END-TO-END DATA ANALYSIS PROJECT

NGUYEN CONG LUC
nguyencongluc.82@gmail.com
Luc Nguyen | LinkedIn
0329206845

INDEX

Preface

I am pleased to share a data analysis project that I've worked on, where I built a product analytics dashboard to track revenue, profits, and trends.

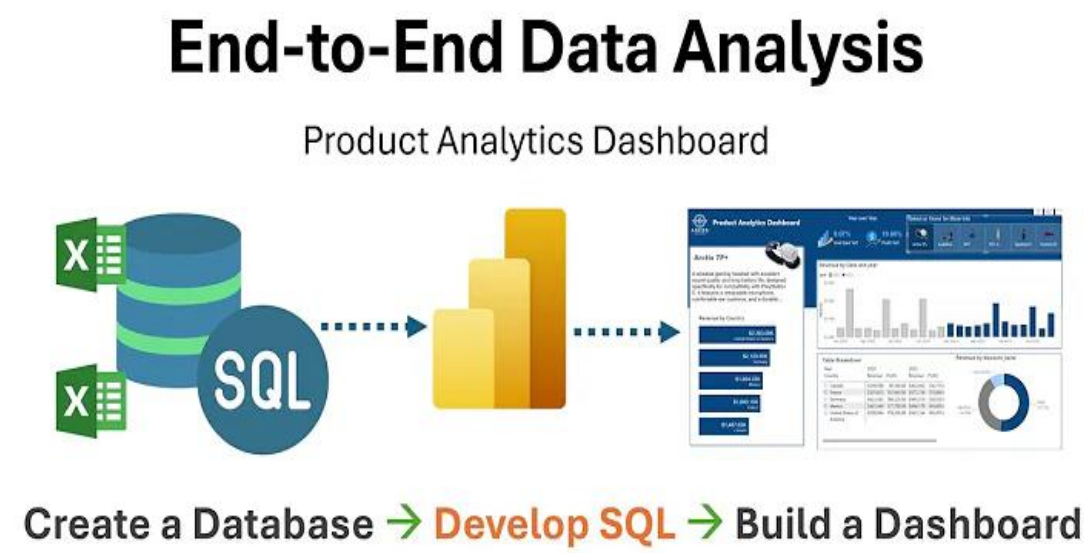
This project gave me the opportunity to practice key skills such as database design, SQL query writing, and using Power BI to create simple yet effective visualizations.

Although I'm still learning and refining my skills, I believe this project helped me better understand the practical steps involved in data analysis, and I hope it demonstrates my growing ability to apply these skills in real-world scenarios.

CHAPTER 1. Introduction

1.1 Overview of the Project

The goal of this project is to develop an end-to-end Product Analytics Dashboard that offers actionable insights into key business metrics such as revenue, profits, and sales trends. This dashboard was created using real-world business data, allowing me to apply various data analysis and visualization techniques to a practical use case.



1.2 Objective

The primary objective of this project is to build a comprehensive dashboard that enables stakeholders to monitor and understand product performance over time. By analyzing revenue, profit, and sales trends, the dashboard serves as a valuable tool for business decision-making.

1.3 Tools Used

SQL Server: To store and manage the data, and to perform queries for extracting meaningful insights.

Power BI: For visualizing the data and creating interactive reports that can be used by business leaders to track performance.

Excel: Used for initial data cleaning and basic analysis before transferring the data into SQL.

CHAPTER 2. Project data setup

2.1 Overview of the Data

The project uses real-world datasets: Product Data, Product Sales, and Discount Data, capturing product details, sales transactions, and discounts.

2.2 Data sources

Product Data: Product categories and prices.

Product ID	Product	Category	Cost Price	Sale Price	Brand	Description	Image url
SR1001	MV7	Dynamic M	$174	$199	Shure	A versatile	https://absentdata.com/wp-content/uploads/2024/11/m7_gif.gif
RR10020	NT1-A	Condenser	$160	$229	Rode	A popular c	https://absentdata.com/wp-content/uploads/2024/11/nt-1_image-removebg-preview.png
FR10021	Scarlett 2i	Audio Inter	$118	$169	Focusrite	A compact	https://absentdata.com/wp-content/uploads/2024/11/focusrite-scarlett-2i2-2-removebg-preview.png
PS10022	AudioBox (	Recording	$90	$199	PreSonus	A complete	https://absentdata.com/wp-content/uploads/2024/11/AudioBox-USB-96-Studio.png
HY1004	QuadCast	Streaming	$97	$139	HyperX	A USB cor	https://absentdata.com/wp-content/uploads/2024/11/hyper_x_color_gif.gif
AR1200	Arctis 7P+	Wireless G	$104	$149	SteelSerie	A wireless	https://absentdata.com/wp-content/uploads/2024/11/Arc_Headset-removebg-preview.png

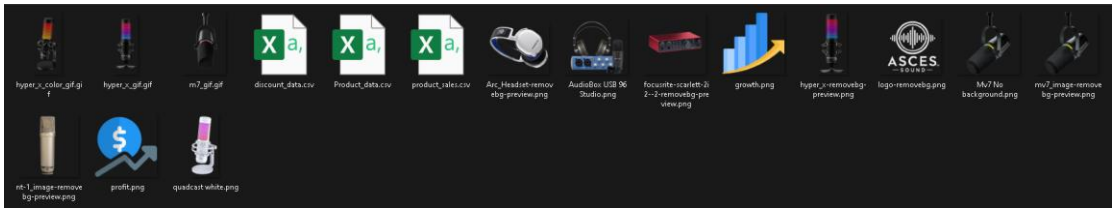
Product Sales: Sales volumes and revenue.

Date	Customer	Country	Product	Discount I	Units Sold
1/12/2023	Governme	Germany	SR1001	None	15
1/12/2023	Education	United Stat	RR10020	None	6
1/12/2023	Governme	Canada	FR10021	None	18
1/12/2023	Governme	Germany	FR10021	None	15
1/12/2023	Governme	Canada	HY1004	None	18
1/12/2023	Education	United States of Amer		None	6
1/12/2023	Governme	France	SR1001	Low	22
1/12/2023	Governme	France	FR10021	Low	22
1/12/2023	Governme	Canada	SR1001	Low	29

Discount Data: Discounts and rates.

	B	C
1	Discount E	Discount
2	none	0
3	low	6
4	medium	10
5	high	5
6	none	0
7	low	6
8	medium	11
9	high	5
10	none	0
11	low	6

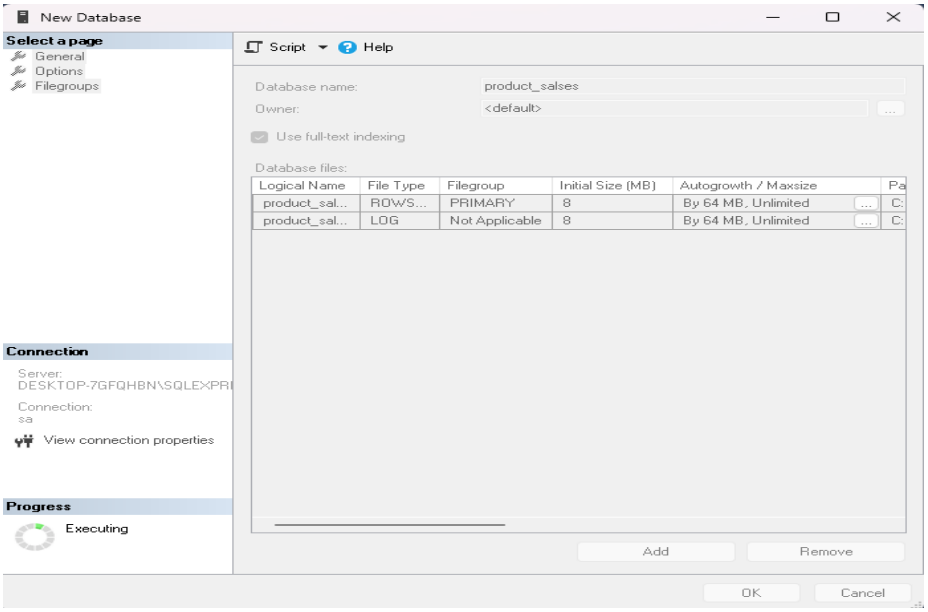
Icons: Growth and profit icons for visualizations.



