Sri Lanka Institute of Information Technology

Data warehousing and Business Intelligence (IT3021)

Assignment 02

Report

Dasanayake D. R. L

IT22350800

Table of Contents

1. Introduction

This report documents the implementation of Assignment 02 for the module IT3021 – Data Warehousing and Business Intelligence, conducted as part of the BSc (Hons) in Information Technology specializing in Data Science. Building upon the data warehouse developed in Assignment 01, this assignment involves the design and implementation of an SSAS cube, performing OLAP operations using Excel, and developing interactive Power BI reports to analyze and visualize business data. The goal is to apply business intelligence techniques to derive meaningful insights and demonstrate proficiency in data modeling, multidimensional analysis, and visualization using industry-standard tools such as SQL Server Analysis Services (SSAS), Microsoft Excel, and Power BI.

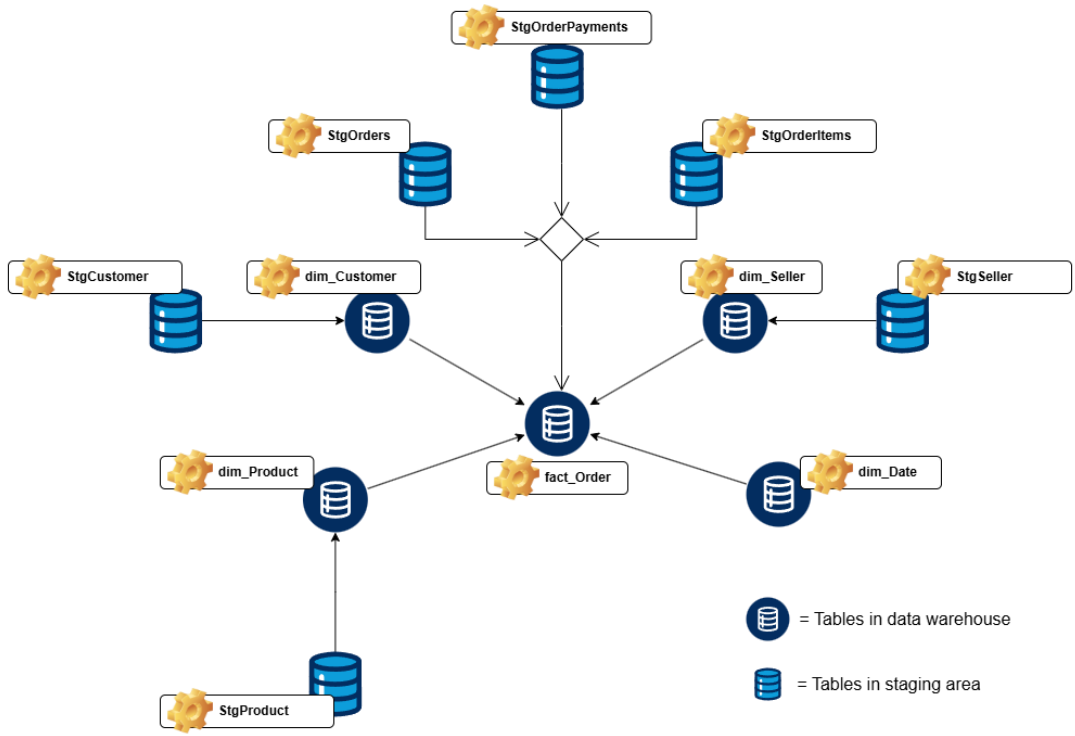
2. Data source

Dataset Chosen: Data Warehouse implemented using the Brazilian E-Commerce Public Dataset by Olist as source dataset.

Source: Kaggle (https://www.kaggle.com/datasets/olistbr/brazilian-ecommerce)

Description: The data used in this assignment was obtained from a structured relational database developed in Assignment 01. It contains information related to sales, customers, products, time, and geography. This data was cleaned, transformed, and loaded into a data warehouse, which serves as the source for building the SSAS cube and generating business intelligence reports.

Entity relationship diagram of the data source is given below:



Detailed structure of the tables in data warehouse is presented below.

Table Name	Description	Key Columns / Notable Attributes
dim_Customer	Customer dimension table with Slowly Changing Dimension (SCD) Type 2	CustomerKey (PK), CustomerID, CustomerUniqueID, CustomerZipCodePrefix, CustomerCity, CustomerState, EffectiveStartDate, EffectiveEndDate, IsCurrent
dim_Product	Product dimension table storing product-specific attributes	ProductKey (PK), ProductID, ProductCategoryName, ProductNameLength, ProductDescriptionLength, ProductWeightGrams, ProductPhotosQty, ProductLengthCM, ProductHeightCM, ProductWidthCM
dim_Seller	Seller dimension table storing location and identity info	SellerKey (PK), SellerID, SellerZipCodePrefix, SellerCity, SellerState
dim_Date	Date dimension table for time-based analysis	DateKey (PK), FullDate, Day, Month, Year, Quarter, DayName, MonthName, IsWeekend
fact_Order	Fact table capturing transaction-level order and payment data	OrderKey (PK), OrderID, CustomerKey (FK), SellerKey (FK), ProductKey (FK), OrderStatus, OrderPurchaseDateKey (FK), OrderApprovedDateKey (FK), OrderDeliveredCustomerDateKey (FK), OrderEstimatedDeliveryDateKey (FK), PaymentType, PaymentInstallments, PaymentValue, FreightValue, ProductPrice, accm_txn_create_time, accm_txn_complete_time, txn_process_time_hours

3. SSAS Cube Implementation

3.1 Overview

To enable multidimensional analysis and support Online Analytical Processing (OLAP), a cube was implemented using SQL Server Analysis Services (SSAS). The process was carried out in several structured steps as outlined below:

Step 1: Create a New SSAS Project

A new Analysis Services Multidimensional Project named **Brazilian_E_Commerce_DW_Cube** was created using SQL Server Data Tools (SSDT).

Step 2: Connect to the Data Warehouse

The data warehouse designed in Assignment 01 (Brazilian_E_Commerce_DW) was used as the data source. A service account was used to establish the connection to the data warehouse.

Step 3: Create a Data Source View (DSV)

A Data Source View (DSV) was created to include the necessary fact and dimension tables. Relationships among these tables were defined to reflect the schema structure accurately.

Step 4: Design the Cube

The Cube Wizard was used to design the cube.

- Required measures were selected from the fact.
- Dimensions were added based on related tables to meet business analysis requirements.

Step 5: Add Hierarchies

To enhance analytical capabilities, hierarchies were added:

- Customer Dimension: State → City
- Date Dimension: Year → Quarter → Month → Day

These hierarchies support drill-down functionality for more detailed analysis.

Step 6: Configure and Finalize the Cube

All cube elements were reviewed to ensure proper connections between measures and dimensions. Attribute properties and visibility were adjusted for usability.

