**Superstore Techround**

This project involves the use of PowerBI, SQL, and DAX for data analysis.

**Data Connection**

The connection to the dataset was made using the credentials provided directly through a SQL Server Import connection on PowerBI. The dataset appears to be currently static.

**Data Cleaning**

On PowerQuery, it was possible to check all data types for each row and confirm that we have only valid rows for all the columns. The data is only for the United States, and we have two columns to identify our client (“Customer ID” and “Customer Name”). Therefore, the columns “Country” and “Customer Name” were removed.

The column “Discount” was filtered to avoid numbers higher than 1, as it does not make sense to have more than 100% of discount.

Duplicates were removed from all columns to avoid duplicated data in future analysis.

Two new columns “Days for Shipment” and “Cost” were created.

**SQL Queries**

The same connection and select were executed using DBeaver to demonstrate that it’s possible to execute the same through SQL queries.

**SELECT** **DISTINCT**

-- Select all columns except "Customer Name" and "Country"

-- Add new calculated columns "Days for Shipment" and "Cost"

[Order ID],

[Segment],

[Ship Date],

[Ship Mode],

[State],

[Product ID],

[Category],

[Sub-Category],

[Product Name],

[Sales],

[Quantity],

[Discount],

[Profit],

**DATEDIFF**(**day**, [Order Date], [Ship Date]) **AS** *[Days for Shipment]*,

[Sales] - [Profit] **AS** *[Cost]*

**FROM**

sales\_data

**WHERE**

-- Only include rows where "Discount" is less than or equal to 1

[Discount] <= 1

SQL can also be used to check if we have duplicated categorical values because of some typo:

**SELECT** **DISTINCT** [Segment] **FROM** sales\_data;

**SELECT** **DISTINCT** [Ship Mode] **FROM** sales\_data;

**SELECT** **DISTINCT** [State] **FROM** sales\_data;

**SELECT** **DISTINCT** [Category] **FROM** sales\_data;

**SELECT** **DISTINCT** [Sub-Category] **FROM** sales\_data;

**SELECT** **DISTINCT** [Product Name] **FROM** sales\_data;

**DAX Functions**

A “**dCalendar**” table was created using DAX for dynamic calendar functionality.

dCalendar = CALENDAR (MIN ('fSales\_Data'[Order Date]), MAX ('fSales\_Data'[Order Date]))

Day = dCalendar[Date].[Day]

Month = dCalendar[Date].[MonthNo]

Month\_Name = dCalendar[Date].[Month]

MonthMMM = FORMAT(dCalendar[Date].[Date], "mmm")

Year = dCalendar[Date].[Year]

A “**Measures**” folder was also created to keep all the measures together.

Profit = SUM(fSales\_Data[Profit])

Profit Margin = DIVIDE([Profit],[Cost])

NCustomer = DISTINCTCOUNT(fSales\_Data[Customer ID])

Cost = SUM(fSales\_Data[Cost])

YearlyProfit = TOTALYTD([Profit],dCalendar[Date])

MonthlyProfit = TOTALMTD([Profit],dCalendar[Date])

LastYearProfit = CALCULATE([YearlyProfit],PREVIOUSYEAR(dCalendar[Date]))

LastMonthProfit = CALCULATE([MonthlyProfit],DATEADD(dCalendar[Date],-1,MONTH))

LastYearMonthProfit = CALCULATE([MonthlyProfit],DATEADD(dCalendar[Date],-12,MONTH))

%YoY Profit =

IF (

    ISBLANK([YearlyProfit]) || ISBLANK([LastYearProfit]),

    0,

    DIVIDE([YearlyProfit] - [LastYearProfit], [LastYearProfit])

)

%MoM Profit =

IF (

    ISBLANK([MonthlyProfit]) || ISBLANK([LastMonthProfit]),

    "No Data",

    DIVIDE([MonthlyProfit] - [LastMonthProfit], [LastMonthProfit])

)

%MoMoftheLastYearProfit = IF (

    ISBLANK([MonthlyProfit]) || ISBLANK([LastYearMonthProfit]),

    "No Data",

    DIVIDE([MonthlyProfit] - [LastYearMonthProfit], [LastYearMonthProfit])

)

TooltipStateTitle = SELECTEDVALUE(fSales\_Data[State])

**Colour Palette**

The colour palette was created in the website <https://coolors.co/>

