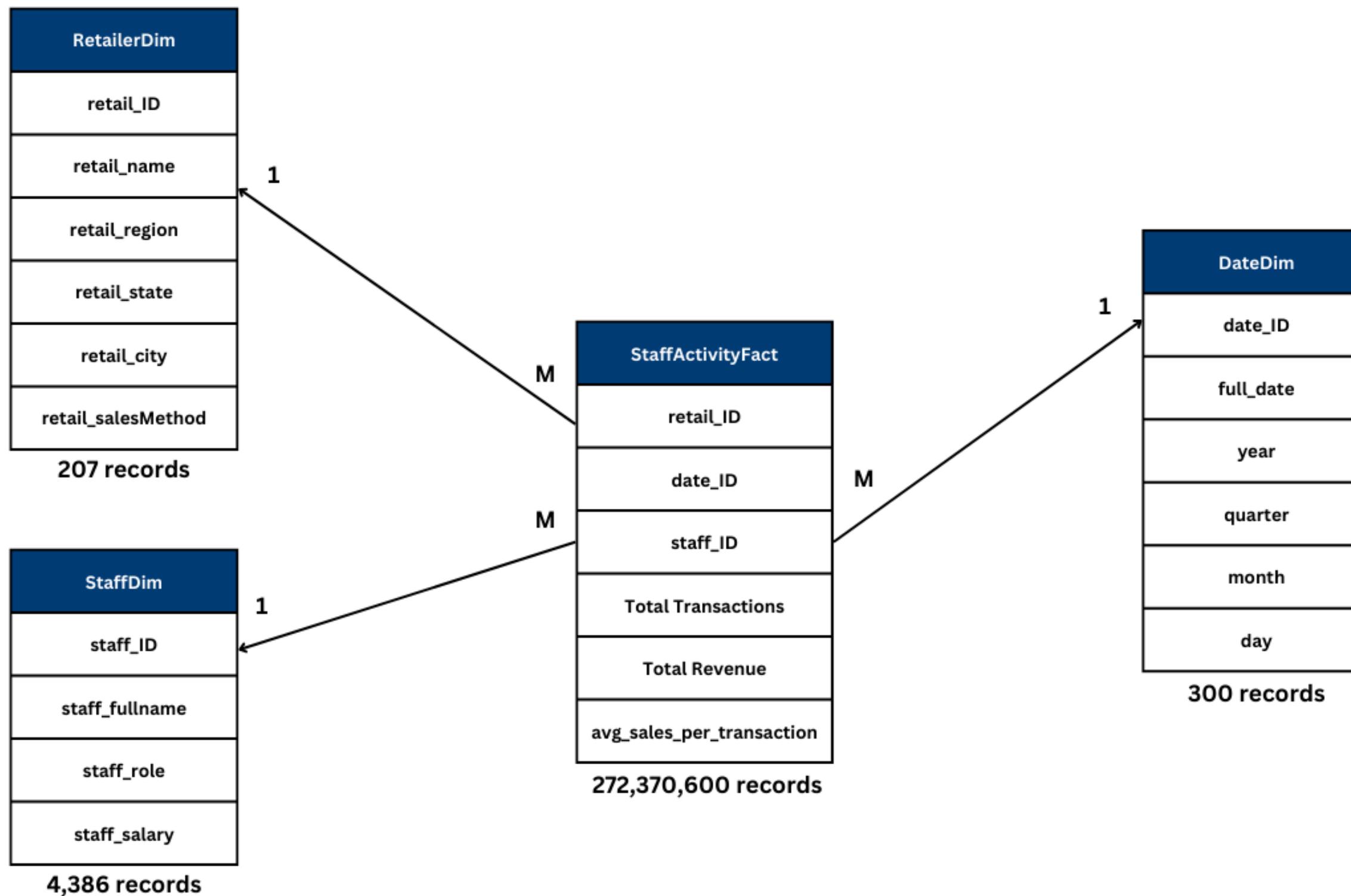
# PurchaseFact (Customer Purchase Behavior Cube)



# BestSellingFact (Top-Selling Products Cube)



# StaffPerformanceFact (Staff Productivity Cube)



# ETL PROCESS



# MAPPING BETWEEN OLTP AND OLAP

---

OLTP Table	OLTP Column	Mapped OLAP Table	OLAP Column	Transformation
Product	prod_SKU	ProductDim	prod_SKU	Direct Mapping
	prod_name		prod_name	Direct Mapping
	prod_price		prod_price	Direct Mapping
	prod_category		prod_category	Direct Mapping
	prod_description		prod_description	Direct Mapping
	prod_image_url		prod_image_url	Direct Mapping
	retail_id		retail_id	Direct Mapping
Retailer	retail_name	RetailerDim	retail_name	Direct Mapping
	retail_region		retail_region	Direct Mapping
	retail_state		retail_state	Direct Mapping
	retail_city		retail_city	Direct Mapping
	retail_salesMethod		retail_salesMethod	Direct Mapping
	cus_citizenID	CustomerDim	cus_citizen_ID	Direct Mapping
Customer	cus_fullname		cus_fullname	Direct Mapping
	cus_age		cus_age	Direct Mapping
	cus_gender		cus_gender	Direct Mapping
	cus_email		cus_email	Direct Mapping
	cus_phone		cus_phone	Direct Mapping
	cus_address		cus_address	Direct Mapping
	cus_registration_date		cus_registration_date	Direct Mapping
	cus_loyaltyPoints		cus_loyaltyPoints	Direct Mapping

