



Sri Lanka Institute of Information Technology

Data Warehousing and Business Intelligence IT3030
[2022/FEB]

ASSIGNMENT – 02

Submitted by:

Registration Number:

Name: Gobisan A

Batch:

Date of Submission: 17/05/2022

This report is submitted in partial fulfilment of the requirement of the Data Warehousing and Business Intelligence module.

Declaration

I declare that this project report does not incorporate, without acknowledgment, any material previously submitted for a degree in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text.

Name of the Candidate: Gobisan A.

Table of Contents

Declaration.....	ii
List of Figures.....	v
1. Data Source.....	7
1.1. Data Source Description	7
1.2. Data Warehouse Schema	7
1.3. Dimensions and Facts Details	8
1.3.1. Hierarchical Data	8
1.3.2. Slowly Changing Dimension	8
1.3.3. Accumulating Fact Table	9
2. SSAS Cube Implementation	10
2.1. Overview.....	10
2.2. Data Source Configuration.....	11
2.3. Data Source View	11
2.4. Cube Creation	12
2.4.1. Measures	12
2.4.2. Dimensions	13
2.4.3. Hierarchies	13
2.5. KPI Creation	15
2.6. Cube Deployment	15
3. Demonstration of OLAP Operations	16
3.1. Overview of OLAP Operations.....	16
3.2. Roll-up and Drill-down Operations	16
3.3. Slice.....	18
3.4. Dice	19
3.5. Pivot.....	19
4. SSRS Reports	21
4.1. Overview of SSRS Reports	21
4.2. Report with a Matrix.....	21
4.2.1. Partner Bank wise YOY Transactions Report	21
4.2.2. Query used to retrieve Dataset.....	22
4.3. Report with more than one parameter	22

4.3.1. Location wise YOY Transactions Report	22
4.3.2. Queries used to retrieve Datasets	23
4.4. SSRS Drill-down Report.....	23
4.4.1. Transaction Method-wise YOY Transactions Report.....	23
4.4.2. Query to retrieve data set	24
4.5. SSRS Drill-through Report	24
4.5.1. Location wise Transaction Analysis	24
4.5.2. Query to retrieve Level - 01 Dataset.....	25
4.5.3. Query to Retrieve Level – 02 Dataset.....	25

List of Figures

Figure 1. 1Snowflake Schema.....	7
Figure 1. 2 Data Source Description	8
Figure 2. 1 Overall Project Structure	10
Figure 2. 2 Data Source View Design.....	11
Figure 2. 3 Cube Structure	12
Figure 2. 4 Measures.....	13
Figure 2. 5 Dimensions	13
Figure 2. 6 Location Hierarchy - Customer Address	14
Figure 2. 7 Loan Hierarchy – Loan Status	14
Figure 2. 8 Date Hierarchy.....	14
Figure 2. 9 KPI Values.....	15
Figure 2. 10 Deployment Progress.....	15
Figure 3. 1 Roll-up to States	16
Figure 3. 2 Drill-down to Cities	17
Figure 3. 3 Roll-up to years	17
Figure 3. 4 Drill-down to Quarters.....	18
Figure 3. 5 State wise Slice Operation	18
Figure 3. 6 Dice Operation.....	19
Figure 3. 7 Pivot Operation.....	20
Figure 4. 1 SSRS Web Portal	21
Figure 4. 2Matrix Report	21
Figure 4. 3 State Selection	22
Figure 4. 4 City Selection	22
Figure 4. 5 City-wise YoY Transactions Report.....	22

Figure 4. 6 Transaction Method-wise YPY Transactions Report	23
Figure 4. 7 Level 01 – Report: State wise Transaction Analysis	24
Figure 4. 8 Level 02 – Report: City wise Transaction Analysis	25

1. Data Source

1.1. Data Source Description

The data warehouse created for Assignment 01 was used as the data source for implementing the SSAS cube. The Data Warehouse represents the data of a Bank that grants loans. A snowflake schema was designed to implement an efficient Data Warehousing and Business solution to analyze the data.

1.2. Data Warehouse Schema

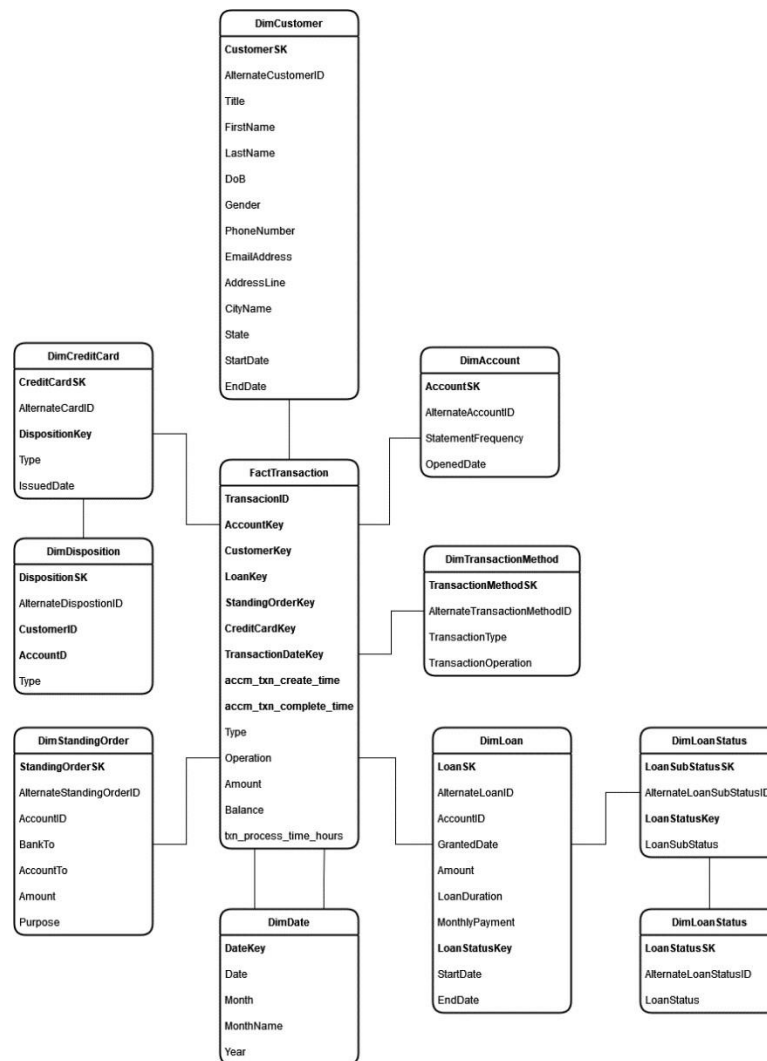


Figure 1. 1Snowflake Schema

1.3. Dimensions and Facts Details

▪ Source: DWBI_Bank_DW	▪ Source Type: SQL Database
▪ Schema: dbo	▪ Object Type: Table