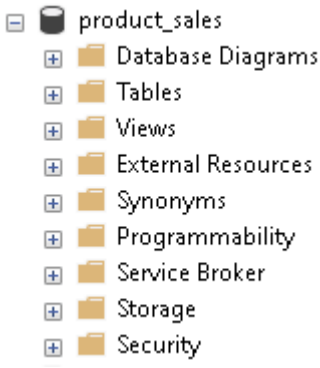
CHAPTER 3. Building the Database

3.1 Database Design and Structure

Create database product_sales to stored data



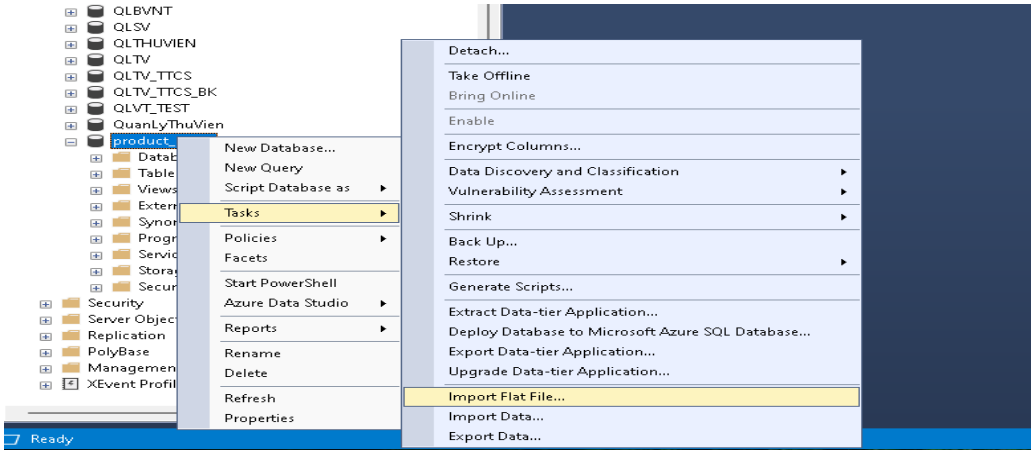
Result

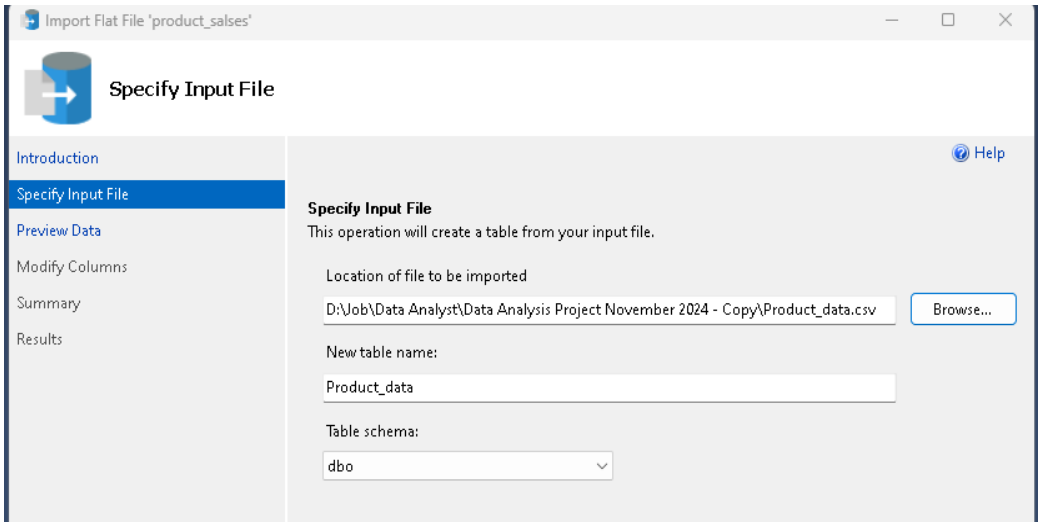


3.2 Data Import and Preprocessing

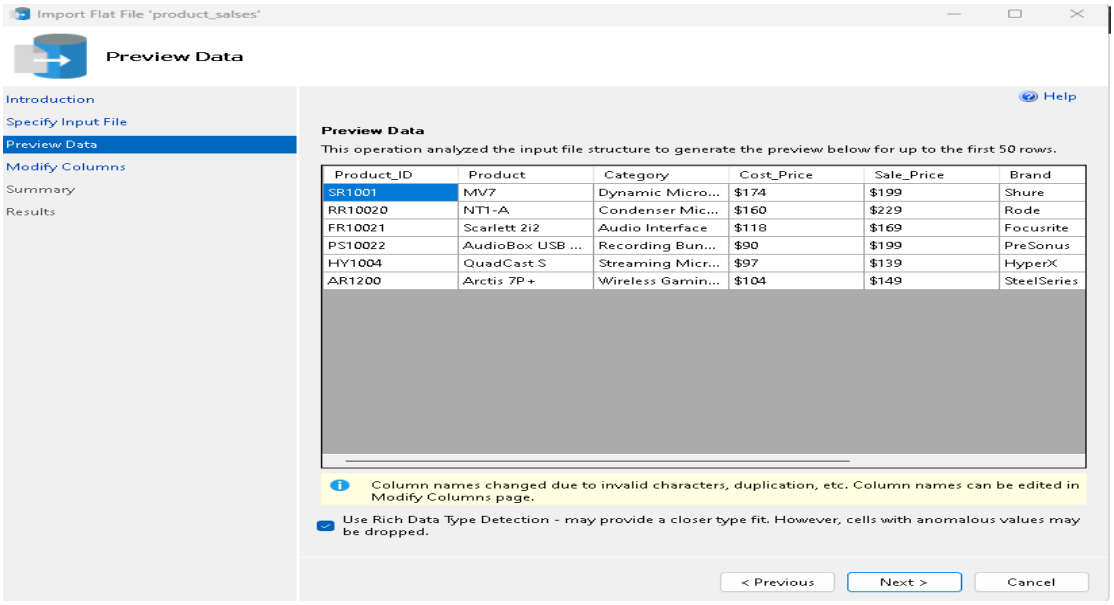
3.2.1. Import file product_data.csv

File import operation

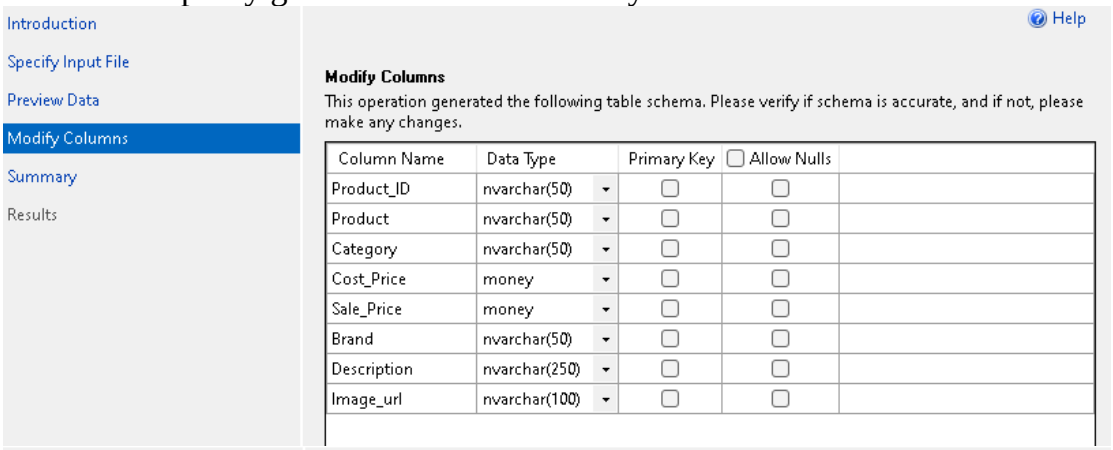


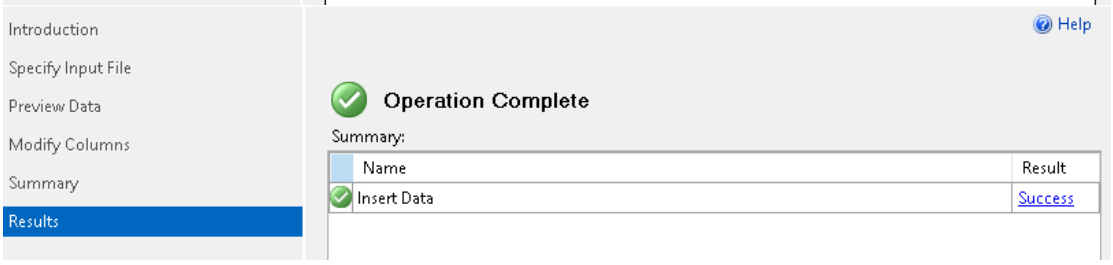


Review data

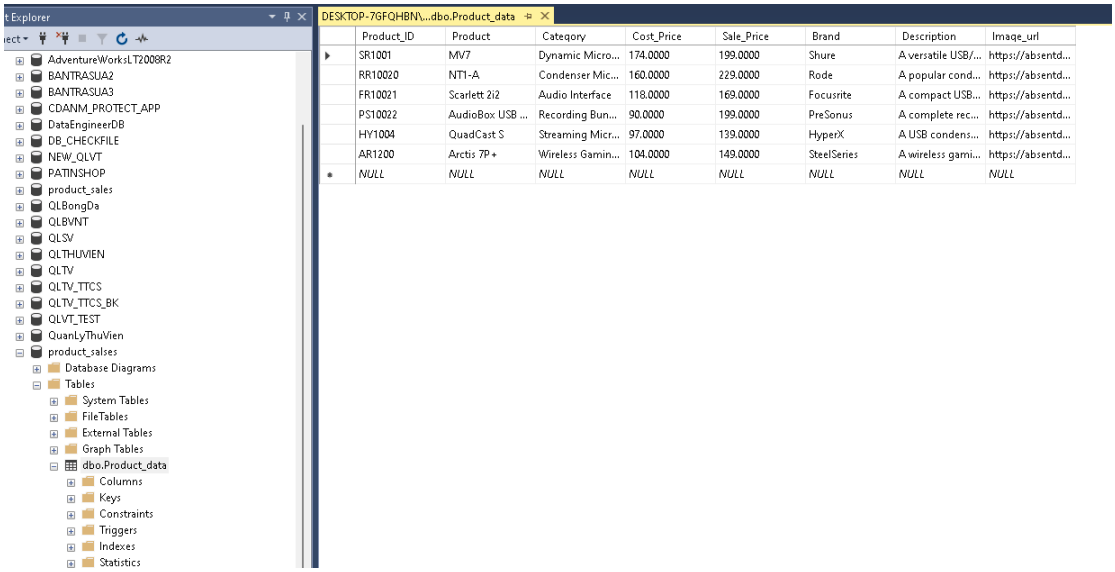


This table is pretty good so no need to Modify Columns for now



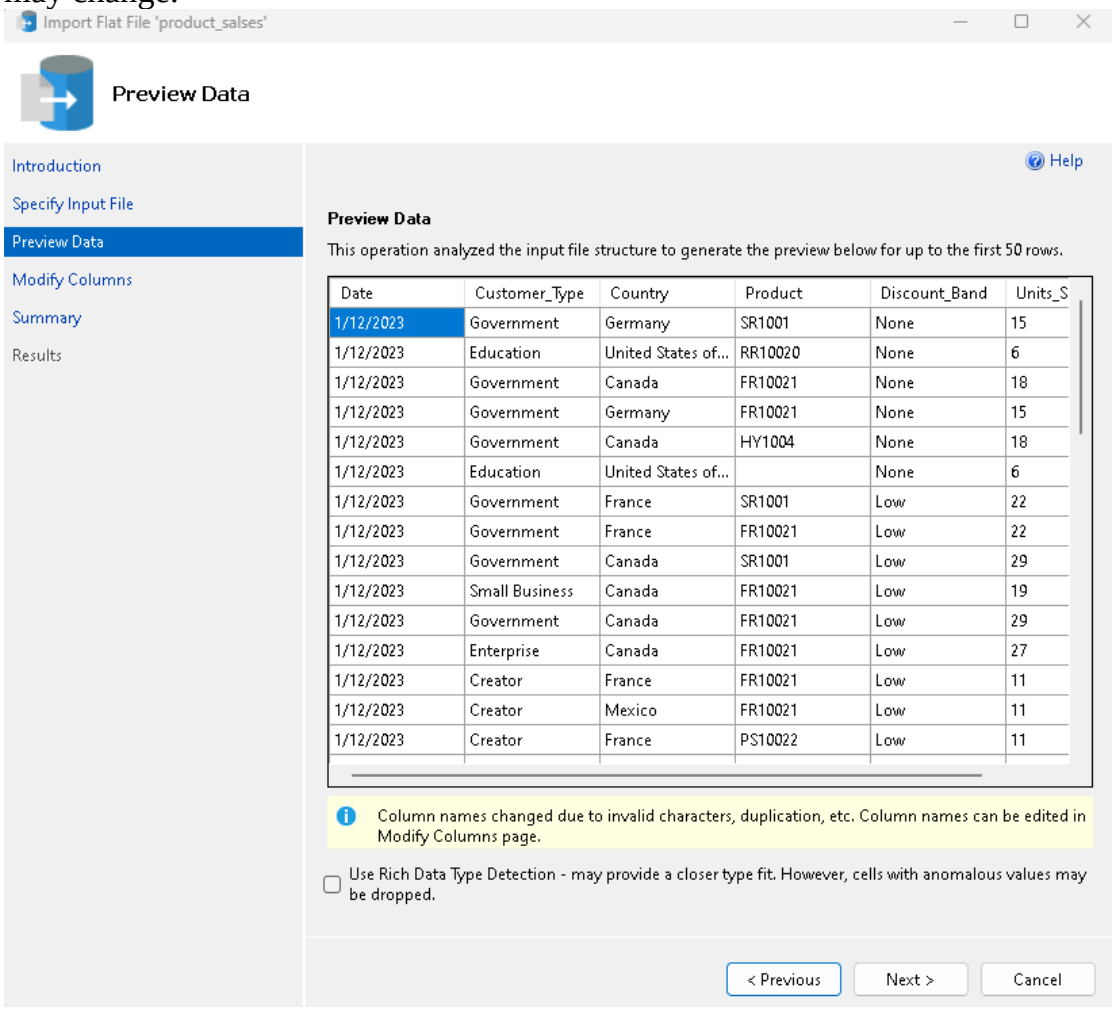


Result

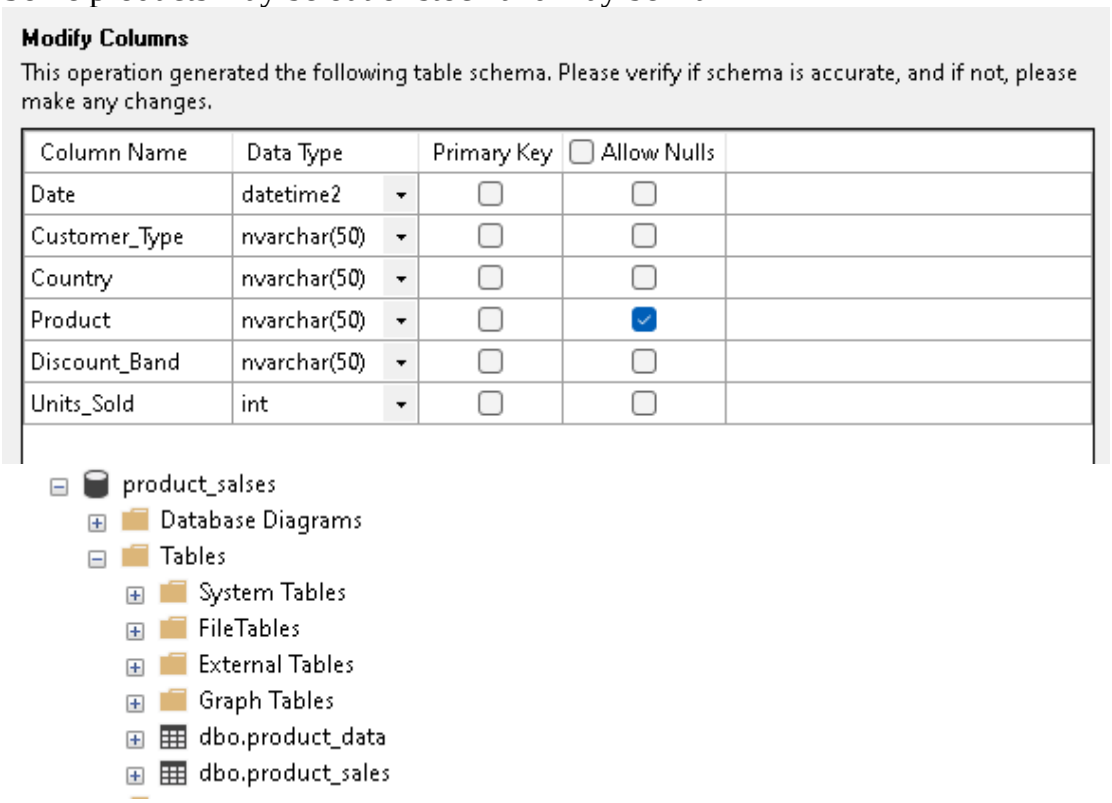


3.2.2 Import file product_sales.csv

Do not select Use Rich Data Type Detection because the Date column data may change.