Step 7: Deploy and Process the Cube

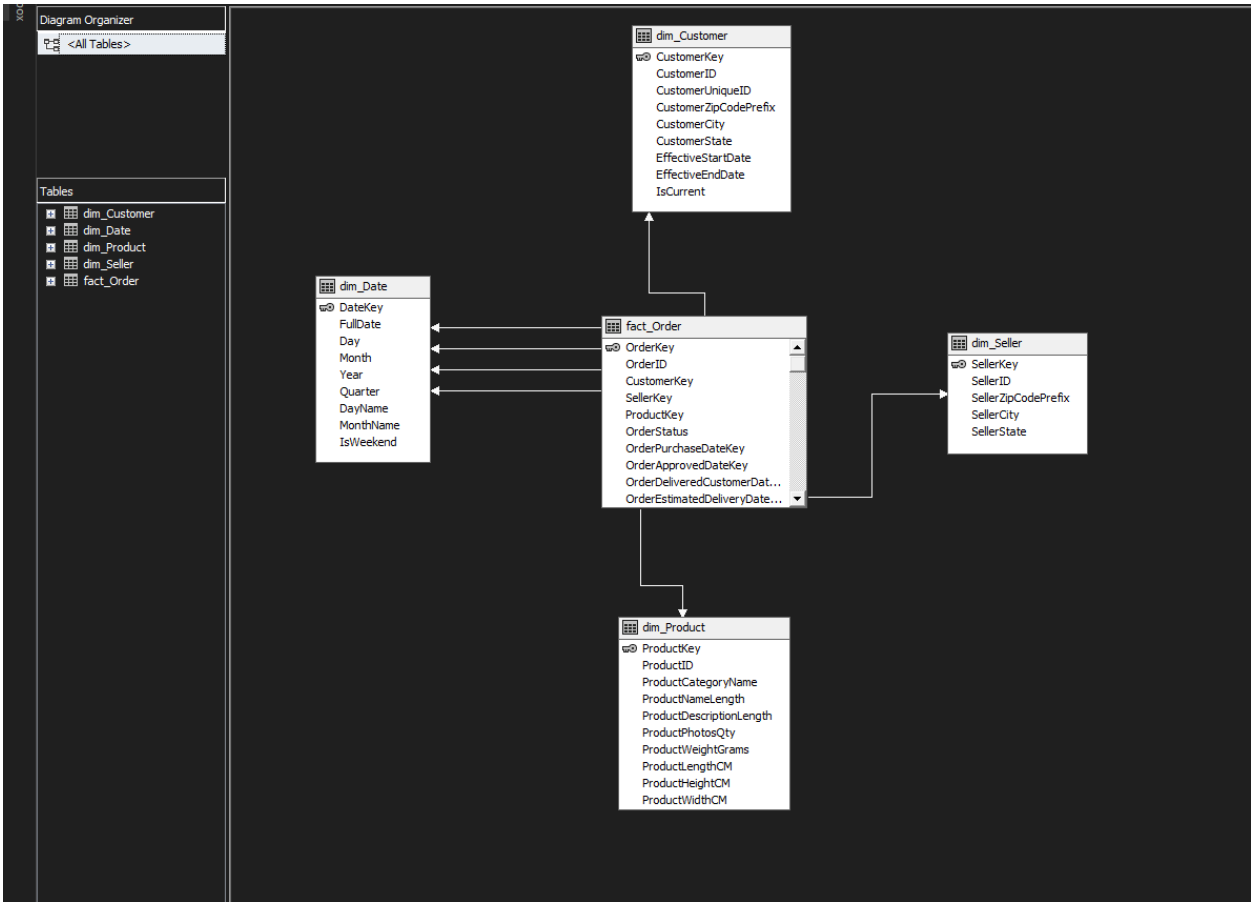
The cube was deployed to the SSAS server. It was then processed to populate it with data from the data warehouse.

Step 8: Validate the Cube

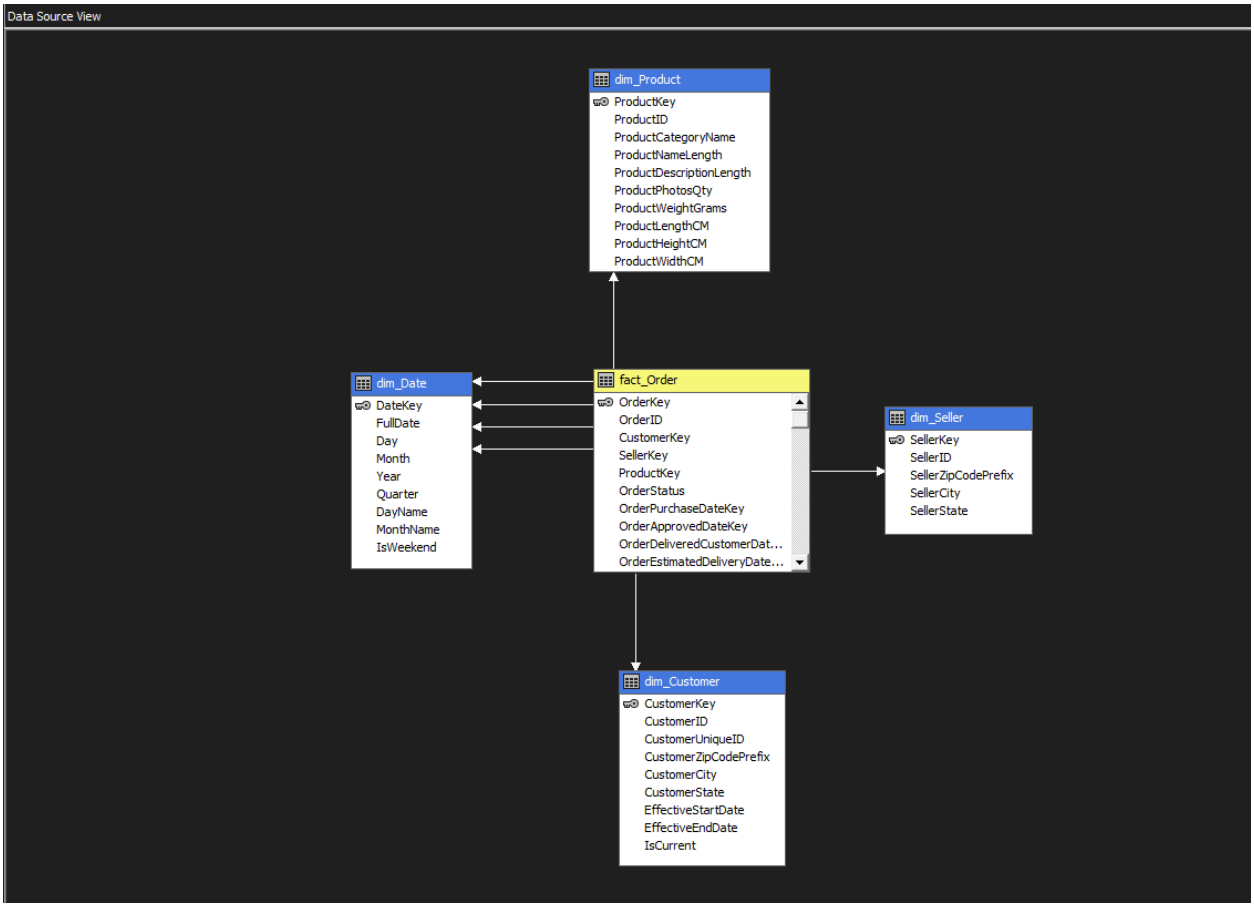
Post-deployment, the cube was tested using tools such as Microsoft Excel and SQL Server Management Studio (SSMS) to ensure that it returns accurate results and supports expected analytical operations.

Below are the screenshots demonstrating processes of creating the SSAS cube.