Object Name	Description
DimAccount	Includes all account details
DimCreditCard	Includes details of credit card issued to accounts
DimCustomer	Includes all customer information
DimDate	Includes data related details
DimDisposition	Includes information about rights of customer to operate accounts
DimLoan	Includes information about loan granted for accounts
DimLoanStatus	Includes loan status information
DimLoanSubStatus	Includes loan sub status information
DimStandingOrder	Includes details of standing orders instructed for accounts
FactTransaction	Includes details of transactions made on accounts

Figure 1. 2 Data Source Description

1.3.1. Hierarchical Data

The Customer Dimension (DimCustomer) contains hierarchical data on customer address. (AddressLine, City, State).

Loan Dimension have reference to Loan Sub Status Dimension and Loan Sub Status Dimension have reference to Loan Status Dimension. Therefore, these three dimensions together forms a hierarchy (Loan, Sub-Status, Status)

1.3.2. Slowly Changing Dimension

The Customer Dimension was modeled as a Slowly Changing Dimension (SCD), to accommodate the history management requirement of customer data. In Customer dimension, phone is a changing attribute (Type 1), and address line, city and state are historical attributes (Type 2).

The Loan dimension (DimLoan) also modeled as a SCD, because Loan Status (Historical Attribute - Type 2) column in Loan Dimension should be maintained for historical data analysis.

1.3.3. Accumulating Fact Table

The Account Transaction Fact table (FactTransaction) was modeled as an accumulating fact table to record the lifetime of a transaction from the time it is created to the time it is completed. The grain of the fact table is the transaction made on a particular account on a particular date.

2. SSAS Cube Implementation

2.1. Overview

An OLAP cube was created and deployed using SQL Server Analysis Services (SSAS). The cube was designed including necessary measures in fact table, connecting dimensions and hierarchies. The cube was built to allow near-instantaneous analysis with aggregated data.

The following figure shows a snapshot of solution explorer from Visual Studio that illustrates the main parts that are required for online analytical processing (OLAP) cubes. These parts are the data sources, data source views, cubes, and dimensions.

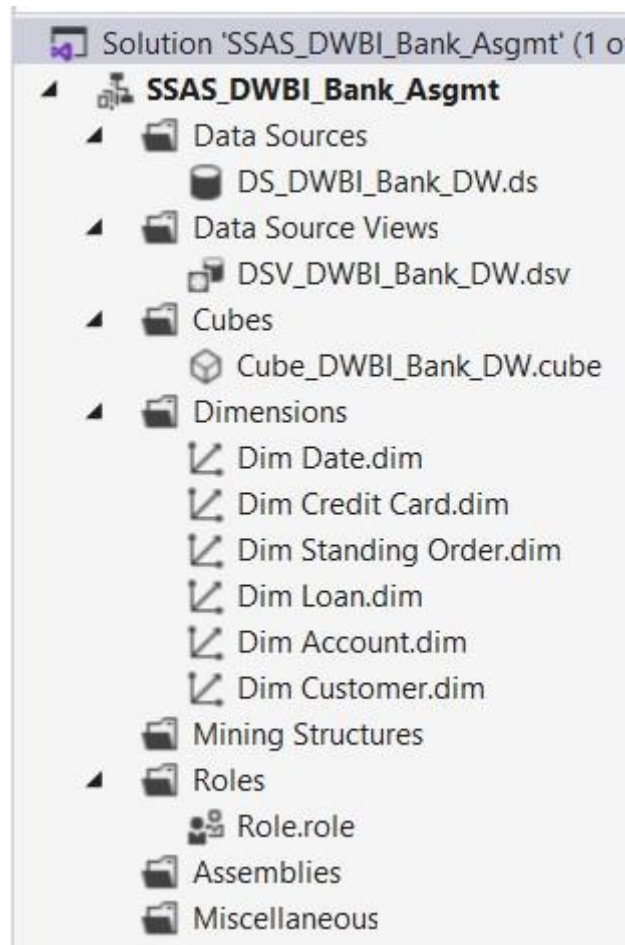


Figure 2. 1 Overall Project Structure

2.2. Data Source Configuration

As the first step data source was configured. DWBI_Bank_DW data warehouse was used as the data source for the cube solution.

The configured data source is the origin of all the data that is contained within the OLAP cube. The OLAP cube connects to the data warehouse to read and process raw data to perform aggregations and calculations for its associated measures. Authentication information about the data source was stored in SSAS to establish the correct level of permissions.

2.3. Data Source View

A data source view (DSV) was created to provide a single, unified view of the metadata from the specified dimensions and fact that the data warehouse defines in the project. This enables to work with the metadata without an open connection to underlying data warehouse.

The DSV is a collection of views that represent the dimension and fact tables from the data source. The DSV contains all the relationships between tables, such as primary and foreign keys. The relationships were defined between fact and dimension tables, within the source relational database.

