

Sri Lanka Institute of Information Technology



Data Warehousing and Business Intelligence - IT3021

Assignment 2

IT22283894 – Dakshina P.D.S.D

Table of Contents

1	Declaration.....	3
2	Data Source for Assignment 2.....	3
3	SSAS Cube Implementation.....	5
3.1	Data Source View	6
3.2	Cube Structure.....	6
3.3	Hierarchies.....	8
3.4	KPI Creation.....	11
3.5	Role.....	12
3.6	Cube deployment.....	12
4	OLAP Operation Demonstration.....	13
4.1	SLICE - OLAP Operation.....	15
4.2	DICE – OLAP OPERATION.....	15
4.3	PIVOT – OLAP OPERATION	16
4.4	DRILL-DOWN AND ROLL UP – OLAP OPERATION.....	17
5	POWER BI REPORTS.....	19
5.1	REPORT 01	20
5.2	REPORT 02	20
5.3	REPORT 03	21
5.4	REPORT 04	23

1 Declaration

I declare that this project report or part of it was not a copy of document done by any organization, university and other institute or a previous student project at SLIIT and was not copied from the internet or other resources.

2 Data Source for Assignment 2

The Data Warehouse created for Assignment 1 is the main data source used to carry out the necessary steps of the task. SQL Server Management Studio 2016 and SQL Server Data tools 2015 (Visual Studio) are the major tools used in Source data set compilation and ETL orchestration in order to create the Data Warehouse.

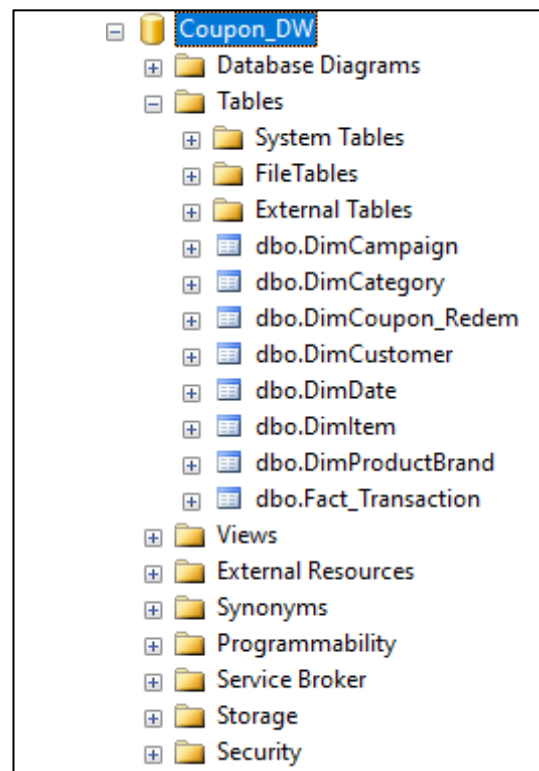


Figure 1 Data Source table structure

The Data Warehouse represents the data of a Brick-and-Mortar Retail shop that uses discount marketing as their major strategy in carrying out successful business operations. By analyzing the business entities, a snowflake schema was designed in order to carry out efficient Data Warehousing and Business Intelligence principles in the business process.

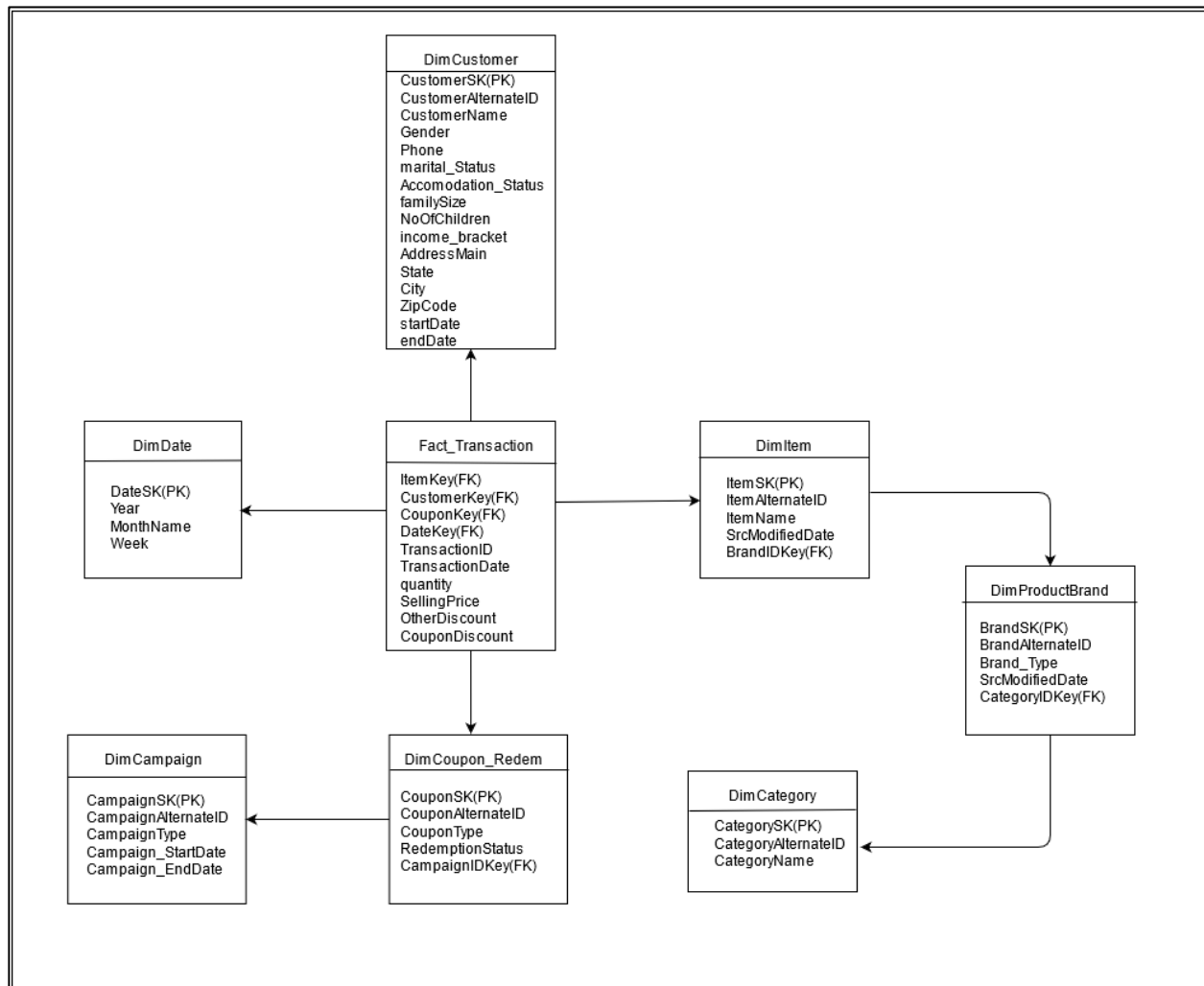


Figure 2 Datawarehouse Snow Flake Schema

The schema consists of one fact table and seven-dimension table including date dimension. It is observed that the entities are in a normalized form because of the snowflake design.

Hierarchical implementations are found in this schema

1. DimProductBrand and DimCategory are hierarchical dimensions of DimItem
2. DimCustomer has a customer location wise hierarchy.
3. DimDate has a date wise hierarchy.

DimCustomer dimension is a **slowly changing dimension** with historical attributes and changing attributes where Type 2 and Type 1 implementations are being enforced, respectively. Transaction of a customer for a particular date is considered as the **grain** of the Fact_Transaction fact table.

3 SSAS Cube Implementation

SQL server data tools was used to create an *Analysis Services Multidimensional and Data Mining Project* in order to create the Cube Structure with the Source data.(Datawarehouse)

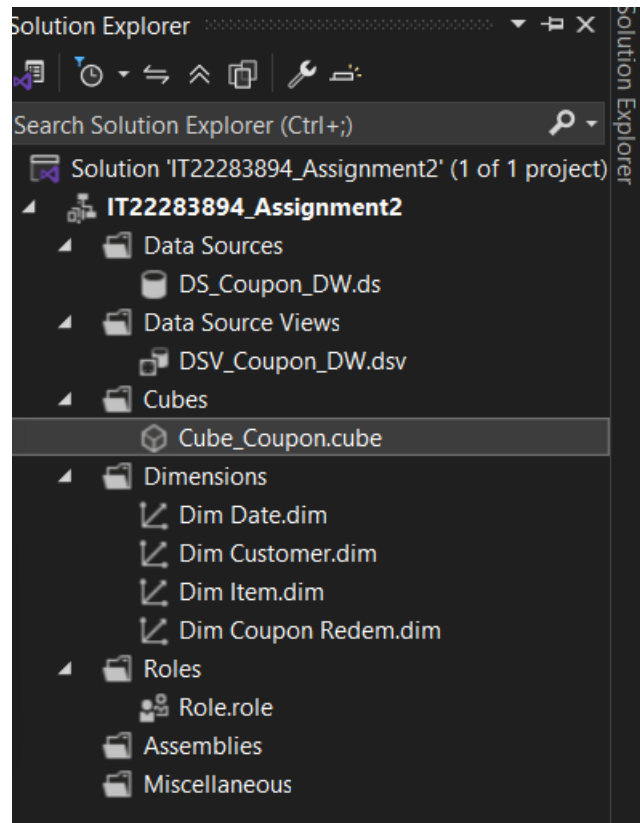


Figure 3Cube structure Solution Explorer

In order to create a working SSAS cube as shown above , sequence of tasks were carried out such as

1. Data source Creation. (**DS_Coupon_DW**)
2. Data Source View Creation.(**DSV_Coupon_DW**)
3. Cube creation.(**Cube_Coupon**)

3.1 Data Source View

Once the data source was created a data source view was created by making use of the Source data warehouse Coupon_DW.