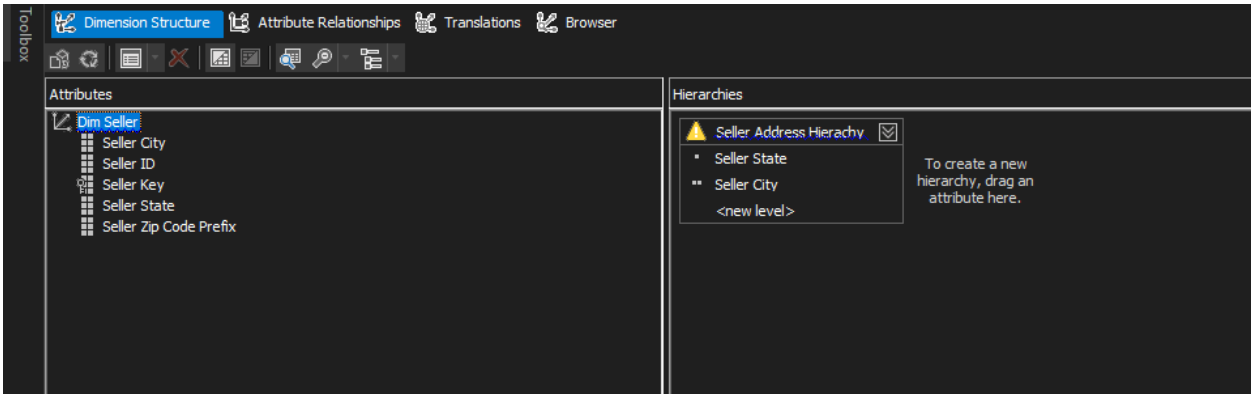
3.2 Data Source View

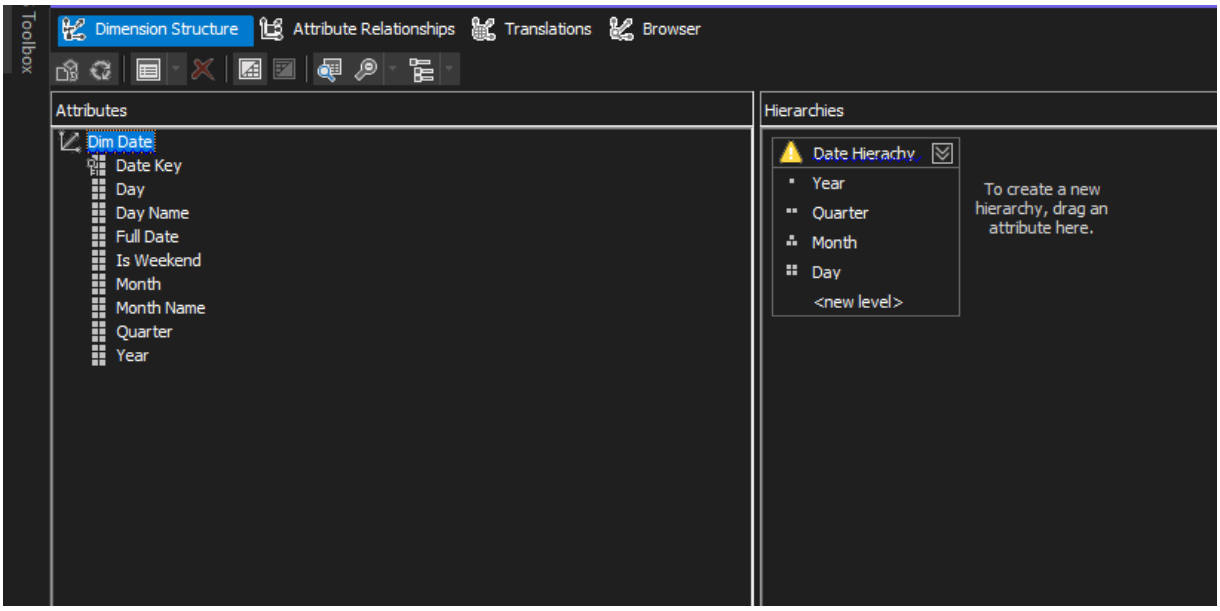


3.3 Cube structure

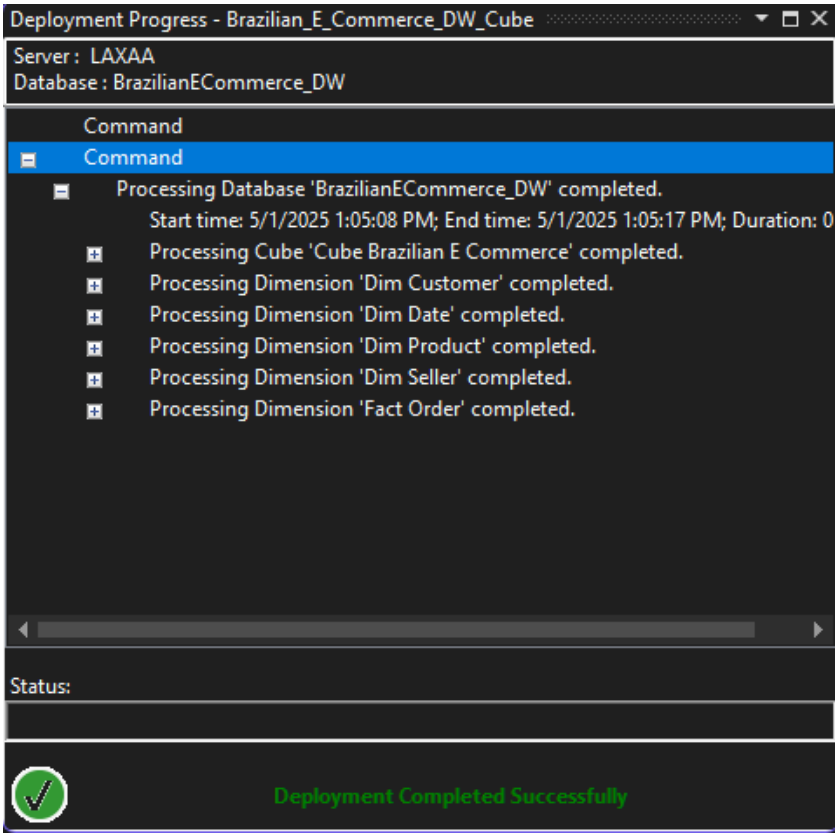


3.4 Creating hierarchies for the Dim Date and Dim Customer dimensions.





3.5 Cube deployment

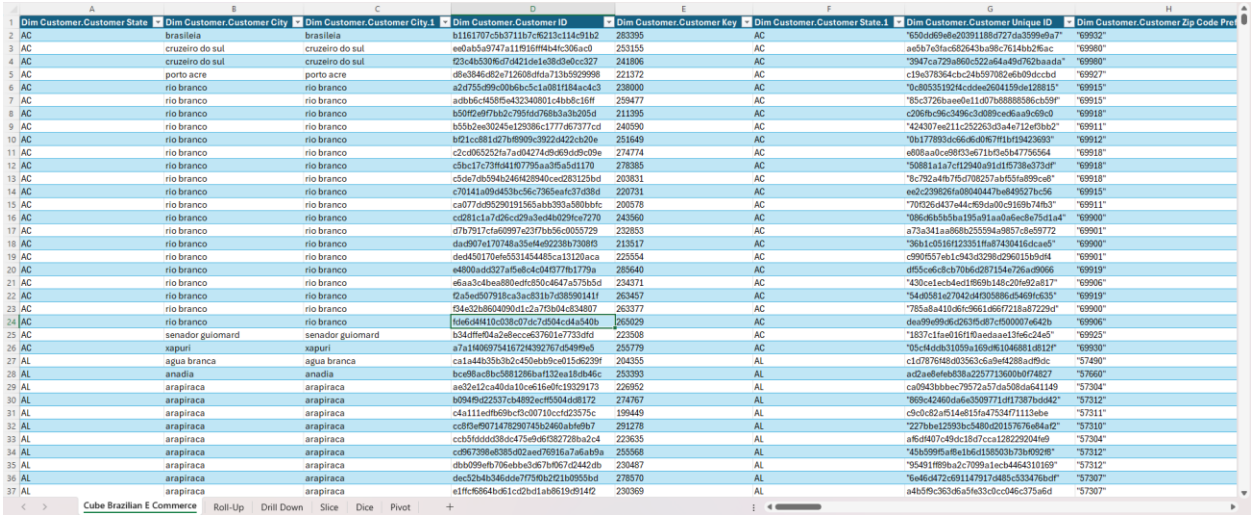


4. Demonstration of OLAP

To demonstrate OLAP functionalities, Microsoft Excel was connected to the deployed SSAS cube using the built-in "From Analysis Services" option under the Data tab. A PivotTable was created to explore the cube data interactively. OLAP operations were demonstrated by performing roll-up and drill-down actions using hierarchies, slicing the data with filters, dicing by selecting multiple dimensions simultaneously, and pivoting to rearrange rows and columns for different analytical perspectives. These visualizations showcased how multidimensional analysis can be applied effectively using Excel's capabilities.

Below are the screenshots demonstrating the OLAP operations performed using Excel.