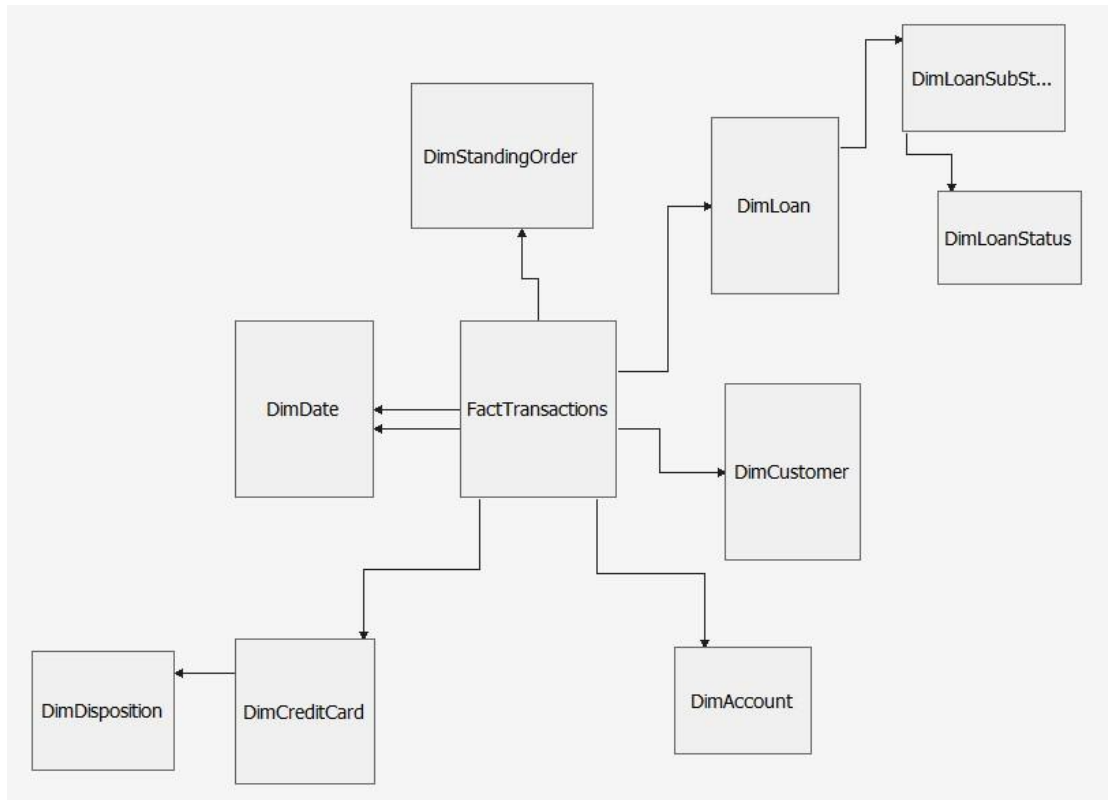


Figure 2. 2 Data Source View Design

2.4. Cube Creation

An OLAP cube was created to provide rapid analysis of data. The OLAP cube supports role up, slice, and dice operations as needed to handle the widest variety of questions.

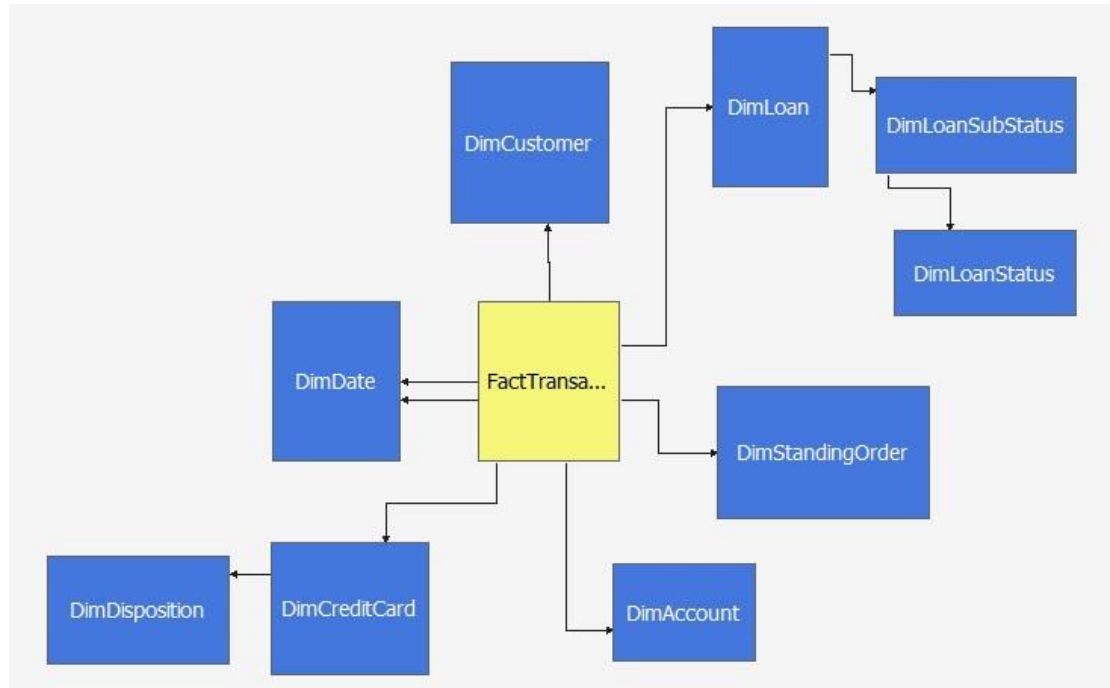


Figure 2. 3 Cube Structure

2.4.1. Measures

Measures contains the numeric values that are needed to be sliced, diced, aggregated, and analyzed; they are one of the fundamental reasons to build OLAP cubes using data warehousing infrastructure.

An example of a measure that exists in the data warehouse is TxnProcessTimeHours. TxnProcessTimeHours is a measure from FactTransactions that represents the time in hours that each transaction took to process from creation to completion.

Other measures are Amount, that is the amount of money transferred and Balance, that is the account balance after transaction.

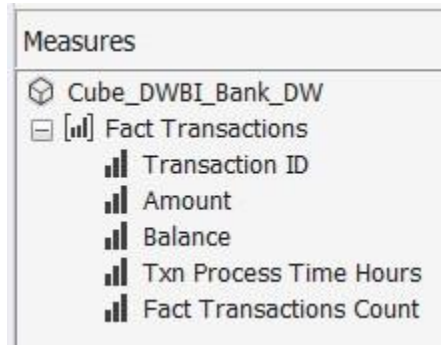


Figure 2. 4 Measures

2.4.2. Dimensions

A dimension in SSAS references a dimension from the Service Manager data warehouse. Dimensions allow the filtering, grouping, and labeling of data.