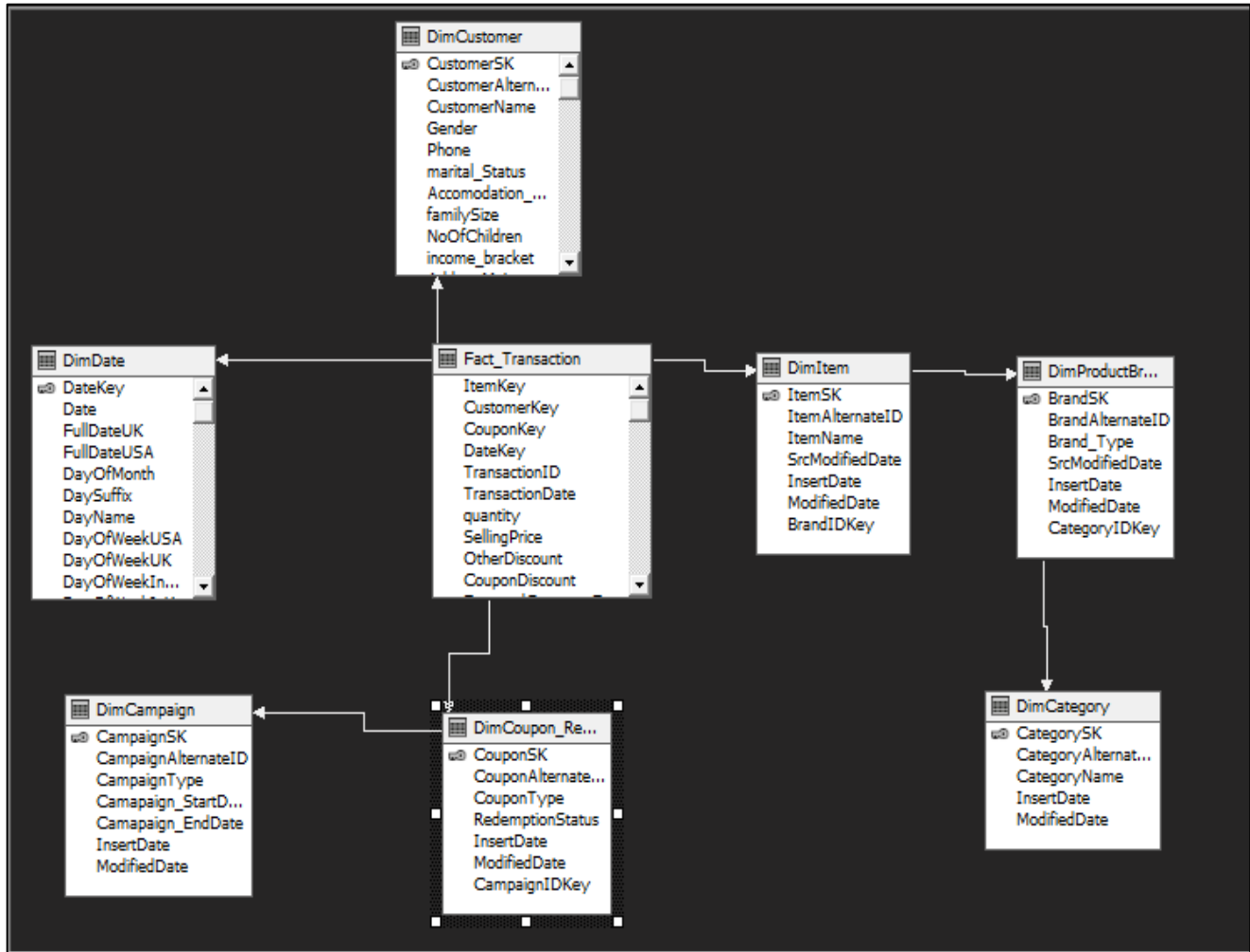
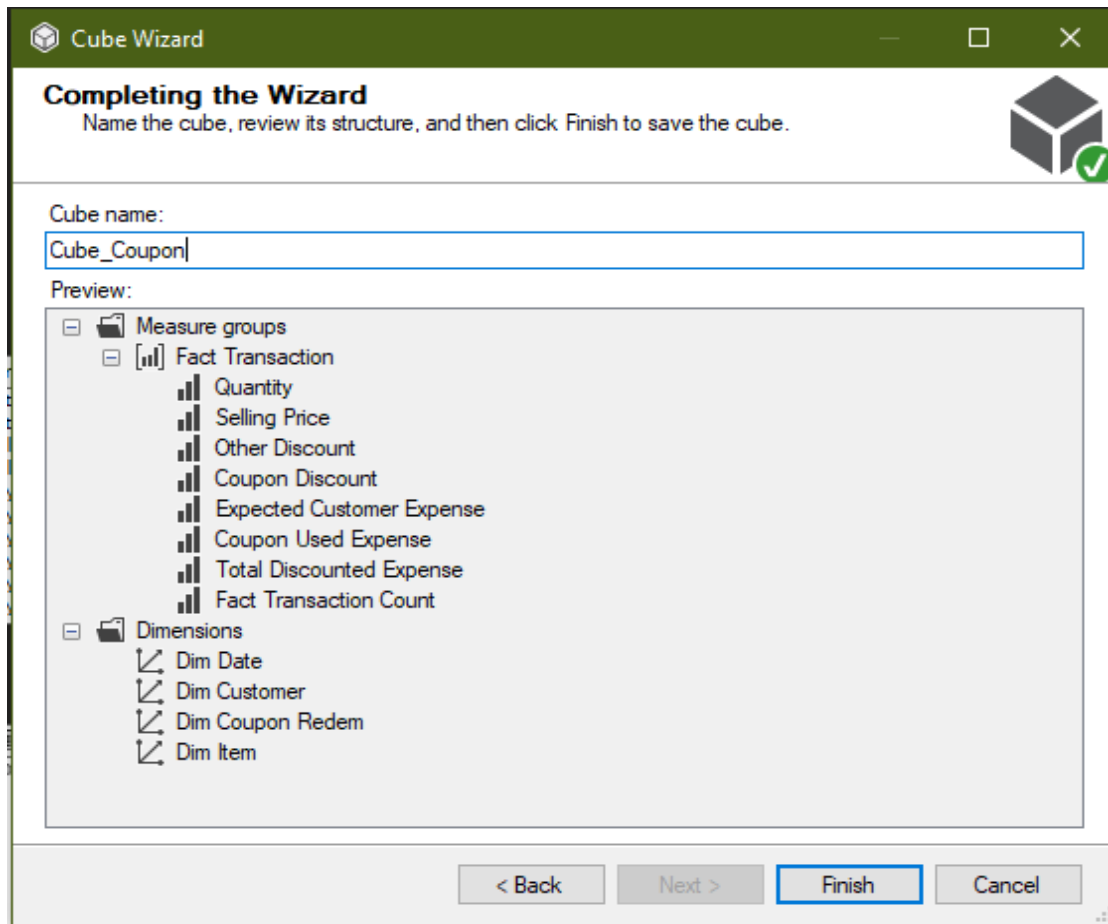


Figure 4 Data Source view(DSV_Coupon_DW)

Relationships among the facts and dimensions were manually created with the use of Surrogate keys of dimensions and foreign keys in the fact table.

3.2 Cube Structure

As the final step the cube structure (**Cube_Coupon**) is designed by using the data source view created in the previous step.



The final cube structure created is as shown below. 7 main dimensions and a fact table are used in creating this cube structure

Dimensions and Fact tables used – DimCustomer , DimDate , DimCampaign , DimCoupon_Redemption , DimCategory , DimProductBrand , DimItem and Fact_Transaction.

The 4 dimensions DimCustomer , DimDate , DimCoupon_Redemption ,and DimItem are directly connected with the fact table whereas DimCategory , DimProductBrand and DimCampaign are hierarchical implementations.

