Sri Lanka Institute of Information Technology

IT3021 – Data Warehousing and Business Intelligence

Year 3 – Semester 01

Individual Project

Assignment 02

Submitted by:

Name - Liyanage S.R

Registration Number - IT20005726

Batch Number – Y3.S1.WD.DS.05.1.G1

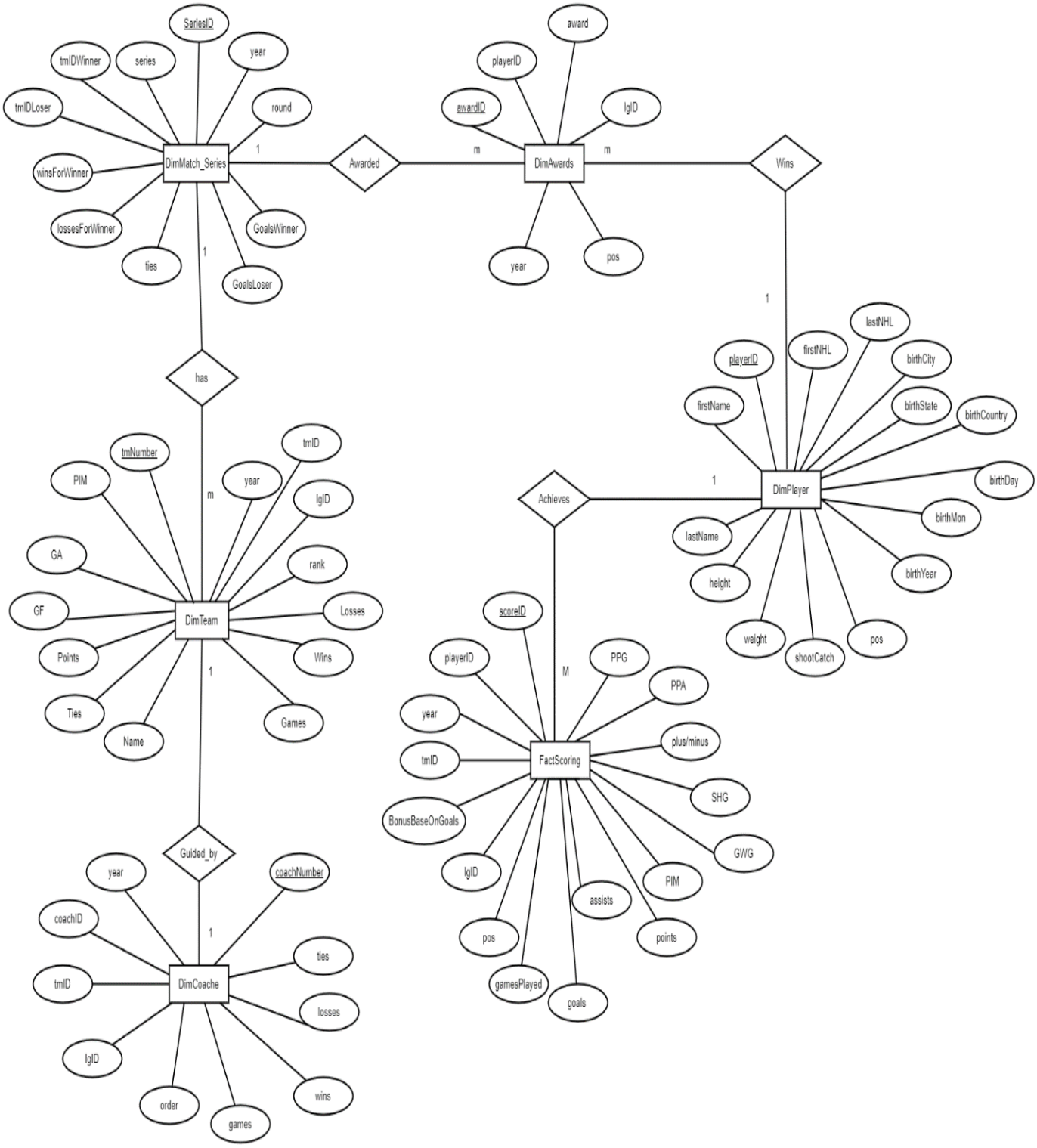
Step 01: Data source for the assignment 2

The information set I have chosen is about Professional Hockey Database. It carries facts on hockey players, teams, and coaches from 1909 to 2011. The Hockey Database is a group of statistics about Professional hockey teams in North the United States content. The data set has been chosen from Kaggle.com. The original dataset consists of a large number of information files which includes players' biological information, scoring, awards, and so forth. However, I selected 6 files out of them specifically master, AwardsPlayers, Scoring, Coaches, teams, and SeriesPost.

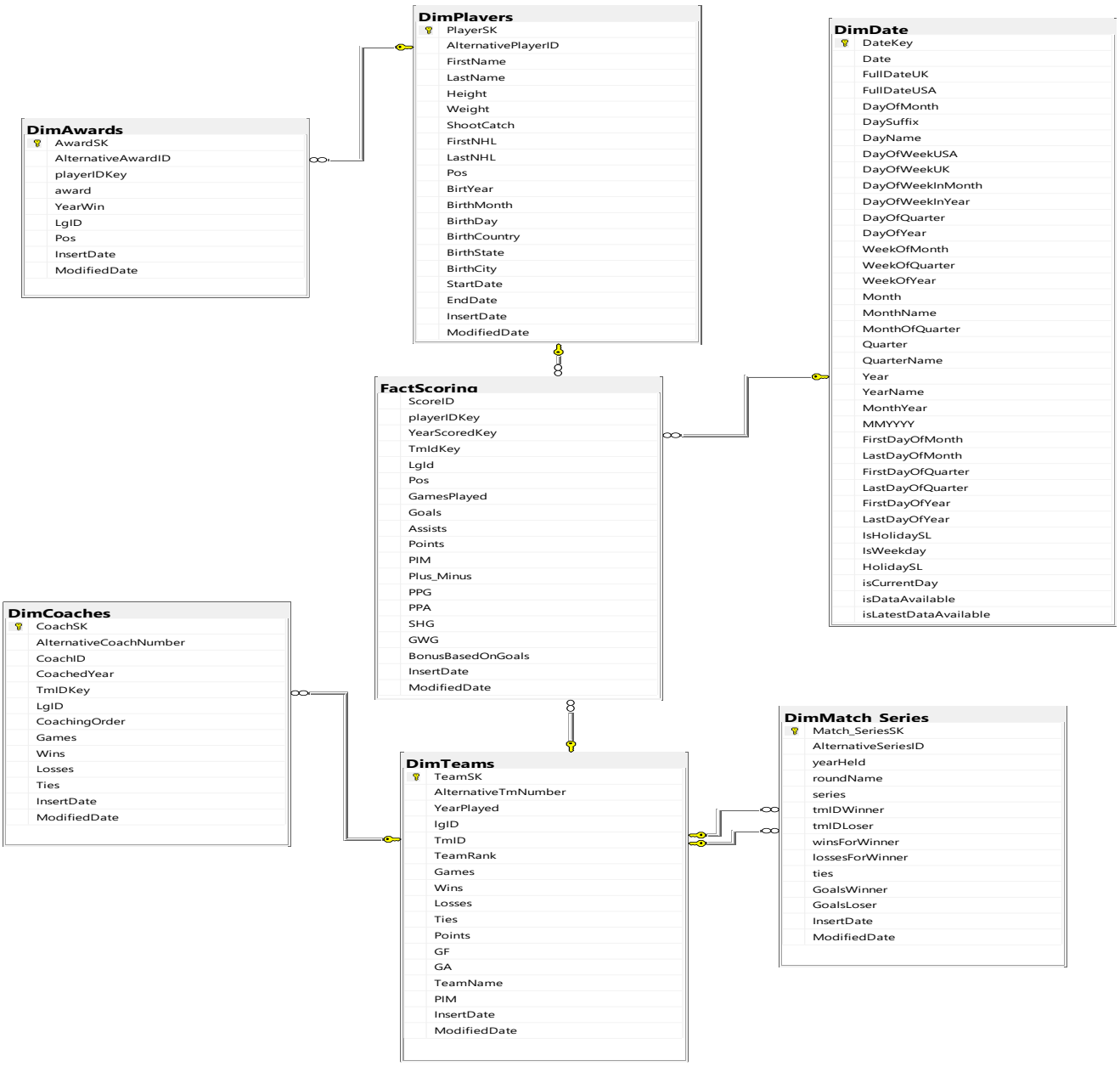
Then I adjusted those files based on my requirements. So there were 6 tables namely players, teams, awards, coaches, scoring, and finally match_series. Then I extracted and load those data to staging DB and thereafter I extracted, transformed, and load the data to the Data warehouse.