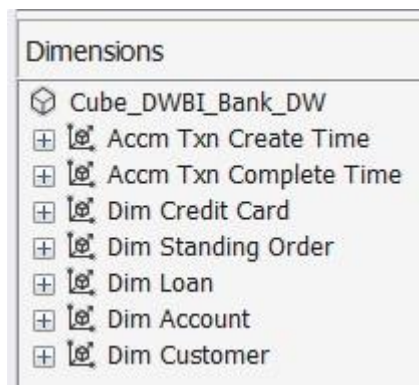


Figure 2. 5 Dimensions

2.4.3. Hierarchies

Hierarchies are useful in visual reporting tools to show the Parent/Child relationship between attributes. That means hierarchies provide the ability to drill down to detailed level data and roll up to aggregated data thus making business intelligence efficient. There several hierarchies were obtained from the data available in the data source, they are mentioned below.

a. Location Hierarchy

A location hierarchy was obtained in customer data. In that hierarchy higher level is state, where a state contains multiple cities, and a region contains multiple address lines. (State → City → AddressLine)

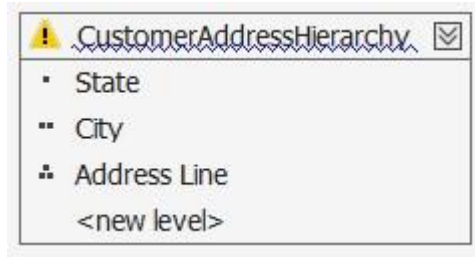


Figure 2. 6 Location Hierarchy - Customer Address

b. Loan Hierarchy

A hierarchy was obtained in loan sub-status and status. In that hierarchy higher level is status, where a status contains multiple sub-statuses. (LoanStatus → LoanSubStatus)



Figure 2. 7 Loan Hierarchy – Loan Status

c. Date Hierarchy

A date hierarchy was created to perform day, week, month, quarter, and year wise drill down and rollup operations.



Figure 2. 8 Date Hierarchy

2.5. KPI Creation

KPI's were created based on the business requirements. Below are the KPI's created to monitor and measure transaction values:

KPI	Goal Expression	Description
KPI Amount	[Measures].[Amount] > 5000000	If goal is true target transaction amount achieved
KPI Balance	[Measures].[Balance] < 10000000	If goal is true customer at debt
KPI Process Time	[Measures].[TxnProcessTimeHours] > 96	If goal is true transaction is unsuccessful or overdue
KPI Transaction Count	[Measures].[Fact Transactions Count] > 500	If goal is true target transaction achieved

Figure 2. 9 KPI Values

2.6. Cube Deployment

The Cube was deployed to support analysis purposes. Deployed, SSAS Cube was made available for analysis under SSAS databases accessible via SSMS.

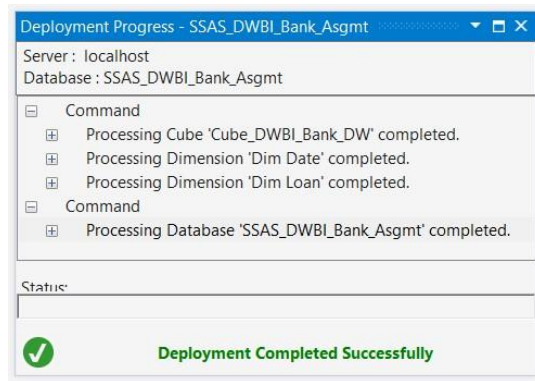


Figure 2. 10 Deployment Progress

3. Demonstration of OLAP Operations

3.1. Overview of OLAP Operations

An Excel workbook was connected to the Cube, in POWERPIVOT mode, to perform OLAP operations. Auto generated MDX queries were used to access data in the cube. pivot tables and pivot charts were used demonstrate OLAP operations.

3.2. Roll-up and Drill-down Operations

The roll-up operation performs aggregation on a data cube, by climbing up concept hierarchies. The drill-down operation navigates from less detailed record to more detailed data.

3.2.1. Location wise Analysis Report

The report presents the total transaction amount, total balances, number of transactions, process time and the achievement of the KPI goal for each year, based on the roll-ups and drill downs of cities and states.