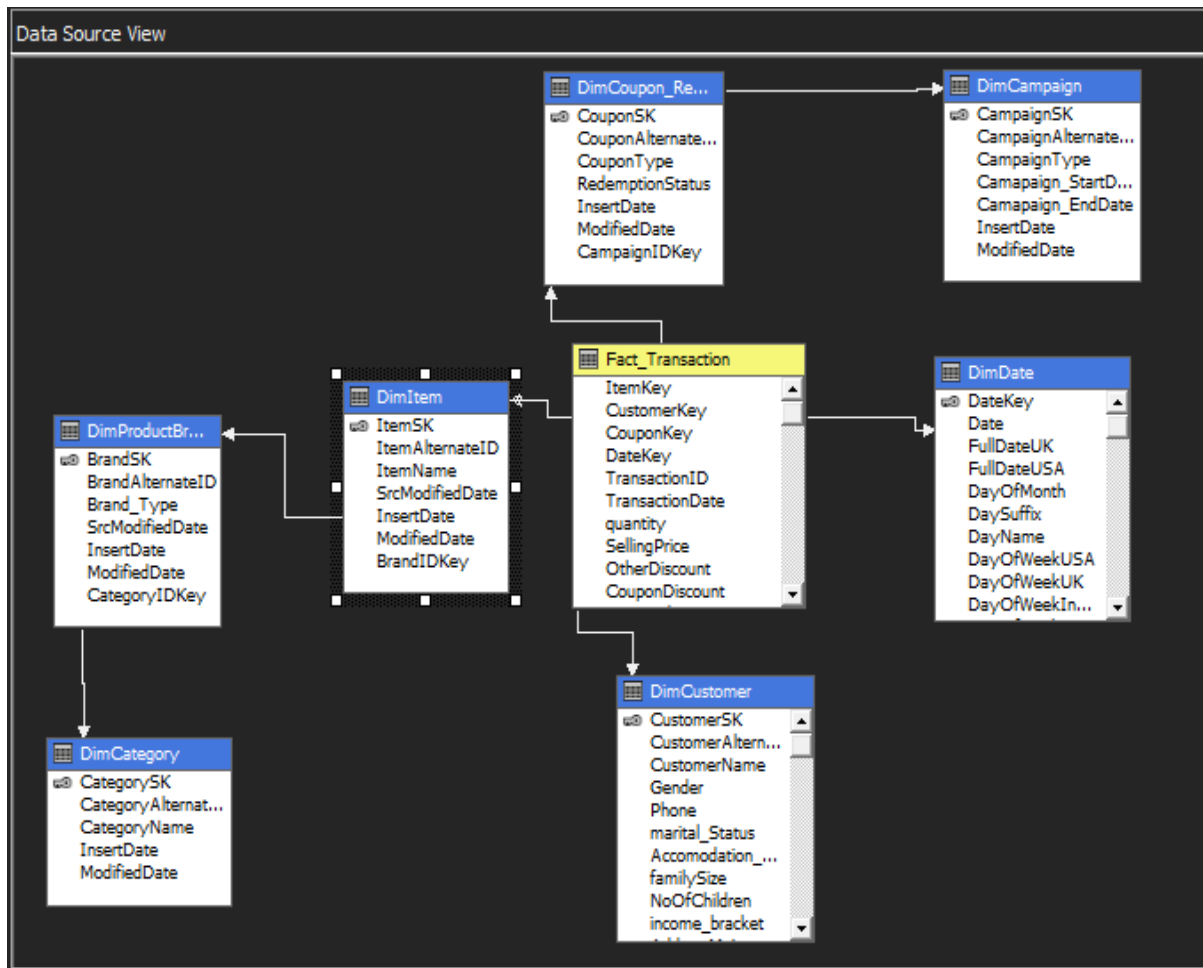


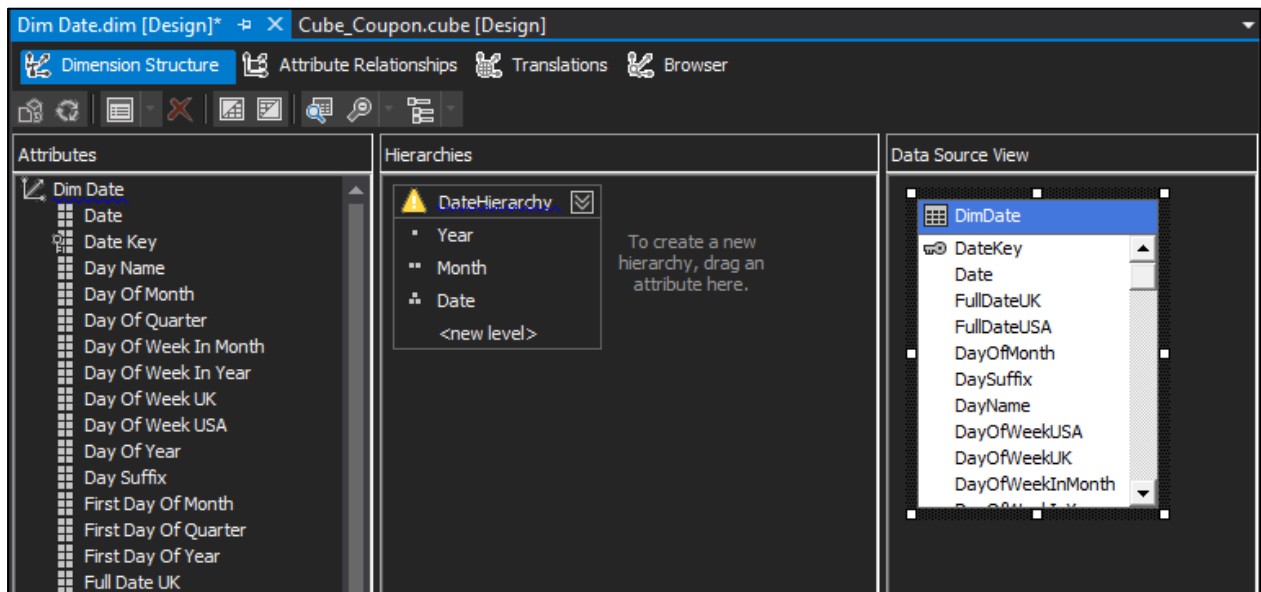
Figure 5 Cube (Snowflake Structure)

3.3 Hierarchies.

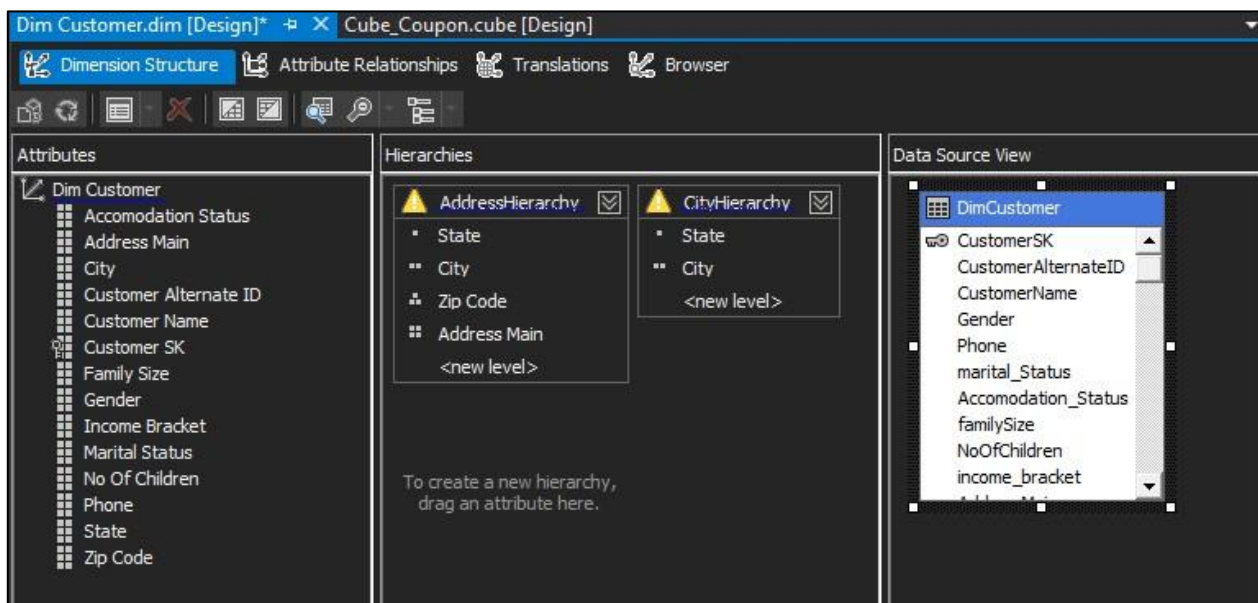
According to the data loaded in data warehouse, the built snowflake schema reflects many hierarchical relationships between data in data warehouse. Hierarchies are created for DimCustomer, DimItem and DimDate dimensions, respectively. The hierarchies created are mainly based on location of the Customers, Category details of Items sold in the retail shop and Date of business processes. In addition, hierarchical relationships between two dimensions DimCampaign and DimCouponRedem also exists. The main purpose of hierarchy creation is to provide the ability to drill down to detailed level data and to roll up to create aggregated data thus making business analysis purposes efficient.

Hierarchies are built in the **Dimension Structure** of the dimensions.

a. Date hierarchy.