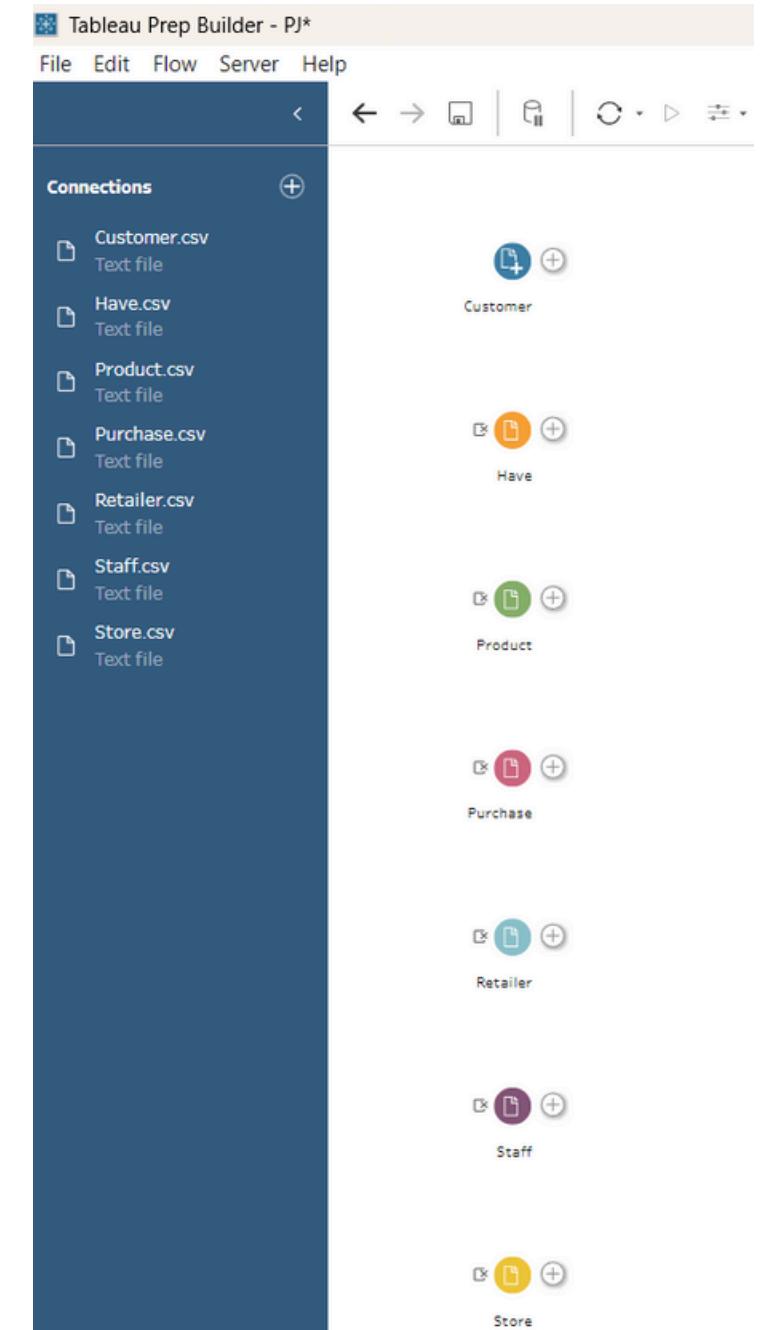
Purchase	pur_ID	PurchaseFact	pur_ID	Direct Mapping
	pur_date	DateDim	date_ID	Date transformation (YYYY-MM-DD → date_ID)
	cus_citizenID	PurchaseFact	cus_citizenID	Direct Mapping
	staff_ID		staff_ID	Direct Mapping
	retail_ID		retail_ID	Direct Mapping
Have	pur_ID	PurchaseFact	pur_ID	Join with Purchase Table
	prod_SKU		prod_SKU	Join with Product Table
	quantity		quantity	Direct Mapping
Store	prod_SKU, retail_ID	RetailerSalesFact	prod_SKU, retail_ID	Aggregate total sales
Staff	staff_ID	StaffDim	staff_ID	Direct Mapping
	staff_fullname		staff_fullname	Direct Mapping
	staff_role		staff_role	Direct Mapping
	staff_age		staff_age	Direct Mapping
	staff_salary		staff_salary	Direct Mapping
	staff_email		staff_email	Direct Mapping
	staff_phone		staff_phone	Direct Mapping
	staff_hireDate		staff_hireDate	Direct Mapping
	retail_ID		retail_ID	Direct Mapping

# ETL PROCESS USING TABLEAU PREP

## 1 Extract (E) - Data Collection

1. Open Tableau Prep Builder.
2. Connect to the data source (e.g., Excel file, CSV, or a database) that contains OLTP tables.
3. Load the following tables into your flow:
  - Product, Retailer, Store, Customer, Staff, Purchase, and Have
4. Drag and drop all tables into the Flow pane.

Since **no dedicated Date table exists in OLTP**, we extract *pur\_date* from the *Purchase* table to create a **Date Dimension**.



# ETL PROCESS USING TABLEAU PREP

## 2 Transform (T) - Data Cleaning and Structuring

### 2.1 Processing Dimension Tables

#### 2.1.1 Customer Dimension (CustomerDim)

- **Data Cleaning:**

- Change *cus\_citizenID* → **Integer to String**.
- Change *cus\_email* → **String to Email**.
- Change *cus\_phone* → **Integer to String**.

- **Ensure Uniqueness:**

- Aggregate on *cus\_citizenID*.

- **Output:** *CustomerDim*.

#### 2.1.2 Product Dimension (ProductDim)

- **Data Cleaning:**

- Change *prod\_image\_URL* → String to URL.

- **Ensure Uniqueness:**

- Aggregate on *prod\_SKU*.

- **Output:** *CustomerDim*.

Type	Field Name	Changes
Abc	<i>cus_citizenID</i>	A <sub>v</sub>
Abc	<i>cus_fullname</i>	
#	<i>cus_age</i>	
Abc	<i>cus_gender</i>	
Abc	<i>cus_email</i>	A <sub>v</sub>
Abc	<i>cus_address</i>	
⊕	<i>cus_registration_date</i>	
Abc	<i>cus_phone</i>	A <sub>v</sub>
#	<i>cus_loyaltyPoints</i>	

Type	Field Name	Changes
Abc	<i>prod_SKU</i>	
Abc	<i>prod_name</i>	
Abc	<i>prod_color</i>	
Abc	<i>prod_category</i>	
Abc	<i>prod_description</i>	
Abc	<i>prod_image_URL</i>	A <sub>v</sub>
#	<i>prod_price</i>	A <sub>v</sub>

# ETL PROCESS USING TABLEAU PREP

## 2.1.3 Retailer Dimension (RetailerDim)

- **Data Cleaning:**

- Change *retail\_ID* → Integer to String.

- **Ensure Uniqueness:**

- Aggregate on *retail\_ID*.

- **Output:** *RetailerDim*.

## 2.1.4 Staff Dimension (StaffDim)

- **Data Cleaning:**

- Change *staff\_ID* → Integer to String.
  - Change *staff\_salary* → Integer to String.
  - Change *staff\_email* → String to Email.
  - Change *staff\_phone* → Integer to String.

- **Ensure Uniqueness:**

- Aggregate on *staff\_ID*.

- **Output:** *StaffDim*.

Type	Field Name	Changes
Abc	retail_ID	A
Abc	retail_name	
Abc	retail_region	
Abc	retail_state	
Abc	retail_city	
Abc	retail_salesMethod	

Type	Field Name	Changes
Abc	staff_ID	∅ × ... A
Abc	staff_fullname	
Abc	staff_role	
#	staff_age	
#	staff_salary	
Abc	staff_email	A
Abc	staff_phone	A
Abc	staff_hireDate	
Abc	staff_gender	
Abc	retail_ID	A