So as the source data for the second assignment the data which were loaded to the Data Warehouse were used. Accordingly, the data warehouse became the source of the second assignment.

ER Diagram :


SeriesID
tmIDWinner
series
year
tmIDLoser
round
DimMatch_Series
winsForWinner
lossesForWinner
ties
GoalsWinner
GoalsLoser
1
Awarded
m
award
playerID
awardID
lgID
DimAwards
year
pos
Wins
1
has
m
tmNumber
PIM
year
tmID
lgID
rank
GA
GF
DimTeam
Losses
Points
Wins
Ties
Games
Name
1
Guided_by
year
coachNumber
coachID
1
tmID
DimCoache
ties
losses
lgID
order
games
wins
Achieves
1
DimPlayer
lastNHL
firstNHL
playerID
birthCity
firstName
birthState
birthCountry
birthDay
birthMon
lastName
birthYear
height
weight
shootCatch
pos
scoreID
M
playerID
PPG
PPA
year
plus/minus
tmID
FactScoring
SHG
BonusBaseOnGoals
GWG
lgID
PIM
assists
pos
points
gamesPlayed
goals

The data warehouse was designed using the snowflake schema as a model.



Step 02: SSAS Cube implementation

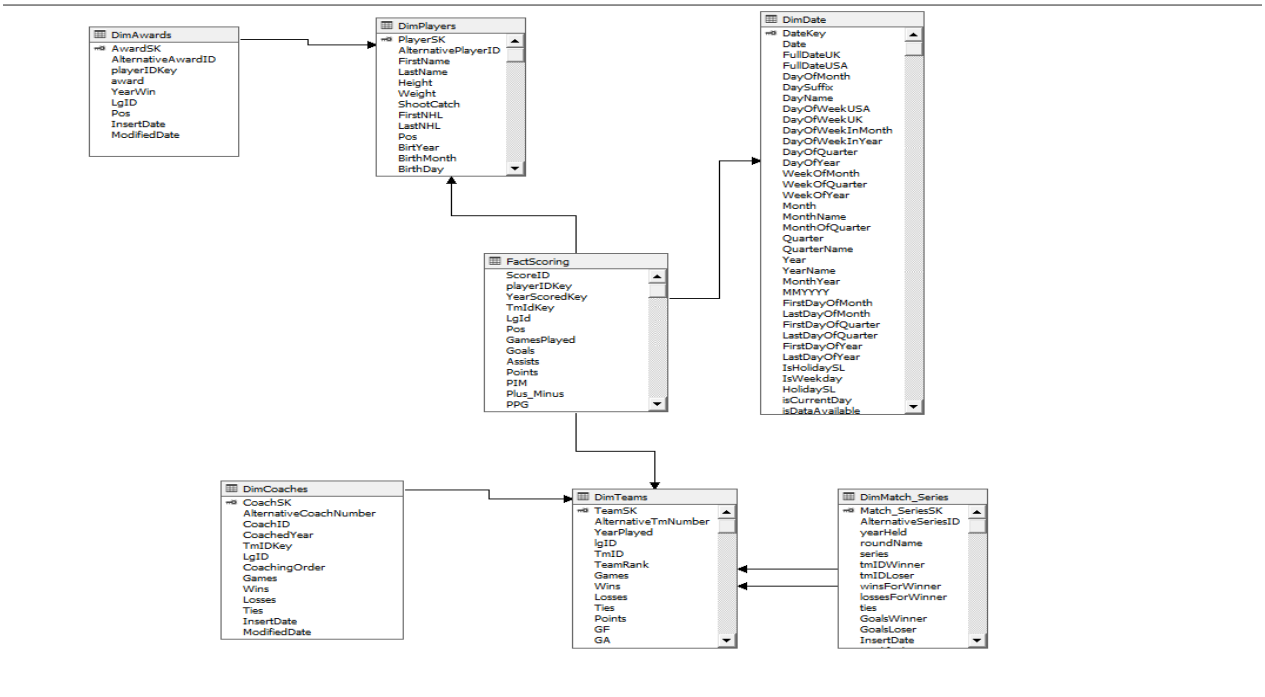
In this phase SSAS Cubes have developed.

Data Sources:

Here Data warehouse uses as the data source as described above. SQL Server authentication is used to connect with the Data warehouse. (Professional_Hockey_DW). The Professional_Hockey_ DW_Src.ds file which is under Data Sources manages the connection and extracts data from the Data Warehouse.

Data Source Views:

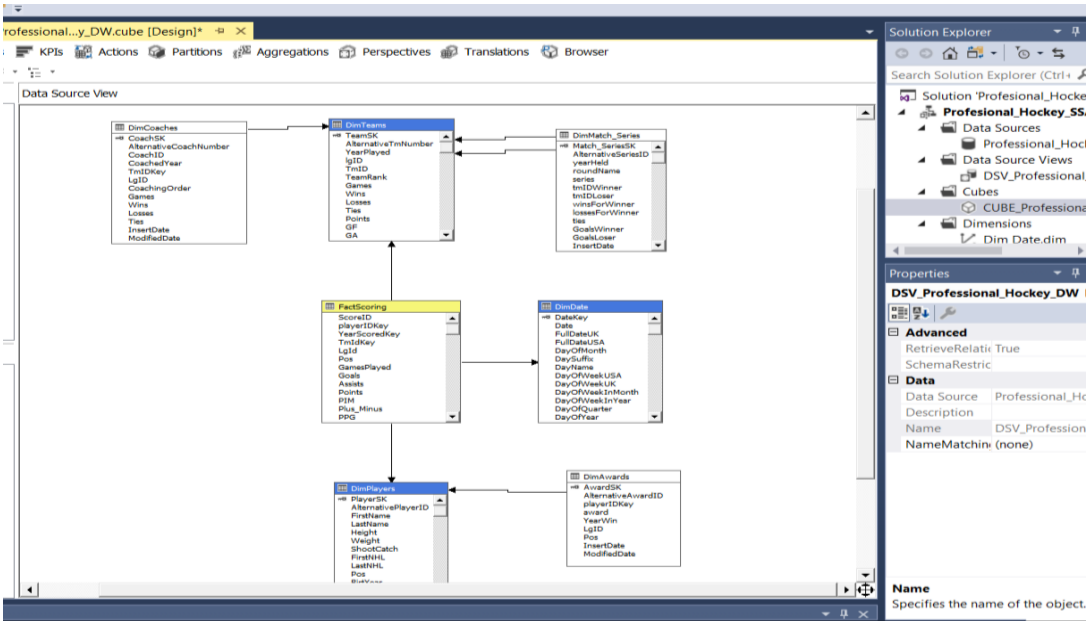
In the data source view, the overall view about the Professional_Hockey_DW has made while using the data source.



Cubes :

In this step, the cube has developed its name as CUBE_Professional_Hockey_DW. In that case, dimensions were created automatically when the cube was created and those dimensions were edited necessarily.

After creating the cube, the cube structure looked like this.



Then the dimensions were modified as below.