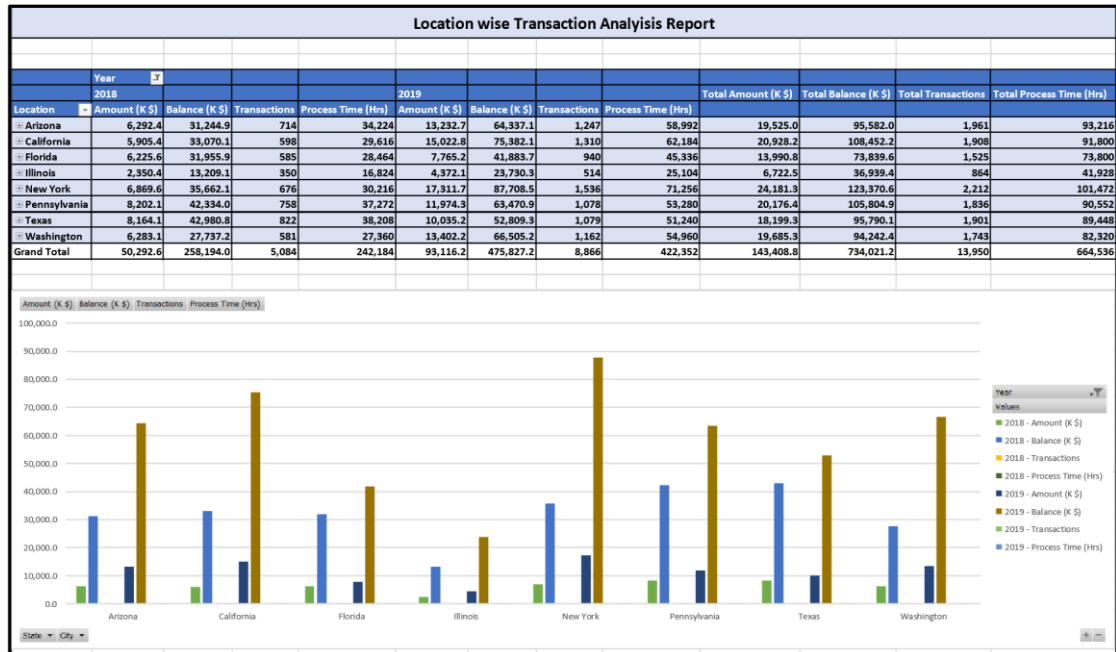


Figure 3. 1 Roll-up to States

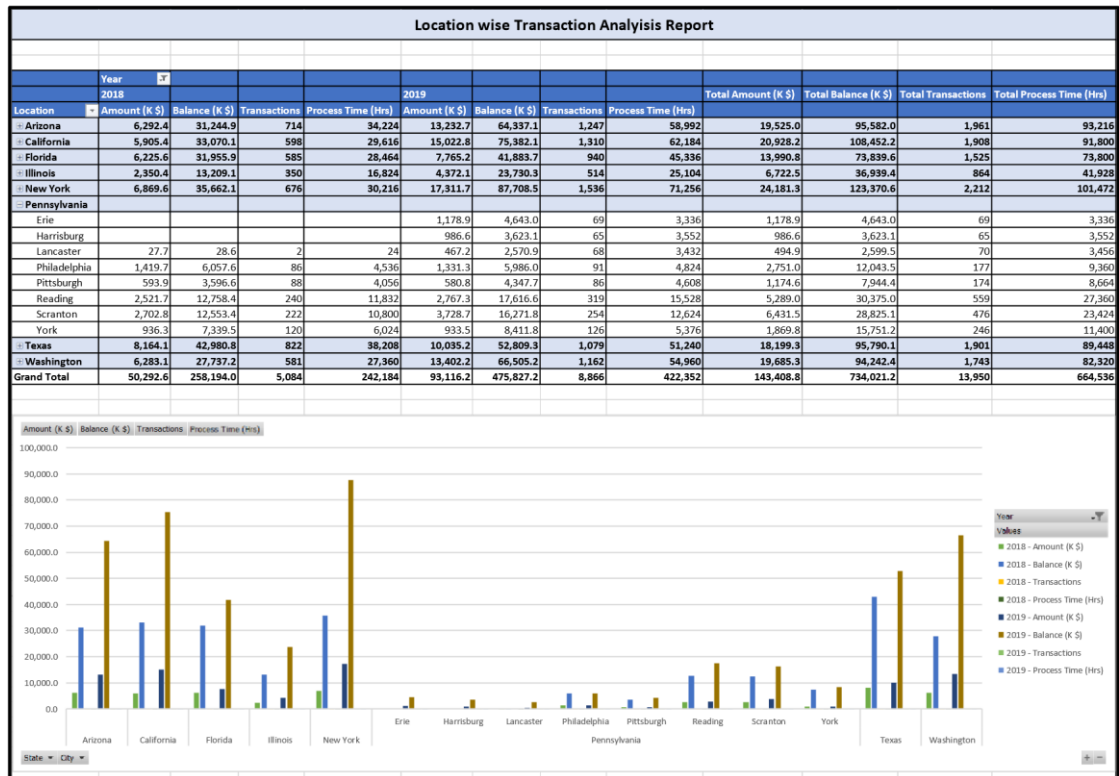


Figure 3. 2 Drill-down to Cities

3.2.1. Year wise Analysis Report

The report presents the total transaction amount, total balances, number of transactions, process time and the achievement of the KPI goal for each transaction method, based on the roll-ups and drill downs of years, quarters and months.

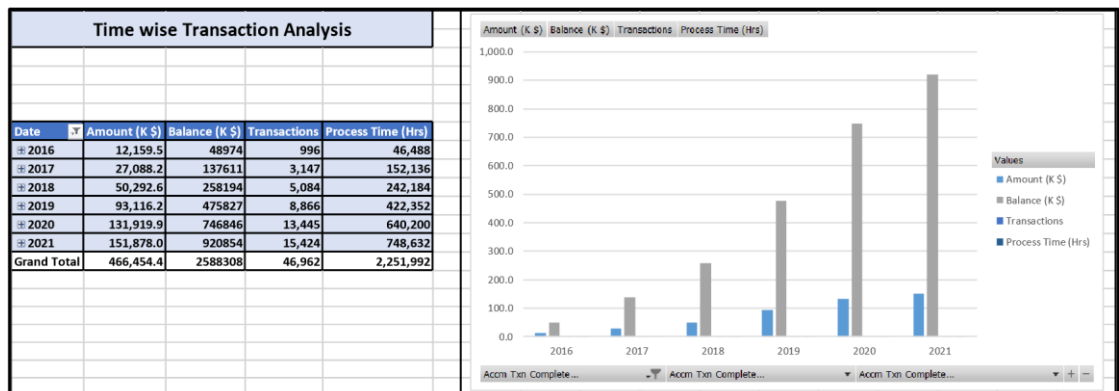


Figure 3. 3 Roll-up to years

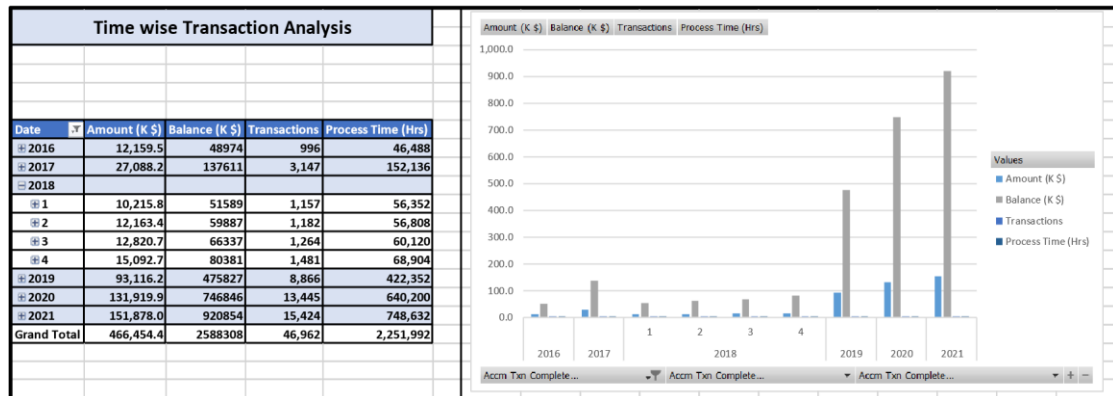


Figure 3. 4 Drill-down to Quarters

3.3. Slice

A slice is a subset of the cubes corresponding to a single value for one or more members of the dimension.