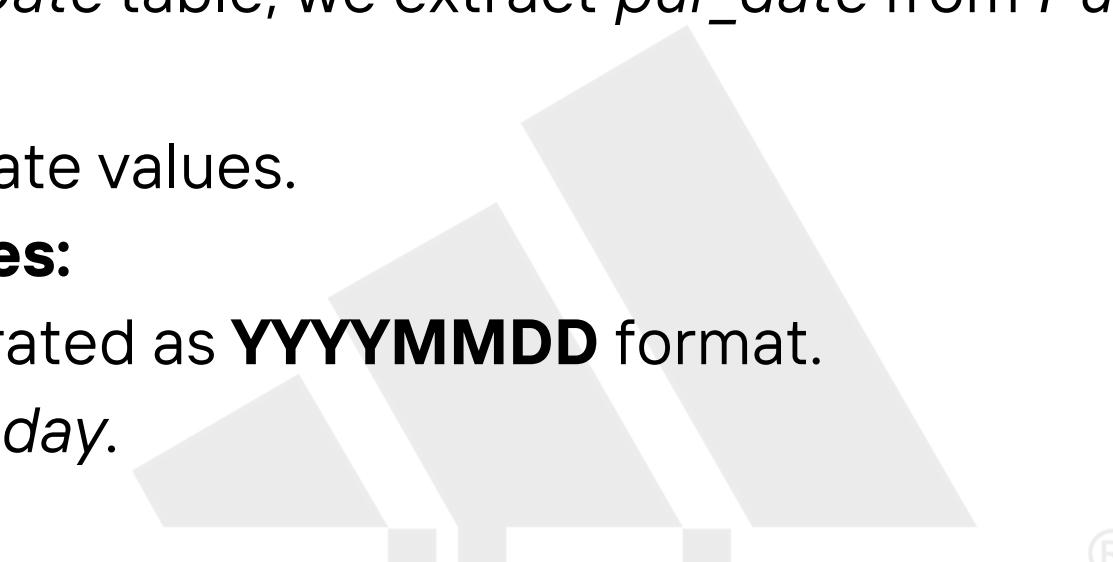
# ETL PROCESS USING TABLEAU PREP

---

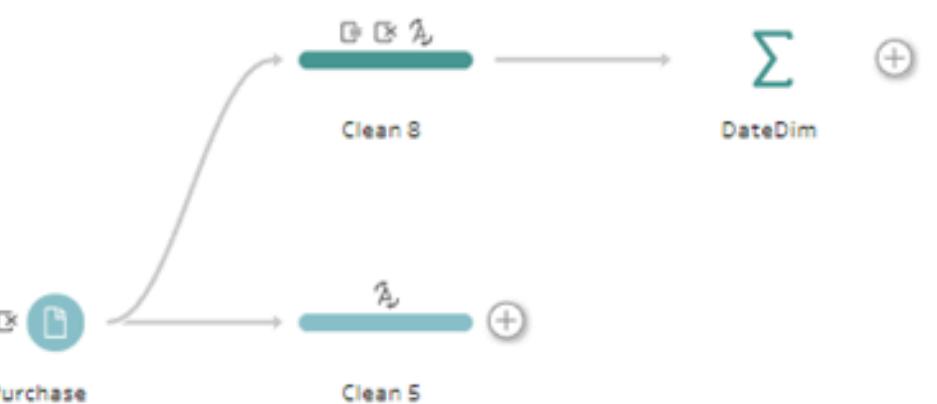
## 2.1.5 Date Dimension (**DateDim**)

Since there is no dedicated *Date* table, we extract *pur\_date* from *Purchase*.

- **Extract Unique Dates:**
  - Identify unique *pur\_date* values.
- **Generate Date Attributes:**
  - *date\_ID* (PK) → Generated as **YYYYMMDD** format.
  - *year, quarter, month, day*.
- **Output:** *DateDim*.



Type	Group Type	Field Name	Changes
Abc	GROUP	<i>date_ID</i>	
Abc	GROUP	Year	
#	GROUP	Quarter	
#	GROUP	Month	
#	GROUP	Day	



# ETL PROCESS USING TABLEAU PREP

## 2.2 Processing Fact Tables

### 2.2.1 Purchase Fact Table (PurchaseFact)

- **Join Tables:**

- Purchase **JOIN** Have on *pur\_ID*.

- **Select Relevant Columns:**

- *cus\_citizenID, retail\_ID, pur\_date, staff\_ID, prod\_SKU, quantity, total\_amount* (calculated as quantity \* prod\_price).

- **Output:** *PurchaseFact*.

### 2.2.2 Best Selling Fact Table (BestSellingFact)

- **Aggregate Data:**

- Group by *retail\_ID, prod\_SKU, pur\_date*.
  - Calculate *total\_sales* = SUM(*amount \* prod\_price*).

- **Select Relevant Columns:**

- *retail\_ID, pur\_date, prod\_SKU, total\_sales*.

- **Output:** *RetailerSalesFact*.

Type	Field Name	Changes
Abc	staff_ID	
Abc	date_ID	
Abc	cus_citizenID	
Abc	retail_ID	
Abc	prod_SKU-1	
Abc	prod_SKU	
#	total_sales	
#	quantity	
#	prod_price	

Save to BestSellingFact.csv		
Type	Field Name	
Abc	retail_ID	
Abc	prod_SKU	
#	amount	
#	quantity	
#	prod_price	
#	total_sales	

# ETL PROCESS USING TABLEAU PREP

---

## 2.2.3 Staff Performance Fact Table (StaffPerformanceFact)

- **Aggregate Data:**

- Group by *staff\_ID, pur\_date*.
- Calculate:
  - $total\_transactions = COUNT(pur\_ID)$ .
  - $total\_sales = SUM(total\_amount)$ .
  - $avg\_sales\_per\_transaction = total\_sales / total\_transactions$ .

- **Select Relevant Columns:**

- *staff\_ID, retail\_ID, pur\_date, total\_transactions, total\_sales, avg\_sales\_per\_transaction.*

- **Output:** *StaffPerformanceFact*.