4.1 Main pivot table in excel



Dim Customer.Customer State	Dim Customer.Customer City	Dim Customer.Customer City.1	Dim Customer.Customer ID	Dim Customer.Customer Key	Dim Customer.Customer State.1	Dim Customer.Customer Unique ID	Dim Customer.Customer Zip Code Pref
AC	brasileia	brasileia	b1161707c5b3711b7cf6213c114c91b2	283395	AC	"650dd69e8e20391188d727da3599e9a7"	"69932"
AC	cruzeiro do sul	cruzeiro do sul	ee0ab5a9747a11f916fff4b4fc306ac0	253155	AC	ae5b7e3fac682643ba98c7614bb2f6ac	"69980"
AC	cruzeiro do sul	cruzeiro do sul	f23c4b530f6d7d421de1e38d3e0cc327	241806	AC	"3947ca729a860c522a64a49d762baada"	"69980"
AC	porto acre	porto acre	d8e3846d82e712608dfda713b5929998	221372	AC	c19e378364cbc24b597082e6b09dccbd	"69927"
AC	rio branco	rio branco	a2d755d99c00b6bc5c1a081f184ac4c3	238000	AC	"0c80535192f4cddee2604159de128815"	"69915"
AC	rio branco	rio branco	adbb6cf458f5e432340801c4bb8c16ff	259477	AC	"85c3726baee0e11d07b88888586cb59f"	"69915"
AC	rio branco	rio branco	b50ff2e9f7bb2c795fdd768b3a3b205d	211395	AC	c206fbc96c3496c3d089ced6aa9c69c0	"69918"
AC	rio branco	rio branco	b55b2ee30245e129386c1777d67377cd	240590	AC	"424307ee211c252263d3a4e712ef3bb2"	"69911"
AC	rio branco	rio branco	bf21cc881d27bf8909c3922d422cb20e	251649	AC	"0b177893dc66d6d0f67ff1bf19423693"	"69912"
AC	rio branco	rio branco	c2cd065252fa7ad04274d9d69ddd9c09e	274774	AC	e808aa0ce98f33e671bf3e5b47756564	"69918"
AC	rio branco	rio branco	c5bc17c73ffd41f07795aa3f5a5d1170	278385	AC	"50881a1a7cf12940a91d1f5738e373df"	"69918"
AC	rio branco	rio branco	c5de7cb594b2464284840ced283125bd	203831	AC	"8c792a4fb7f5d708257abf55fa899ce8"	"69918"
AC	rio branco	rio branco	c70141a09d453bc56c7365eafc37d38d	220731	AC	ee2c239826fa08040447be84f9527bc56	"69915"
AC	rio branco	rio branco	ca077dd952901915 65abb393a580bbfc	200578	AC	"70f326d437e44cf69da00c9169b74fb3"	"69911"
AC	rio branco	rio branco	cd281c1a7d26cd29a3ed4b029fce7270	243560	AC	"086d6b5b5ba195a91aa0a6ec8e75d1a4"	"69900"
AC	rio branco	rio branco	d7b7917cfa60997e23f7bb56c0055729	232853	AC	a73a341aa868b255594a9857c8e59772	"69901"
AC	rio branco	rio branco	dad907e170748a35ef4e92238b7308f3	213517	AC	"36b1c0516f123351ffa87430416dcae5"	"69900"
AC	rio branco	rio branco	ded450170efe6531454485ca13120aca	225554	AC	c990f557eb1c943d3298d296015b9df4	"69901"
AC	rio branco	rio branco	e4800add327af5e8c4c04f377fb1779a	285640	AC	df55ce6c8cb70b6d267154e726ad9066	"69919"
AC	rio branco	rio branco	e6aa3c4bea880edfc850c4647a575b5d	234371	AC	"430ce1ecb4ed1f869b148c20fe92a817"	"69906"
AC	rio branco	rio branco	f2a5ed507918ca3ac831b7d38590141f	263457	AC	"54d0581e27042d4f305886d5469fc635"	"69919"
AC	rio branco	rio branco	f34e32b8604090d1c2a7f3b04c834807	263377	AC	"785a8a410d6fc9661d66f7218a87229d"	"69900"
AC	rio branco	rio branco	fde6d4f410c038c07dc7d504cd4a540b	265029	AC	dea99e99d6d2635d87cf500007e642b	"69906"
AC	senador guiomard	senador guiomard	b34dffef04a2e8ecce637601e7733dfd	223508	AC	"1837c1fae616f1f0aedaae13fe6c24e5"	"69925"
AC	xapuri	xapuri	a7a1f4069754167 24392767d5498fe5	255779	AC	"05cf4dc8b31059a169df6104688 1d812f"	"69930"
AL	agua branca	agua branca	ca1a44b35b3b2c450ebb9ce615d6239f	204355	AL	c1d7876f48d03563c6a9ef4288adf9dc	"57490"
AL	anadia	anadia	bce98acfbc5881286baf132ea1fdb46c	253393	AL	ad2aa8efeb838a2257713600b0f74827	"57660"
AL	arapiraca	arapiraca	ae32e12ca40da10ce616e0fc19329173	226952	AL	ca0943bbbec79572a57da598da641149	"57304"
AL	arapiraca	arapiraca	b094f9d22537cb4892ecff5504dd8172	274767	AL	"869c42460da6a3509771df17387bdd42"	"57312"
AL	arapiraca	arapiraca	c4a111edfb69bcf3c0710ccfd23575c	199449	AL	c9c0c82af514e815fa4753 47f1113ebe	"57311"
AL	arapiraca	arapiraca	cc8f3ef907147829074 5b2460abfe9b7	291278	AL	"227bbe12593bc5480d20157676e84af2"	"57310"
AL	arapiraca	arapiraca	ccb5fdddd38dc475e9d6f082728ba2c4	223635	AL	af6df407c49dc18d7cca128229204fe9	"57304"
AL	arapiraca	arapiraca	cd967398e8385d02aed76916a7a6ab9a	255568	AL	"45b599f5af8e1b6d158503b73bf092f8"	"57312"
AL	arapiraca	arapiraca	dbb099efb706ebbe3d67bf067d2442db	230487	AL	"95491ff89ba2c7099a1ecb4464310169"	"57312"
AL	arapiraca	arapiraca	dec52b4b346dde7f75f0b2f21b0955bd	278570	AL	"6e46d472c691147917d485c533476bdf"	"57307"
AL	arapiraca	arapiraca	e1ffcf6864bd61cd2bd1ab8619d914f2	230369	AL	a4b5f9c363d6a5fe33c0cc046c375a6d	"57307"

Cube Brazilian E Commerce | Roll-Up | Drill Down | Slice | Dice | Pivot

4.2 Roll-Up Operation in Excel

Roll-up: This operation aggregates data to a higher level in the hierarchy, such as viewing total orders by state instead of by city. Below is the screenshot demonstrating the Roll-up operation.