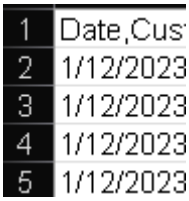
Some products may be out of stock and may be Null



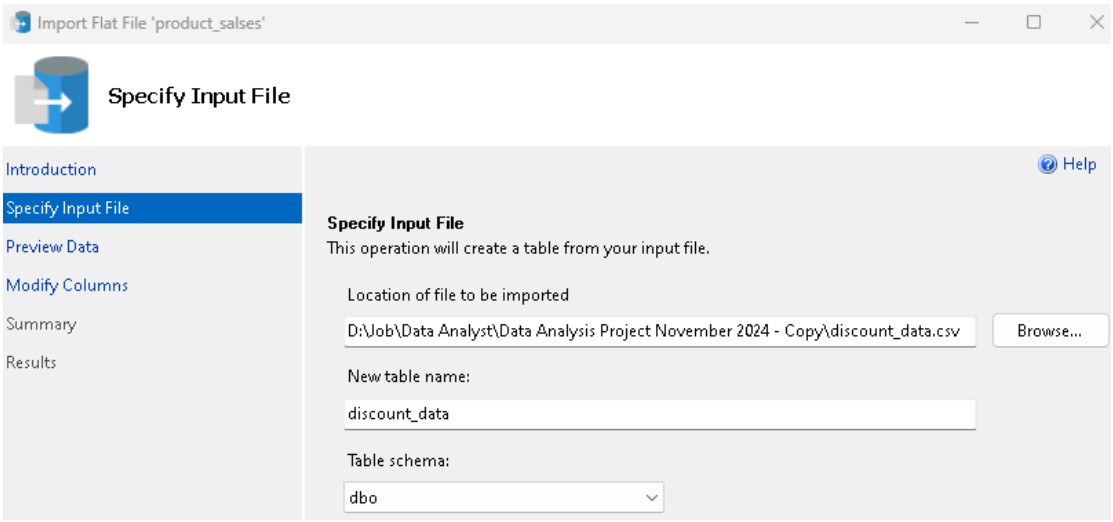
3.2.3 Import file product_sales.csv

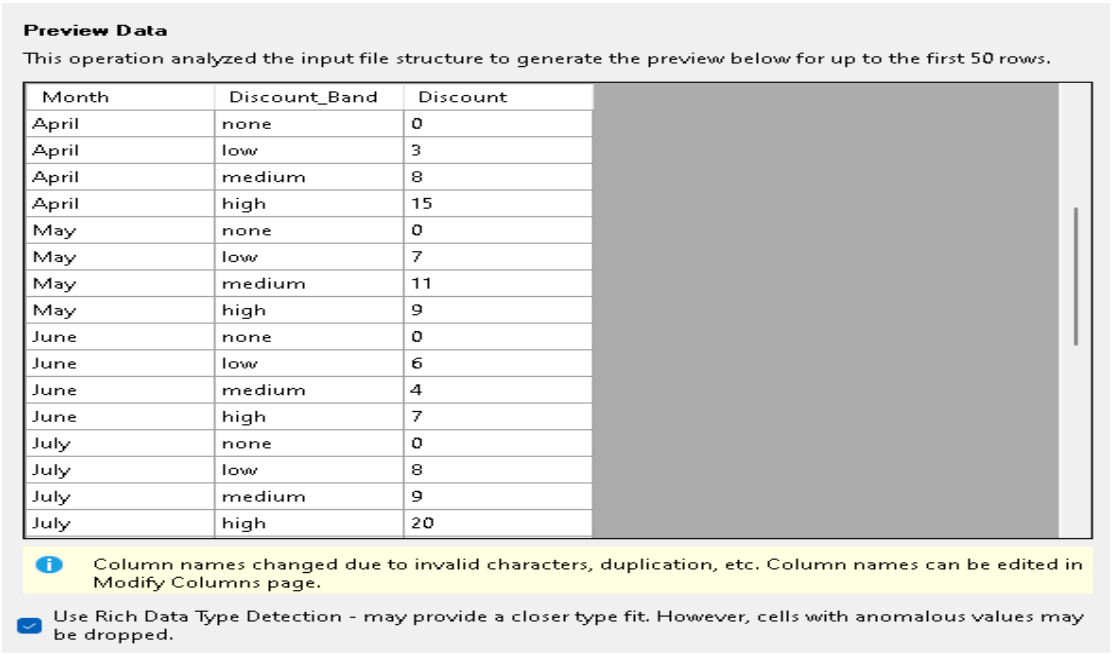
By default, SQL Server uses system regional settings to interpret date formats (e.g., DMY for Vietnam, MDY for US), then converts them to YMD format when storing.

-To handle DMY format, run SET DATEFORMAT DMY before importing.




```
SET DATEFORMAT DMY;
```





Preview Data

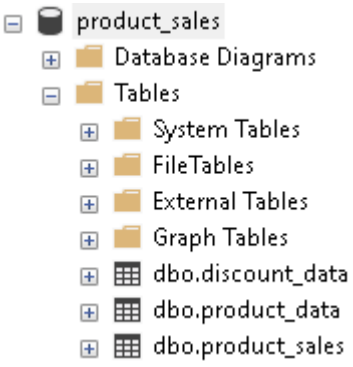
This operation analyzed the input file structure to generate the preview below for up to the first 50 rows.

Month	Discount_Band	Discount
April	none	0
April	low	3
April	medium	8
April	high	15
May	none	0
May	low	7
May	medium	11
May	high	9
June	none	0
June	low	6
June	medium	4
June	high	7
July	none	0
July	low	8
July	medium	9
July	high	20

Column names changed due to invalid characters, duplication, etc. Column names can be edited in Modify Columns page.

Use Rich Data Type Detection - may provide a closer type fit. However, cells with anomalous values may be dropped.

Final result

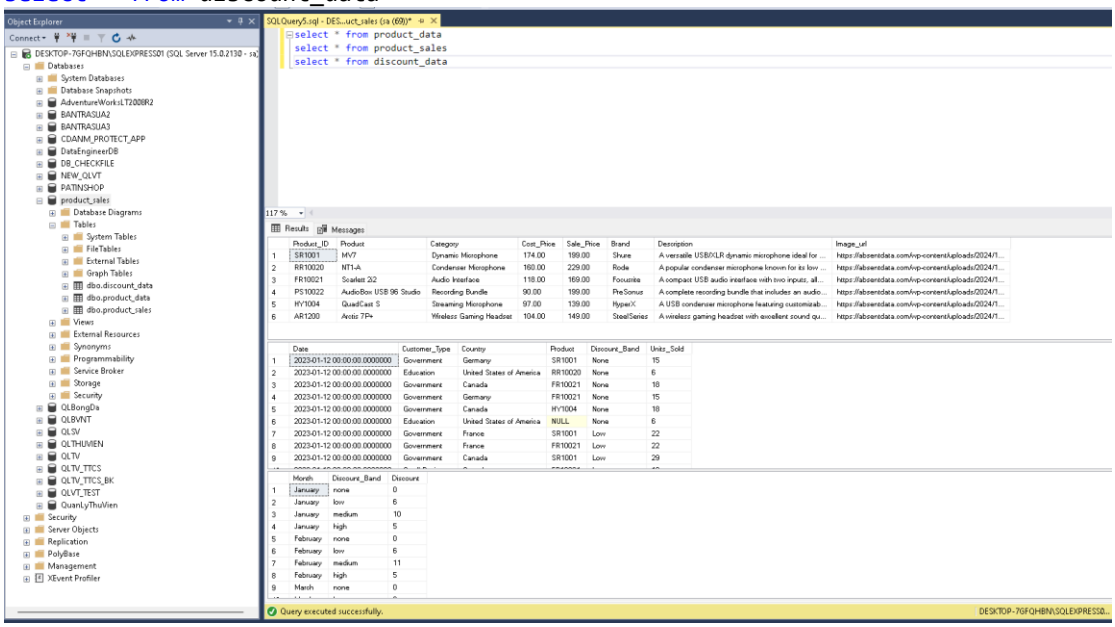


CHAPTER 4. Developing SQL Queries

4.1 Writing SQL Queries to Extract Data

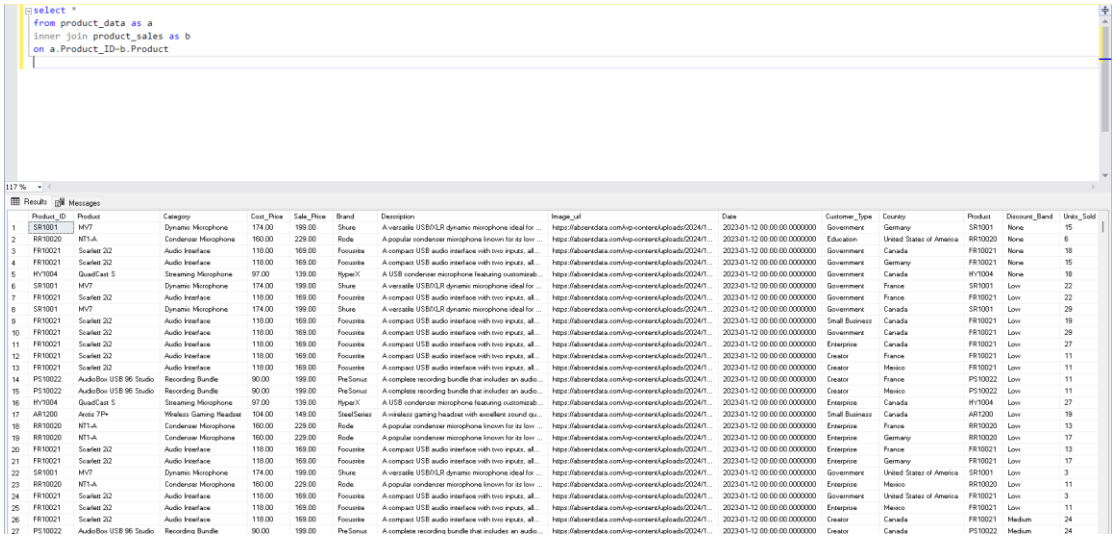
Check data in tables

```
select * from product_data
select * from product_sales
select * from discount_data
```



Check which products have been sold

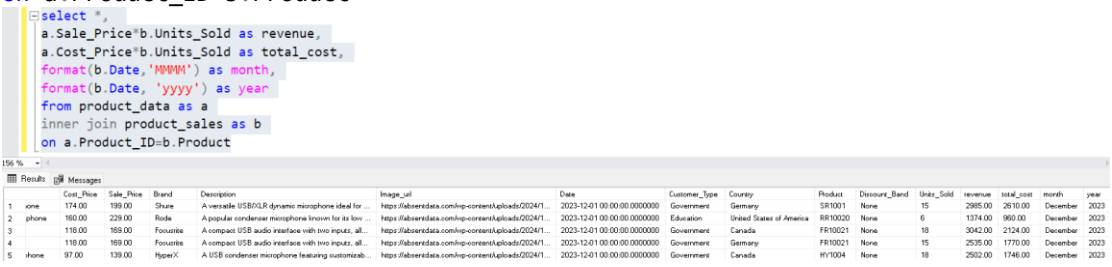
```
select *
from product_data as a
inner join product_sales as b
on a.Product_ID=b.Product
```