Save to StaffPerformanceFact.csv

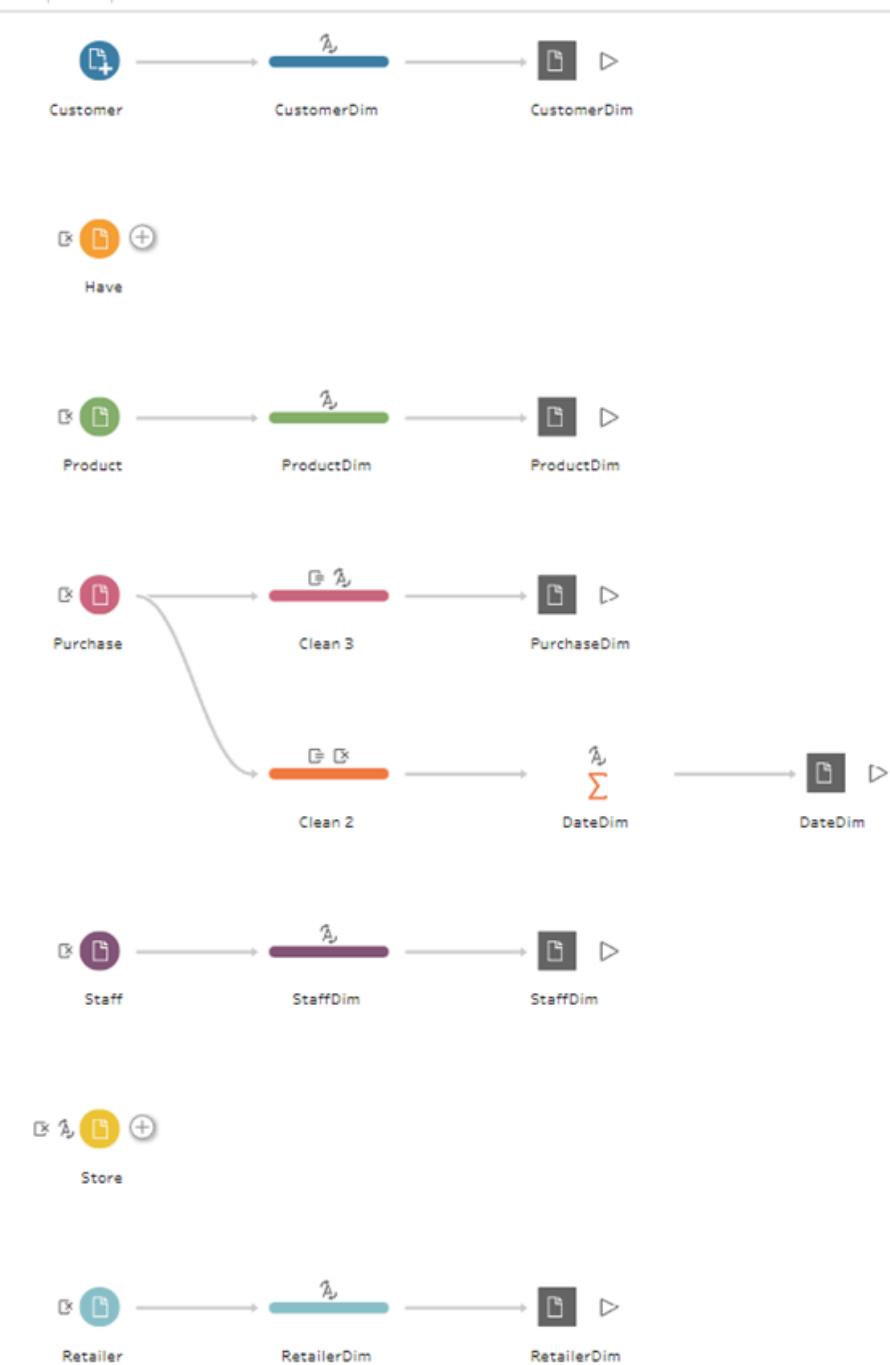
Type	Field Name
Abc	staff_ID
Abc	retail_ID
Abc	date_ID
#	Total Transactions
#	Total Revenue
#	avg_sales_per_transaction

# ETL PROCESS USING TABLEAU PREP

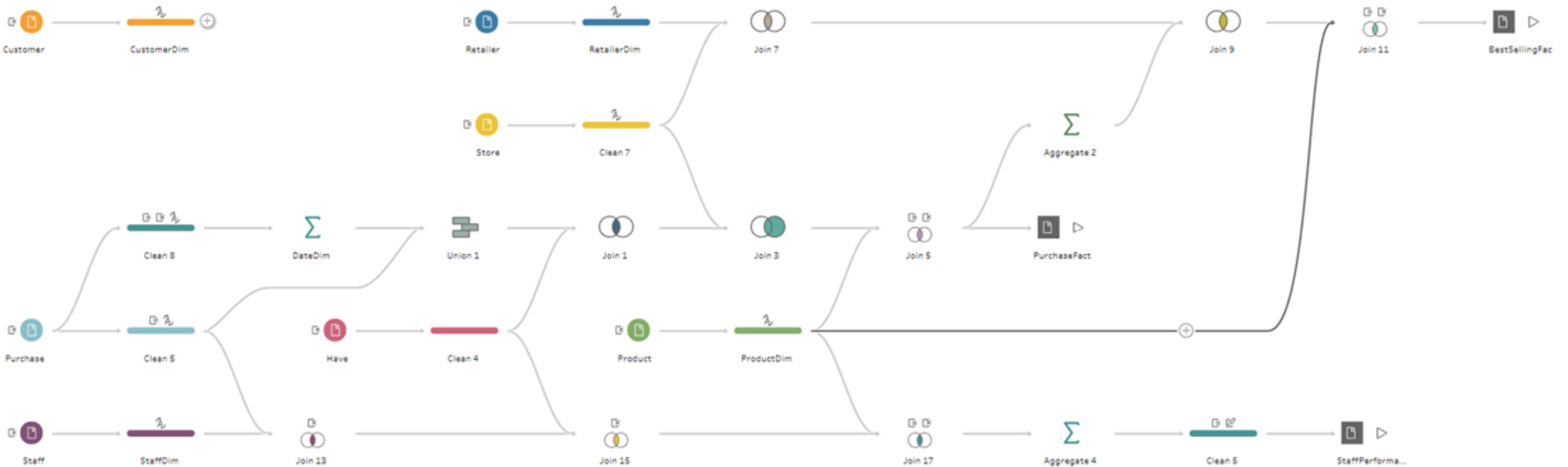
## 3 Load (L) - Data Warehouse Population

Once the data is structured and transformed, we load it into the OLAP data warehouse for business intelligence purposes.

- **Target Database:** PostgreSQL / MySQL (Star Schema).
- **Stored Tables:**
  - Dimension Tables: *CustomerDim*, *ProductDim*, *RetailerDim*, *StaffDim*, *DateDim*.
  - Fact Tables: *PurchaseFact*, *RetailerSalesFact*, *StaffPerformanceFact*.