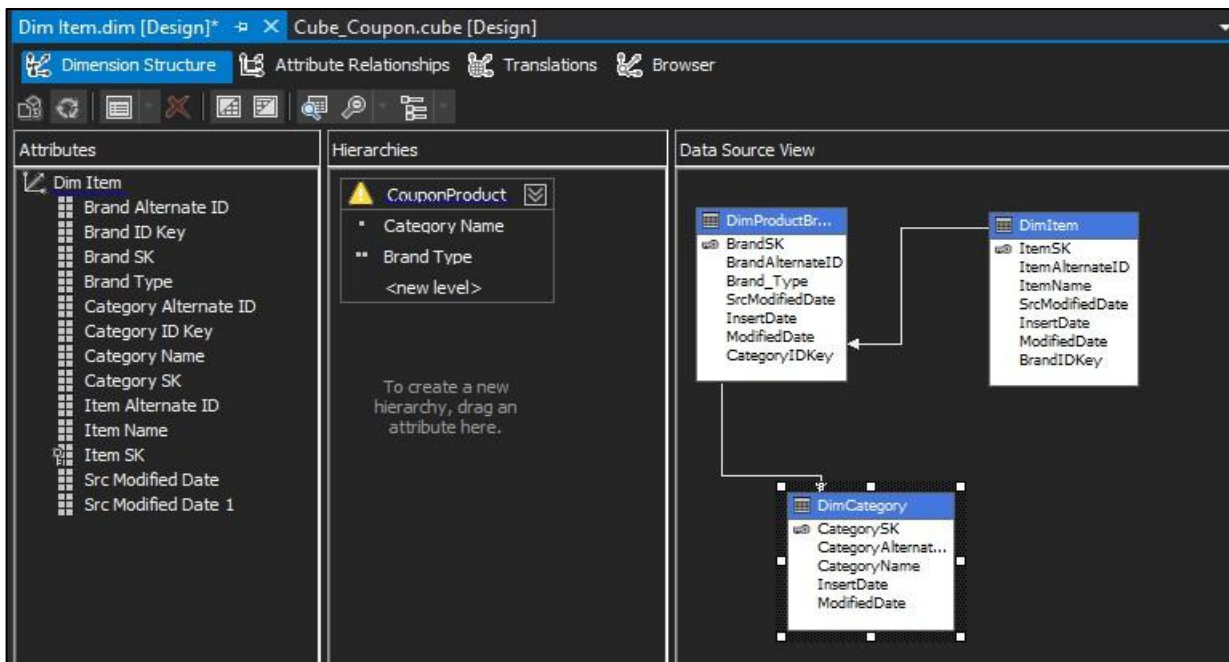


b. Location hierarchy.



Customer location is analyzed in many aspects as per business requirements. Therefore 2 location hierarchies are created in the same dimension providing two different levels of detail.

c. Product hierarchy.



Products in retail shop are analyzed in terms of category and brands where each item belongs to.

Once the hierarchies are created, they can be viewed in the cube browser as shown here.

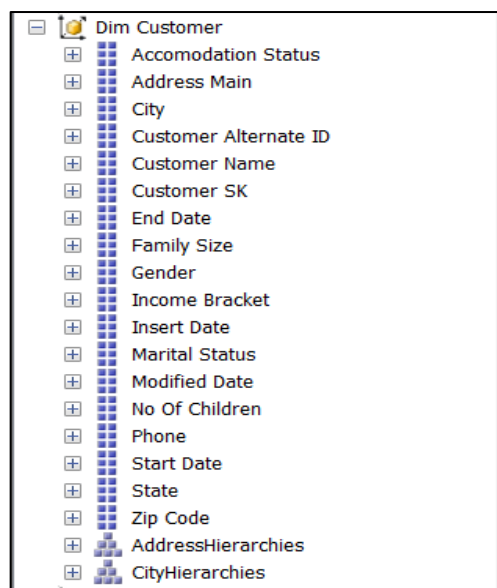
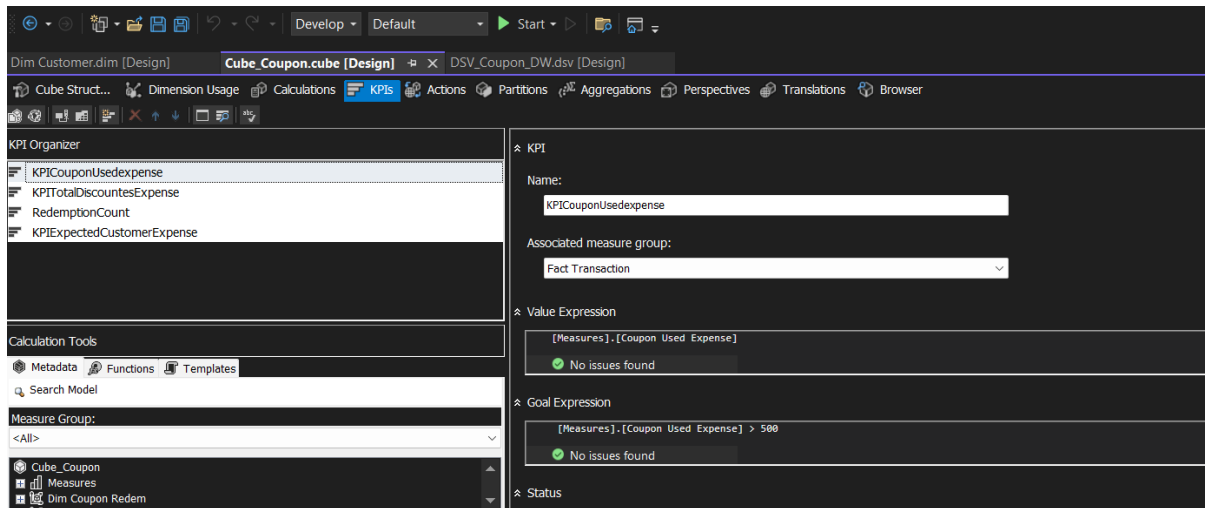


Figure 6Cube Browser view

3.4 KPI Creation.

KPI (Key Performance Indicator) shows the amount of progress that is done to achieve a business goal. The fact table Fact_Transaction is mainly used to create the KPIs here



KPI	GOAL	Description
KPICouponUsedExpense	[Measures].[Coupon Used Expense] > 500	If goal is false, then there is no expected coupon usage.
RedemptionCount	[Dim Coupon Redem].[Redemption Status] > 200	If goal is true, then the coupon marketing target successful.
KPITotalDiscountExpense	[Measures].[Total Discounted Expense] > 5000	If goal is true business discount transactions successful,
KPISExpectedCustomerExpense	[Measures].[Expected Customer Expense]>5000	If goal is true target is successful.

3.5 Role.

Roles are created to provide control access to users in order to ensure who can do what. Role has a set of customized control access options to permit or deny permissions to objects within a particular database or cube. In this step Full control (Administrator) permission is provided to the role.

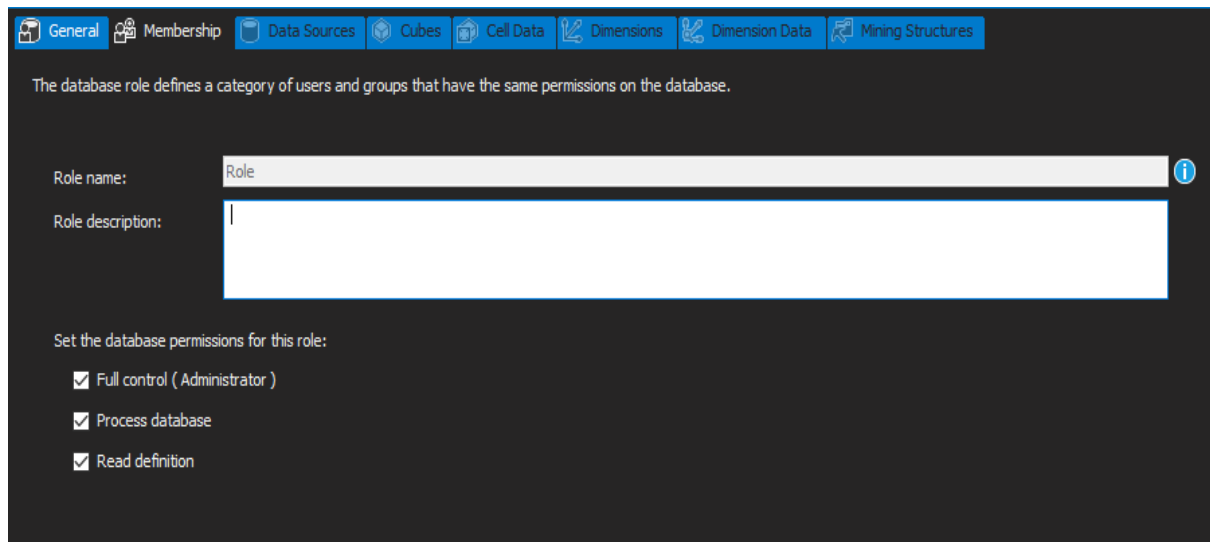
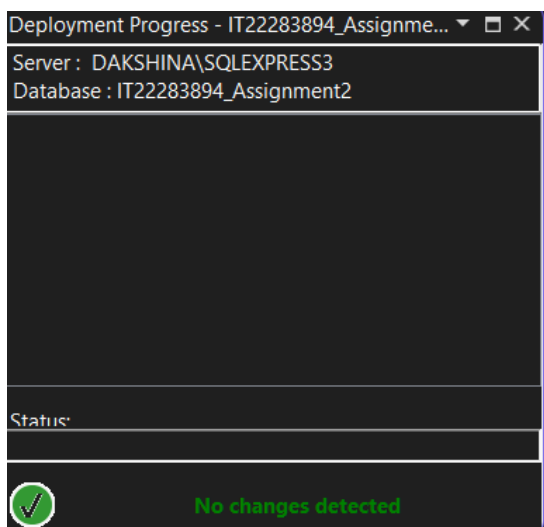


Figure 7User Role

3.6 Cube deployment.

Once the sub tasks are over finally cube is deployed in order to carry out the analysis purpose.



this image is taken after 2nd successful deployment.

Figure 8Cube deployment

4 OLAP Operation Demonstration.

Once the cube is deployed successfully, deployed cube can be viewed in SQL Server Analysis service of SQL Server Management Studio.