Row Labels	Sum of Fact Order Count	Sum of Payment Value
AM	1	984
BA	190	103328
CE	32	10457
DF	298	40200
ES	108	19603
GO	171	25566
MA	148	18487
MG	2840	437747
MS	18	4332
MT	44	6781
PA	4	712
PB	12	5634
PE	164	39257
PI	4	1101
PR	2835	540022
RJ	1586	344182
RN	22	6109
RO	4	841
RS	758	164095
SC	1322	254715
SE	2	574
SP	25563	3651435
Grand Total	36126	5676162

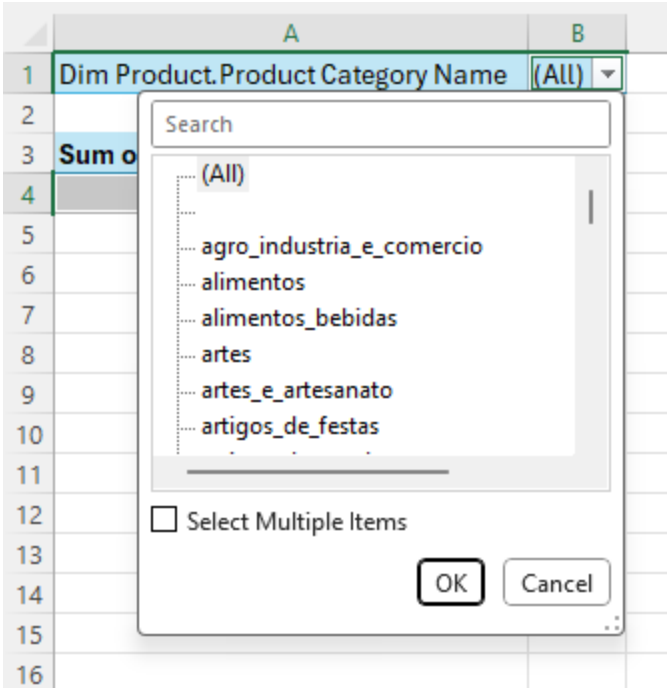
4.2 Drill-Down Operation in Excel

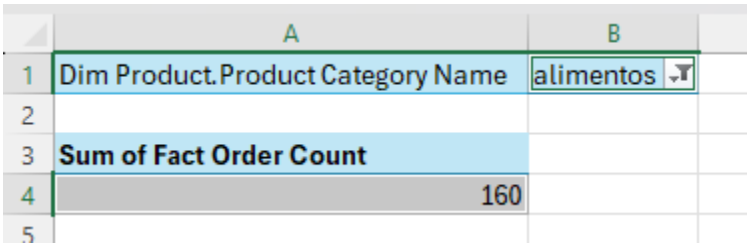
Drill-down: This allows exploring data at a more detailed level, breaking down yearly sales into quarterly or monthly or daily figures. Below is the screenshot demonstrating the Drill-Down operation.

Row Labels	Sum of Freight Value	Sum of Fact Order Count	Sum of Payment Value	Sum of Payment Installments	Sum of Product Price
2016	**2016**	**103**	**18677**	**408**	**15970**
4	**2016**	**103**	**18677**	**408**	**15970**
October	1625	84	15947	339	13633
November	391	19	2730	69	2337
2017	**288700**	**15336**	**2380655**	**47528**	**1912447**
1	**27658**	**1509**	**233924**	**4364**	**190980**
January	1525	104	11814	286	9619
February	9421	519	79863	1489	64147
March	16712	886	142247	2589	117214
2	**62120**	**3324**	**506359**	**10709**	**418968**
April	12596	681	115262	2037	96371
May	26402	1388	209547	4549	175562
June	23122	1255	181550	4123	147035
3	**81599**	**4357**	**651302**	**13825**	**517639**
July	24997	1338	192687	4323	155178
August	28768	1583	225846	5014	177168
September	27834	1436	232769	4488	185293
4	**117323**	**6146**	**989070**	**18630**	**784860**
October	33725	1709	288458	5024	226988
November	34241	1786	289768	5280	230885
December	49357	2651	410844	8326	326987
2018	**420098**	**20687**	**3276830**	**58045**	**2592379**
1	**135978**	**7246**	**1088454**	**20103**	**867480**
January	46818	2491	380457	6931	307362
February	42187	2222	327677	6205	259372
March	46973	2533	380320	6967	300746
2	**165040**	**8085**	**1318612**	**22806**	**1041510**
April	61965	2968	486132	8259	382495
May	51853	2615	419243	7435	334504
June	51222	2502	413237	7112	324511
3	**119080**	**5356**	**869764**	**15136**	**683389**
July	49582	2167	355898	6200	273205
August	69068	3170	510658	8874	407419
September	430	19	3208	62	2765
Grand Total	**710814**	**36126**	**5676162**	**105981**	**4520796**

4.3 Slice Operation in Excel

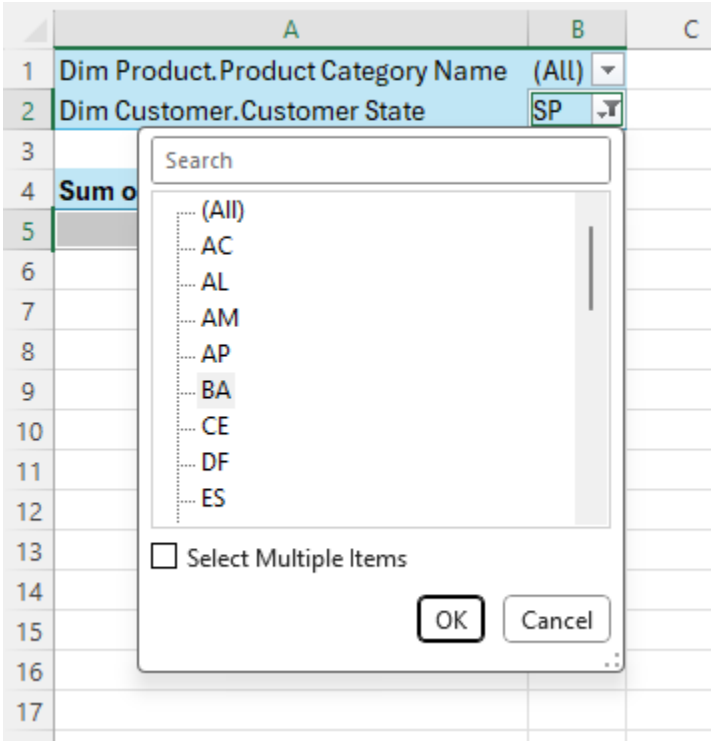
Slice: Filters data by a single dimension, such as showing orders for a selected product category name. Below is the screenshot demonstrating the Slice operation.


A
B
1
Dim Product.Product Category Name
(All)
2
3
Sum o
4
5
6
7
8
9
10
11
12
13
14
15
16
Search
(All)
agro_industria_e_comercio
alimentos
alimentos_bebidas
artes
artes_e_artesanato
artigos_de_festas
Select Multiple Items
OK
Cancel


A
B
1
Dim Product.Product Category Name
alimentos
2
3
Sum of Fact Order Count
4
160
5

4.4 Dice Operation in Excel

Applies filters on multiple dimensions to analyze specific combinations. Specifically using category name and customer state to filter order count in this scenario. Below is the screenshot demonstrating the Dice operation.



	A	B
1	Dim Product.Product Category Name	agro_industria_e_comercio
2	Dim Customer.Customer State	BA
3		
4	**Sum of Fact Order Count**	
5	2	
6		

4.5 Pivot Operation in Excel

Pivot: Rearrange data to view it from different dimensions. Below is the screenshot demonstrating the Pivot operation.

Sum of Fact Order Count	Column Labels			
Row Labels	2016	2017	2018	Grand Total
		303	213	516
agro_industria_e_comercio		18	47	65
alimentos		34	126	160
alimentos_bebidas		41	41	82
artes		13	66	79
artes_e_artesanato		1	9	10
artigos_de_festas		4	10	14
artigos_de_natal		13	28	41
audio	1	56	72	129
automotivo	3	516	908	1427
bebes	4	418	608	1030
bebidas		18	82	100
beleza_saude	15	1208	1964	3187
brinquedos	9	874	522	1405
cama_mesa_banho	1	1657	1840	3498
casa_conforto		89	54	143
casa_conforto_2		9	4	13
casa_construcao		38	151	189
cds_dvds_musicais		5		5
cine_foto		1	19	20
climatizacao	2	36	49	87
consoles_games	3	222	132	357
construcao_ferramentas_construcao		35	247	282
construcao_ferramentas_ferramentas		7	29	36
construcao_ferramentas_iluminacao		8	87	95
construcao_ferramentas_jardim		20	43	63
construcao_ferramentas_seguranca		9	52	61
cool_stuff	3	793	511	1307
dvds_blu_ray		15	10	25
eletrodomesticos		88	192	280
eletrodomesticos_2		29	54	83
eletronicos	1	283	630	914
eletroportateis		113	107	220
esporte_lazer	8	1322	1488	2818
fashion_bolsas_e_acessorios	2	385	292	679

Sum of Fact Order Count	Column Labels															
Row Labels	agro_industria_e_comercio	alimentos	alimentos_bebidas	artes	artes_e_artesanato	artigos_de_festas	artigos_de_natal	audio	automotivo	bebes	bebidas	beleza_saude	brinquedos	cama_mesa
2016								1	3	4		15	9	
2017	303	18	34	41	13	1	4	13	56	516	418	18	1208	874	
2018	213	47	126	41	66	9	10	28	72	908	608	82	1964	522	
Grand Total	**516**	**65**	**160**	**82**	**79**	**10**	**14**	**41**	**129**	**1427**	**1030**	**100**	**3187**	**1405**	

5. Power BI Reports

The SSAS cube was connected to Power BI Desktop using the "Analysis Services" connector, allowing live data exploration. For each report, data fields were selected from dimensions and fact tables and added to the report canvas using visuals such as matrix tables, bar charts, line charts, cards, gauges and slicers.