4.2 Calculations: Revenue, Total Cost, Profit, Sales Trends

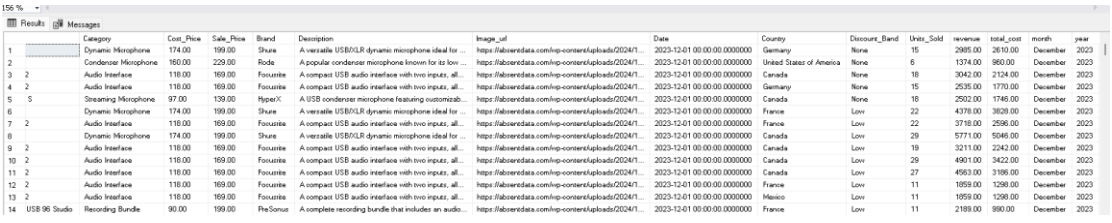
View Revenue and Total Cost by month and year

```
select *,
a.Sale_Price*b.Units_Sold as revenue,
a.Cost_Price*b.Units_Sold as total_cost,
format(b.Date,'MMMM') as month,
format(b.Date, 'yyyy') as year
from product_data as a
inner join product_sales as b
on a.Product_ID=b.Product
```



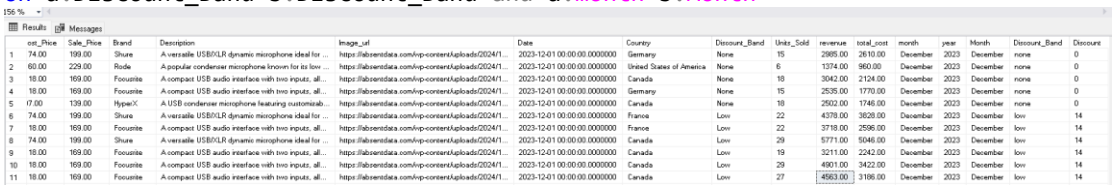
Filtering and Aggregating Data

```
select
a.Product,
a.Category,
a.Cost_Price,
a.Sale_Price,
a.Brand,
a.Description,
a.Image_url,
b.Date,
b.Country,
b.Discount_Band,
b.Units_Sold,
a.Sale_Price*b.Units_Sold as revenue,
a.Cost_Price*b.Units_Sold as total_cost,
format(b.Date,'MMMM') as month,
format(b.Date, 'yyyy') as year
from product_data as a
inner join product_sales as b
on a.Product_ID=b.Product
```



Combines data from three tables (product_data, product sales, discount data) to overall view of revenue and discounts for each month.

```
with cte as(
select
a.Product,
a.Category,
a.Cost_Price,
a.Sale_Price,
a.Brand,
a.Description,
a.Image_url,
b.Date,
b.Country,
b.Discount_Band,
b.Units_Sold,
a.Sale_Price*b.Units_Sold as revenue,
a.Cost_Price*b.Units_Sold as total_cost,
format(b.Date,'MMMM') as month,
format(b.Date, 'yyyy') as year
from product_data as a
inner join product_sales as b
on a.Product_ID=b.Product )
select * from cte as a
inner join discount_data as b
on a.Discount_Band=b.Discount_Band and a.month=b.Month
```



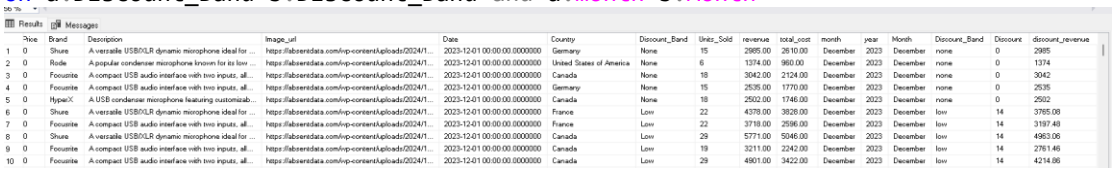
156 %

Results | Messages

	ost_Price	Sale_Price	Brand	Description	Image_url	Date	Country	Discount_Band	Units_Sold	revenue	total_cost	month	year	Month	Discount_Band	Discount
1	74.00	199.00	Shure	A versatile USB/XLR dynamic microphone ideal for ...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Germany	None	15	2985.00	2610.00	December	2023	December	none	0
2	60.00	229.00	Rode	A popular condenser microphone known for its low ...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	United States of America	None	6	1374.00	960.00	December	2023	December	none	0
3	18.00	169.00	Focusrite	A compact USB audio interface with two inputs, all...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	None	18	3042.00	2124.00	December	2023	December	none	0
4	18.00	169.00	Focusrite	A compact USB audio interface with two inputs, all...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Germany	None	15	2535.00	1770.00	December	2023	December	none	0
5	17.00	139.00	HyperX	A USB condenser microphone featuring customizab...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	None	18	2502.00	1746.00	December	2023	December	none	0
6	74.00	199.00	Shure	A versatile USB/XLR dynamic microphone ideal for ...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	France	Low	22	4378.00	3628.00	December	2023	December	low	14
7	18.00	169.00	Focusrite	A compact USB audio interface with two inputs, all...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	France	Low	22	3718.00	2596.00	December	2023	December	low	14
8	74.00	199.00	Shure	A versatile USB/XLR dynamic microphone ideal for ...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	Low	29	5771.00	5046.00	December	2023	December	low	14
9	18.00	169.00	Focusrite	A compact USB audio interface with two inputs, all...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	Low	19	3211.00	2242.00	December	2023	December	low	14
10	18.00	169.00	Focusrite	A compact USB audio interface with two inputs, all...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	Low	29	4901.00	3422.00	December	2023	December	low	14
11	18.00	169.00	Focusrite	A compact USB audio interface with two inputs, all...	https://absentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	Low	27	4563.00	3186.00	December	2023	December	low	14

Calculate revenue after discount %

```
with cte as(
select
a.Product,
a.Category,
a.Cost_Price,
a.Sale_Price,
a.Brand,
a.Description,
a.Image_url,
b.Date,
b.Customer_Type,
b.Country,
b.Discount_Band,
b.Units_Sold,
a.Sale_Price*b.Units_Sold as revenue,
a.Cost_Price*b.Units_Sold as total_cost,
format(b.Date,'MMMM') as month,
format(b.Date, 'yyyy') as year
from product_data as a
inner join product_sales as b
on a.Product_ID=b.Product )
select
*,
(1-CAST(b.Discount as float)/100)*revenue as discount_revenue
from cte as a
inner join discount_data as b
on a.Discount_Band=b.Discount_Band and a.month=b.Month
```



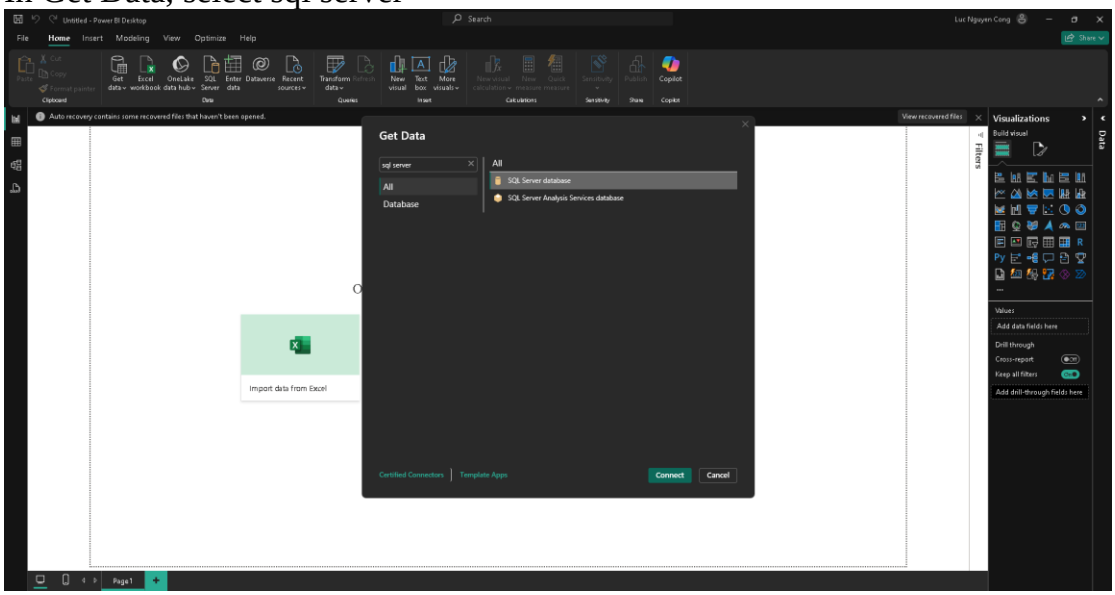
Results Messages

	Price	Brand	Description	Image_url	Date	Country	Discount_Band	Units_Sold	revenue	total_cost	month	year	Month	Discount_Band	Discount	discount_revenue
1	0	Shure	A versatile USB/XLR dynamic microphone ideal for ...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Germany	None	15	2985.00	2610.00	December	2023	December	none	0	2985
2	0	Rode	A popular condenser microphone known for its low ...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	United States of America	None	6	1374.00	960.00	December	2023	December	none	0	1374
3	0	Focusrite	A compact USB audio interface with two inputs, all...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	None	18	3042.00	2124.00	December	2023	December	none	0	3042
4	0	Focusrite	A compact USB audio interface with two inputs, all...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Germany	None	15	2535.00	1770.00	December	2023	December	none	0	2535
5	0	HyperX	A USB condenser microphone featuring customizab...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	None	18	2502.00	1746.00	December	2023	December	none	0	2502
6	0	Shure	A versatile USB/XLR dynamic microphone ideal for ...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	France	Low	22	4378.00	3826.00	December	2023	December	low	14	3765.08
7	0	Focusrite	A compact USB audio interface with two inputs, all...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	France	Low	22	3718.00	2596.00	December	2023	December	low	14	3197.48
8	0	Shure	A versatile USB/XLR dynamic microphone ideal for ...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	Low	29	5771.00	5046.00	December	2023	December	low	14	4963.06
9	0	Focusrite	A compact USB audio interface with two inputs, all...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	Low	19	3211.00	2242.00	December	2023	December	low	14	2761.46
10	0	Focusrite	A compact USB audio interface with two inputs, all...	https://labsentdata.com/wp-content/uploads/2024/1...	2023-12-01 00:00:00.0000000	Canada	Low	29	4901.00	3422.00	December	2023	December	low	14	4214.86