- Dim Players :

DimPlayers dimension was contain lots of detail about players' biological data and insert and modified dates about the time when data loaded to the Data warehouse. Among them, necessary data filed for analytical purposes are chosen as attributes.

Moreover, there were special attributes that can take as a hierarchy. Accordingly, three hierarchies were made.

01.HiearchyOfBirthPlace –

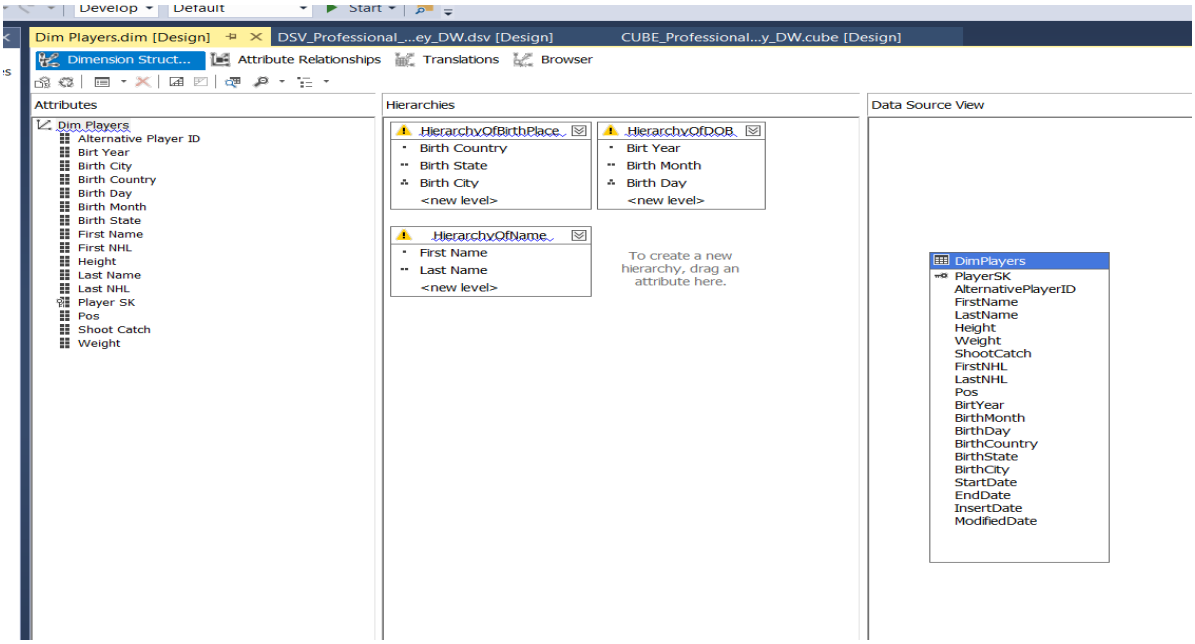
- About the birthplace of players.
- Birth Country, Birth State, and Birth City

02.HiearachyOfDOB

- About the Date of birth of players.
- Birth Year, Birth Month, and Birthday

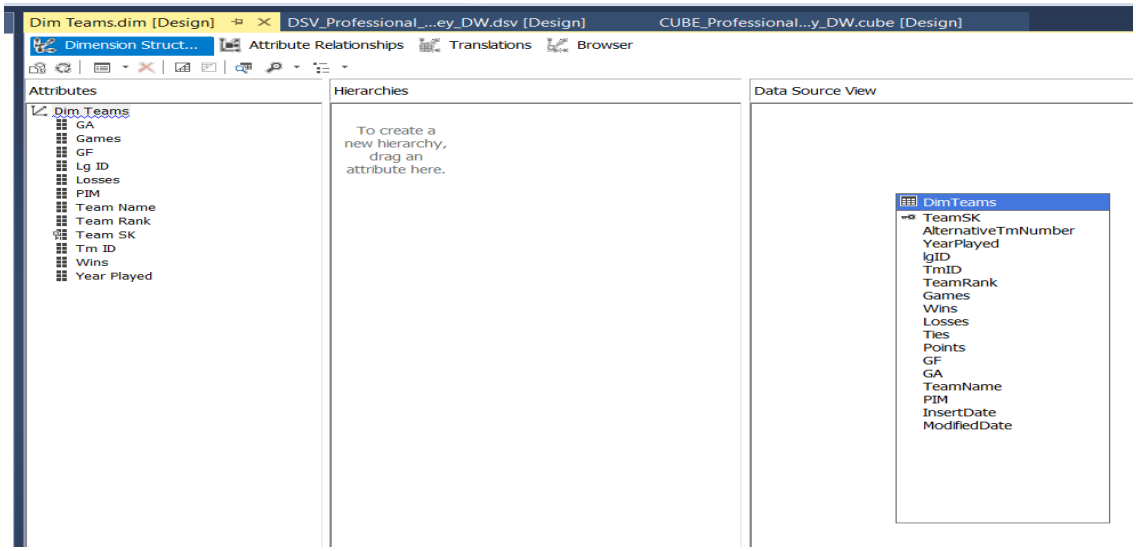
03. HiearchyOfName

- About the full name of players.
- First name and Last Name



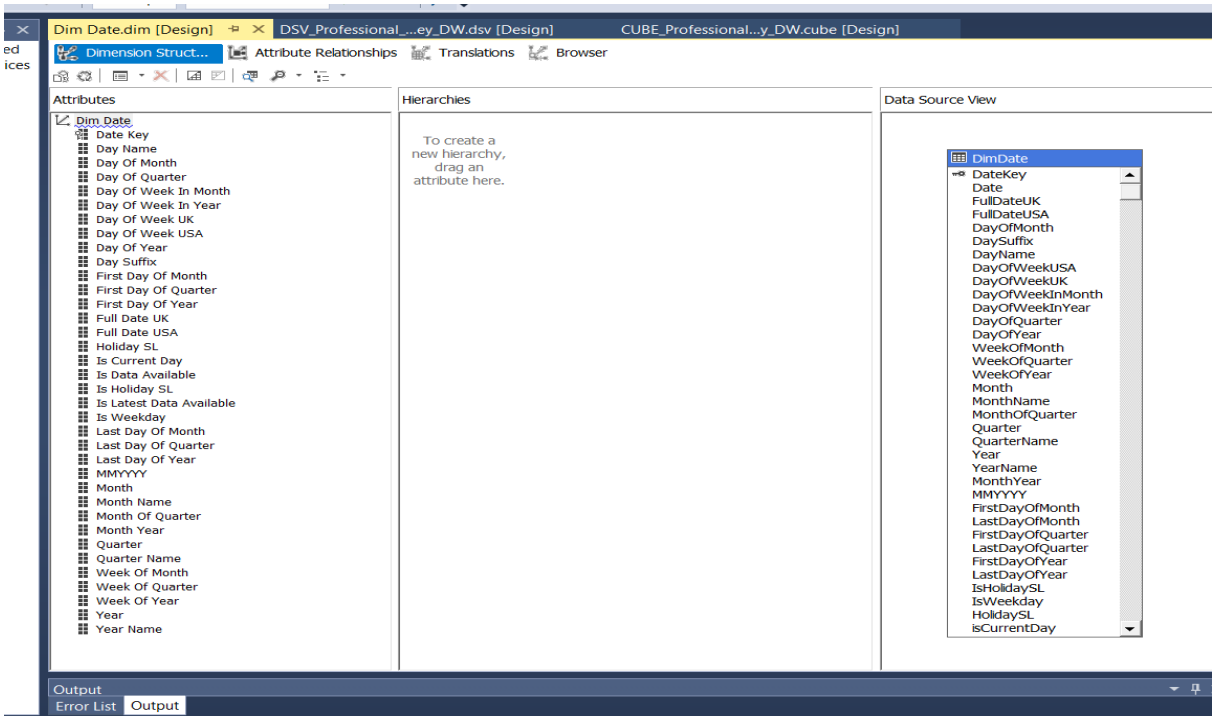
- Dim Teams :

Teams dimension contained details about the teams which were played in the match series. In that case, some columns were helpful for analytical purposes. Those useful columns have been taken as attributes.



