Sri Lanka Institute of Information Technology

Assignment 2

Data Warehouse & Business Intelligence – IT3021

B.Sc. (Hons) in Information Technology

IT22562456

H.M.T.W Dilshan

Step 1: Data source for the Assignment 2

The data source for this assignment is the Olympic Data Warehouse (Olympic_DW), which I created for Assignment 01. This data warehouse stores Olympic Games data, focusing on medal achievements, and uses a snowflake schema for efficient data organization. The structure includes one fact table and several dimension tables, described below.

- Fact Table: MedalsList

  Measures: MedalCount (total number of medals), MedalID

  Keys: Links to AthleteID, CountryID, EventID, YearID, MedalTypeID

- Dimension Tables:

  Athlete: AthleteID, AthleteName, Gender

  Country: CountryID, CountryName, Continent

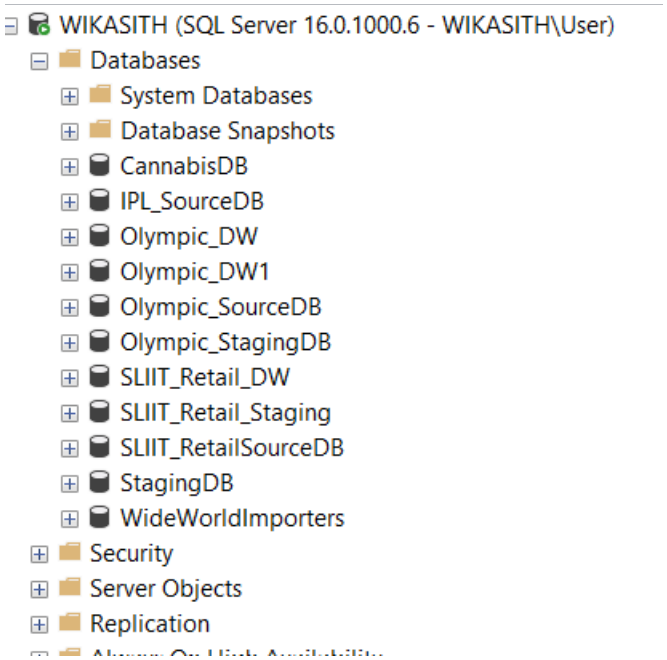
  Event: EventID, EventName, DisciplineID

  Year: YearID, YearValue

  MedalType: MedalTypeID, MedalName (Gold, Silver, Bronze)

  Discipline: DisciplineID, DisciplineName

The snowflake schema normalizes dimension tables, such as linking Discipline to Event, to reduce redundancy and enhance query performance. The ER-diagram below shows the relationships between these tables.



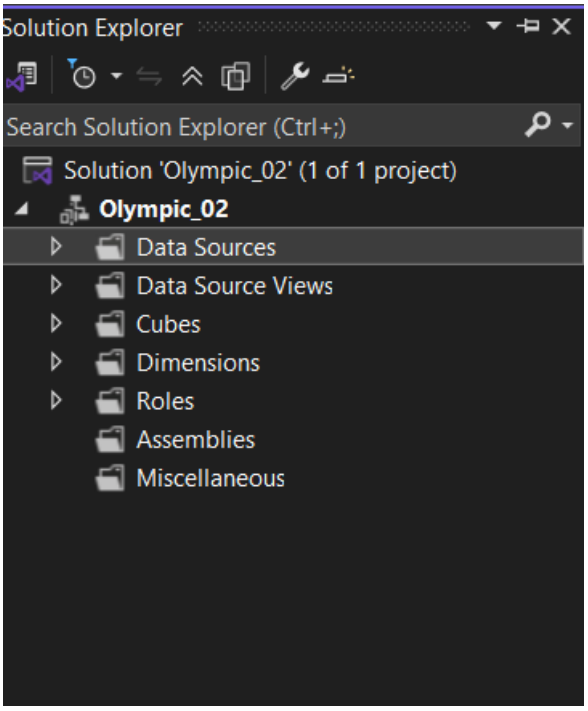
Step 2 – SSAS Cube Implementation

This section details the creation of an SSAS cube using the Olympic_DW data warehouse as the data source. The cube was built using SQL Server Data Tools (SSDT), SQL Server Management Studio (SSMS), and SQL Server Analysis Services (SSAS).

Tools Used –

- SQL Server Data Tools (SSDT) for Visual Studio
- SQL Server Management Studio (SSMS)
- SQL Server Analysis Services (SSAS)

Snapshot of the solution explorer.