3.3.1. State wise Transactions Report

A slice operation was executed on the customer location that represents total transaction amount, total balances, number of transactions, process time and the achievement of the KPI goal for each year, sliced on states.

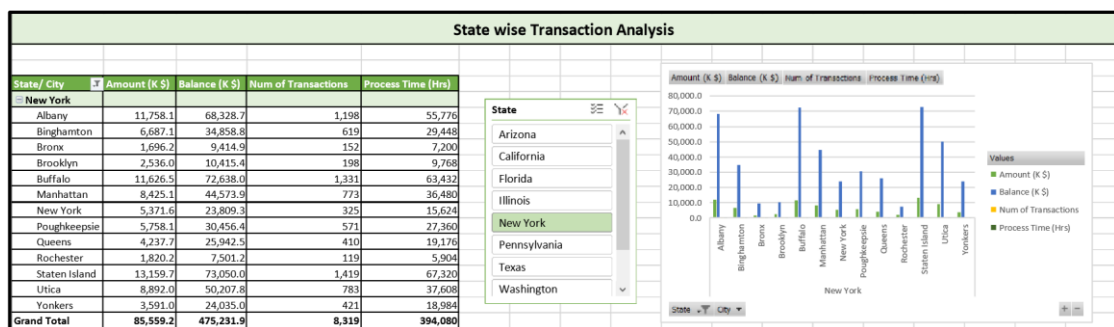


Figure 3. 5 State wise Slice Operation

3.4. Dice

The dice operation describes a sub-cube by applying a selection on two or more dimension.

3.4.1. Year, State and Operation wise Report

The dice operation was implanted on year, state, and transaction method to analyze total transaction amount, total balances, number of transactions and process time.

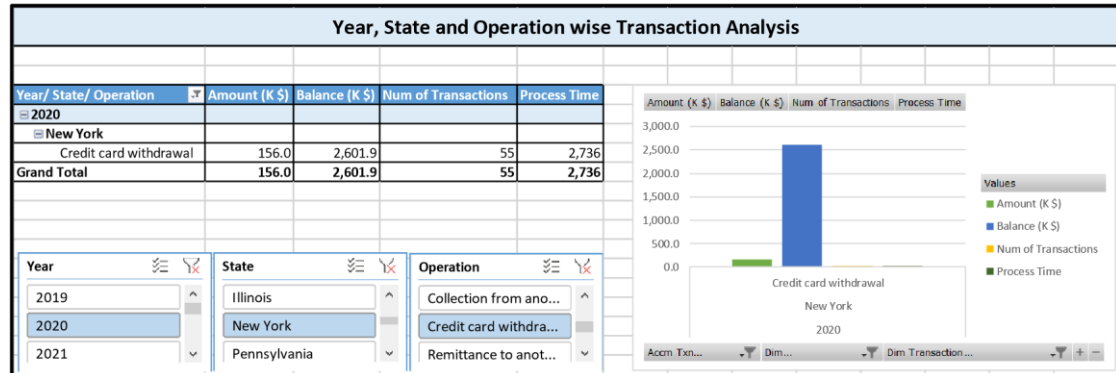


Figure 3. 6 Dice Operation

3.5. Pivot

Pivot OLAP operation rotates the data axes in view to provide an alternative presentation of the data. It may contain swapping the rows and columns or moving one of the row dimensions into the column dimensions.

3.5.1. Location wise Transaction Report

The report presents the total transaction amount, total balances, and number of transactions, for each year, based on the states.

State	Amount (M \$) - 2018	Amount (M \$) - 2019	Balance (M \$) - 2018	Balance (M \$) - 2019	Transactions - 2018	Transactions - 2019
Arizona	6.00	13.00	31.00	64.00	0.00	0.00
California	6.00	15.00	33.00	75.00	0.00	0.00
New York	7.00	17.00	36.00	88.00	0.00	0.00
Texas	8.00	10.00	43.00	53.00	0.00	0.00
Washington	6.00	13.00	28.00	66.00	0.00	0.00

Figure 3. 7 Pivot Operation

4. SSRS Reports

4.1. Overview of SRRS Reports

SQL Server Reporting Service and Report Builder were used for creating SSRS reports.

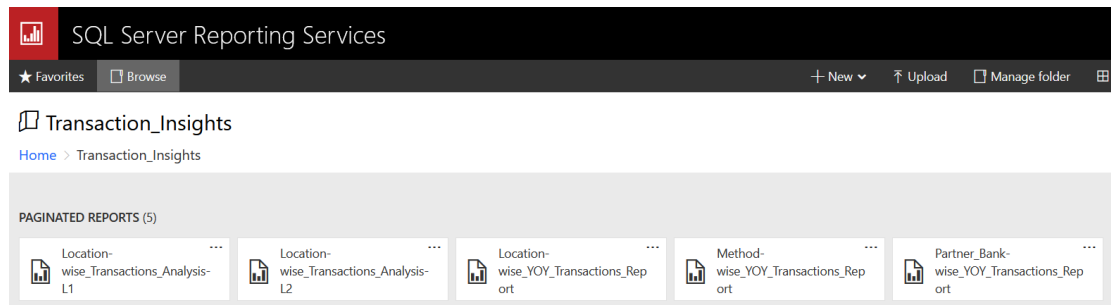


Figure 4. 1 SSRS Web Portal

4.2. Report with a Matrix

A report with matrix was configured to show data grouped by columns and rows, with aggregate data at the intersection.

4.2.1. Partner Bank wise YOY Transactions Report

The report presents the total amount transferred, total balance after transactions and total processing time of transactions for respective Years. the data was filtered for year just 2020 and 2021.