Below are the screenshots demonstrating the reports created using Power BI.

5.1 Report 1: Matrix Visual

Report 1 – Matrix Visual: Displays detailed tabular data with row and column groupings. Below is the screenshot of Report 1 in Power BI.

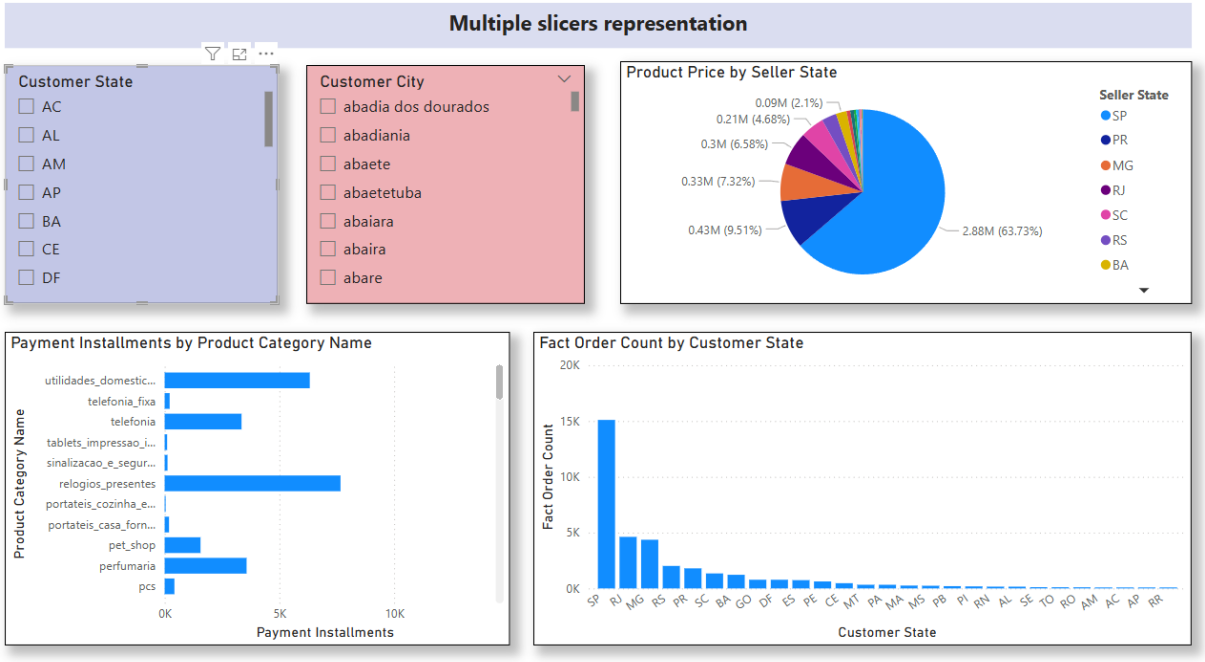
Matrix view representation

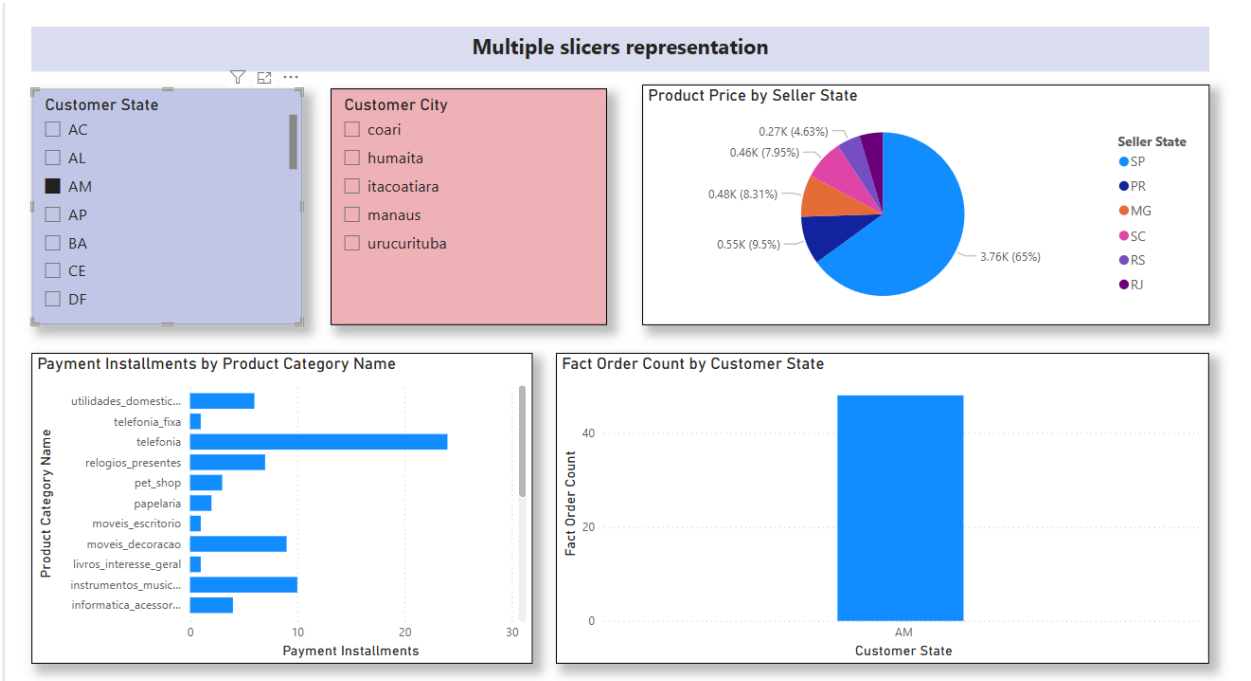
Customer Product Sales Matrix

Customer State		agro_industria_e_comercio	alimentos	alimentos_bebidas	artes	artes_e_artesanato	artigos_de_festas	artigos_de_natal	audio	auto
AC										
AL	144.00	1,518.00	215.00						261.00	2
AM	638.00									
AP		657.00			227.00					
BA	3,163.00	1,359.00	669.00	250.00	306.00			76.00	1,415.00	10
CE	1,533.00								830.00	5
DF	1,349.00		81.00		336.00	151.00	269.00	276.00	448.00	8
ES	983.00		591.00		76.00			24.00	141.00	3
GO	1,711.00	1,024.00	132.00		123.00				1,129.00	6
MA	545.00			103.00					760.00	2
MG	8,422.00	4,568.00	1,086.00	347.00	812.00	222.00	1,019.00	204.00	2,984.00	30
MS	358.00	555.00	265.00		151.00					1
MT	701.00	445.00			97.00			109.00	72.00	4
PA	1,231.00	821.00			142.00				608.00	4
PB	2,083.00		46.00		154.00					1
PE	3,211.00	5,601.00	5.00	301.00				425.00	315.00	4
PI	196.00	301.00							217.00	4
PR	4,668.00	938.00	500.00	464.00	103.00		273.00	29.00	1,293.00	10
RJ	7,796.00	4,225.00	932.00	1,168.00	2,019.00		268.00	169.00	3,796.00	32
Total	**72,982.00**	**31,584.00**	**12,339.00**	**6,730.00**	**8,327.00**	**1,136.00**	**2,756.00**	**3,663.00**	**21,709.00**	**252**