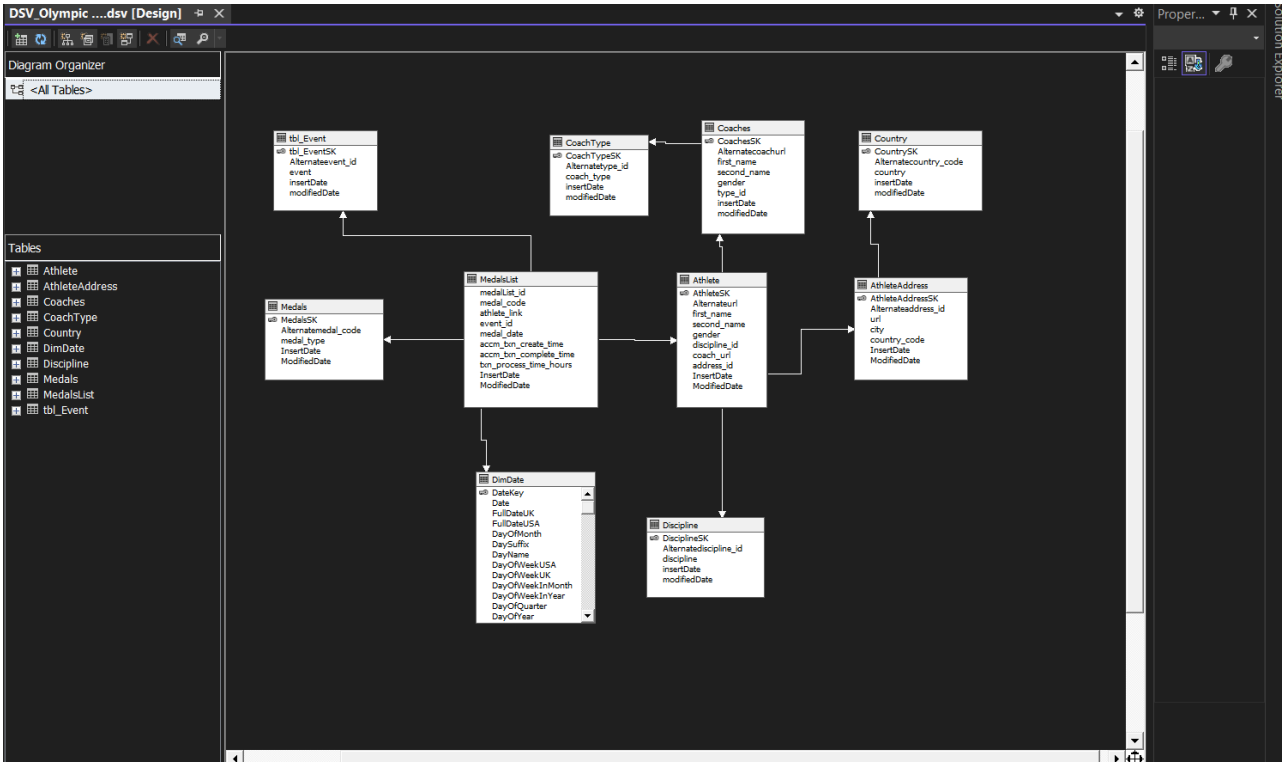


1.Create a Data source

I created a data source to connect to the Olympic_DW database. In SSDT, under the Data Sources folder, I added a new data source named DS_Olympic_DW. The connection used the SQL Server Native Client, linking to the Olympic_DW database on my local server.

2. Create a Data Source View

Next, I created a data source view to define the schema. Under the Data Source Views folder in SSDT, I added DSV_Olympic_DW, including the MedalsList fact table and dimension tables (Athlete, Country, Event, Year, MedalType, Discipline).

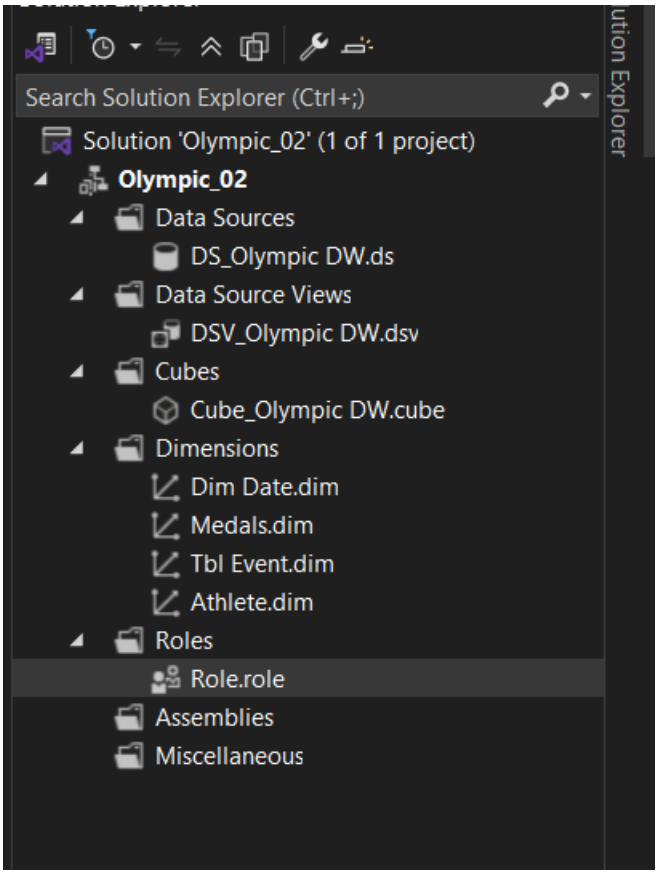


3. Create a Cube

I created the cube using the data source view. Under the Cubes folder in SSDT, I added Cube_Olympic_DW with the following components:

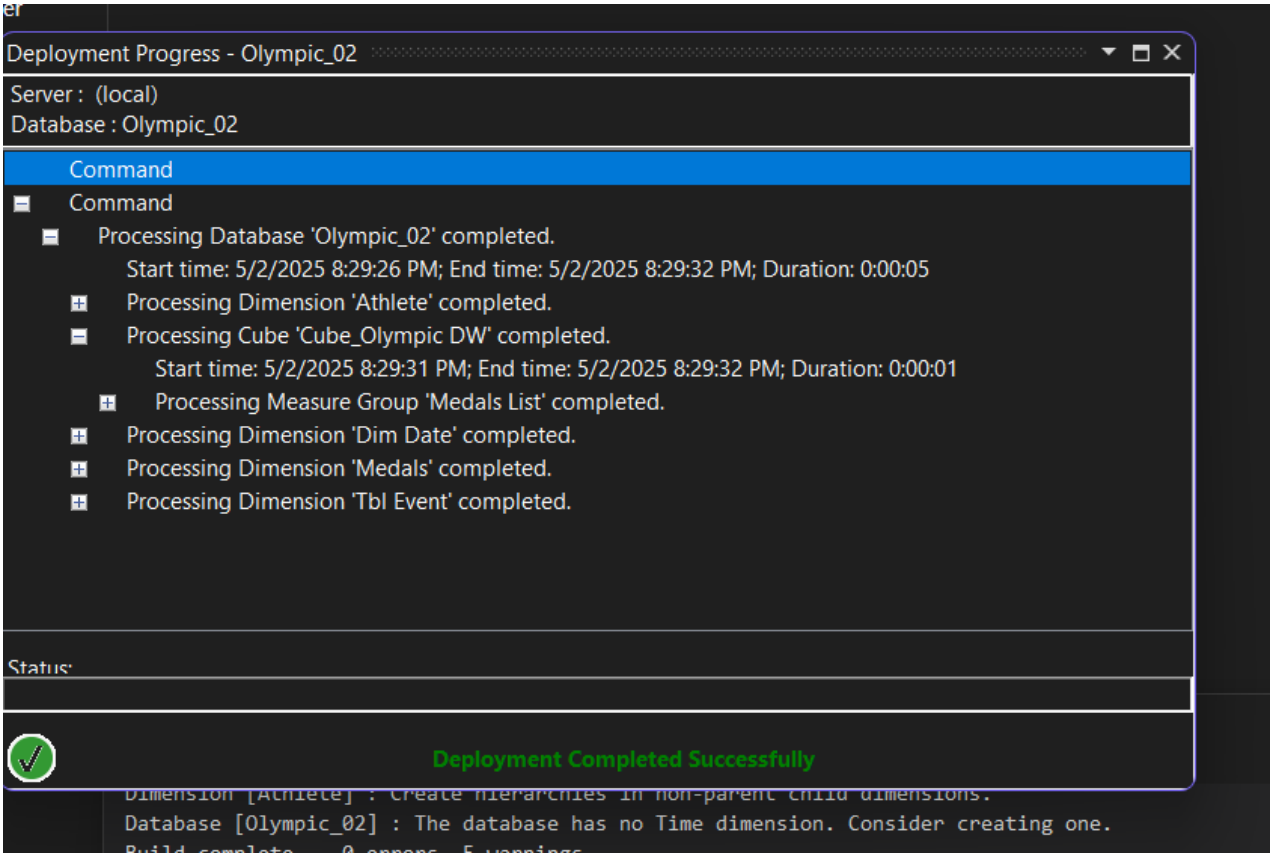
- Measures: Sum of MedalCount (named TotalMedals).
- Dimensions: Athlete, Country, Event, Year, MedalType, Discipline.
- Hierarchy: A Country to Continent hierarchy in the Country dimension, allowing aggregation from country to continent level

Folder structure after been deployed.



4. Deploy the Cube

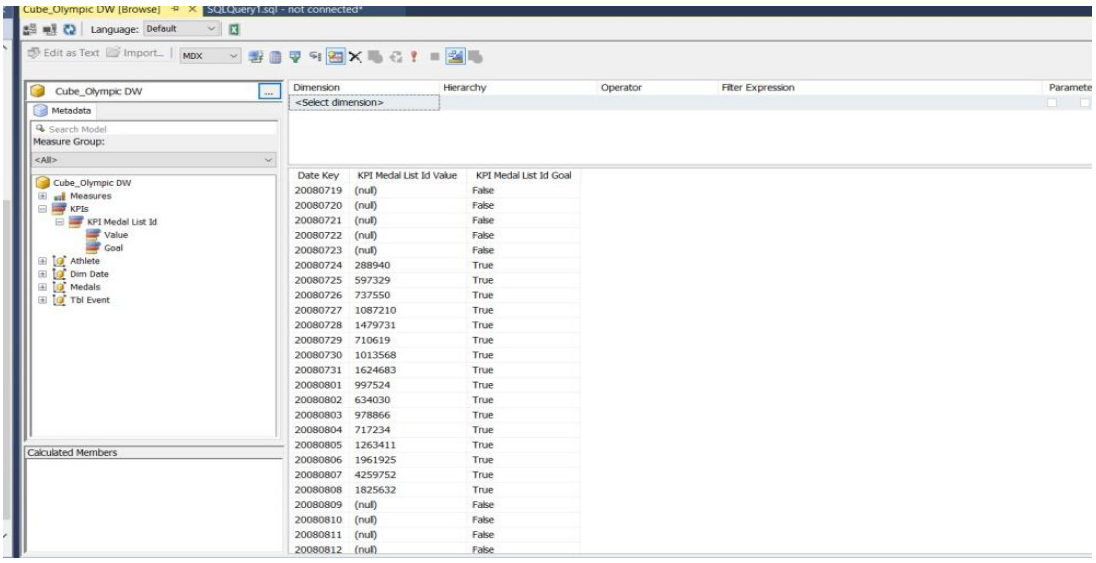
Finally, I deployed the cube to the SSAS server. In SSDT, I right-clicked the project and selected Deploy. The deployment was successful, as verified in SSMS.