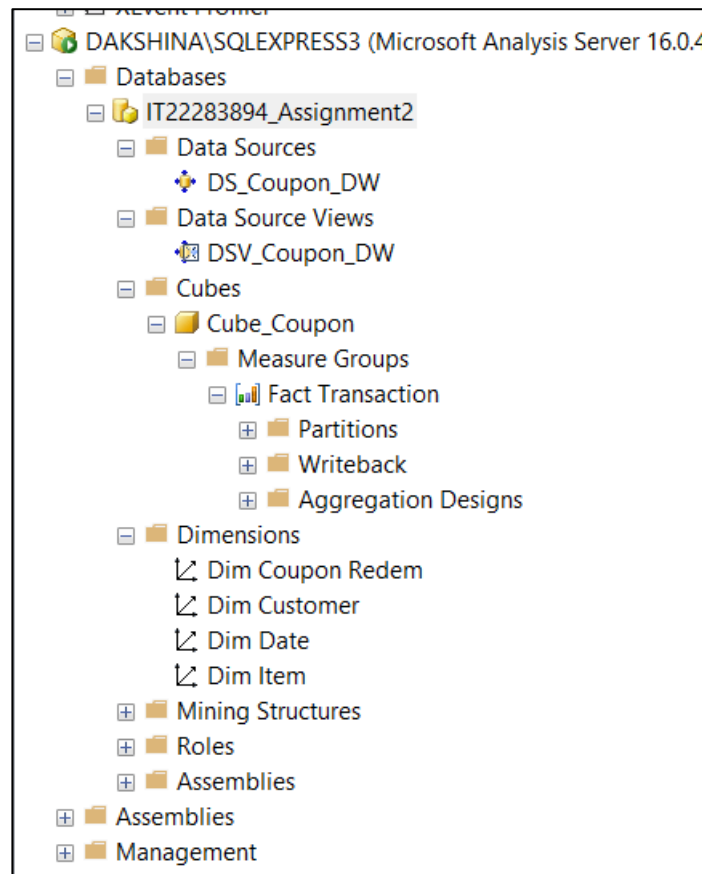


Figure 9SSAS Cube Structure

This Cube Structure can be used to browse data in SQL Server Analysis Service where queries can be passed as MDX query for visualization processes.

The screenshot shows the 'Cube_Coupon [Browse]' window in SQL Server Analysis Services. The window displays a table of data with the following columns: Campaign End Date, Category Name, Brand Type, and Coupon Used Expense. The data is filtered by the 'Fact Transaction' measure group.

Campaign End Date	Category Name	Brand Type	Coupon Used Expense
2012-10-27	Bakery	Local	681.76
2012-10-27	Grocery	Established	119.68
2012-10-27	Grocery	Local	156.9
2012-10-27	Natural Products	Established	156.72
2012-10-27	Packaged Meat	Established	178.1
2012-10-27	Pharmaceutical	Established	1415.54
2012-11-16	Bakery	Local	13835.62
2012-11-16	Dairy juices & ...	Established	372.76
2012-11-16	Grocery	Established	60237.72
2012-11-16	Grocery	Local	28480.28
2012-11-16	Meat	Established	355.48
2012-11-16	Natural Products	Established	5306.12
2012-11-16	Natural Products	Local	89.04
2012-11-16	Packaged Meat	Established	3445.02
2012-11-16	Packaged Meat	Local	2250.48

Figure 10Cube Browse data

There are some main OLAP Operations demonstrated through Excel visualizations

- Slice Operations
- Dice Operations
- Pivot Operations
- Drill-Down Operations
- Roll-Up Operations

To demonstrate all of these operations , Excel workbook was connected to the cube via Data tab. Data from Cube is considered as the data source and is obtained via Analysis services.

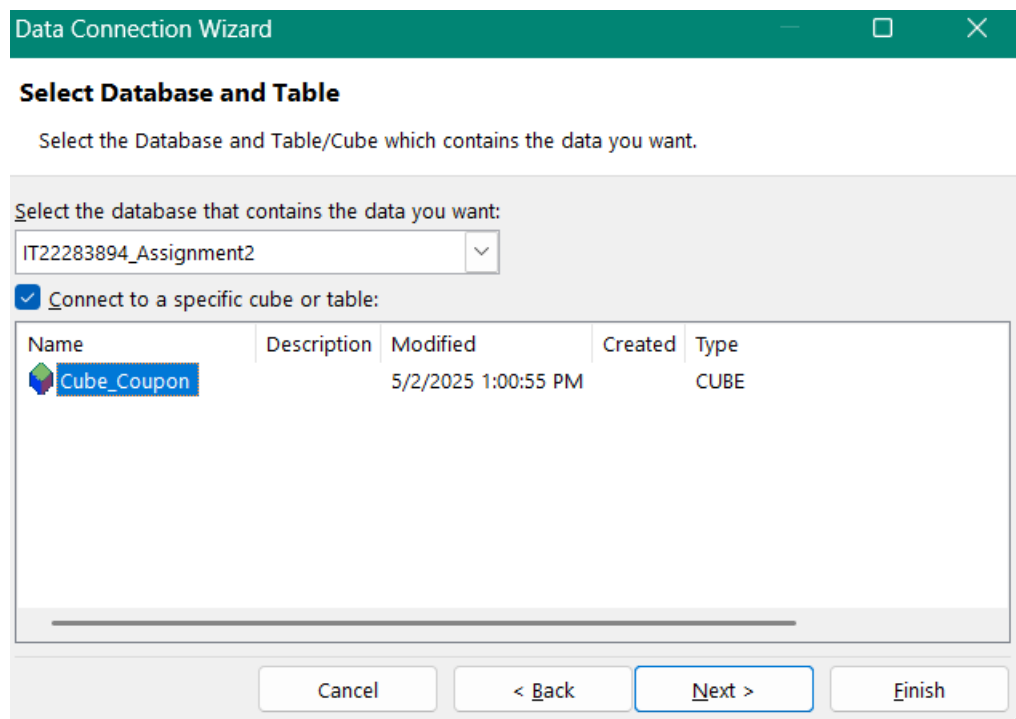


Figure 11Cube Connection Wizard

For all the operations Pivot charts and tables can be generated using Pivot Chart fields which was obtained from the Cube accessed via Analysis Service of Data tab.

4.1 SLICE - OLAP Operation.

In Slice a single dimension is selected from OLAP cube which represents a Sub cube creation. Slice operation changes the overview section in accordance with the passed dimension. In simple terms the original view is sliced to get another different view.

In order to demonstrate the SLICE operation Pivot Chart and Pivot Table was used along with a Slicer.

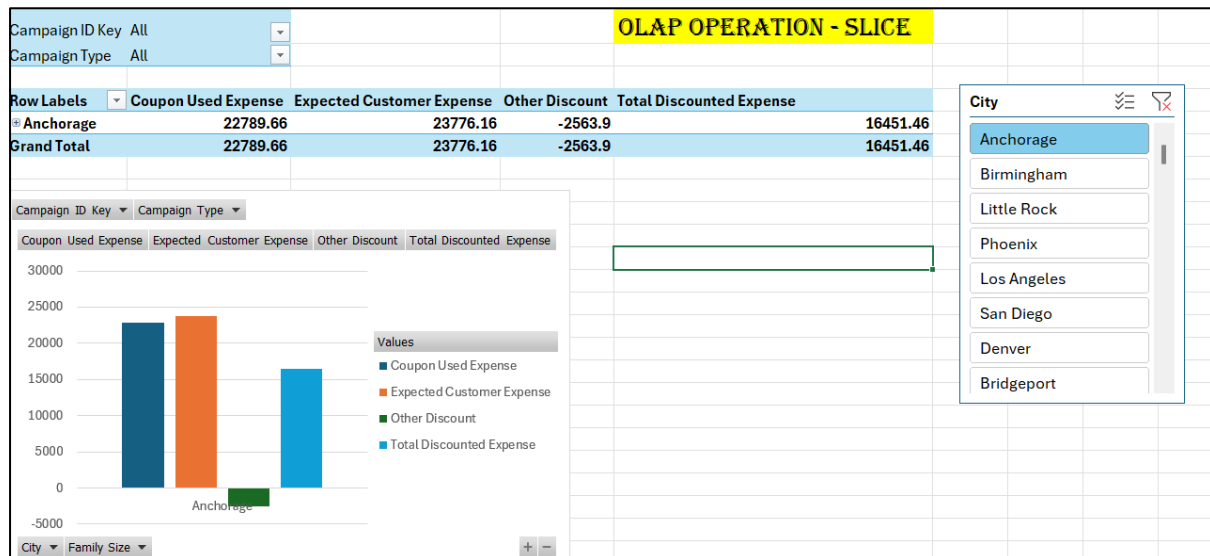


Figure 12 OLAP SLICE

The above Slice Operation Demonstrates the visualization that results by Slicing Cube using Customer City. Expected Customer expense, Other Discount , Coupon Used Expense and Total Discounted Expense filtered by the Type of Campaign is Visualized here by Slicing , we can analyze the city wise statistics.

4.2 DICE – OLAP OPERATION

In Dice two or more dimensions are selected from OLAP cube which represents a Sub cube creation. Dice operation changes the overview section in accordance with the passed dimensions. In simple terms the original view is diced to get another different view.

In order to demonstrate the DICE operation Pivot Chart and Pivot Table was used along with more than 1 Slicer. Here Cube is divided based on many dimensions , by selecting specific values on those dimensions.