# ETL PROCESS USING TABLEAU PREP



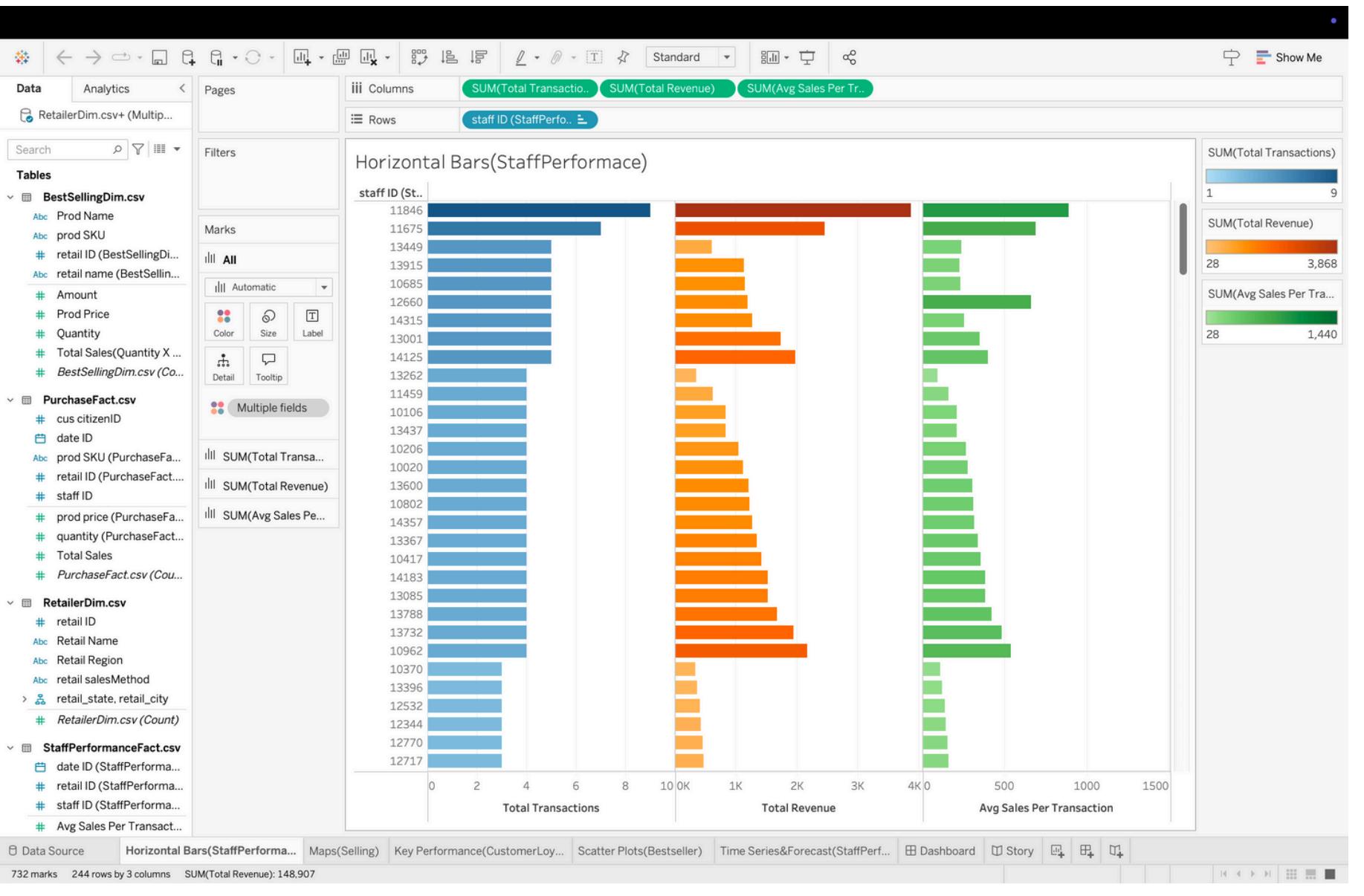
[Back to Agenda](#)

# **ANALYSIS AND VISUALIZATION REPORTS**

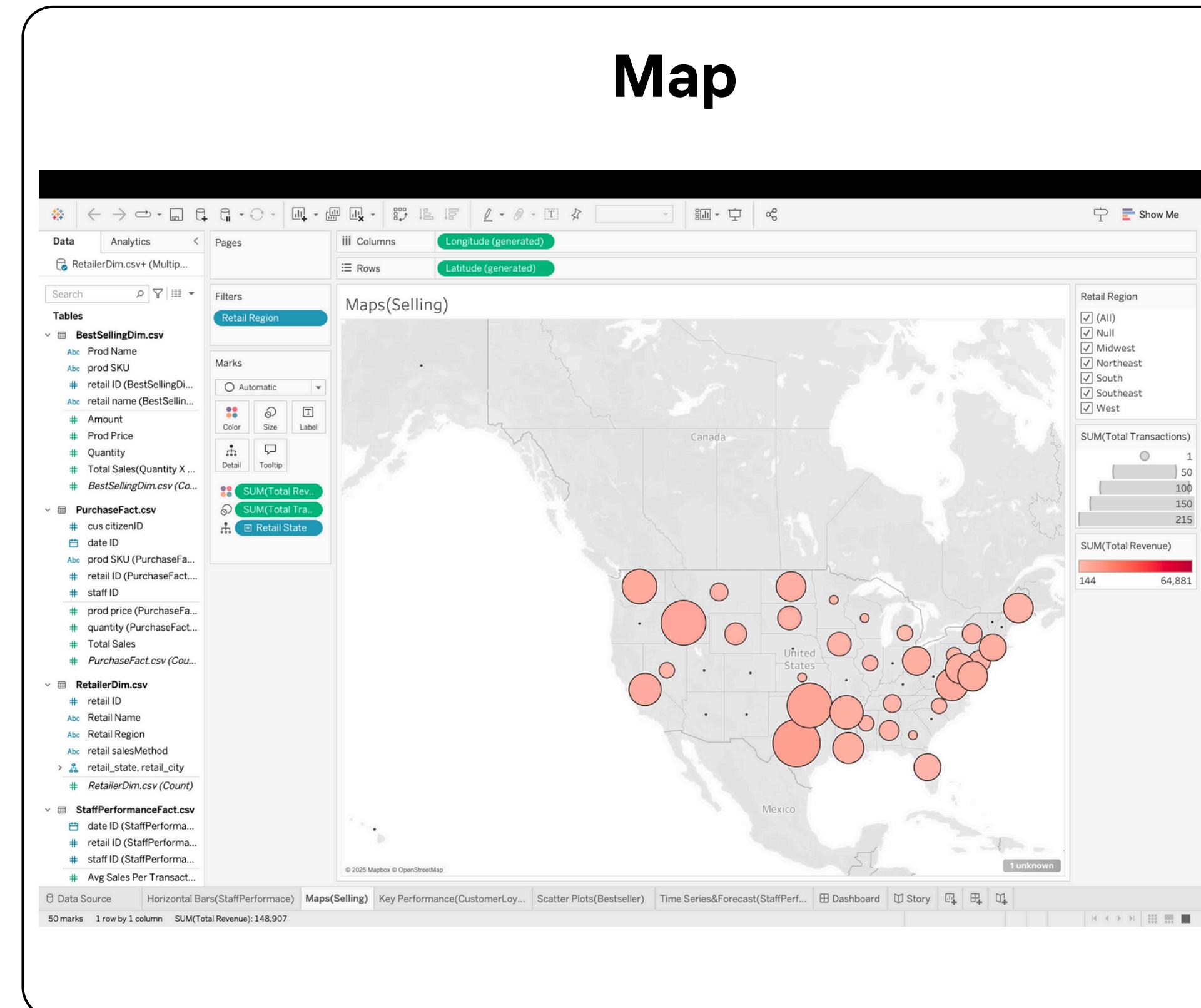
[Back to Agenda](#)