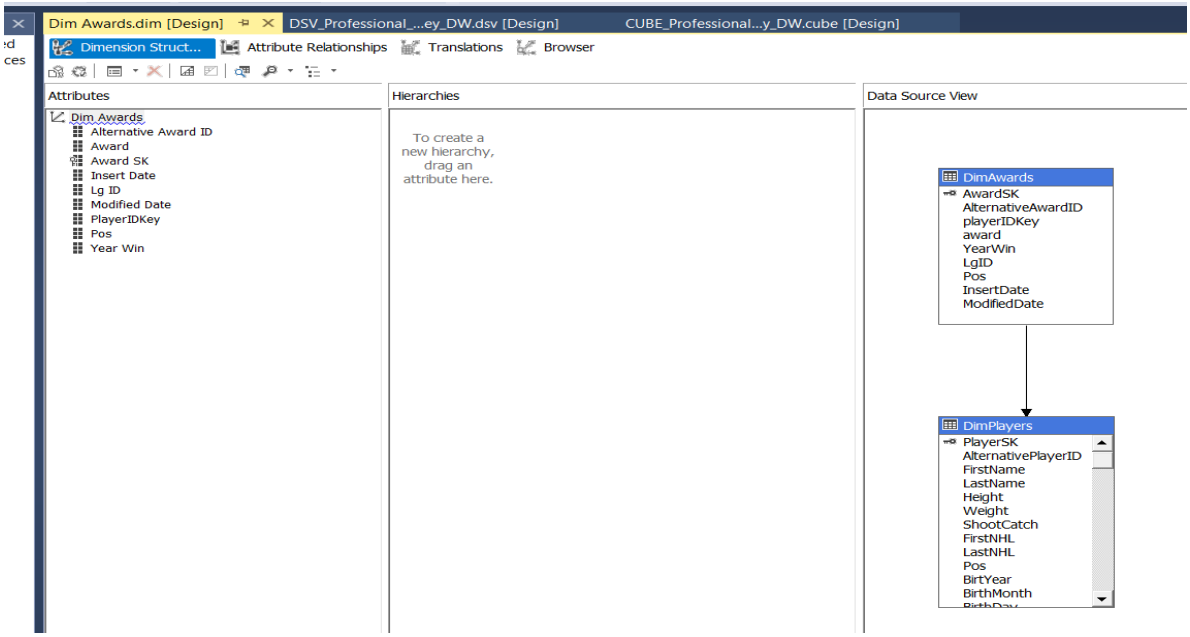
➢ Dim Date :

Date dimension had varieties of dates, months, years, and many more. Since those data can be useful for analytics all of those data have taken as attributes.



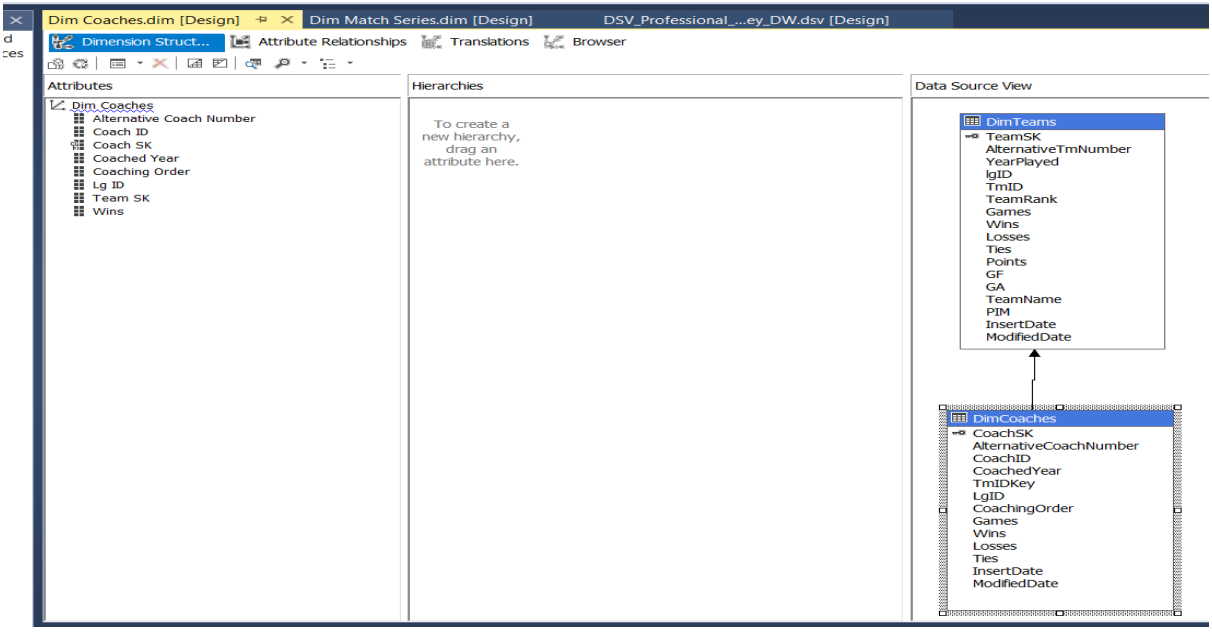
➢ Dim Awards :

The awards dimension contains details about the awards which were won by players. Here also all necessary attributes were chosen to do better analytics about the data.



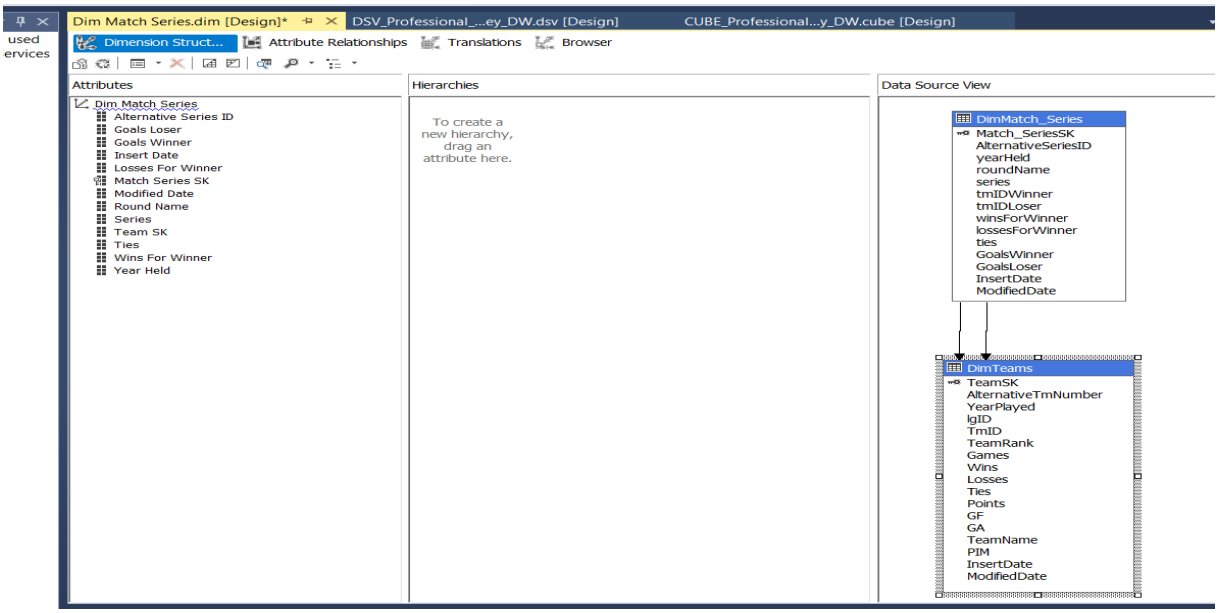
➢ Dim Coaches :

Coaches' dimensions were about the coaches' details who guided the teams. In this step also, the most appropriate details for analytics were selected as attributes.