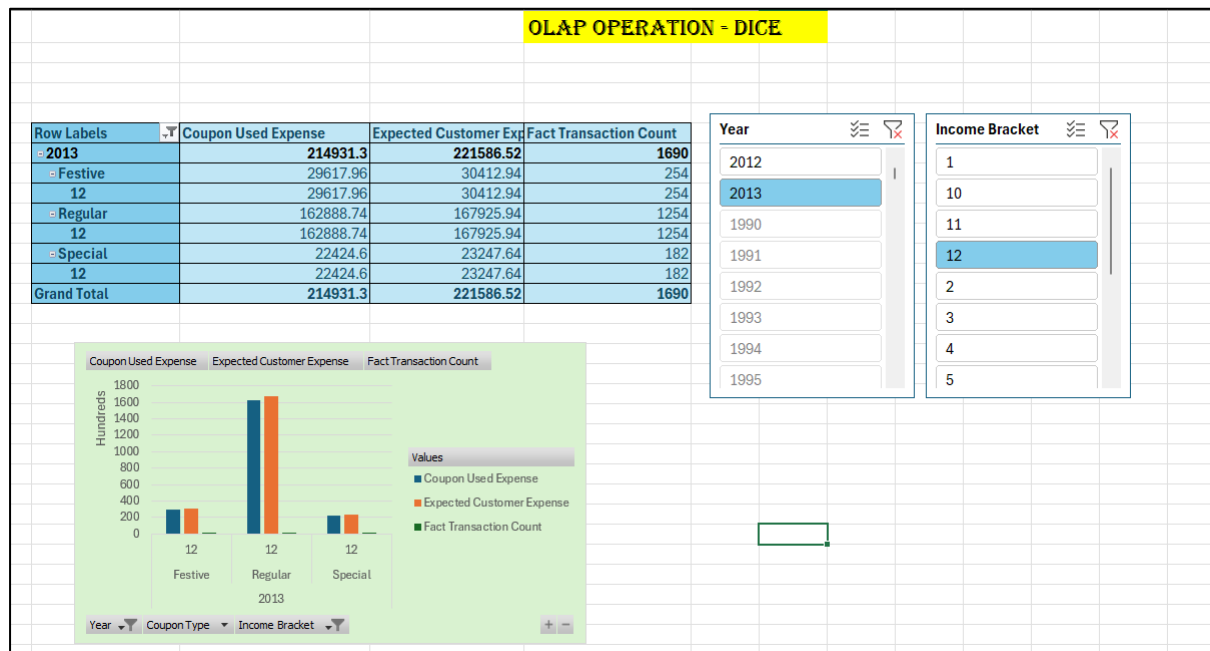
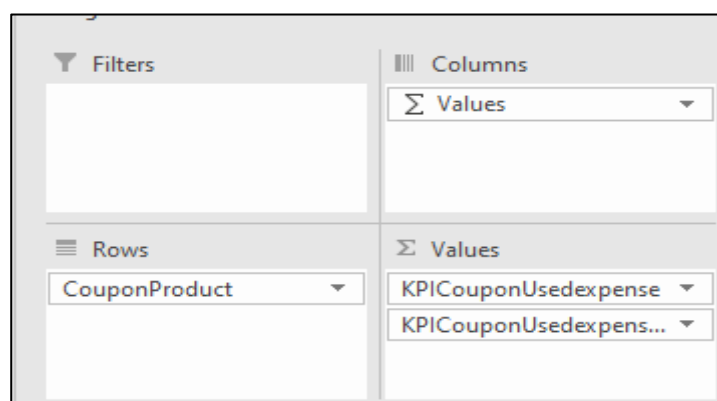


Figure 13 DICE Operation

The above Dice Operation Demonstrates the visualization that results by Dicing Cube using Customer income bracket and Year. Expected Customer expense, Coupon Used Expense and Number of Transactions in that Year with regard to customer's income bracket is visualized here. By dicing, this statistics in the point of view of various dimensions can be created.

4.3 PIVOT – OLAP OPERATION

Pivot is a visualization operation which rotates the data axes to provide an alternative presentation of the data. Simply said it's a summary table, where a single column can represent all the data by particular rows to the column value.



OLAP OPERATION - PIVOT		
Row Labels	KPICouponUsedex	KPICouponUsedexpense Goal
Bakery	572,705.20	TRUE
Local	572,705.20	TRUE
Dairy juices & Snacks	35,170.56	TRUE
Established	35,170.56	TRUE
Grocery	1,540,330.68	TRUE
Meat	1,036.52	TRUE
Miscellaneous		FALSE
Natural Products	213,475.70	TRUE
Packaged Meat	161,356.80	TRUE
Pharmaceutical	653,031.82	TRUE
Prepared Food	574.90	TRUE
Seafood	30,371.24	TRUE
Skin & Hair Care		FALSE
updated_name		FALSE
Unknown	55,092.04	TRUE
Grand Total	3,263,145.46	TRUE

Figure 14 OLAP Pivot

4.4 DRILL-DOWN AND ROLL UP – OLAP OPERATION

The drill down operation is the opposite of roll-up operation. Drill down operation is used to navigate from less detailed data to more detailed data. That is from the Top-level hierarchy to lower most level hierarchy.

It can be visualized by stepping down through small fragments of 1 dimension or introducing additional dimensions. On contrary to drill down roll up climbs up a hierarchy of a dimension to aggregate data.

The below image shows how **Customer Location wise data** is analyzed. Cube is also sliced to visualize the cube based on specific states too. As defined in the hierarchy while designing the cube , this visualization helps to drill down along Customer State , City , Zip Code and Main Address and analyze the number of transactions done , Coupon Used Expense and Total Discounted Expense of Customer with the retail shop.