6. Browse Cube Data

Browsing data is done by using the SSMS. By connecting SSAS to SSMS using instance and MDX queries can generated by selecting the relevant fields from the dimensions. When browsing cube data, a KPI value or measurement value is compulsory. Otherwise, it will not be executed.

Below figure shows how to browse data in SSMS.



Step 3 – Demonstration of OLAP Operations

This section demonstrates OLAP operations (Slice, Dice, Roll-up, Drill-down, Pivot) by connecting an Excel workbook to Cube_Olympic_DW and creating pivot tables and charts.

Tools Used

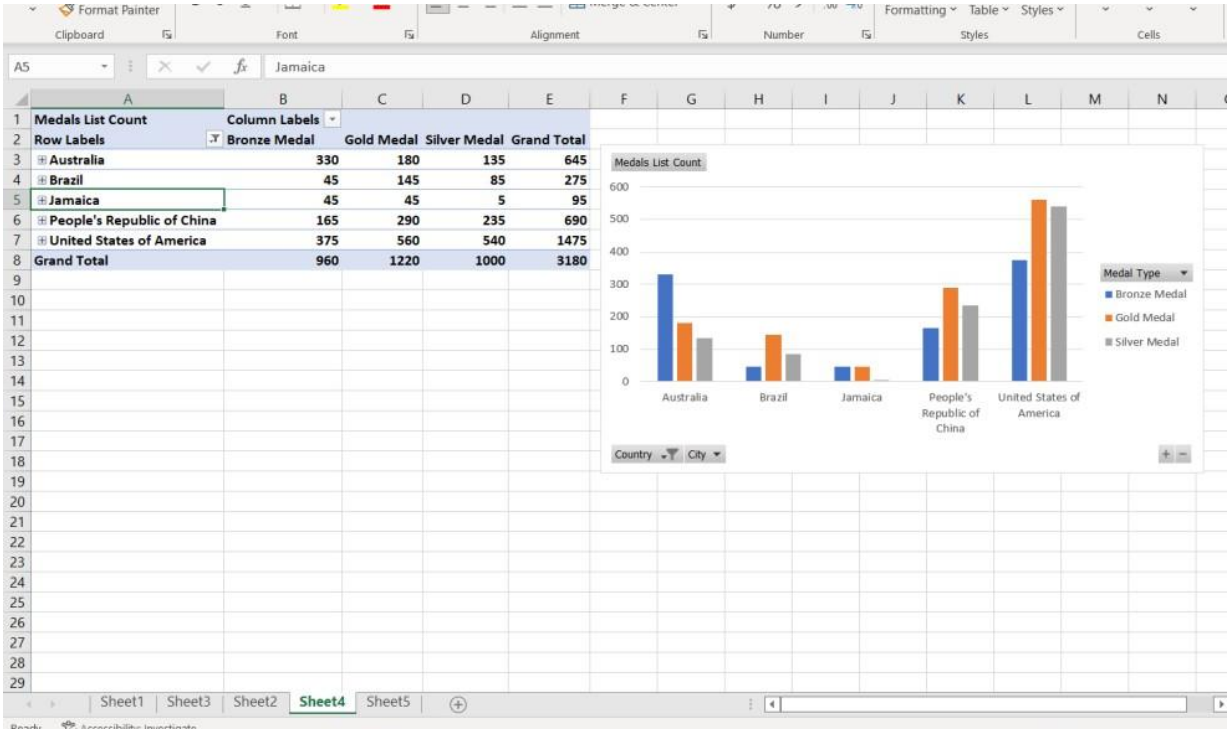
- Microsoft Excel
- SQL Server Management Studio (SSMS)
- SQL Server Analysis Services (SSAS)

Excel Connection

I connected Excel to the cube for data analysis. In Excel, I navigated to Data, selected Get Data, chose From Other Sources, and then From Analysis Services. I selected Cube_Olympic_DW from the SSAS server.