# Horizontal Bars



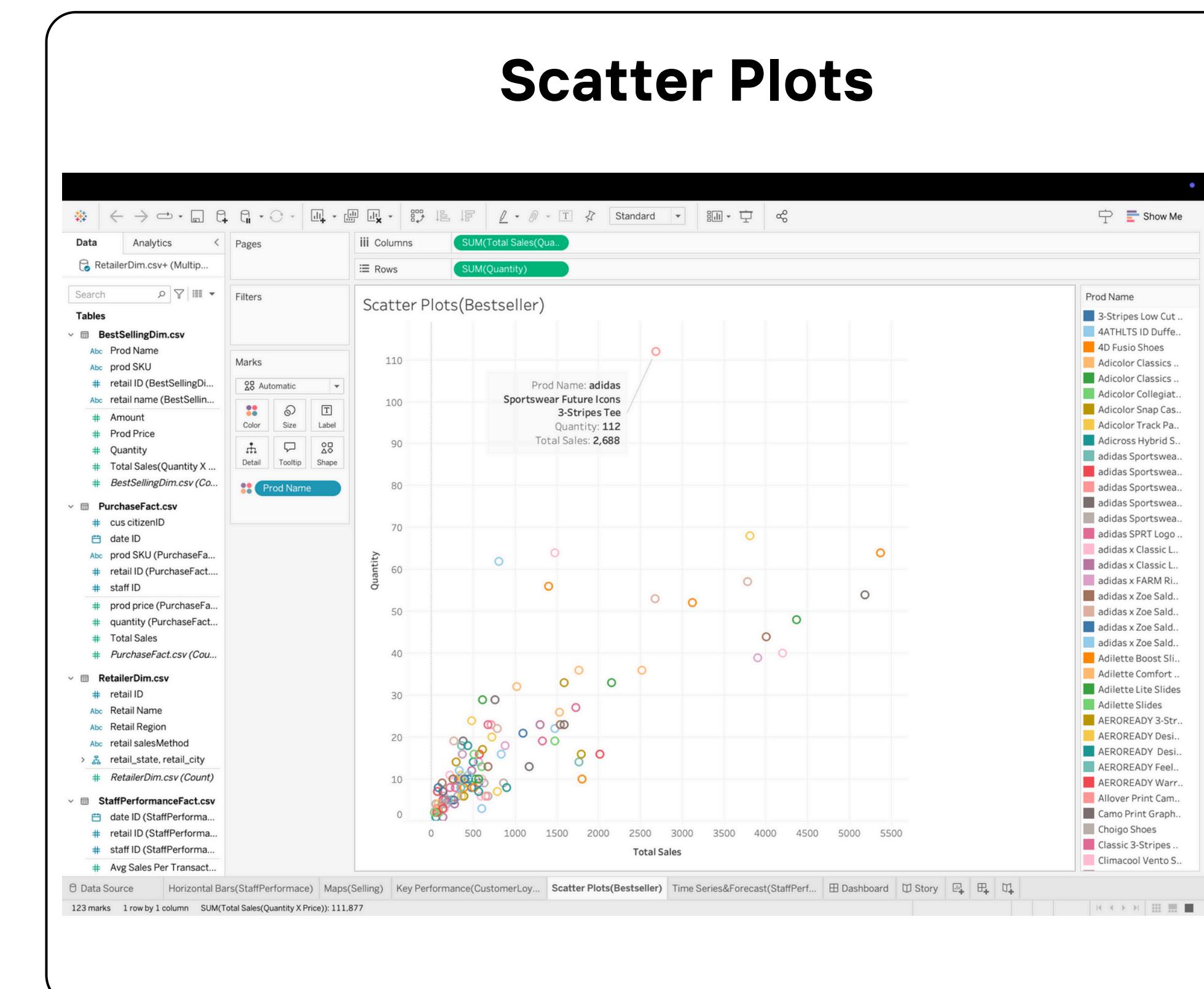
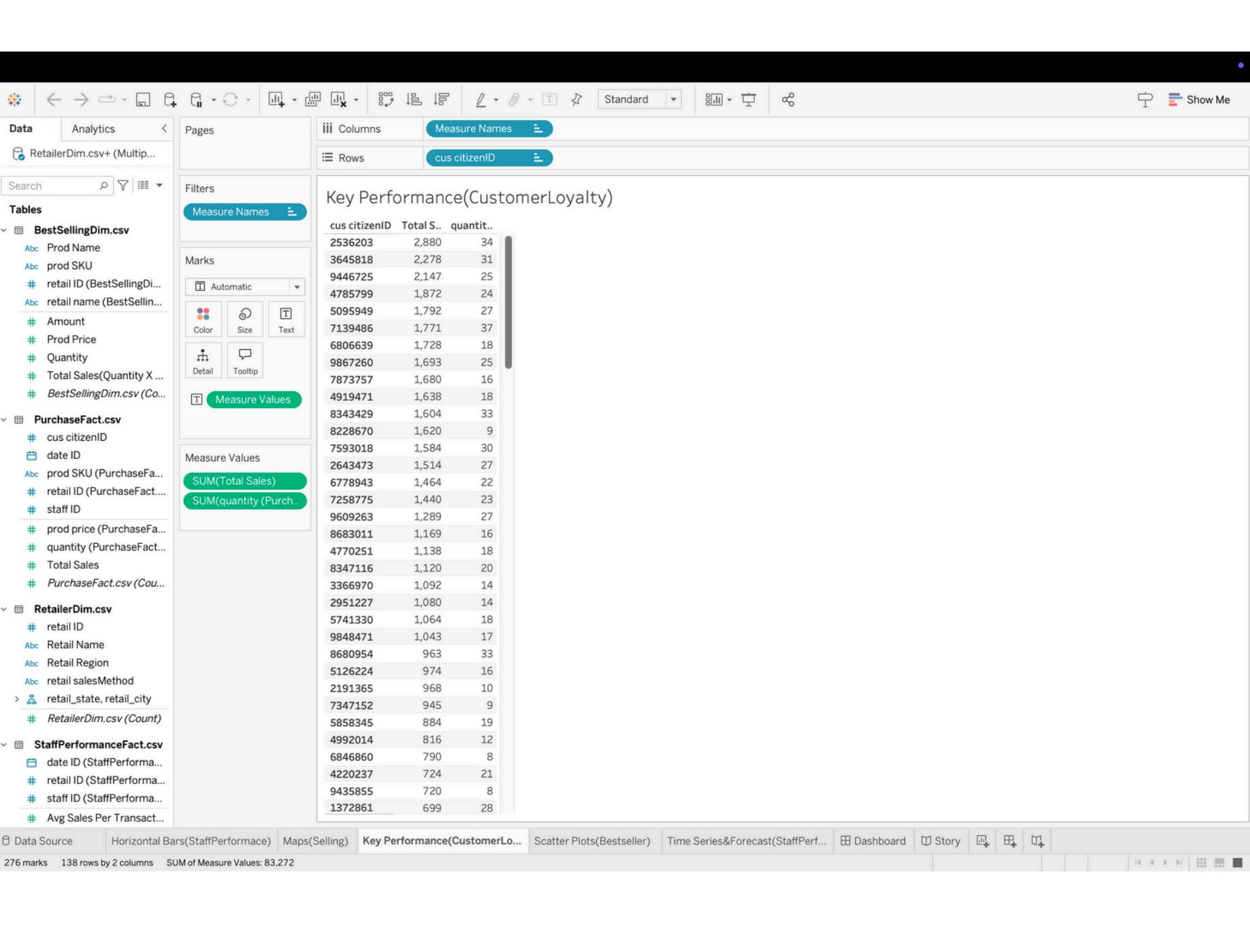
# Map



[Back to Agenda](#)