5.2 Report 2: Slicer with Graphical Presentations

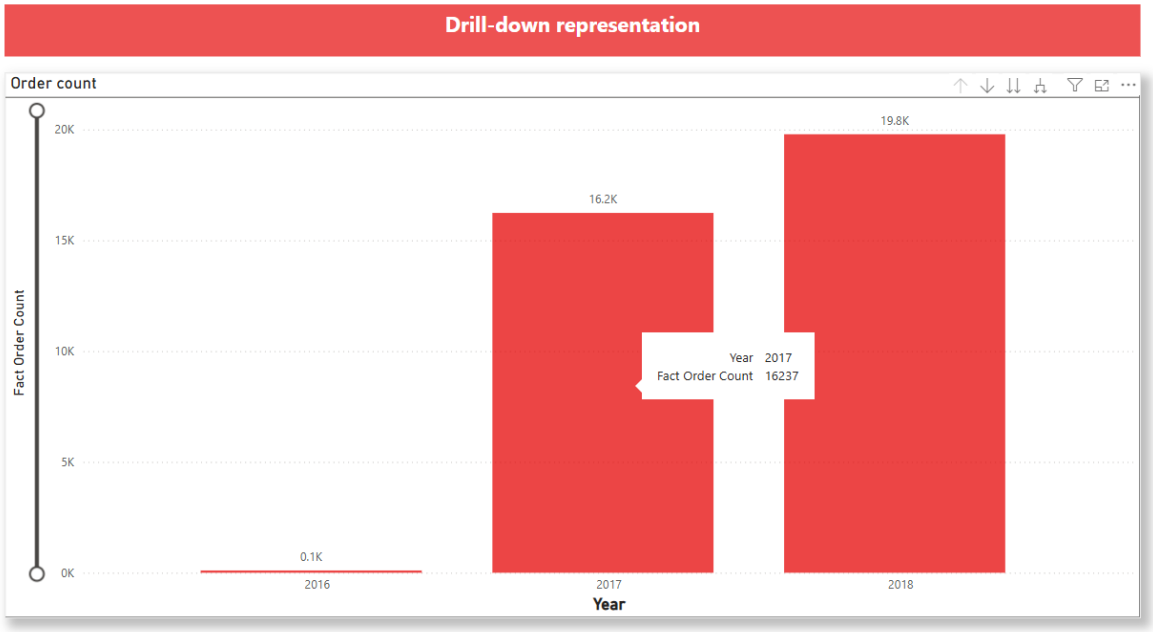
Report 2 – Cascading Slicers: Uses dependent slicers and multiple visuals to show filtered insights. Customer state and customer city are used as filters here. Below is the screenshot of Report 2 in Power BI.



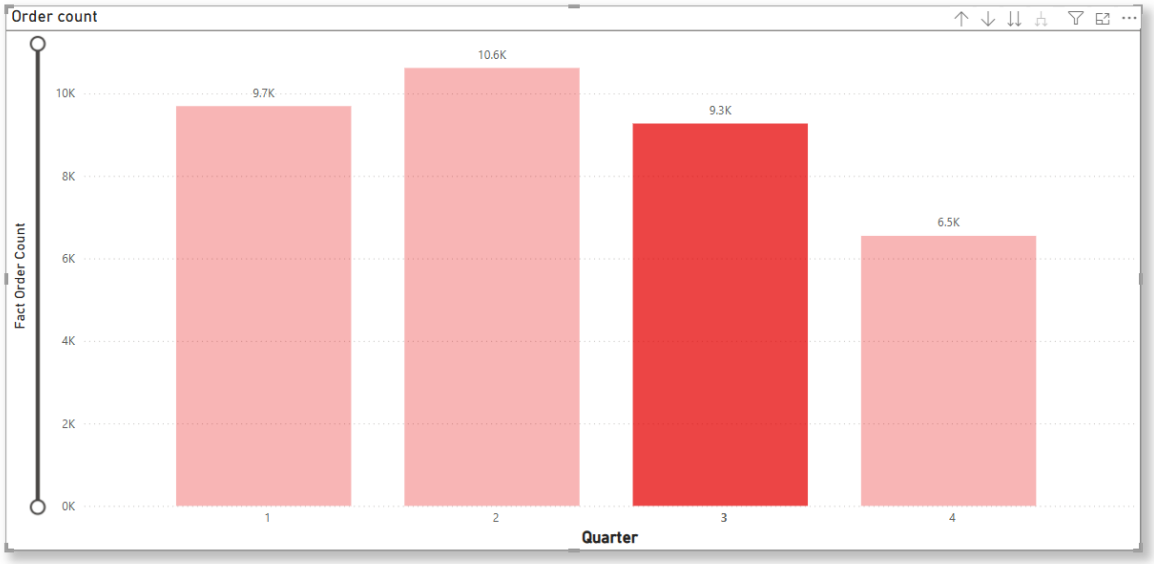


5.3 Report 3: Drill-Down

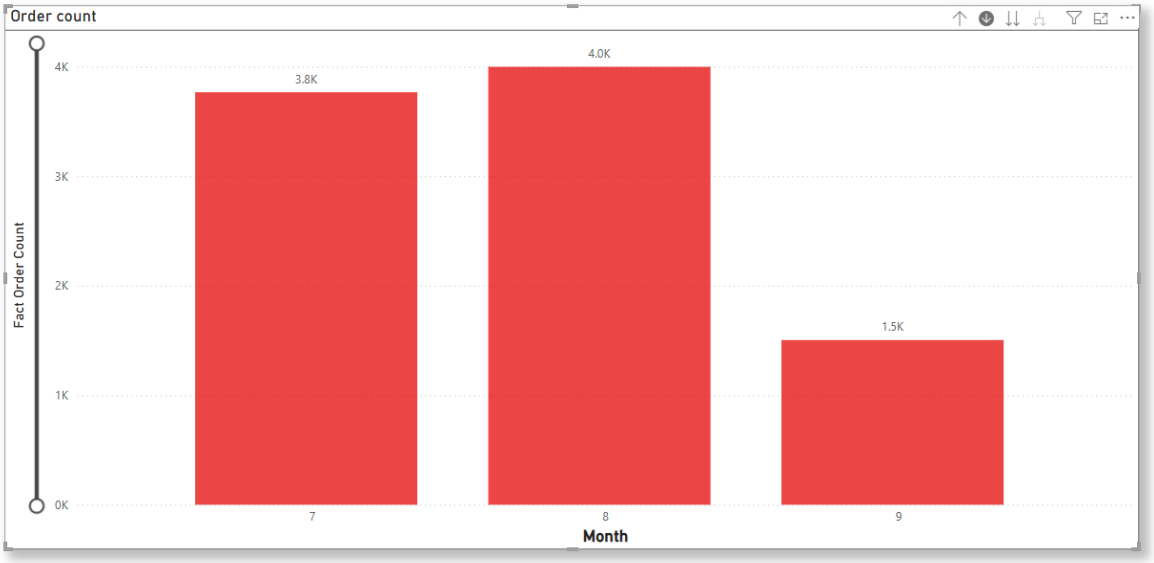
Report 3 – Drill-down Report: Allows exploring data hierarchically. Specifically, from year -> quarter -> month -> day in this scenario. Below is the screenshot of Report 3 in Power BI.