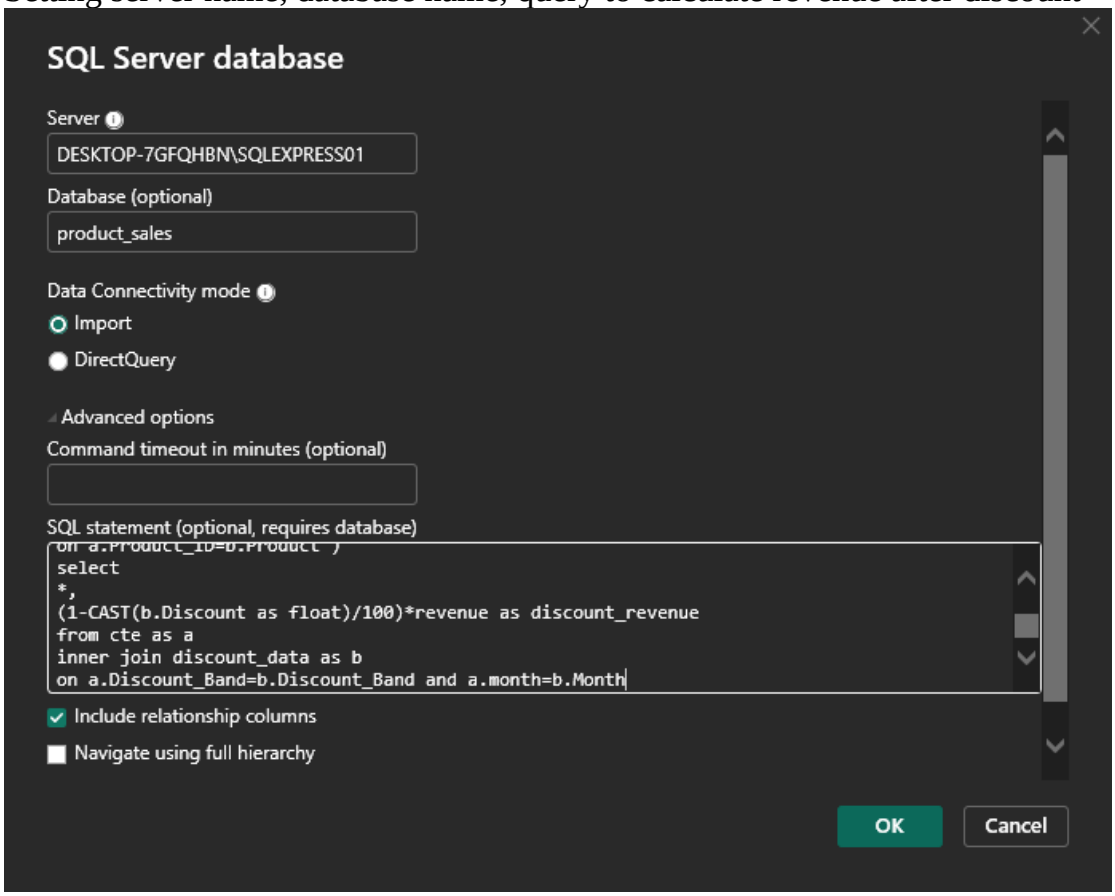
CHAPTER 5. Connecting to Power BI

5.1 Connecting SQL Database to Power BI

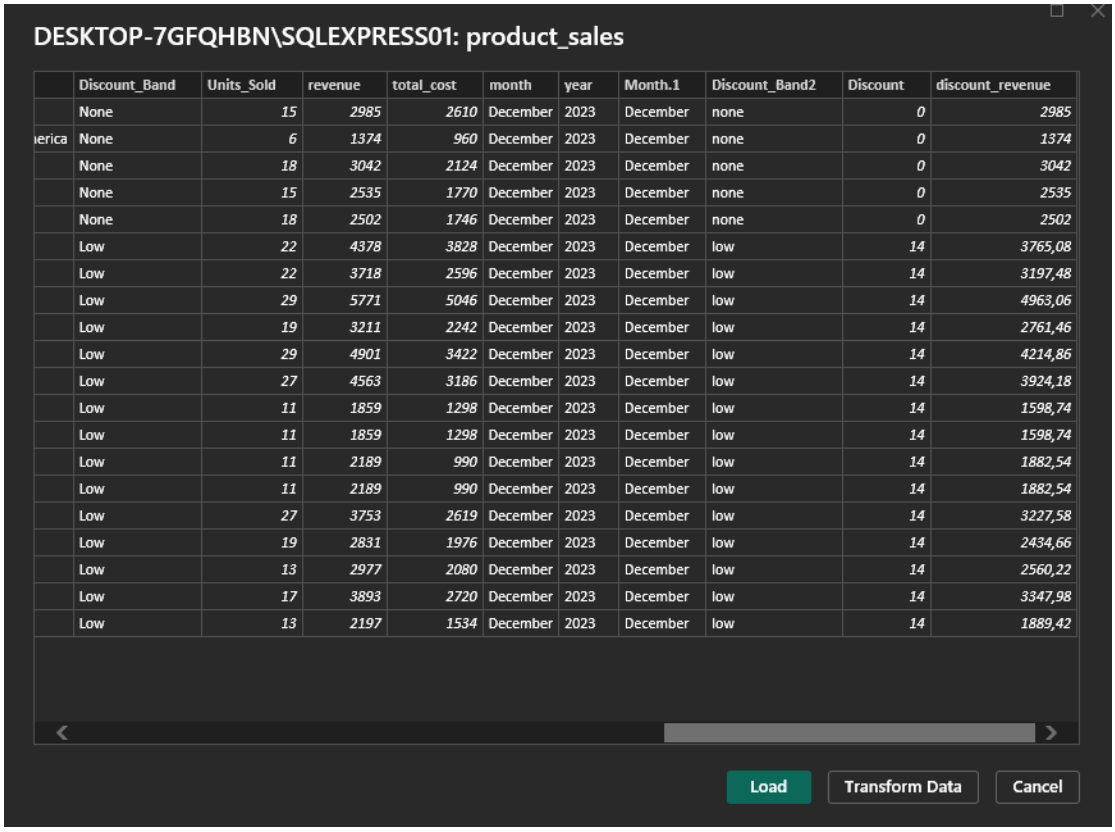
In Get Data, select sql server



Setting server name, database name, query to calculate revenue after discount



5.2 Check and load data

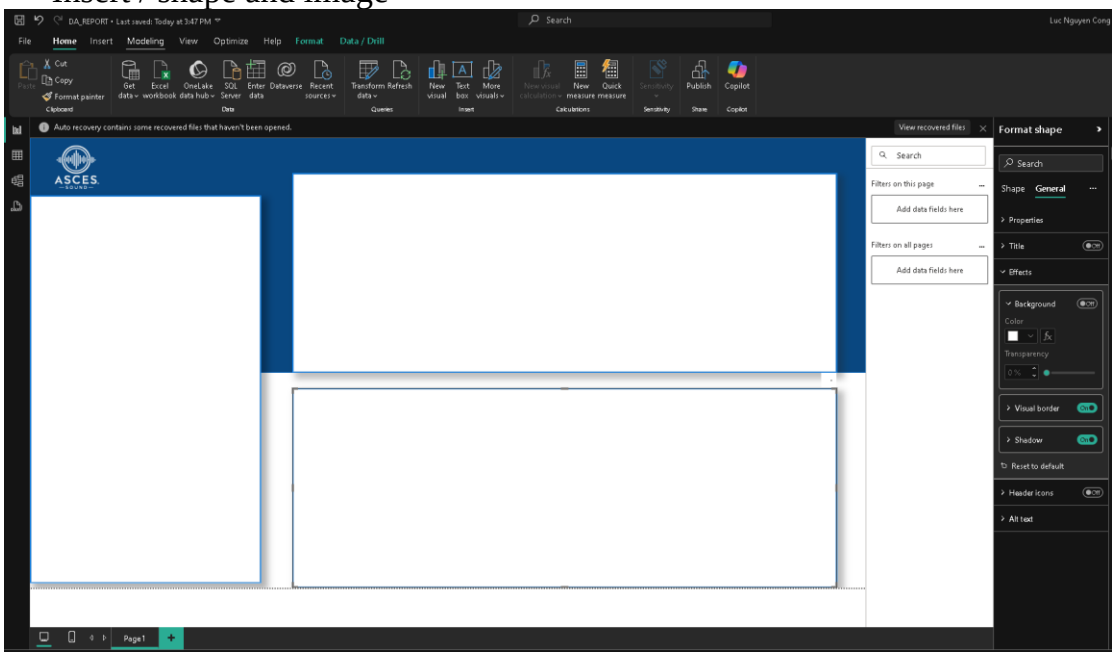


	Discount_Band	Units_Sold	revenue	total_cost	month	year	Month.1	Discount_Band2	Discount	discount_revenue
	None	15	2985	2610	December	2023	December	none	0	2985
ıerica	None	6	1374	960	December	2023	December	none	0	1374
	None	18	3042	2124	December	2023	December	none	0	3042
	None	15	2535	1770	December	2023	December	none	0	2535
	None	18	2502	1746	December	2023	December	none	0	2502
	Low	22	4378	3828	December	2023	December	low	14	3765,08
	Low	22	3718	2596	December	2023	December	low	14	3197,48
	Low	29	5771	5046	December	2023	December	low	14	4963,06
	Low	19	3211	2242	December	2023	December	low	14	2761,46
	Low	29	4901	3422	December	2023	December	low	14	4214,86
	Low	27	4563	3186	December	2023	December	low	14	3924,18
	Low	11	1859	1298	December	2023	December	low	14	1598,74
	Low	11	1859	1298	December	2023	December	low	14	1598,74
	Low	11	2189	990	December	2023	December	low	14	1882,54
	Low	11	2189	990	December	2023	December	low	14	1882,54
	Low	27	3753	2619	December	2023	December	low	14	3227,58
	Low	19	2831	1976	December	2023	December	low	14	2434,66
	Low	13	2977	2080	December	2023	December	low	14	2560,22
	Low	17	3893	2720	December	2023	December	low	14	3347,98
	Low	13	2197	1534	December	2023	December	low	14	1889,42