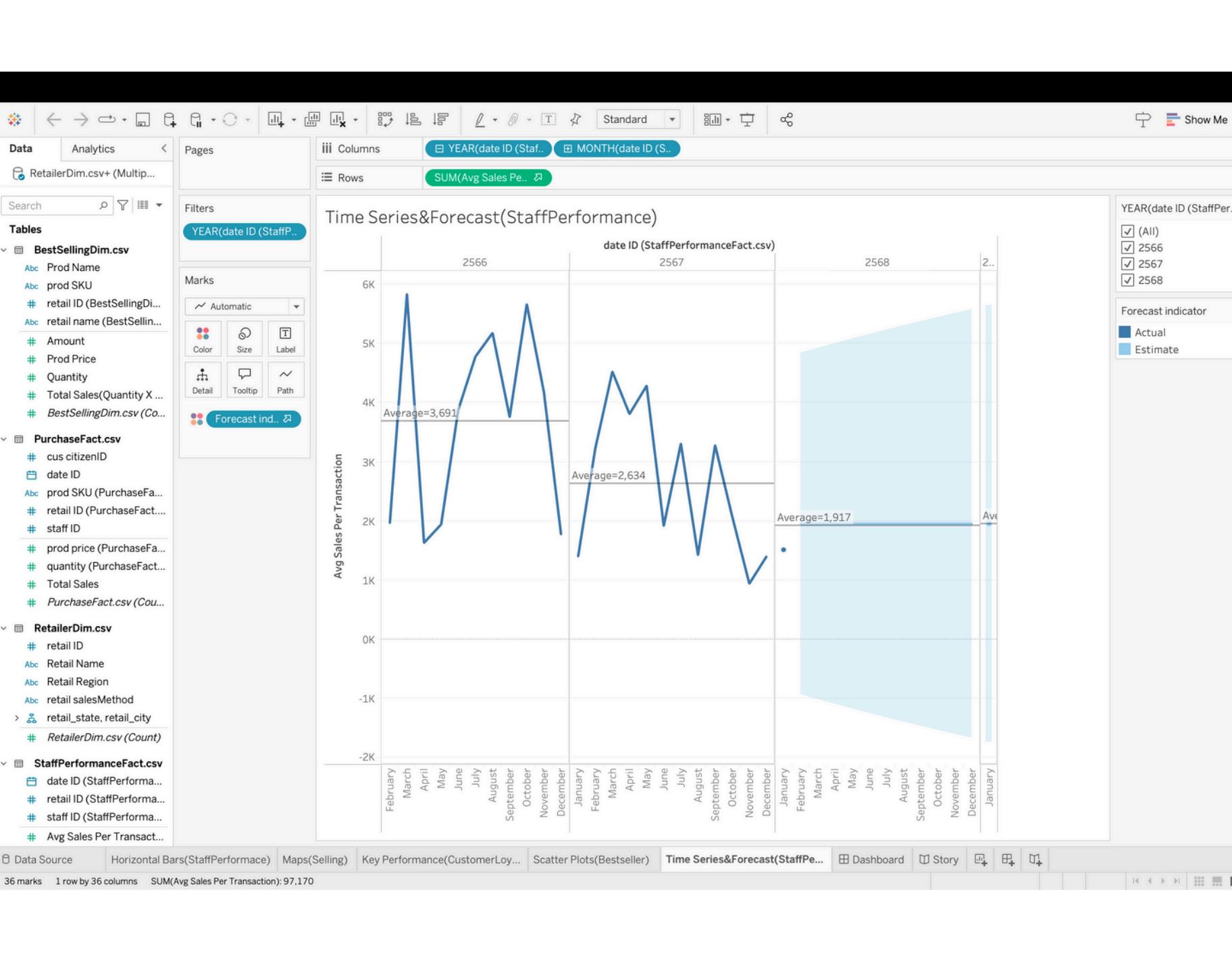
# Key Performance Indicator



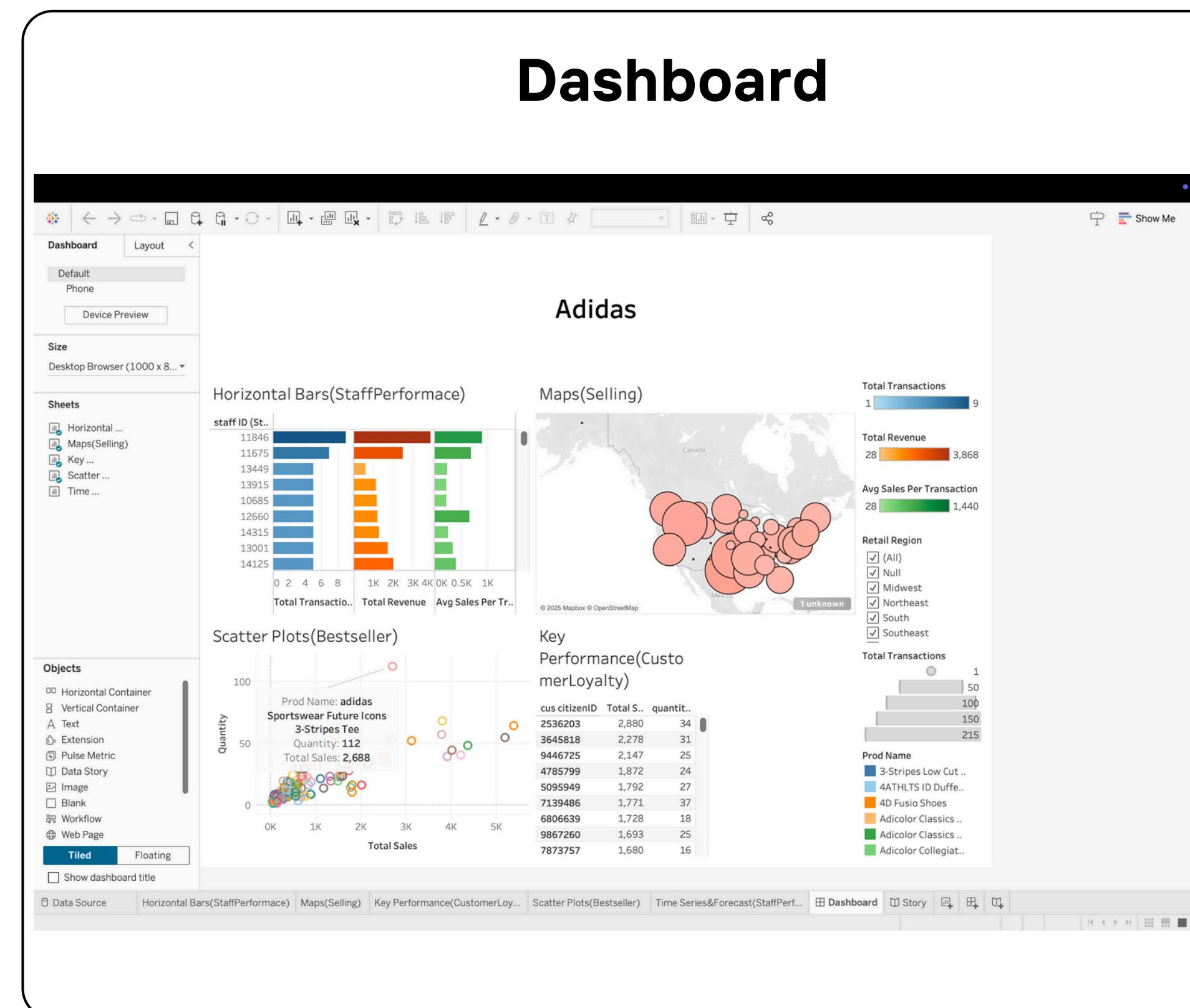
[Back to Agenda](#)