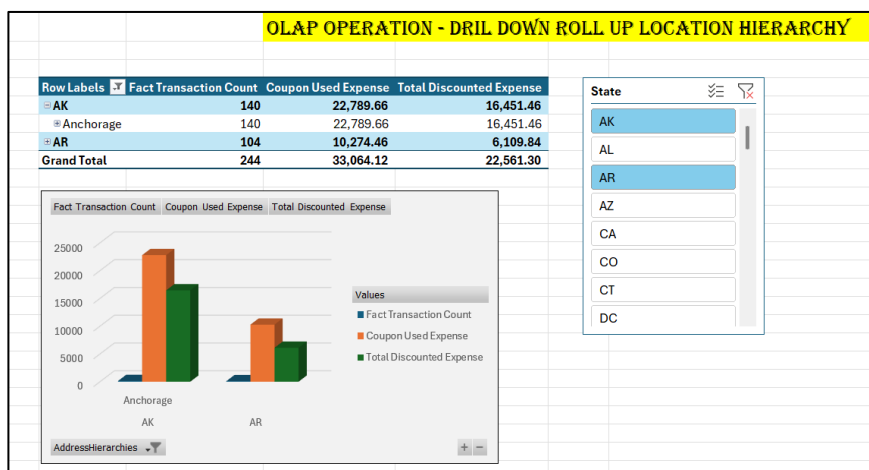


Figure 15 OLAP DRILL DOWN-ROLL UP

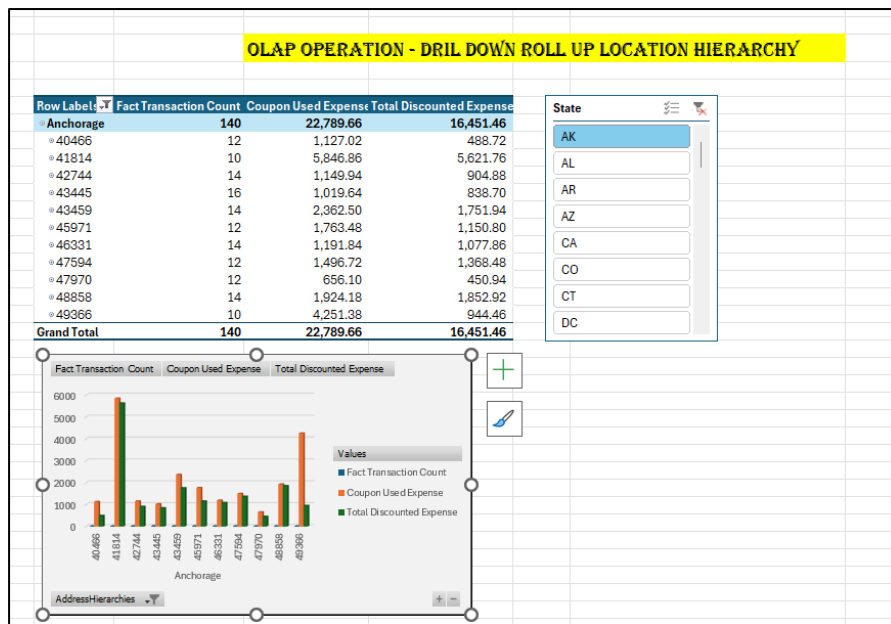


Figure 16 After DRILL DOWN

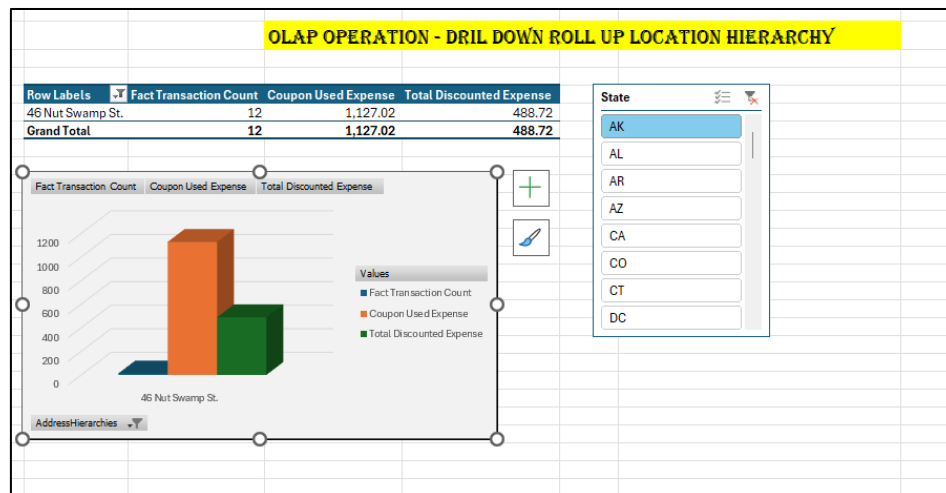


Figure 17 Further DRILL DOWN

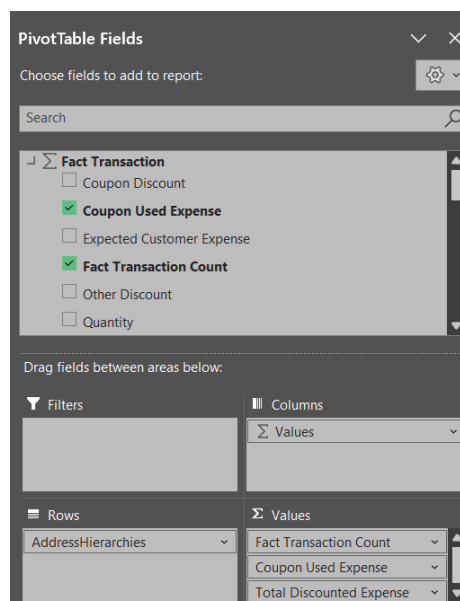
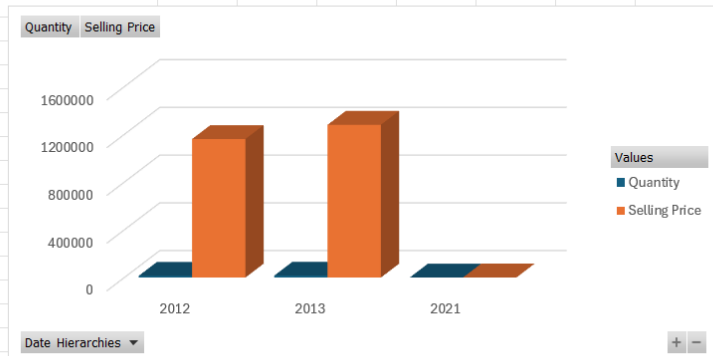


Figure 18 PIVOT Fields

PIVOT -ROLL UP

Row Labels	Quantity	Selling Price
2012	14410	1160994.3
2013	5946	1280140.98
2021	8	716.66
Grand Total	30364	2441851.94



5 POWER BI REPORTS

To demonstrate all of these operations , connected to the POWER BI using Get data tab. Data from Cube is considered as the data source and is obtained via Analysis services.

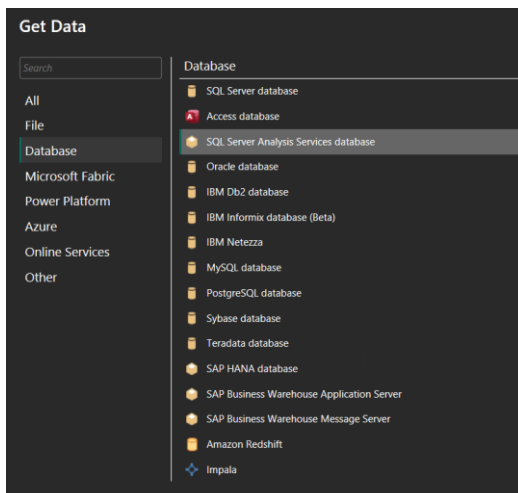
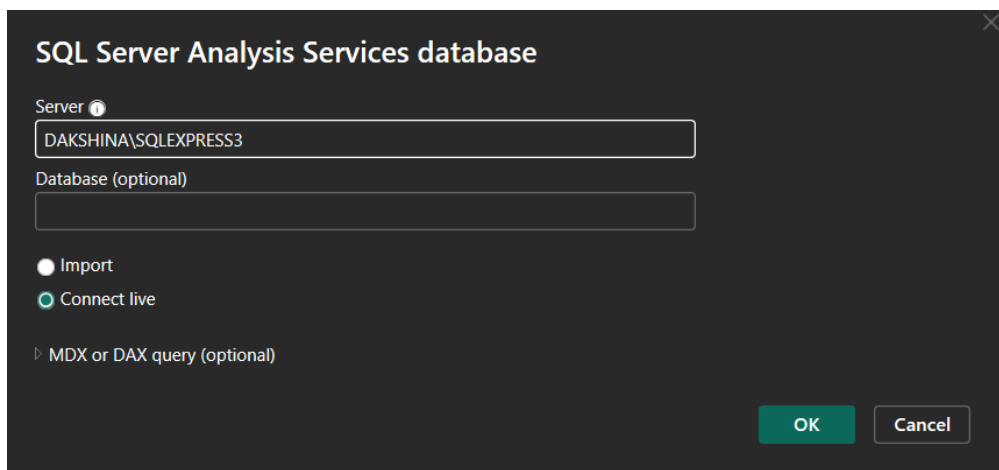
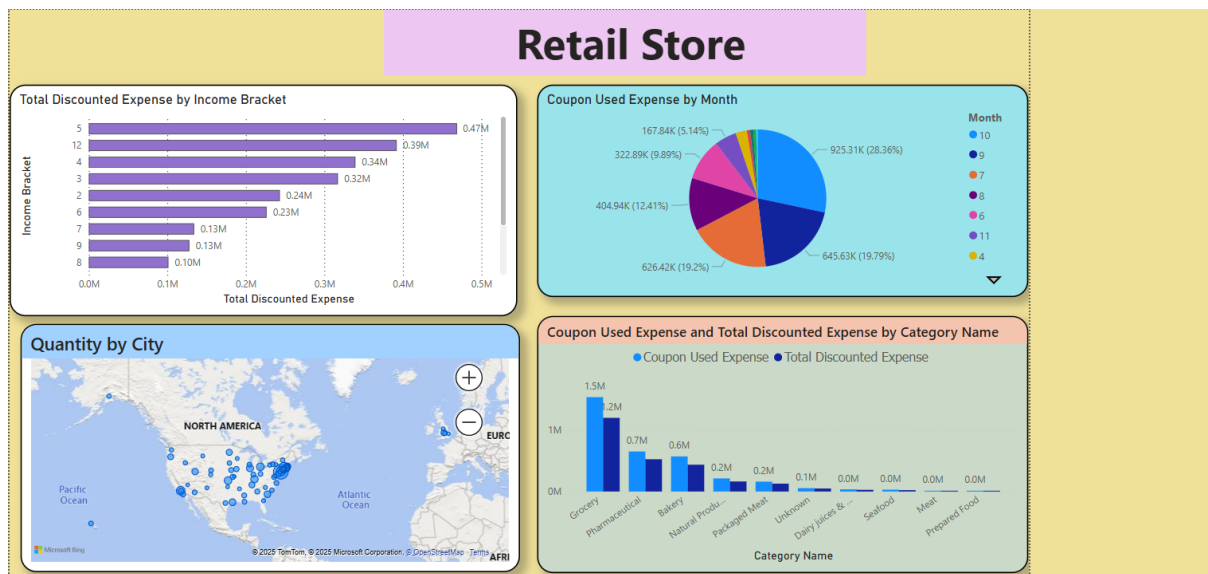


Figure 19 Connecting to the POWER BI



5.1 REPORT 01

This report uses a matrix visual in Power BI for detailed analysis of data. The data is grouped by rows and columns, allowing to view specific values in categories (e.g. Quantity by city, coupon used expense by month etc.). This helps recognize key performers, underperforming areas, and trends across different data dimensions.



5.2 REPORT 02

slicers can easily narrow down data shown using reports and dashboards. With absolutely no need to open drop-down lists to select filter options, it displays all filtering options right there on the report page-the slicers. Perhaps you might use some slicers in your report to filter data based on region, product category, or date.

Used Slicers:

- **Category Name Slicer:**
 - Displays a list of category names.
 - Allows single or multi-select.
- **Year Slicer:**
 - Displays a list of years.
 - **Cascading Behavior:** The years displayed in this slicer are filtered based on the selected Category Name.
 - Allows single or multi-select.