➢ Dim Match_Series :

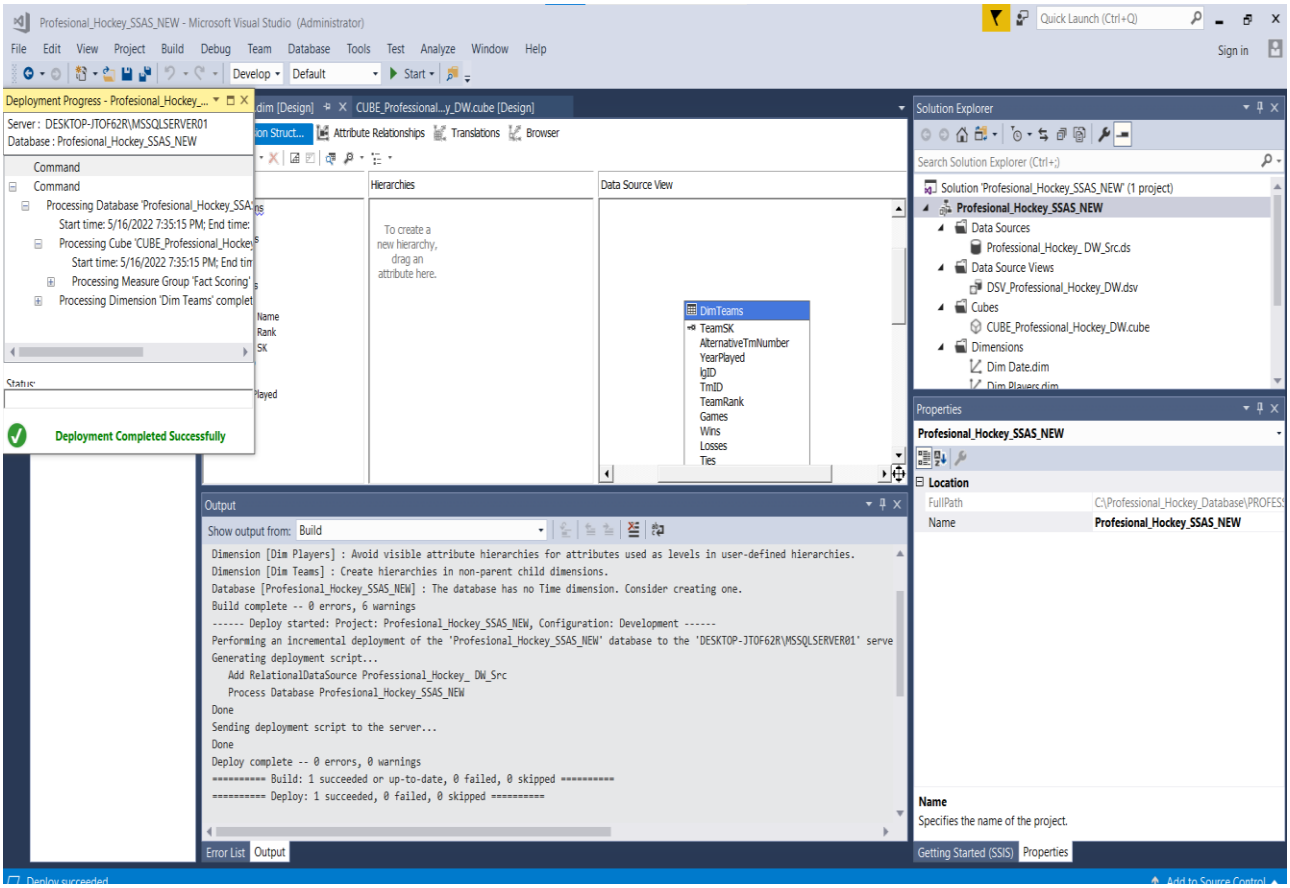
The Match series dimension was about the match series in which those teams were played and awards awarded to players. In this case, more useful data about match series were chosen as attributes.

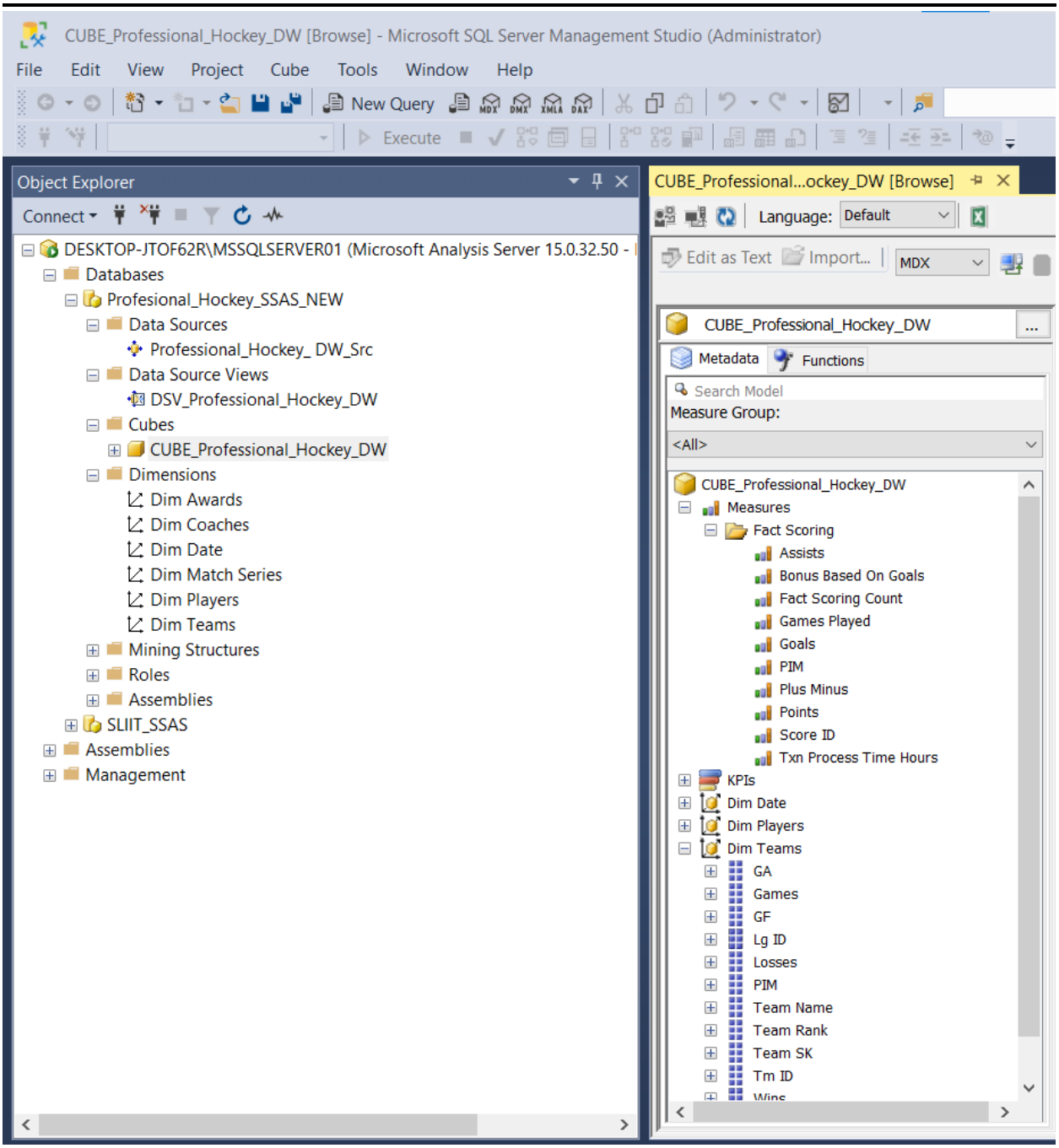


- Fact Scoring :

FactScoring table which was in the Data warehouse was chosen as a measure group to do better analytics about the data while including calculations and aggregations. In Fact Scoring Measure, the most suitable data chose as measures.