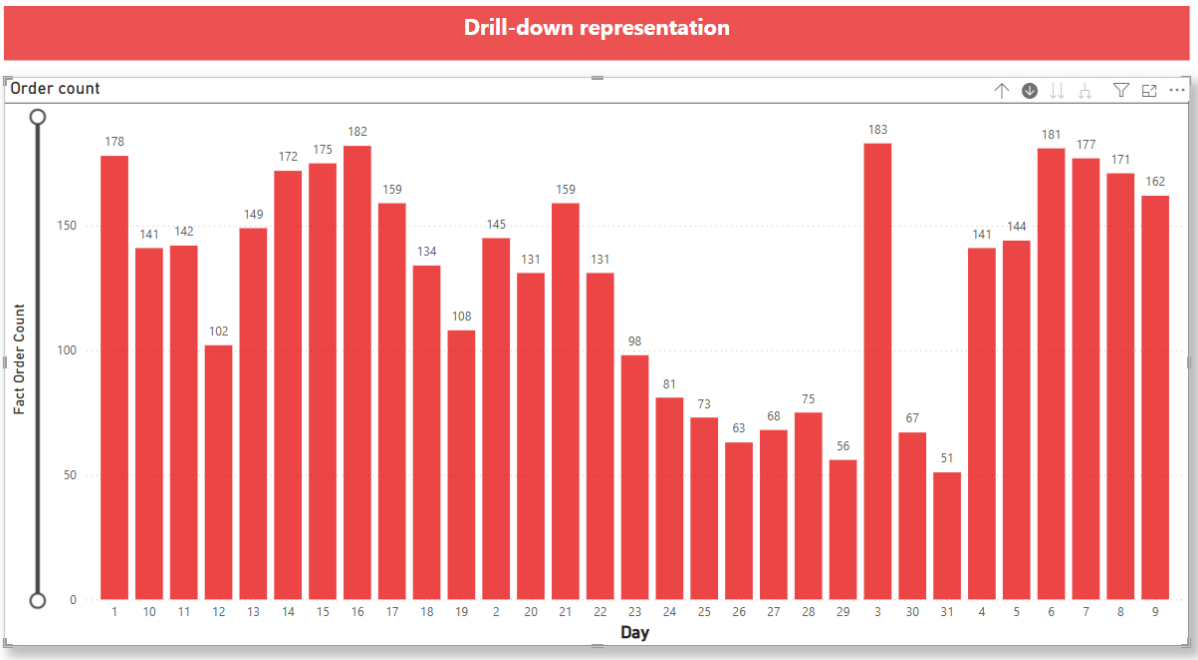


Drill-down representation



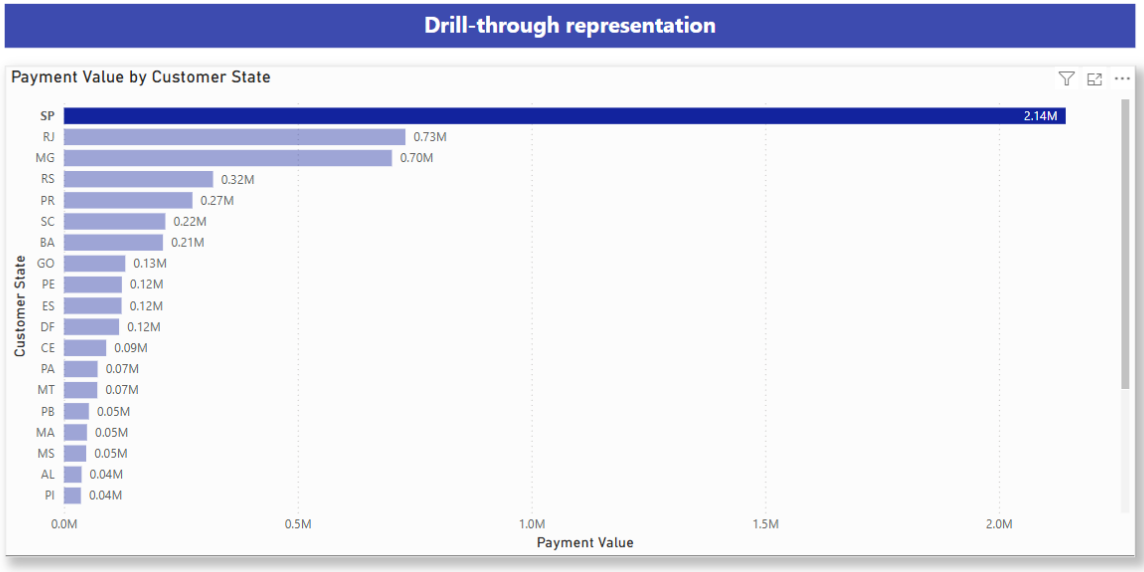
Drill-down representation

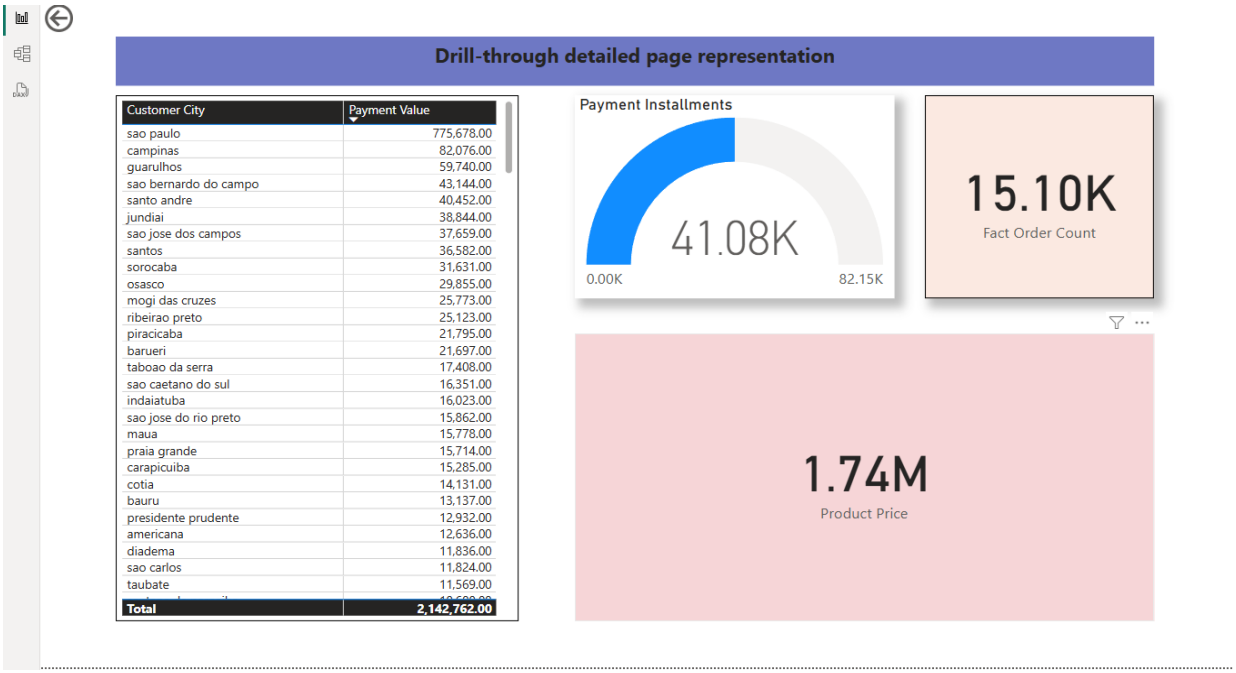




5.4 Report 4: Drill-Through

Report 4 – Drill-through Report: Lets users navigate to detailed pages from summary visuals. Below are the screenshots of Report 4 in Power BI.





6. Conclusion

This assignment demonstrated the practical implementation of a business intelligence solution using a data warehouse, SSAS cube, Excel OLAP operations, and Power BI reports. Each task helped to apply key BI concepts such as multidimensional modeling, interactive data analysis, and visualization. The tools used enabled efficient data exploration and insight generation, supporting informed decision-making.

7. References

[1] Microsoft, "Power BI Documentation," Microsoft Learn, 2024. [Online]. Available: https://learn.microsoft.com/en-us/power-bi/

[2] SolarWinds, "SSIS (SQL Server Integration Services)," *SolarWinds IT Glossary*, 2024. [Online]. Available: https://www.solarwinds.com/resources/it-glossary/ssis-sql-server-integration-services