CHAPTER 6. Building the Dashboard

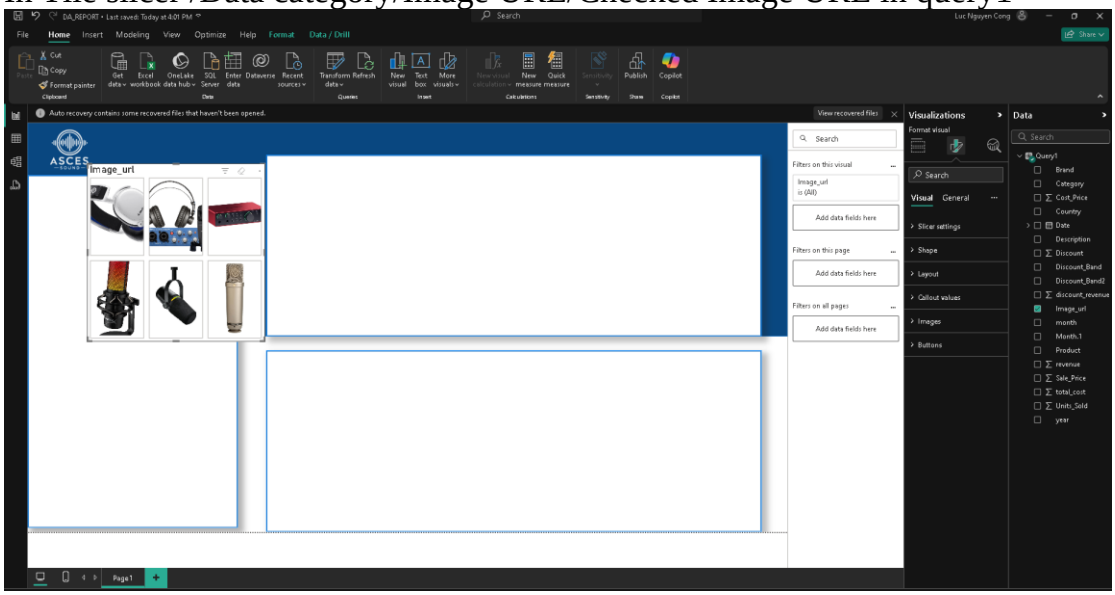
6.1 Building the Dashboard Structure

Insert / shape and image

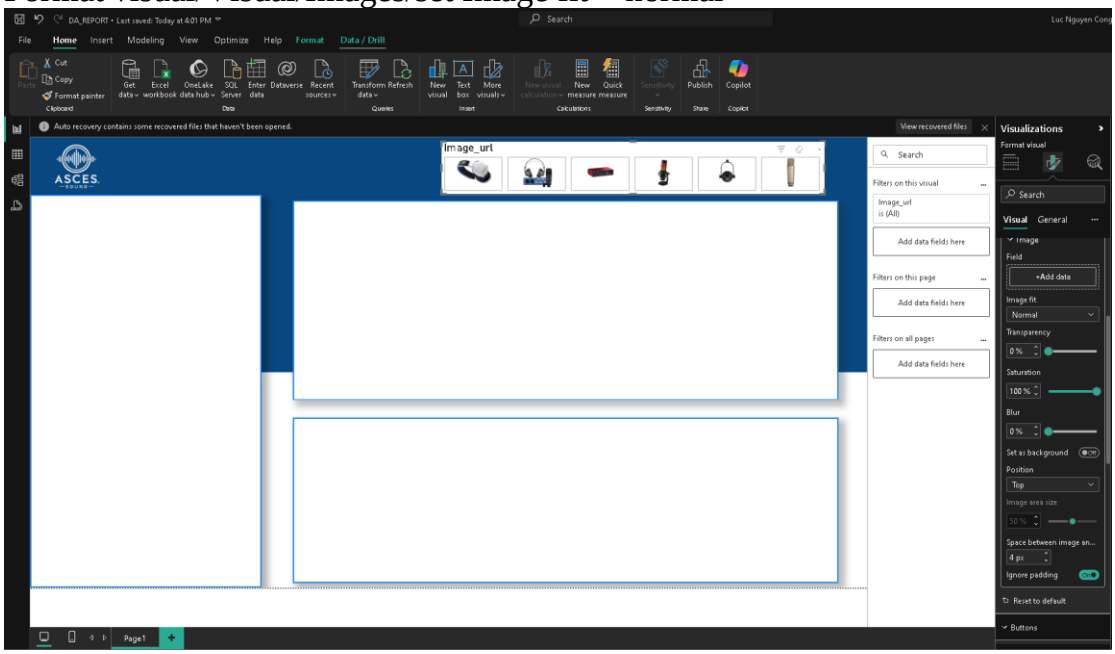


6.2 Image Slicer

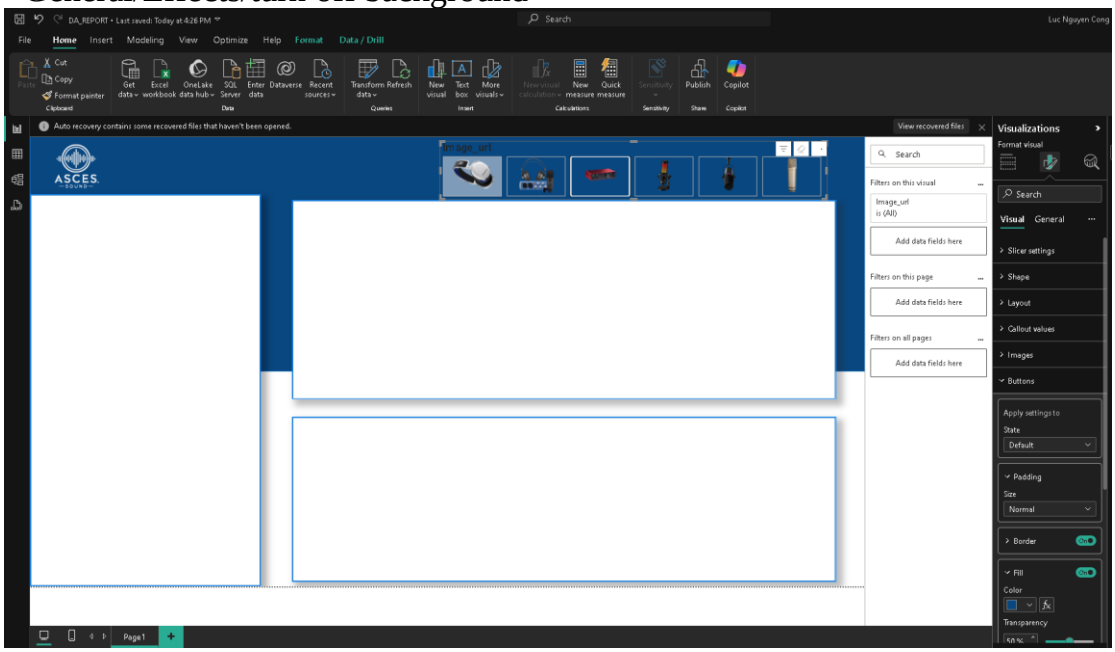
Build visual/ Tile slicer /
In Tile slicer /Data category/Image URL/Checked Image URL in query1



In Format visual/Layout/set Max rows shown = 1/set Columns shown = 6
Format visual/Visual/Images/set Image fit = normal

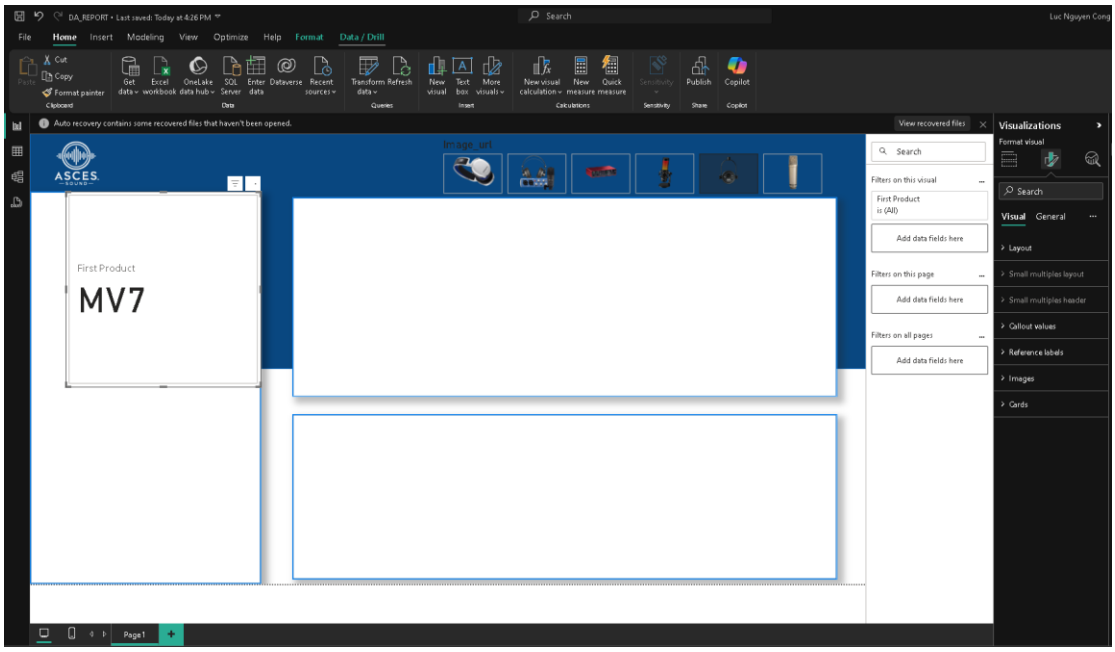


+Visual/button/fill/set color with blue and set Transparency 50 %
+General/Effects/turn off background

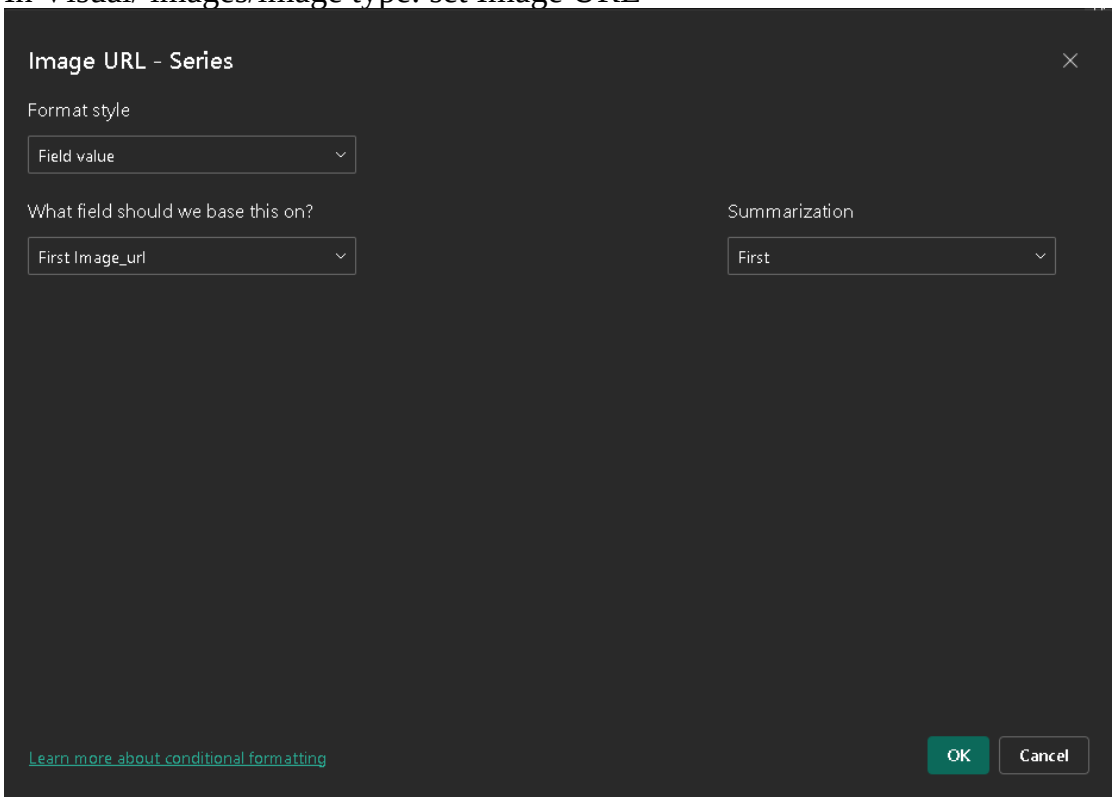


6.3 Create Animated Image

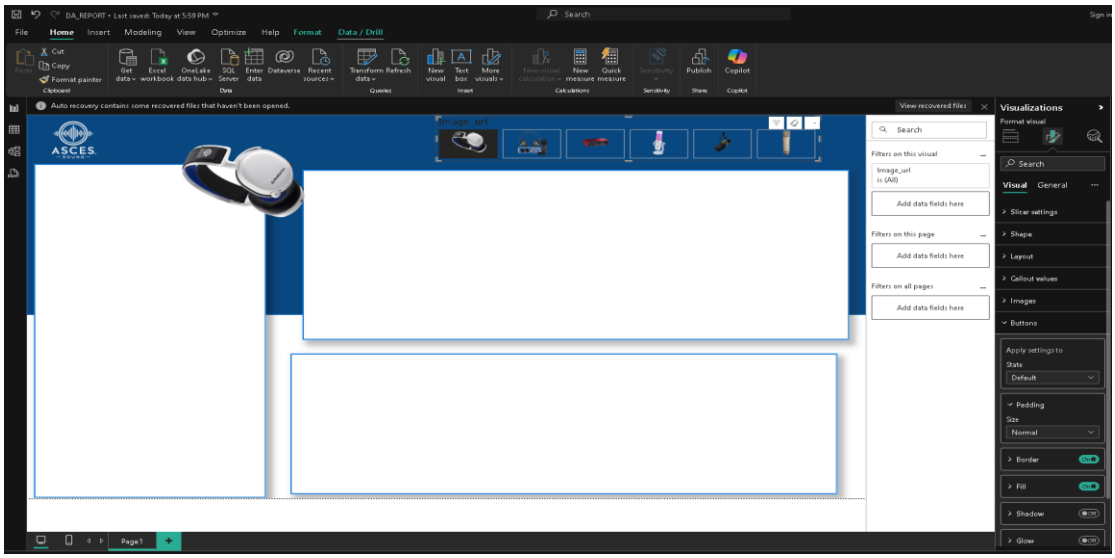
In Build visual click Card -> click column product in Query1->click Card(new) in Build visual



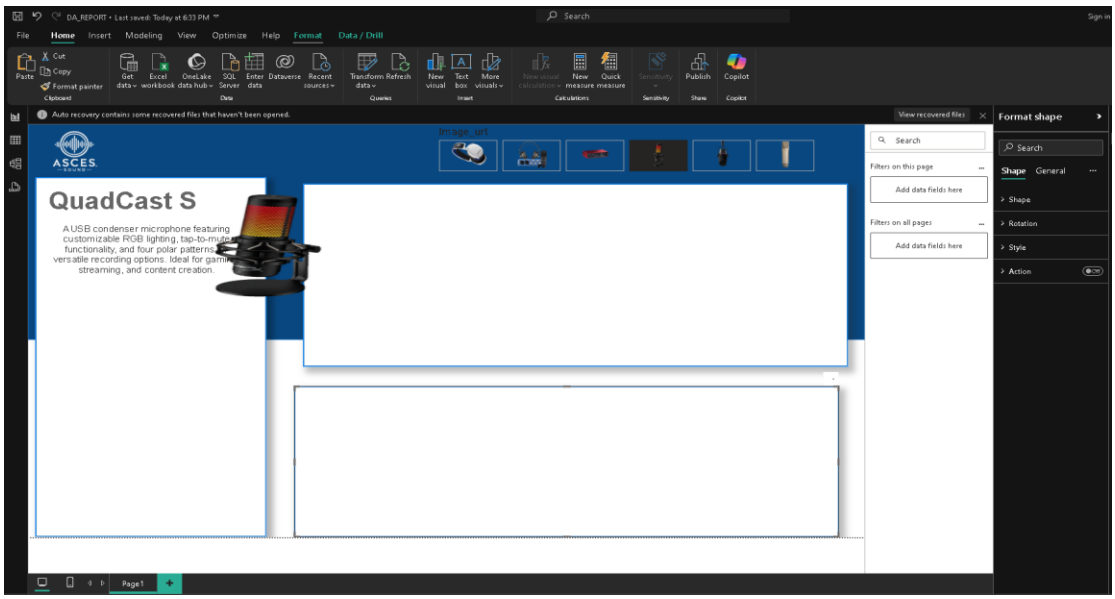
In Visual/ images/image type: set Image URL