# Time Series & Forecast



# Dashboard



[Back to Agenda](#)



New story point

 Blank Duplicate

# Story of Adidas

 DashboardHorizontal  
Bars(StaffPerformanc  
e)

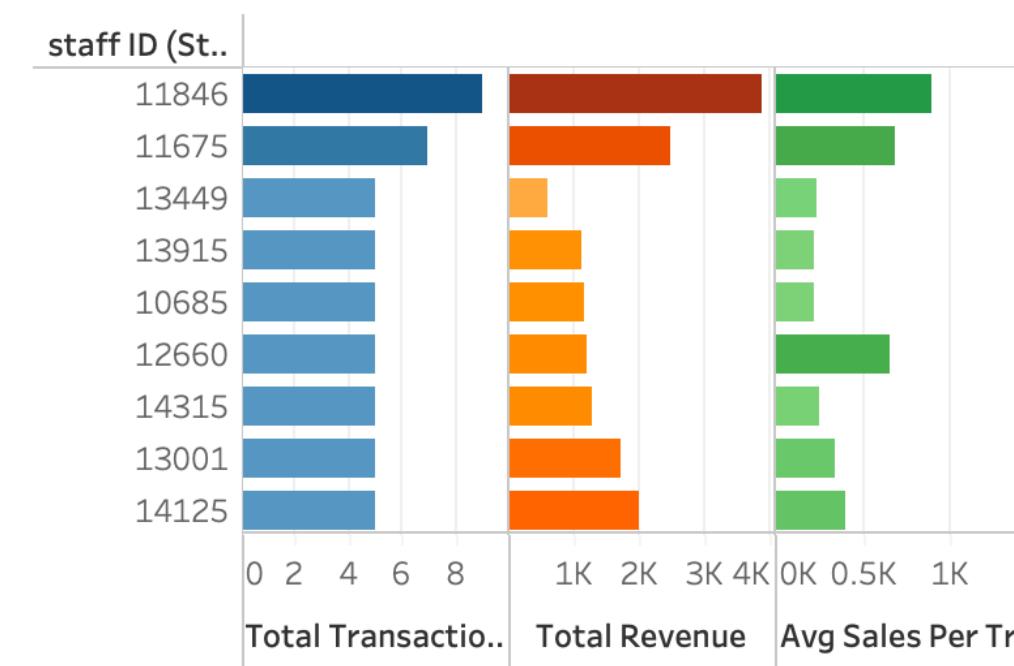
Maps(Selling)

Scatter  
Plots(Bestseller)Key  
Performance(Custom  
erLoyalty)

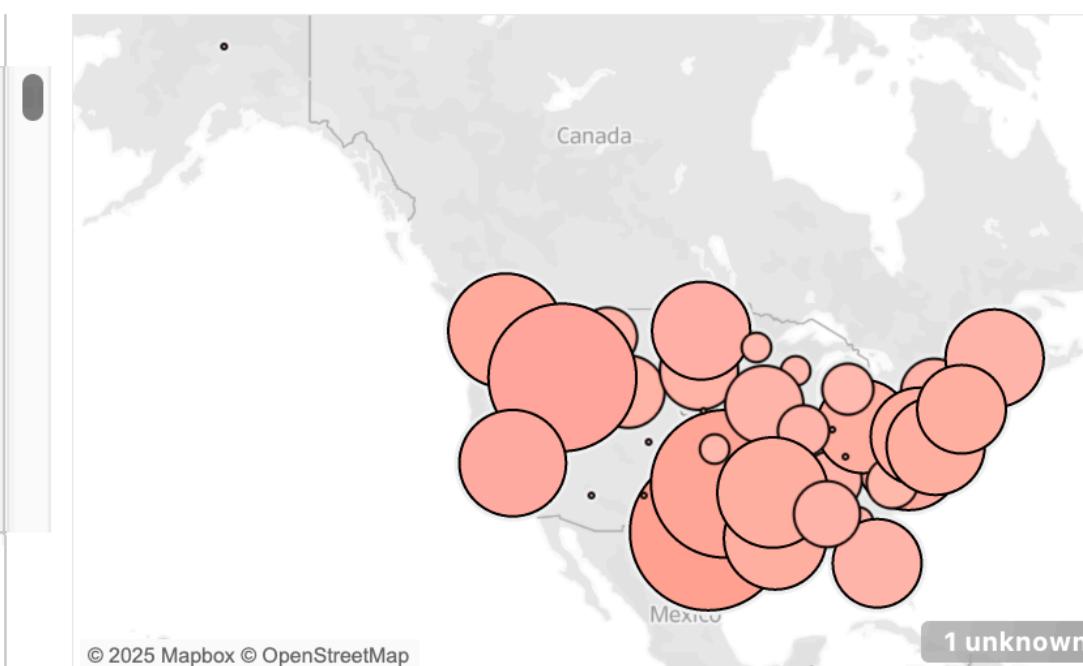
- Horizontal ...
- Maps(Selling)
- Key ...
- Scatter ...
- Time ...
- Dashboard

## Adidas

### Horizontal Bars(StaffPerformance)



### Maps(Selling)



Total Transactions

1 9

Total Revenue

28 3,868

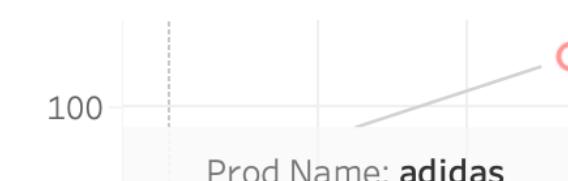
Avg Sales Per Transaction

28 1,440

Retail Region

- (All)
- Null
- Midwest
- Northeast
- South
- Southeast

### Scatter Plots(Bestseller)



### Key Performance(CustomerLoyalty)

cus\_citizenID TotalS quantity

Total Transactions

1 50

100

[Back to Agenda](#)

# DISCUSSION AND CONCLUSION

---



The integration of BI tools and a data warehouse transformed Adidas' approach to data management and analytics. By addressing data fragmentation and inefficiencies, the company streamlined internal processes and gained a competitive edge in the global market. This case highlights the power of advanced data solutions in overcoming business challenges and underscores their long-term impact on success.



# **Q & A SESSION**





# Group 7 Members

Mr. Suppakorn Pojsomphong 6588077

Mr. Thannaton Thongsuk 6588109

Miss Nuttida Wiphaalongkot 6588144

Miss Ananya Sangtanchai 6588191