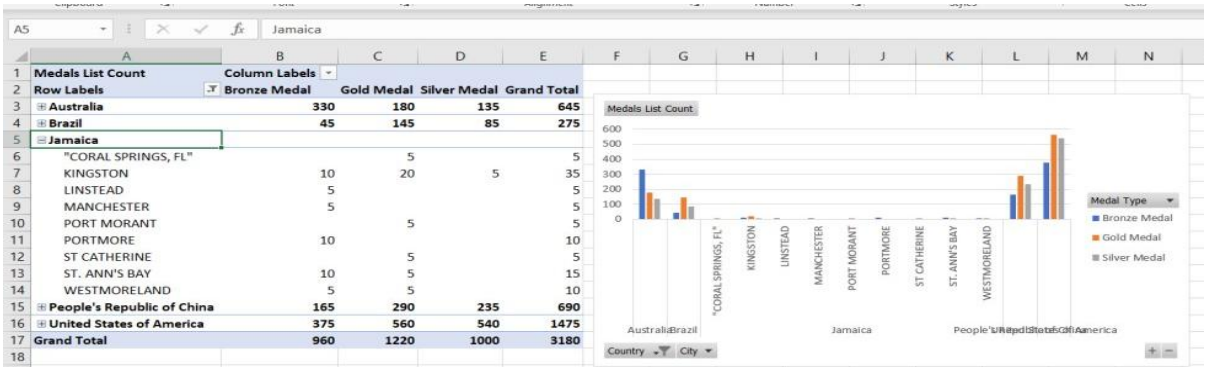
1. Roll-up

I aggregated data up a hierarchy. I created a pivot table with TotalMedals in Values and the Country to Continent hierarchy in Rows. This displayed medal counts aggregated at the Continent level.



2. Drill-down

I explored detailed data by navigating down a hierarchy. In the same pivot table, I expanded the Country to Continent hierarchy to show medal counts at the Country level within each Continent.

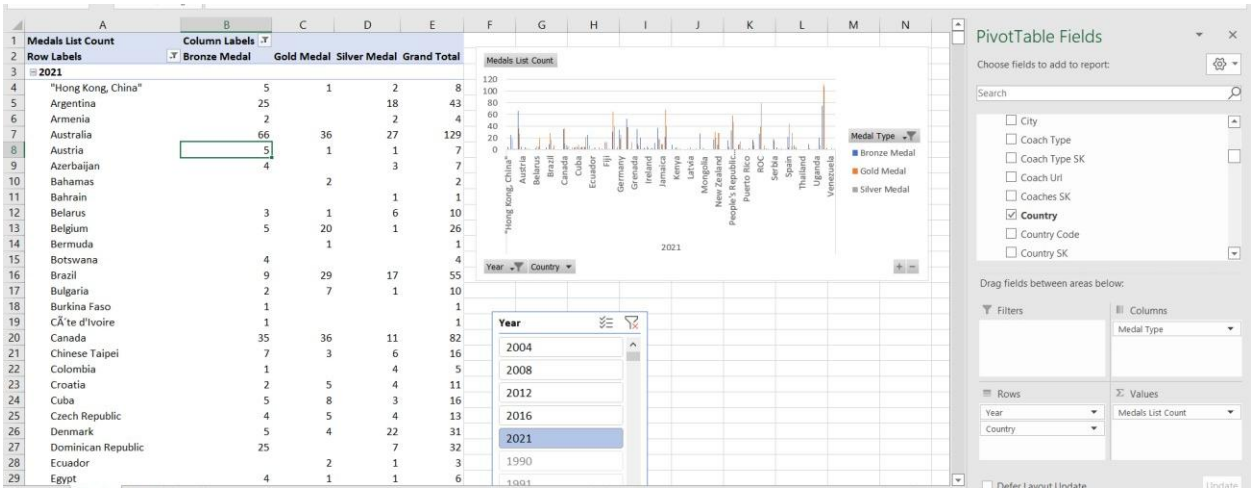


Medals List Count	Column Labels			
Row Labels	Bronze Medal	Gold Medal	Silver Medal	Grand Total
Australia	330	180	135	645
Brazil	45	145	85	275
Jamaica				
"CORAL SPRINGS, FL"		5		5
KINGSTON	10	20	5	35
LINSTEAD	5			5
MANCHESTER	5			5
PORT MORANT		5		5
PORTMORE	10			10
ST CATHERINE		5		5
ST. ANN'S BAY	10	5		15
WESTMORELAND	5	5		10
People's Republic of China	165	290	235	690
United States of America	375	560	540	1475
Grand Total	960	1220	1000	3180

The screenshot shows the pivot table with the hierarchy expanded to Countries.
The screenshot shows the pivot table with the Year slicer set to 2020.