Click ok
Visual/ Callout values/ Turn off values and turn off label
General/Header icons->turn off
Visual/Images/Position=Above Text and Size =250px
Visual/Card/turn off background and turn off border
General/Effects/Turn off background



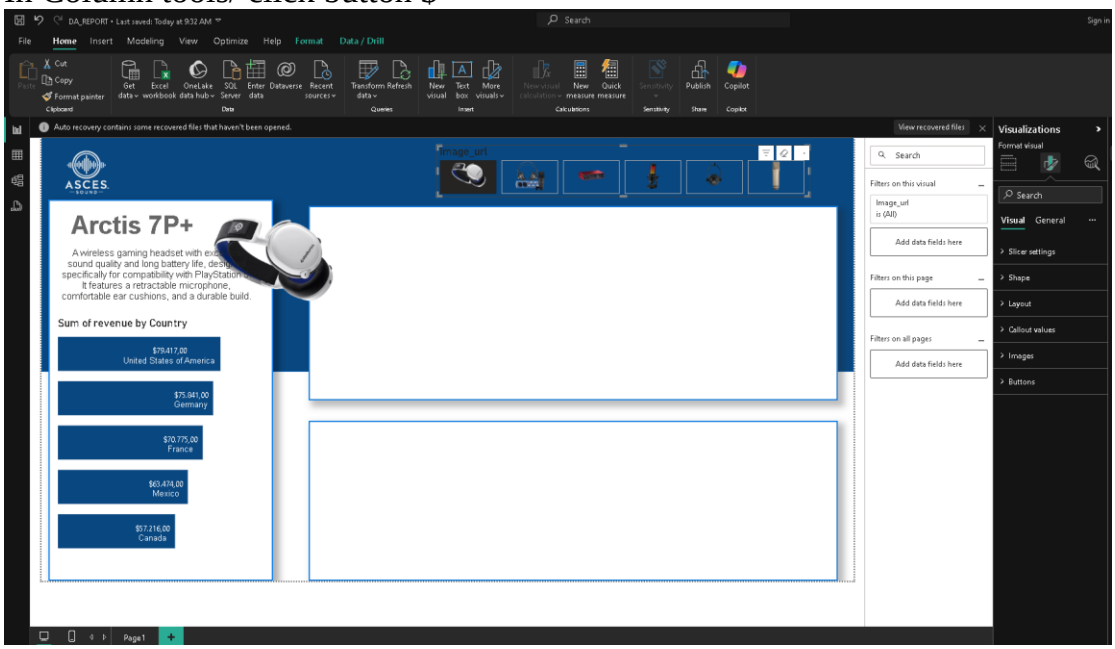
Create description: Build visual /Card/checked Description in query1/click Card(new)
Visual/ Callout values/ Values/ font=Arial and text size =11 and Text wrap = on
General/Effects/Turn off background
Visual/ Callout values/turn off label
Visual/Card/turn of background and border
Create Name: Build visual /Card/checked product in query1/click Card(new) and manual editing



6.4 Bar Charts (Revenue by Country and Over Time)

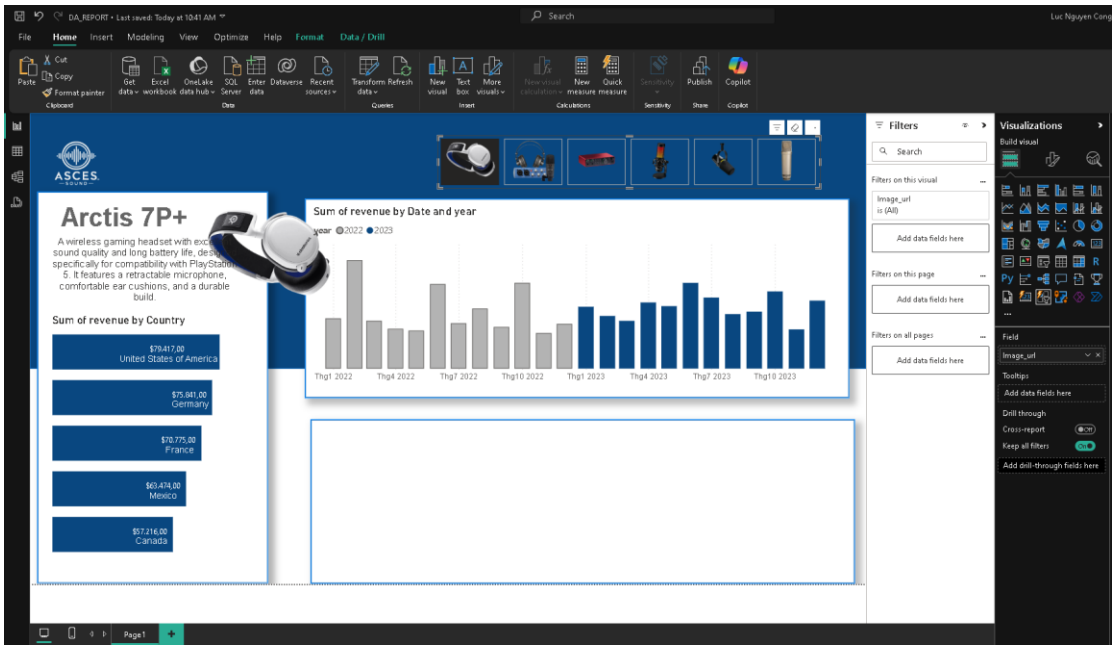
6.4.1. Revenue by Country

Build visual/clustered bar chart/ click column revenue and column Country in query
visual/ X-axis/turn off values and title
visual/ Y-axis/turn off title
Visual/Bars/color=blue
Visual/Data labels = on/ Options/Position = inside end/
Visual/Data labels /Detail/ Date/Add data/Click column contry
Visual/Data labels /Layout/ select multi-line
Visual/Data labels /Detail/ Font/select Arial font and select 10 size font
visual/ Y-axis/turn off values
Visual/Data labels/Value/ Display units =None and Value decimal places = 2/ click column revenue/
In Column tools/ click button $



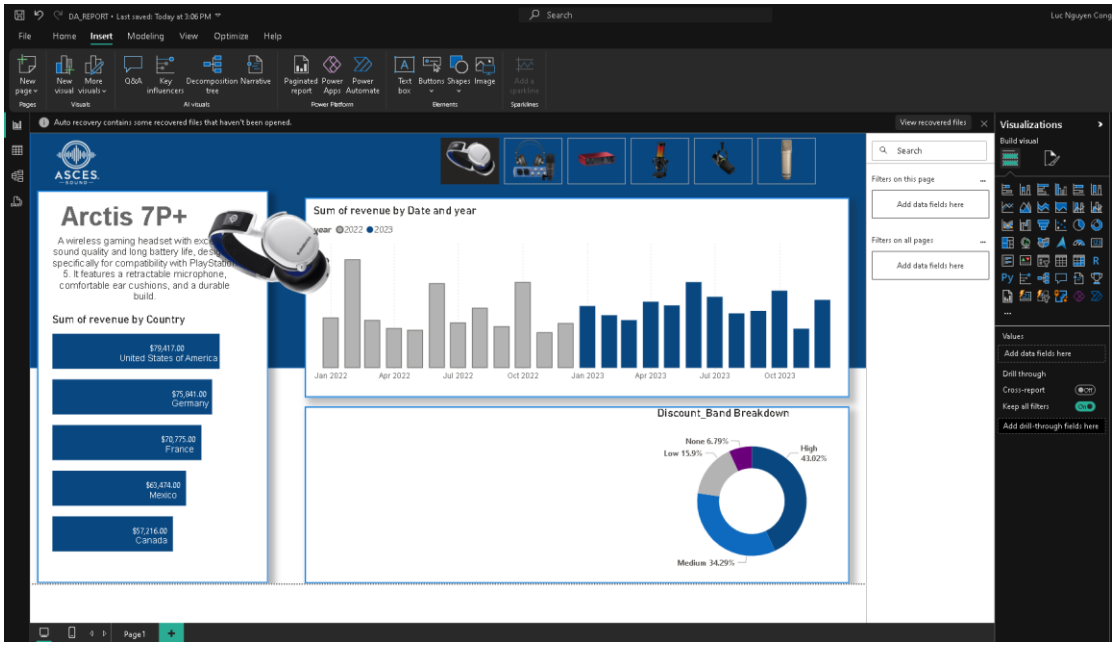
6.4.2. Revenue by Over Time

Build visual/stacked column chart/ X-axis/ Add data fields here / pull Date column to inside / Click right mouse at there/ change Date hierachy to Date .
Build visual/ Y-axis/ Add data fields here / pull Revenue column to inside
Build visual/ Legend/ Add data fields here/ pull Year column to inside
Format visual/Visual/ Y-axis/ Turn off Title and turn off Values
Format visual /Visual/ X-axis/ Turn off Title
Format visual / General / Effects / Background -> set off
Format visual /Visual/ Columns/ Apply settings to Series -> Change to 2022 and Color = Gray and Border = on
Format visual /Visual/ Columns/ Apply settings to Series -> Change to 2023 and Color = Blue



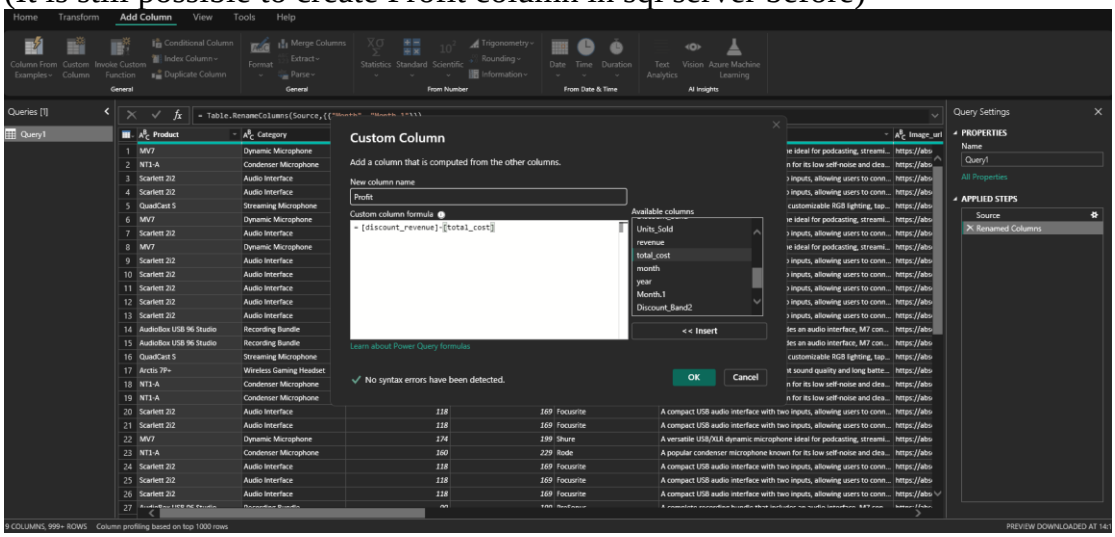
6.5 Pie Charts

Build visual/Donut chart/Click column Revenue and column Discount band
Format visual/Visual/Legend = off
Format visual/Datalabel/ Label contents -> select Category, percent of…
Format visual/General/Effects/Background = off
Change color: Format visual/ Visual/ Slices/ Colors
Format visual/Datalabel/ Values/ Click “B button” to Bold
Format visual/General/