Partner Bank-wise YOY Transactions Report						
Partner Bank	2020			2021		
	Amount (\$)	Balance (\$)	Processing Time (Hrs)	Amount (\$)	Balance (\$)	Processing Time (Hrs)
AB	15,703,992.50	94,290,448.0	80,112.0	18,085,441.10	118,159,652.2	95,664.0
CD	8,431,088.80	42,926,555.9	35,592.0	10,442,882.90	54,091,324.0	45,336.0
EF	9,908,258.30	63,682,272.8	55,872.0	10,123,019.99	68,780,454.4	58,440.0
GH	10,356,339.00	58,917,472.0	50,640.0	16,925,628.20	92,083,271.7	67,656.0
IJ	7,306,232.40	45,786,262.8	34,608.0	5,933,353.10	45,109,980.1	37,320.0
KL	4,851,381.10	30,219,100.3	30,576.0	5,832,071.10	38,683,183.3	35,448.0
MN	13,149,991.30	67,338,666.0	61,464.0	14,101,007.70	80,641,839.4	72,888.0
OP	7,396,628.90	34,220,005.2	30,792.0	7,994,000.00	43,039,360.0	36,864.0
QR	9,885,155.40	55,190,743.5	45,240.0	11,533,599.20	65,862,847.4	50,592.0
ST	5,232,260.90	32,433,202.0	32,664.0	6,411,667.20	42,920,830.4	36,504.0
UV	16,371,646.21	93,475,516.5	77,952.0	19,773,456.90	121,184,668.0	93,096.0
WX	10,069,859.50	52,871,338.2	42,648.0	11,181,192.70	61,821,537.3	49,488.0
YZ	13,257,079.10	75,494,733.5	62,040.0	13,540,654.90	88,475,410.9	69,336.0
Total	131919913.4100	746846316.7376	640200	151877975.0132	920854359.0404	748632

Figure 4. 2 Matrix Report

4.2.2. Query used to retrieve Dataset

```
SELECT dst.BankTo, dd.Year, ft.Amount, ft.Balance, ft.txn_process_time_hours
FROM FactTransactions ft
INNER JOIN DimStandingOrder dst
ON ft.StandingOrderKey = dst.StandingOrderSK
INNER JOIN DimDate dd
ON ft.accm_txn_complete_time = dd.DateKey
```

4.3. Report with more than one parameter

An SSRS report with multiparameter-values to allow users to pass either one or more input parameter values to the report to filter values and analyze deeply.

4.3.1. Location wise YOY Transactions Report

The report allows to select multiple states and respective districts through a drop-down list. The report presents the total amount transferred, total balance after transactions and total processing time of transactions for specific states and respective states. respective Years. the data was filtered for year just 2020 and 2021.



Figure 4. 3 State Selection

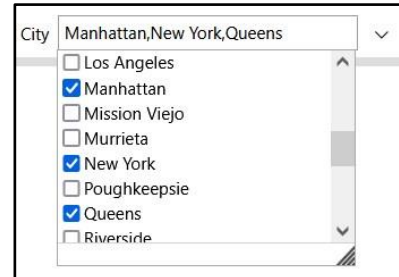


Figure 4. 4 City Selection

City-wise YOY Transactions Report							
State	City	2020			2021		
		Amount (USD)	Balance (USD)	Processing Time (Hrs)	Amount (USD)	Balance (USD)	Processing Time (Hrs)
California	Los Angeles	499,197.8	3,755,867.2	3,864.0	508,507.5	4,313,533.8	3,984.0
	San Diego	1,214,983.8	6,606,112.7	4,704.0	1,522,635.5	8,516,768.7	5,160.0
	San Francisco	1,187,125.2	4,476,881.6	2,856.0	1,211,996.2	5,439,998.7	4,632.0
New York	New York	1,621,711.9	6,551,199.0	4,272.0	1,507,040.1	7,568,108.4	4,728.0
	Queens	810,870.8	5,921,386.1	3,696.0	816,634.7	6,426,850.9	3,888.0
	Yonkers	1,335,355.3	8,602,294.5	6,576.0	1,560,597.6	12,512,964.8	9,768.0

Figure 4. 5 City-wise YoY Transactions Report

4.3.2. Queries used to retrieve Datasets

-- Transaction Data

```
SELECT dcu.State, dcu.City, dd.Year, ft.Amount, ft.Balance,
ft.txn_process_time_hours
FROM FactTransactions ft
INNER JOIN DimCustomer dcu ON ft.CustomerKey = dcu.CustomerSK
INNER JOIN DimDate dd ON ft.accm_txn_complete_time = dd.DateKey
WHERE dcu.City IN (@City)
```

-- State List

```
SELECT DISTINCT State FROM DimCustomer
```

-- City List

```
SELECT DISTINCT City FROM DimCustomer WHERE State IN (@State)
```

4.4. SSRS Drill-down Report

An SSRS drill down report was created to facilitate to step down the hierarchy to analyze detailed data.

4.4.1. Transaction Method-wise YOY Transactions Report

To demonstrate the drill down operation transaction method hierarchy was used. This report presents total transaction amounts, the total balances after transaction and the number of transactions done for each year based on the transaction type and operation.

Transaction Method-wise YOY Transactions Report							
Transaction Type	Transaction Operation	2020			2021		
		Amount (USD)	Balance (USD)	Processing Time (Hrs)	Amount (USD)	Balance (USD)	Processing Time (Hrs)
Deposit	Cash Deposit	43,086,958.1	239,567,854.2	189,408.0	49,169,481.0	286,449,888.6	213,600.0
	Check Deposit	7,444,322.0	30,297,519.4	27,360.0	8,588,499.0	37,110,805.4	33,792.0
	Collection from another bank	23,920,592.0	47,638,757.0	27,336.0	27,476,442.0	58,240,717.5	32,952.0
Withdrawal	Cash Withdrawal	45,050,762.3	276,723,628.3	251,208.0	50,483,369.6	315,550,573.5	275,688.0
	Credit card withdrawal	689,000.0	14,383,550.2	12,768.0	1,476,900.0	33,062,982.0	27,120.0
	Remittance to another bank	11,728,279.0	138,235,007.5	132,120.0	14,683,283.4	190,439,392.0	165,480.0

Figure 4. 6 Transaction Method-wise YPY Transactions Report

4.4.2. Query to retrieve data set

```
SELECT tm.TransactionType, tm.TransactionOperation,
       dd.Month, dd.MonthName, dd.Year,
       ft.Amount, ft.Balance, ft.txn_process_time_hours
FROM FactTransactions ft
INNER JOIN DimTransactionMethod tm
ON ft.TransactionMethodKey = tm.TransactionMethodSK
INNER JOIN DimDate dd
ON ft.accm_txn_complete_time = dd.DateKey
```

4.5. SSRS Drill-through Report

A drill through report was created to provide more efficient analysis by directing the user from a high-level report to more detailed report on a specific subject data.