3. Slice

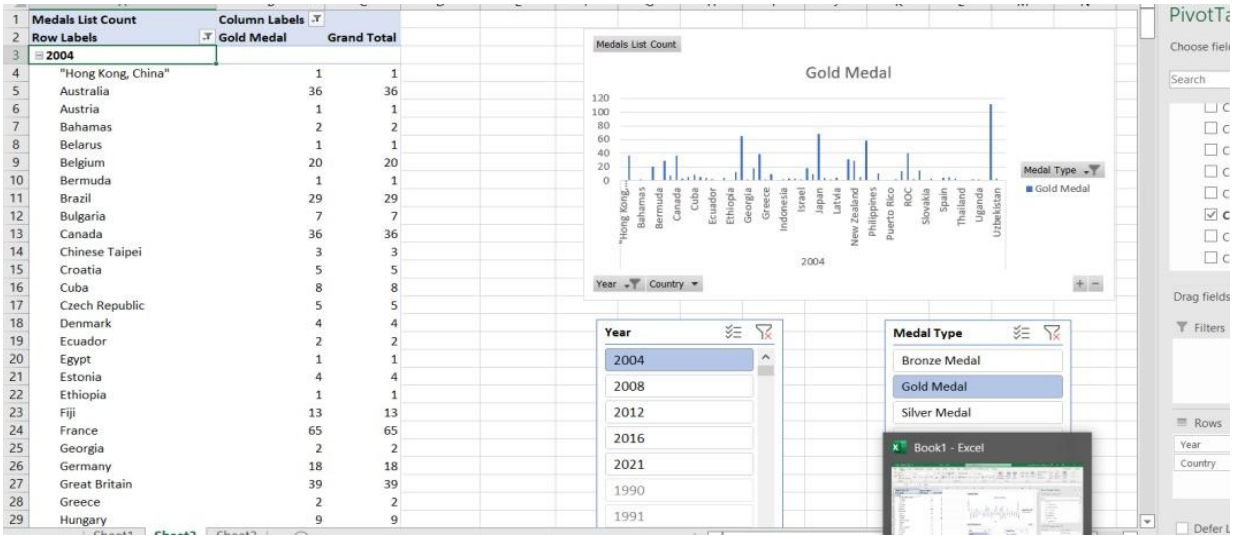
I used a slicer to filter data dynamically. I created a pivot table with TotalMedals in Values, Country in Rows, and Year as a slicer. Selecting a year, such as 2020, filtered the medal counts by country for that year.



Medals List Count	Column Labels			
Row Labels	Bronze Medal	Gold Medal	Silver Medal	Grand Total
2021				
"Hong Kong, China"	5	1	2	8
Argentina	25		18	43
Armenia	2		2	4
Australia	66	36	27	129
Austria	5	1	1	7
Azerbaijan	4		3	7
Bahamas		2		2
Bahrain			1	1
Belarus	3	1	6	10
Belgium	5	20	1	26
Bermuda		1		1
Botswana	4			4
Brazil	9	29	17	55
Bulgaria	2	7	1	10
Burkina Faso	1			1
CÃ´te d'Ivoire	1			1
Canada	35	36	11	82
Chinese Taipei	7	3	6	16
Colombia	1		4	5
Croatia	2	5	4	11
Cuba	5	8	3	16
Czech Republic	4	5	4	13
Denmark	5	4	22	31
Dominican Republic	25		7	32
Ecuador		2	1	3
Egypt	4	1	1	6

4. Dice

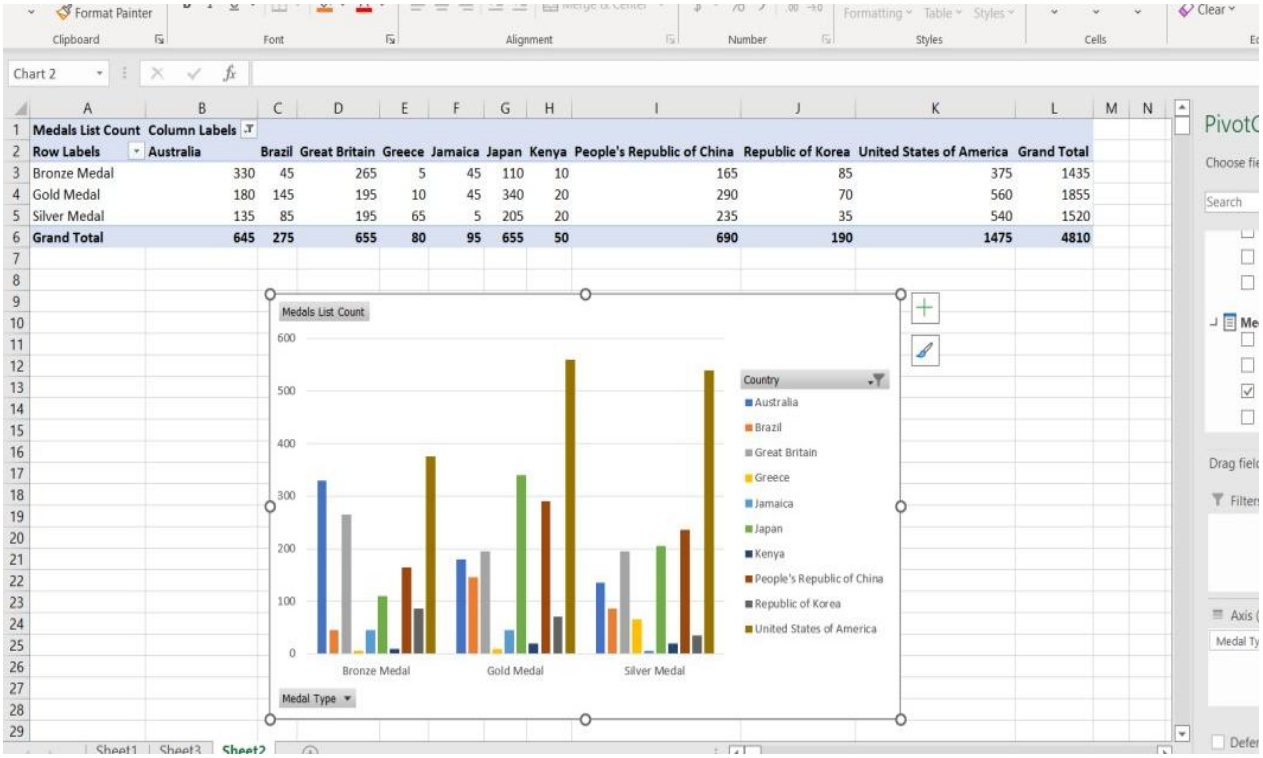
I used multiple slicers to group data. I created a pivot table with TotalMedals in Values, Country in Rows, and added Year and MedalType as slicers. I selected specific years (2016 and 2020) and medal types (Gold) to analyze targeted data.