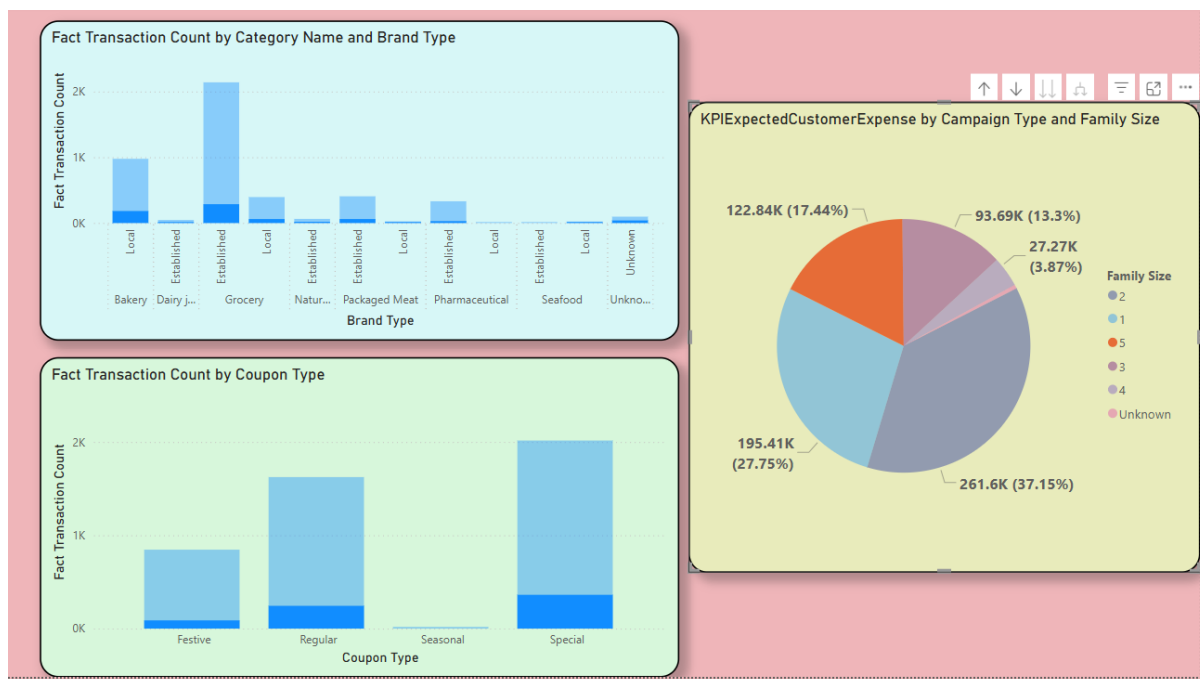
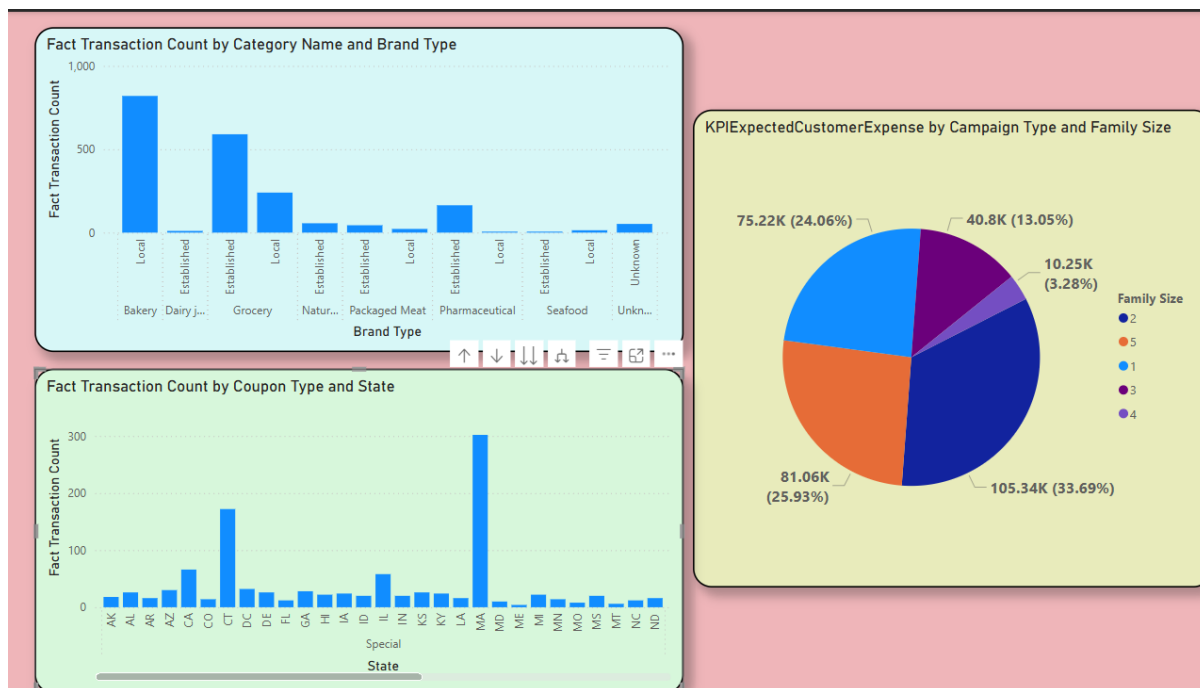


Figure 22After DRILL Down

5.4 REPORT 04

OLAP, or Online Analytical Processing, is a processing methodology specifically for multi-dimensional data analysis-multi-dimensional analysis of data and facilitate users in getting quick answers to complex queries. Data is structured in a multi-dimensional way in which cubes contain aggregated data such as totals by regions, products, and time.

So, in addition to analyzing OLAP data at the high level-you might see a total sales for a region, an option might arise where you would want to know the reason behind it. This is when drilling through comes about.

From that top-level summary, drill-through takes you to the minutiae, the individual records that formed the calculation.

This report is based on DRILL THROUGH operations in OLAP. Here I created 2 separate pages for process this operation. In the first page after click on the drill through option it shows my next page which I was apply drill through operation

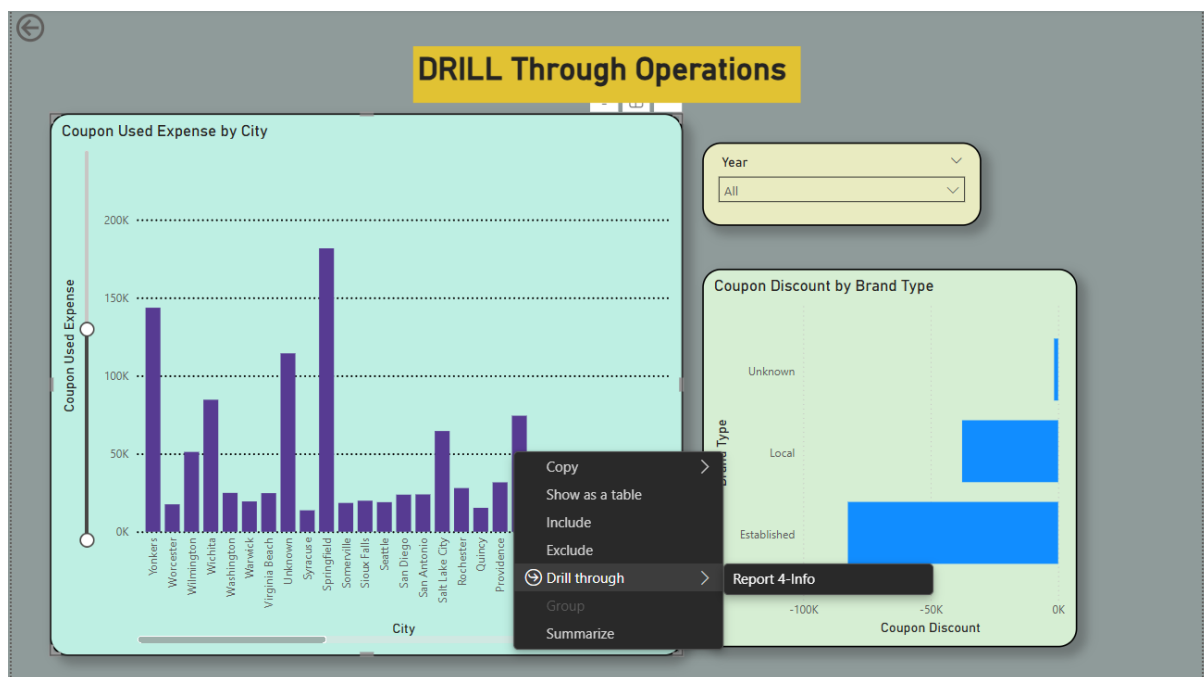


Figure 24 1 page of DRILL THROUGH

Month Name	Festive	Regular	Special	Total
May	106.86			106.86
OR	106.86			106.86
April			248.62	248.62
OR			248.62	248.62
November	141.76	922.54		1,064.30
OR	141.76	922.54		1,064.30
September	2,234.06		317.74	2,551.80
July	1,309.38	1,469.28	1,006.28	3,784.94
OR	1,309.38	1,469.28	1,006.28	3,784.94
June	2,294.32		1,912.32	4,206.64
OR	2,294.32		1,912.32	4,206.64
August	1,668.68	4,904.48	704.56	7,277.72
OR	1,668.68	4,904.48	704.56	7,277.72
October	817.82	2,641.56	40,220.02	43,679.40
Total	8,657.66	9,937.86	44,409.54	63,005.06

ME	Portland	92	-509.44	11,357.66
State	City	Fact Transaction Count	Coupon Discount	Coupon Used Expense
OR	Portland	268	-584.28	63,005.06
State	City	Fact Transaction Count	Coupon Discount	Coupon Used Expense

Figure 25 After DRILL Through Operation