Finally the cube was deployed successfully.

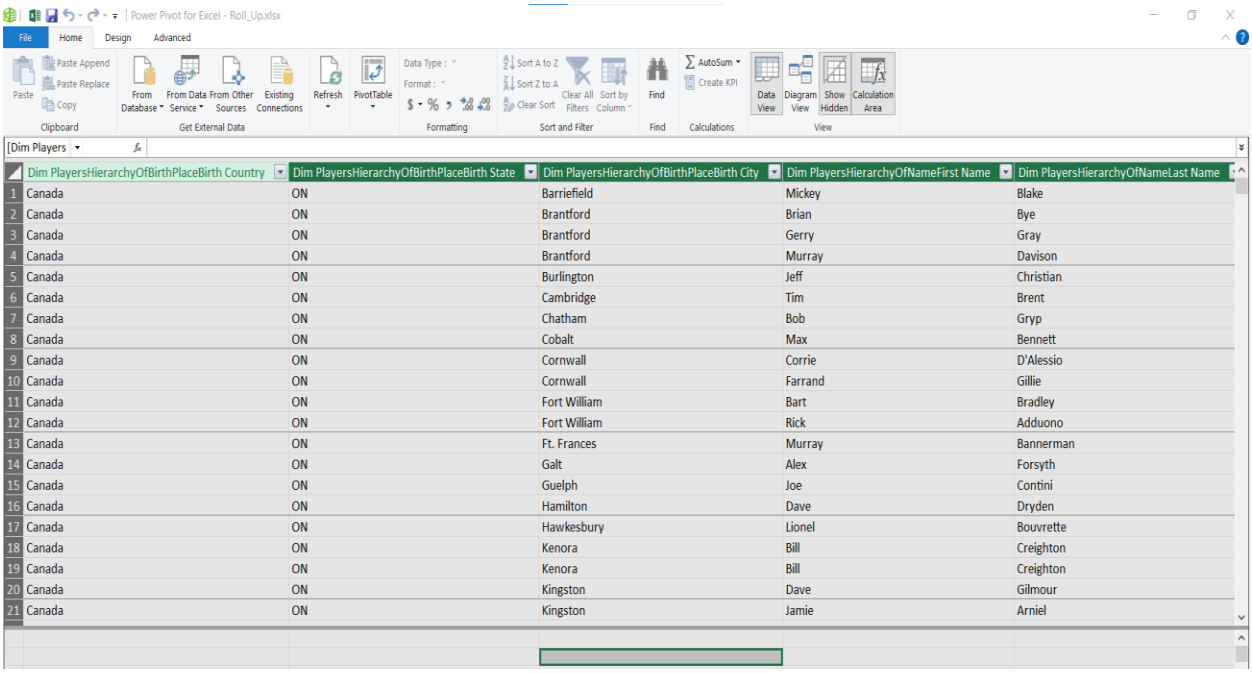




After the deployment the relevant Database created in SSMS under analysis services Databases correctly as below.

Step 03: Demonstration of OLAP operations

In this step I have connected the Excel workbook to the cube using POWERPIVOT mode. Then the excel successfully connected to the analysis services' database (Professional_Hockey_SSAS_NEW) and the MDX query runs and validated. Then it returns data successfully for each OLAP operations.

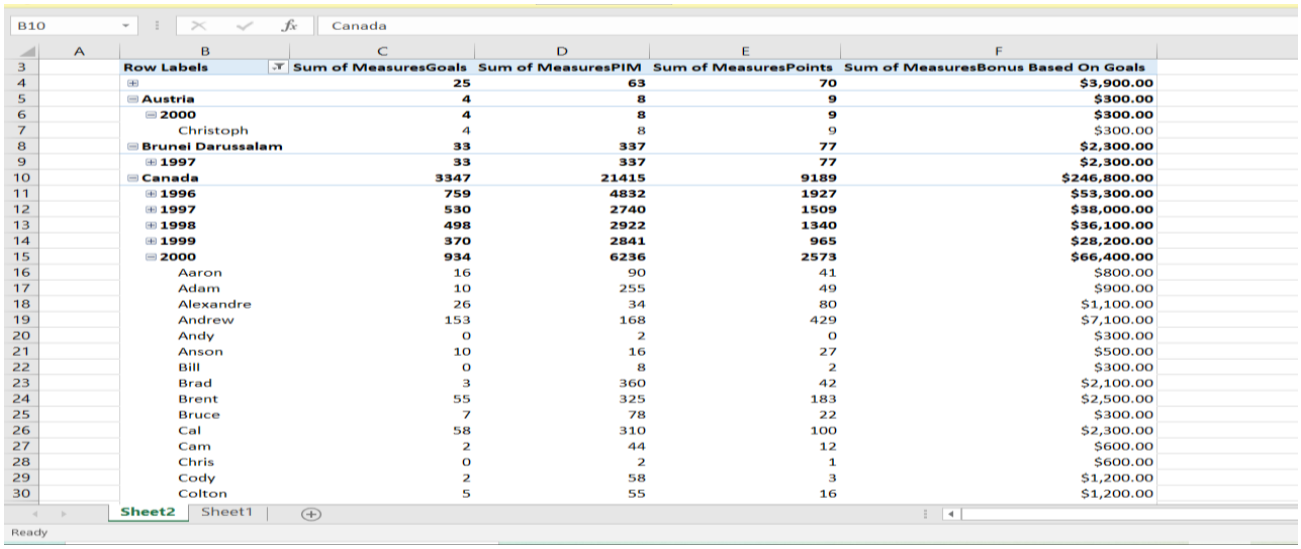


	Dim PlayersHierarchyOfBirthPlaceBirth Country	Dim PlayersHierarchyOfBirthPlaceBirth State	Dim PlayersHierarchyOfBirthPlaceBirth City	Dim PlayersHierarchyOfNameFirst Name	Dim PlayersHierarchyOfNameLast Name
1	Canada	ON	Barriefield	Mickey	Blake
2	Canada	ON	Brantford	Brian	Bye
3	Canada	ON	Brantford	Gerry	Gray
4	Canada	ON	Brantford	Murray	Davison
5	Canada	ON	Burlington	Jeff	Christian
6	Canada	ON	Cambridge	Tim	Brent
7	Canada	ON	Chatham	Bob	Gryp
8	Canada	ON	Cobalt	Max	Bennett
9	Canada	ON	Cornwall	Corrie	D'Alessio
10	Canada	ON	Cornwall	Farrand	Gillie
11	Canada	ON	Fort William	Bart	Bradley
12	Canada	ON	Fort William	Rick	Adduono
13	Canada	ON	Ft. Frances	Murray	Bannerman
14	Canada	ON	Galt	Alex	Forsyth
15	Canada	ON	Guelph	Joe	Contini
16	Canada	ON	Hamilton	Dave	Dryden
17	Canada	ON	Hawkesbury	Lionel	Bouvrette
18	Canada	ON	Kenora	Bill	Creighton
19	Canada	ON	Kenora	Bill	Creighton
20	Canada	ON	Kingston	Dave	Gilmour
21	Canada	ON	Kingston	Jamie	Arniel