The screenshot shows the pivot table with Year and MedalType slicers active.

5. Pivot

I created a pivot chart for data visualization. I built a pivot table with TotalMedals in Values, MedalType in Columns, and Country in Rows. I then generated a bar chart to show medal types by country.



Medals List Count	Column Labels										
Row Labels	Australia	Brazil	Great Britain	Greece	Jamaica	Japan	Kenya	People's Republic of China	Republic of Korea	United States of America	Grand Total
Bronze Medal	330	45	265	5	45	110	10	165	85	375	1435
Gold Medal	180	145	195	10	45	340	20	290	70	560	1855
Silver Medal	135	85	195	65	5	205	20	235	35	540	1520
Grand Total	**645**	**275**	**655**	**80**	**95**	**655**	**50**	**690**	**190**	**1475**	**4810**

The screenshot shows the bar chart visualizing TotalMedals by MedalType and Country.

Step 4 – PowerBI Reports

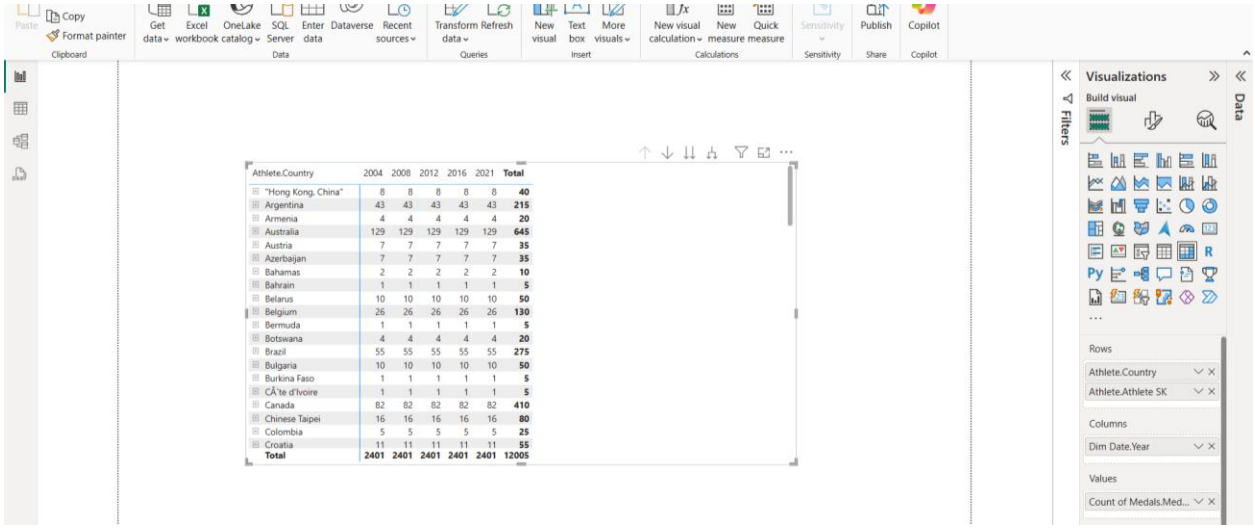
Tools Used

- PowerBI Desktop
- Power BI Service

Published Link: https://app.powerbi.com/links/aGQDdrv8xE?ctid=44e3cf94-19c9-4e32-96c3-14f5bf01391a&pbi_source=linkShare

Report 1: Report with a matrix

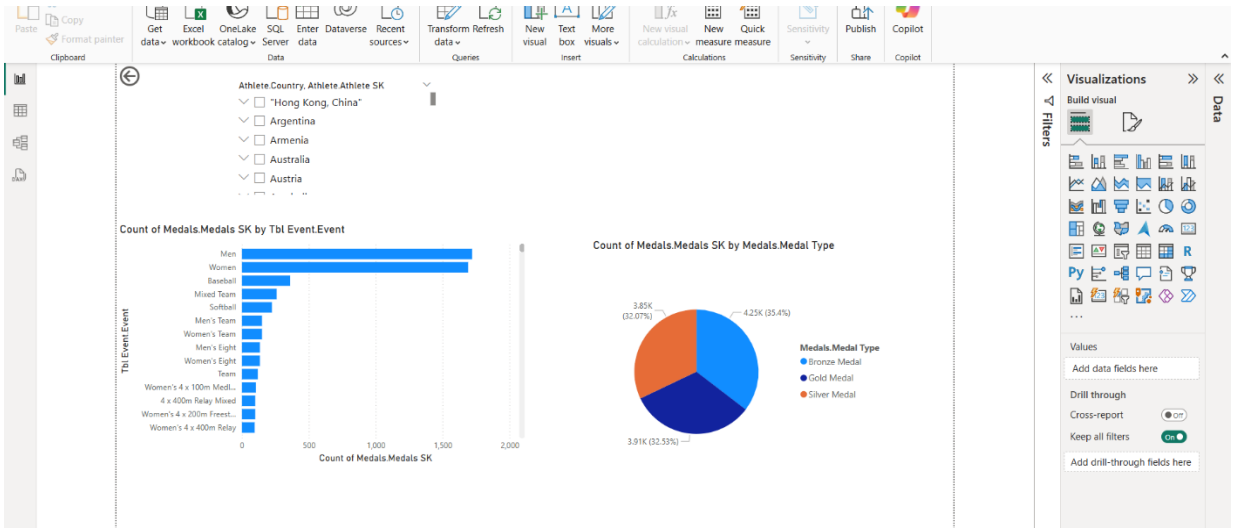
Created a Matrix visual with rows (Country Code, Athlete_Id), columns (Year), and values (TotalMedals). Enabled subtotals and grand totals. Formatted with bold headers and larger text.