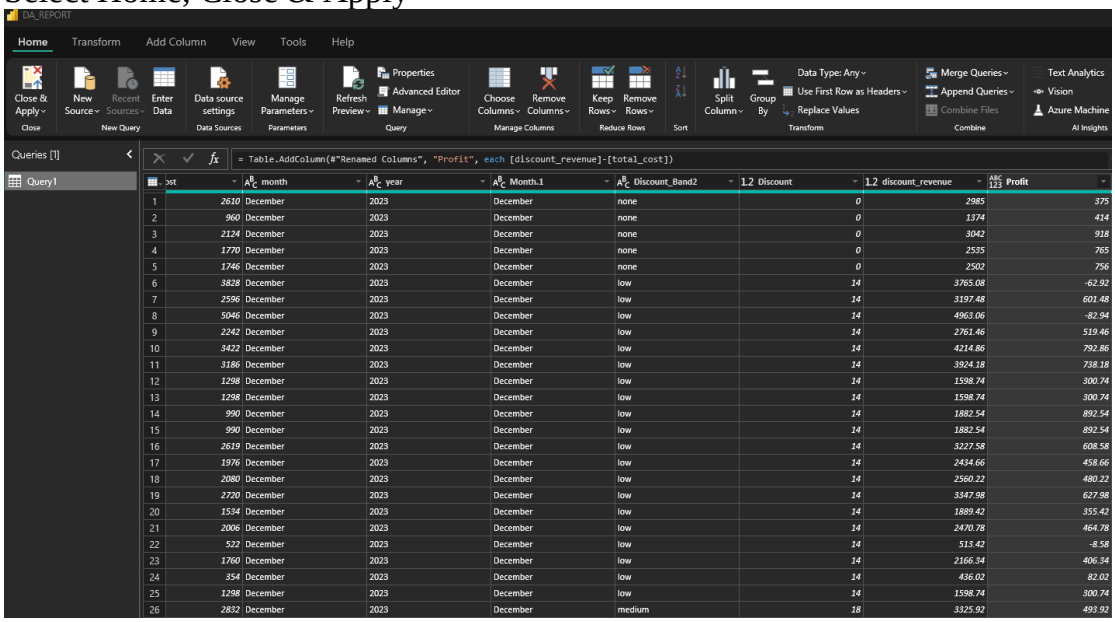


6.6 Add Column and Profit Calculation

Click to Transform Data/ Add Column / Custom column/ Set column = Profit
Profit = Revenue after discount – Total Cost
(It is still possible to create Profit column in sql server before)

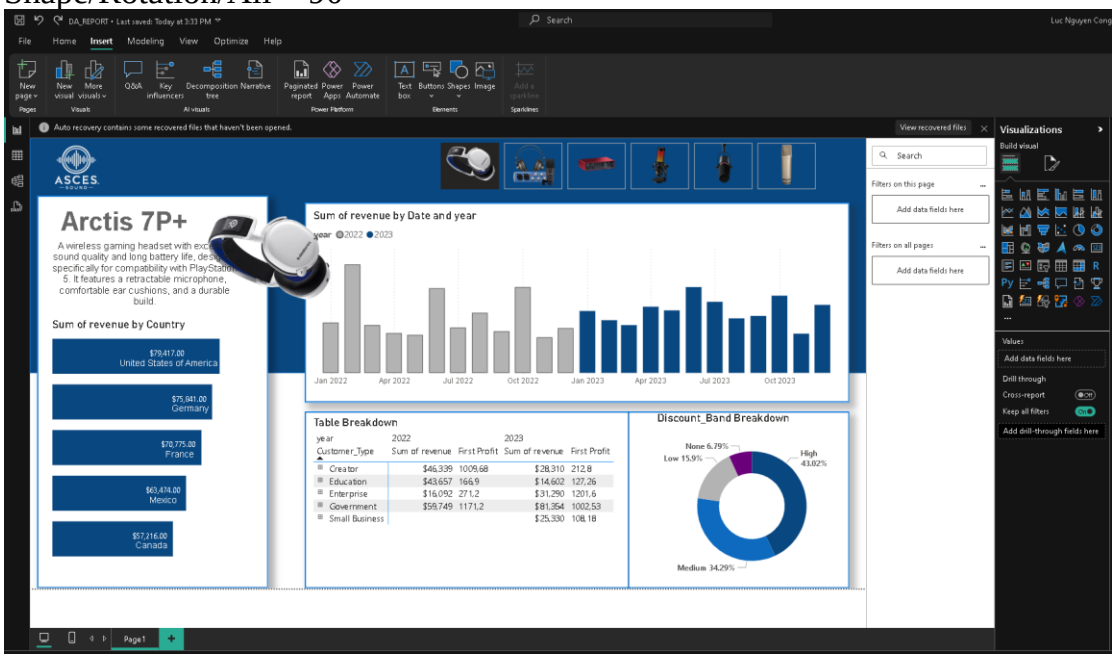


Select Home, Close & Apply



	ost	month	year	Month.1	Discount_Band2	Discount	discount_revenue	Profit
1	2610	December	2023	December	none	0	2985	375
2	960	December	2023	December	none	0	1374	414
3	2124	December	2023	December	none	0	3042	918
4	1770	December	2023	December	none	0	2535	765
5	1746	December	2023	December	none	0	2502	756
6	3828	December	2023	December	low	14	3765.08	-62.92
7	2596	December	2023	December	low	14	3197.48	601.48
8	5046	December	2023	December	low	14	4963.06	-82.94
9	2242	December	2023	December	low	14	2761.46	519.46
10	3422	December	2023	December	low	14	4214.86	792.86
11	3186	December	2023	December	low	14	3924.18	738.18
12	1298	December	2023	December	low	14	1598.74	300.74
13	1298	December	2023	December	low	14	1598.74	300.74
14	990	December	2023	December	low	14	1882.54	892.54
15	990	December	2023	December	low	14	1882.54	892.54
16	2619	December	2023	December	low	14	3227.58	608.58
17	1976	December	2023	December	low	14	2434.66	458.66
18	2080	December	2023	December	low	14	2560.22	480.22
19	2720	December	2023	December	low	14	3347.98	627.98
20	1534	December	2023	December	low	14	1889.42	355.42
21	2006	December	2023	December	low	14	2470.78	464.78
22	522	December	2023	December	low	14	513.42	-8.58
23	1760	December	2023	December	low	14	2166.34	406.34
24	354	December	2023	December	low	14	436.02	82.02
25	1298	December	2023	December	low	14	1598.74	300.74
26	2832	December	2023	December	medium	18	3325.92	493.92

Build visual/Matrix/ Rows -> pull Customer_Type column to inside
and Columns-> pull year column to insde
and Values -> pull revenue column to inside & pull profit column to inside too
Format visual/visual/ Column subtotals = off and Row subtotals = off
Build visual/ Rows -> pull product column inside to add a row under Customer_Type
Format visual/ General/ Title = on / Text = Table Breakdown
Create a boundary between table breakdown and pie chart: Insert/shapes/lines/ Shape/Rotation/All = 90



6.6 Year on Year Calculation and Visual Card

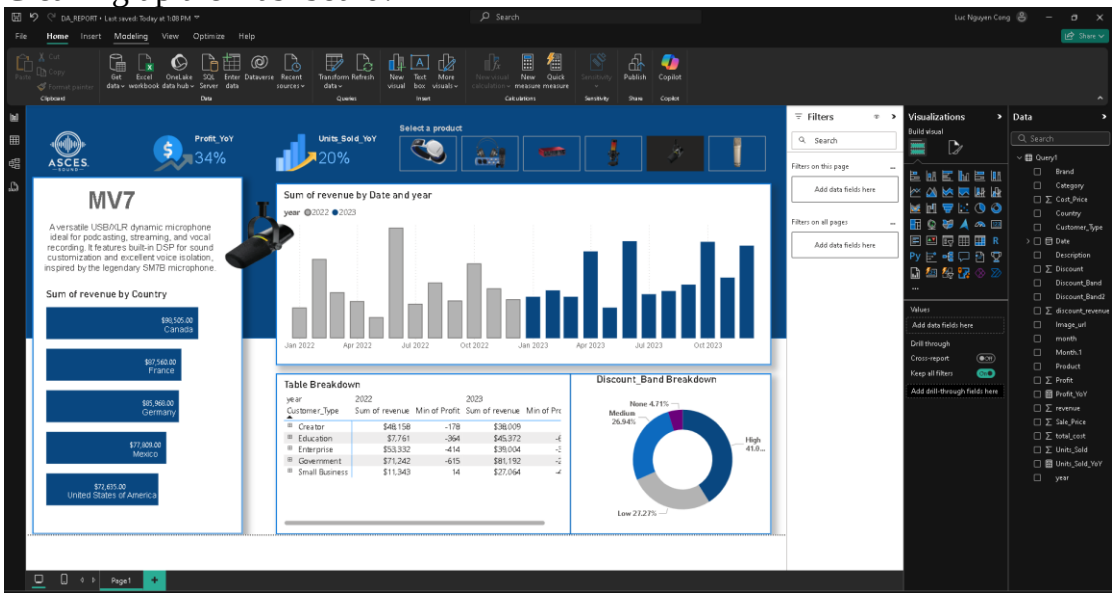
Create Measure Profit_YoY to Calculate the percentage change in profits from the current year compared to the previous year.

```
Profit_YoY =
VAR Profit_Current = SUM(Query1[Profit])
VAR Profit_Previous = CALCULATE(SUM(Query1[Profit]), DATEADD(Query1[Date], -1, YEAR))
VAR FINAL = (DIVIDE(Profit_Current - Profit_Previous, Profit_Previous, 0) * 100)-100
RETURN  FORMAT(FINAL, "0") & "%"
```

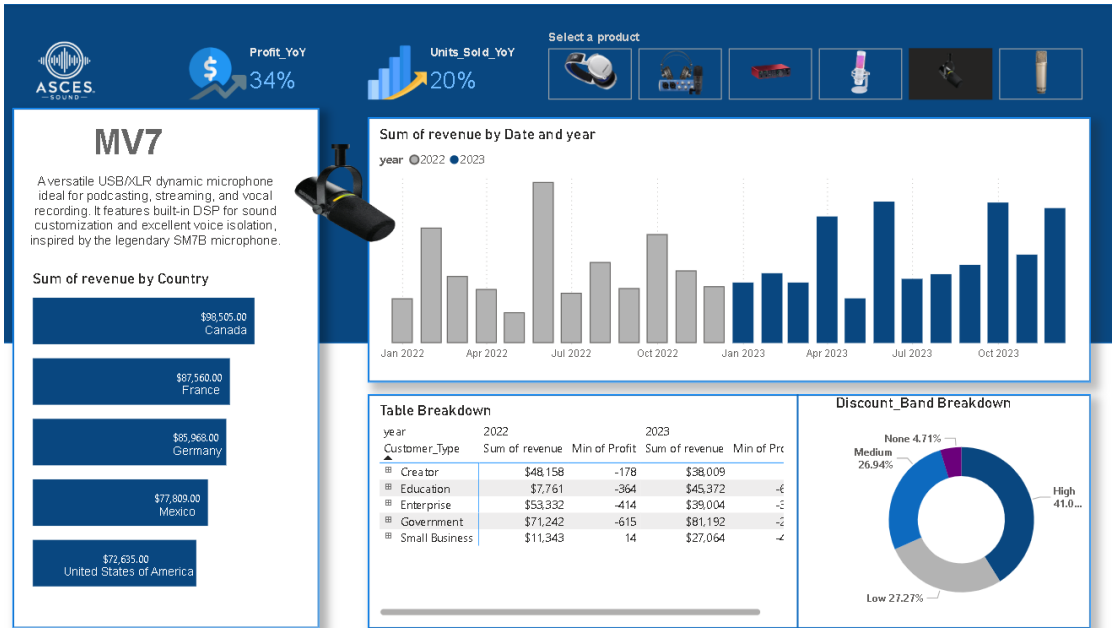
Create Measure Units_Sold_YoY to calculate the percentage change in the number of units sold between the current year and the previous year,

```
Units_Sold_YoY =
VAR Units_Current = SUM(Query1[Units_Sold])
VAR Units_Previous = CALCULATE(SUM(Query1[Units_Sold]), DATEADD(Query1[Date], -1, YEAR))
VAR FINAL = (DIVIDE(Units_Current - Units_Previous, Units_Previous, 0) * 100)-100
RETURN FORMAT(FINAL, "0") & "%"
```

Build visual/Card(New)/ Pull Measure Profit_YoY and Measure Units_Sold_YoY to Data
Format Visual/ Visual/ Images/ Add image to 2 data
Cleaning up the Dashboard!



Final result



Link project: [OneDrive]

CHAPTER 7. Conclusion

This project involved creating a Product Analytics Dashboard, aimed at understanding key business metrics such as revenue, profit, and sales trends. The process allowed me to explore tasks like data setup, database design, SQL query writing, and data visualization using Power BI.

While this project served as a demonstration of the steps involved in data analysis, it was a valuable learning experience in applying foundational skills to real-world scenarios.

Although I am still in the process of refining my abilities, this project helped me gain a better understanding of the practical aspects of data analysis and how such tools can assist in business decision-making. I look forward to improving these skills through future opportunities.