01. Roll-Up Operation

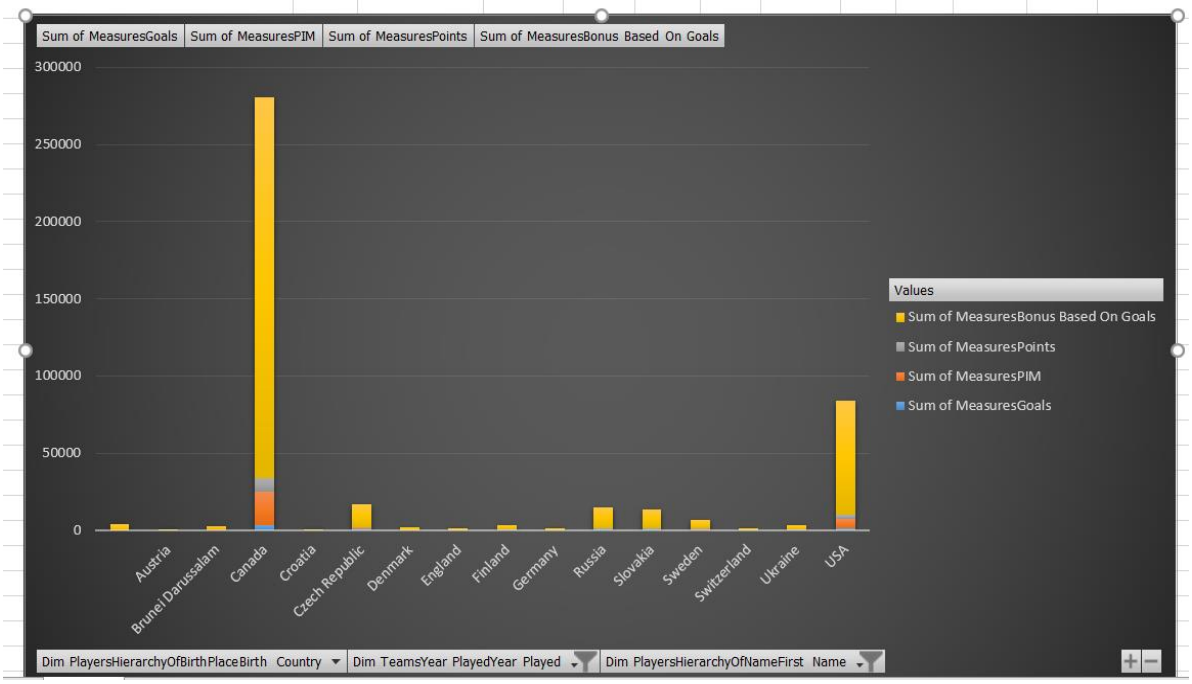
The roll-up operation describes the country-wise players' statistics about the goals, points, penalty minutes (PIM), and bonuses they got based on their goals (1995-2011). Since roll-up refers to the process of viewing data with decreasing detail and calculating totals for a hierarchy, I have chosen a higher level of hierarchy here.

In this case, I chose 'HiearchyOfBirthPlace' and the country name which is the higher level of that hierarchy.



Row Labels	Sum of MeasuresGoals	Sum of MeasuresPIM	Sum of MeasuresPoints	Sum of MeasuresBonus Based On Goals
	25	63	70	$3,900.00
Austria	4	8	9	$300.00
2000	4	8	9	$300.00
Christoph	4	8	9	$300.00
Brunei Darussalam	33	337	77	$2,300.00
1997	33	337	77	$2,300.00
Canada	3347	21415	9189	$246,800.00
1996	759	4832	1927	$53,300.00
1997	530	2740	1509	$38,000.00
1998	498	2922	1340	$36,100.00
1999	370	2841	965	$28,200.00
2000	934	6236	2573	$66,400.00
Aaron	16	90	41	$800.00
Adam	10	255	49	$900.00
Alexandre	26	34	80	$1,100.00
Andrew	153	168	429	$7,100.00
Andy	0	2	0	$300.00
Anson	10	16	27	$500.00
Bill	0	8	2	$300.00
Brad	3	360	42	$2,100.00
Brent	55	325	183	$2,500.00
Bruce	7	78	22	$300.00
Cal	58	310	100	$2,300.00
Cam	2	44	12	$600.00
Chris	0	2	1	$600.00
Cody	2	58	3	$1,200.00
Colton	5	55	16	$1,200.00

Further roll-up data is presented as a Staked Column chart as well.



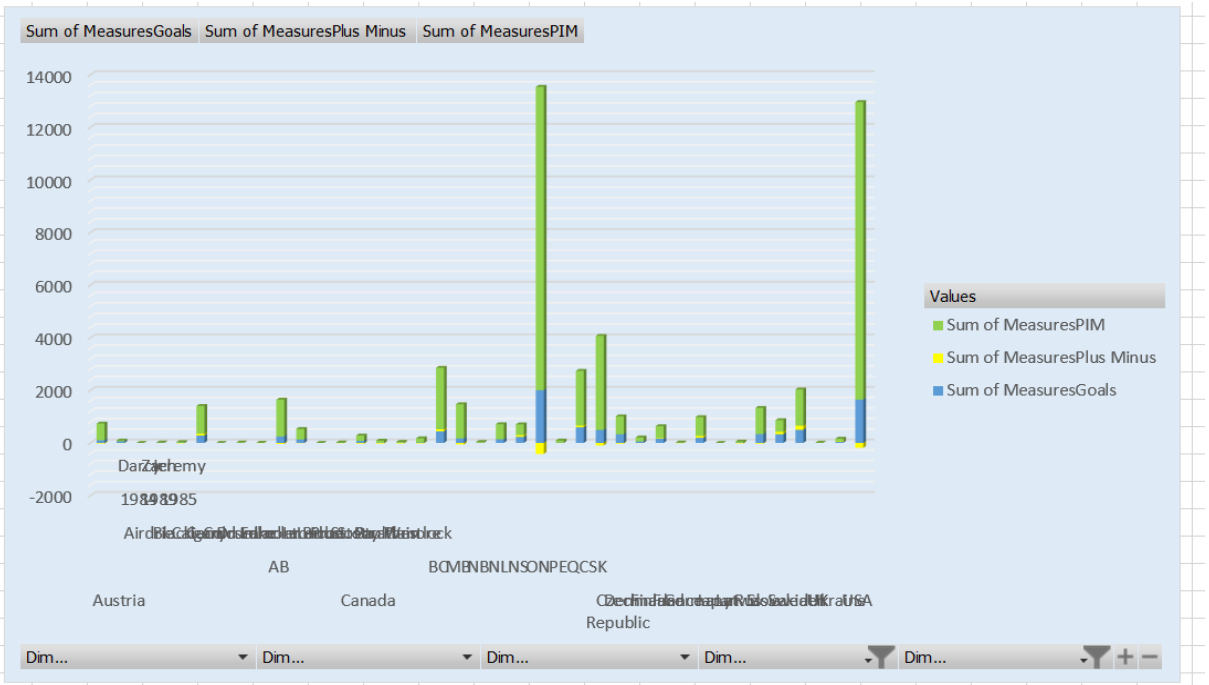
02. Drill down Operation

Here, in the drill_down operation, it describes the birth country, birth state, and birth city-wise Players' details along with their first name and birth years. It gives more details about players' scores including goals, plus-minus points, and penalty minutes (PIM).

In that case, since drill-down describes more details about the data which are in hierarchical order I have chosen 'HiearchyOfBirthPlace' and included players' data according to their birth country, birth state, and birth city-wise respectively.