Athlete.Country	2004	2008	2012	2016	2021	Total
"Hong Kong, China"	8	8	8	8	8	40
Argentina	43	43	43	43	43	215
Armenia	4	4	4	4	4	20
Australia	129	129	129	129	129	645
Austria	7	7	7	7	7	35
Azerbaijan	7	7	7	7	7	35
Bahamas	2	2	2	2	2	10
Bahrain	1	1	1	1	1	5
Belarus	10	10	10	10	10	50
Belgium	26	26	26	26	26	130
Bermuda	1	1	1	1	1	5
Botswana	4	4	4	4	4	20
Brazil	55	55	55	55	55	275
Bulgaria	10	10	10	10	10	50
Burkina Faso	1	1	1	1	1	5
CÃ´te d'Ivoire	1	1	1	1	1	5
Canada	82	82	82	82	82	410
Chinese Taipei	16	16	16	16	16	80
Colombia	5	5	5	5	5	25
Croatia	11	11	11	11	11	55
Total	**2401**	**2401**	**2401**	**2401**	**2401**	**12005**

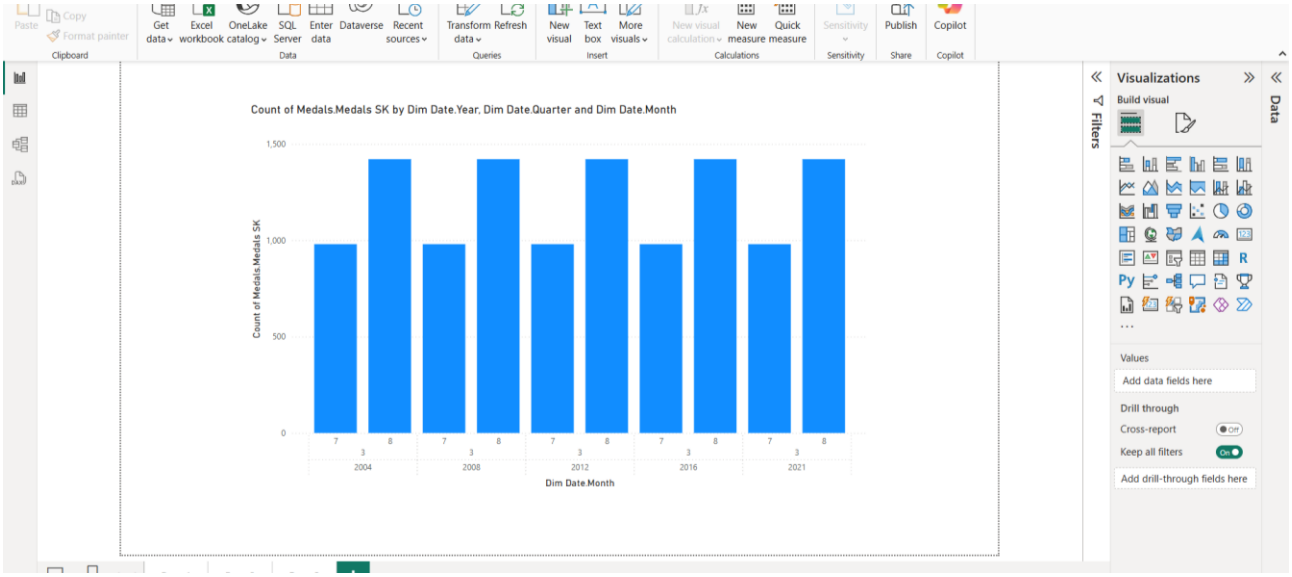
Report 2: Multiple Slicers with Cascading Filters

Added slicers for Country Code and Athlete_Id. Configured cascading by syncing filters in the Filters pane. Created Bar Chart (Event Name, TotalMedals) and Pie Chart (Medal_Type, TotalMedals). Formatted slicers as dropdowns.



Report 3: Drill-down Report

Created a Column Chart with DateHierarchy (Year > Quarter > Month) on Axis and TotalMedals on Values. Enabled drill-down for hierarchical exploration. Added a Table visual with Athlete_Id, Event, TotalMedals. Formatted with blue columns.



Report 4: Drill-Through Report

Created a Bar Chart (Country Code, TotalMedals) on the main page. Added a drill-through page with a Table (Athlete_Id, Event, Medal_Type, TotalMedals) and Country Code as the drill-through filter. Enabled drill-through navigation. Formatted with hidden detail page.