4.5.1. Location wise Transaction Analysis

This report represents two level of data the higher level is state wise transactions data, and the lower level is city wise transactions data.

This report contains two bar charts. Left side chart represents the data for total transaction amounts for each location (state/ city). Right side chart represents the total number of transactions done for each state.

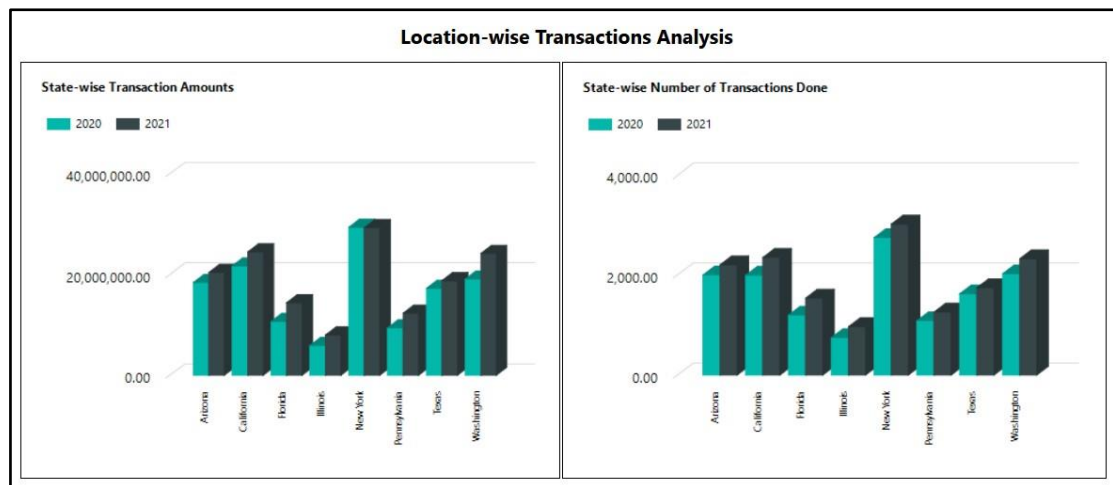


Figure 4. 7 Level 01 – Report: State wise Transaction Analysis

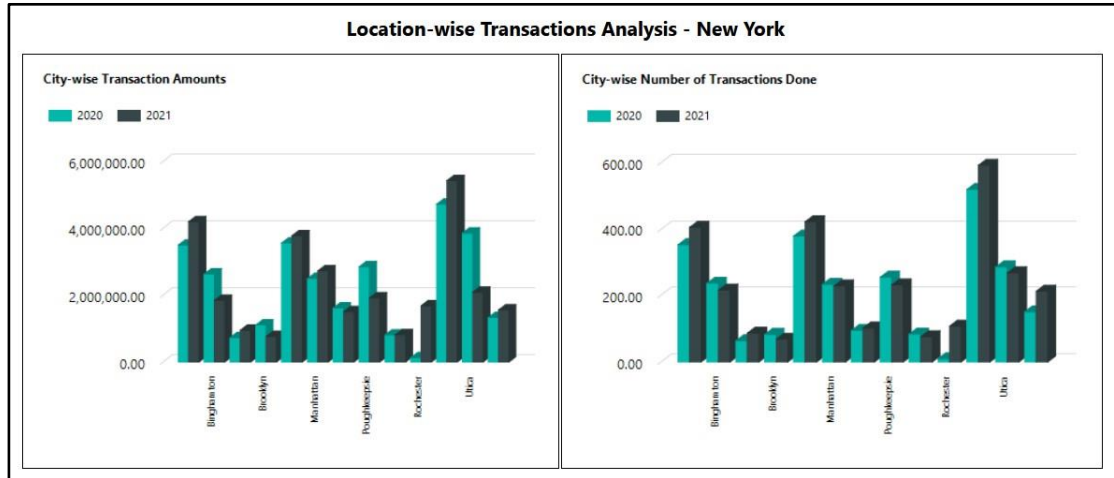


Figure 4. 8 Level 02 – Report: City wise Transaction Analysis

4.5.2. Query to retrieve Level - 01 Dataset

```
SELECT dcu.State, dd.Year, ft.TransactionID, ft.Amount,
       ft.Balance, ft.txn_process_time_hours
FROM FactTransactions ft
INNER JOIN DimCustomer dcu ON ft.CustomerKey = dcu.CustomerSK
INNER JOIN DimDate dd ON ft.accm_txn_complete_time = dd.DateKey;
```

4.5.3. Query to Retrieve Level – 02 Dataset

```
SELECT dcu.State, dcu.City, dd.Year,
       ft.TransactionID, ft.Amount, ft.Balance,
       ft.txn_process_time_hours
FROM FactTransactions ft
INNER JOIN DimCustomer dcu ON ft.CustomerKey = dcu.CustomerSK
INNER JOIN DimDate dd ON ft.accm_txn_complete_time = dd.DateKey WHERE
dcu.State = @State
```

References

Docs.microsoft.com. 2022. *Overview of OLAP cubes for advanced analytics*. [online] Available at: <https://docs.microsoft.com/en-us/system-center/scsm/olap-cubes-overview?view=sc-sm-2022> [Accessed 12 May 2022].

Docs.microsoft.com. 2022. *Defining a Data Source View*. [online] Available at: <https://docs.microsoft.com/en-us/analysis-services/multidimensional-tutorial/lesson-1-3-defining-a-data-source-view?view=asallproducts-allversions> [Accessed 12 May 2022].

Docs.microsoft.com. 2022. *Create and manage hierarchies in Analysis Services tabular models*. [online] Available at: <https://docs.microsoft.com/en-us/analysis-services/tabular-models/create-and-manage-hierarchies-ssas-tabular?view=asallproducts-allversions> [Accessed 13 May 2022].

www.javatpoint.com. 2022. *OLAP Operations - javatpoint*. [online] Available at: <https://www.javatpoint.com/olap-operations> [Accessed 14 May 2022].