Row Labels	Sum of MeasuresGoals	Sum of MeasuresPlus Minus	Sum of MeasuresPIM
	100	**-25**	**634**
Austria	**59**	**-3**	**30**
Canada	**4818**	**-434**	**25299**
AB	**749**	**-77**	**3394**
Airdrie	7	-4	6
1984	**0**	**0**	**0**
Darcy	0	0	0
1989	**7**	**-4**	**6**
Zach	7	-4	6
Blackie	3	-12	26
1985	**3**	**-12**	**26**
Jeremy	3	-12	26
Calgary	277	73	1052
Camrose	0	0	0
Cold Lake	0	0	12
Drumheller	0	-2	2
Edmonton	245	-36	1397
Lac La Biche	126	-6	404
Leduc	0	0	0
Provost	1	-6	13
St. Paul	73	-28	204
Stony Plain	2	-21	83
Strathmore	2	-19	34
Westlock	13	-16	161
BC	**439**	**79**	**2334**
MB	**166**	**-58**	**1303**

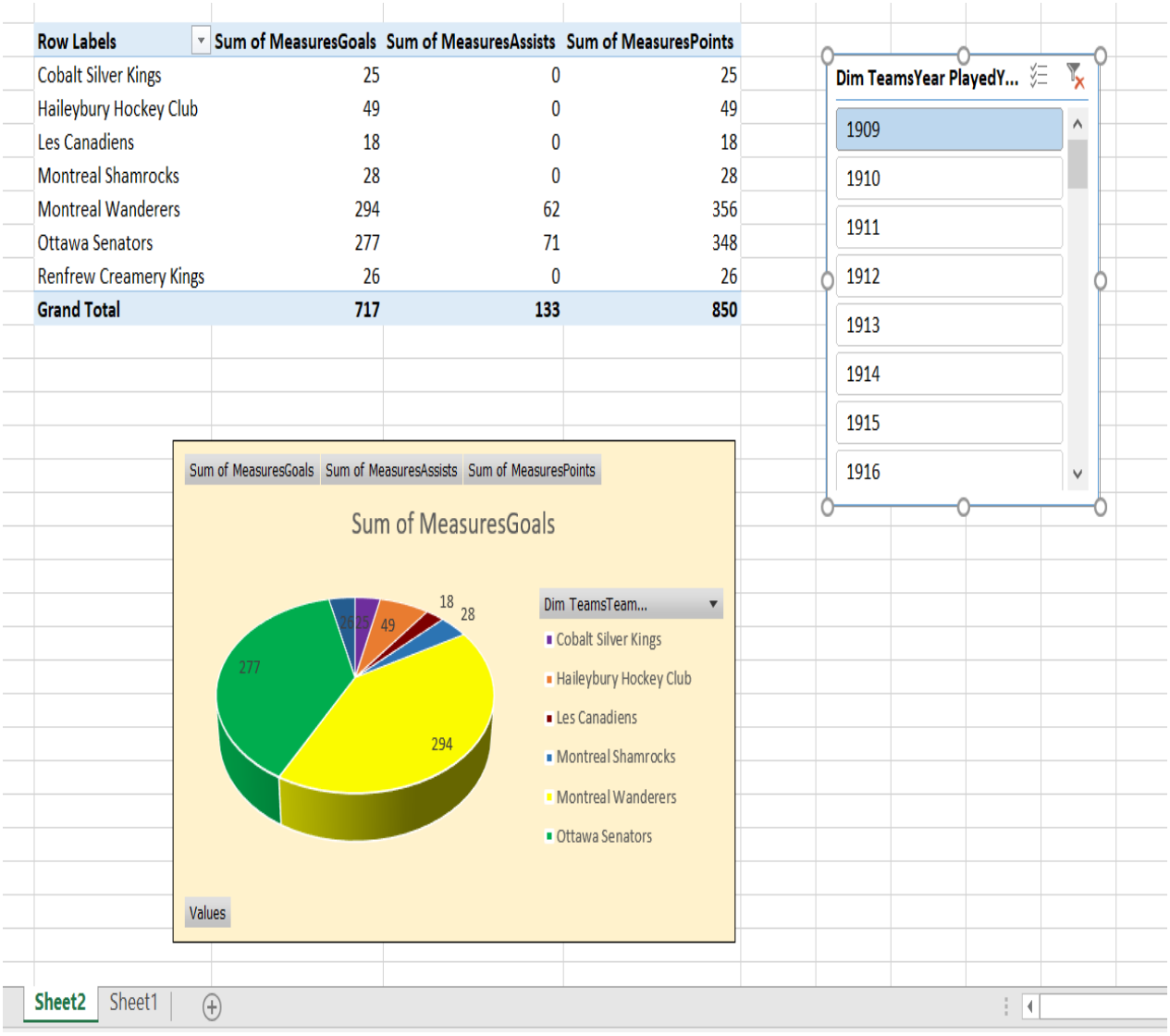
Moreover, drill-down data is presented as 3d- stacked column chart as well.



03. Slice Operation

The Slice operation describes the Teams name along with the scores those teams achieved namely, goals, assists, and points based on year. Here since, a Slice represents the two-dimensional view of an OLAP Cube, I have chosen a slicer. So I chose the year in which those matches were played by those teams.

Additionally, I have represented the Slice operation as a pie chart as well.



04. Dice Operation

The dice operation describes the Teams' names and Players' names who were played under those teams along with scoring statistics namely the sum of goals, assists, points, and average of Bonuses given based on goals.

Since dice describe the data set in two or more dimensions, I selected the Year in which the match was played as those teams and the players' country as slicers.

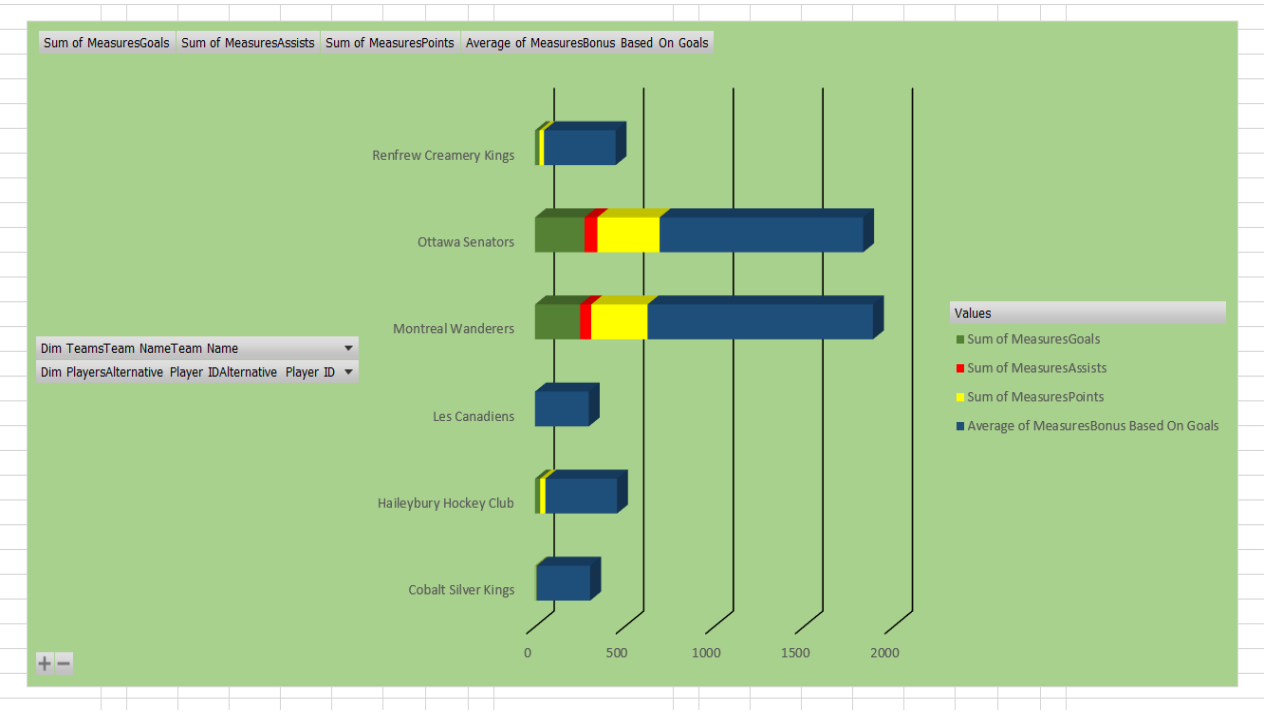
Row Labels	Sum of MeasuresGoals	Sum of MeasuresAssists	Sum of MeasuresPoints	Average of MeasuresBonus Based On Goals
Cobalt Silver Kings	**4**	**0**	**4**	**$300.00**
campban01	4	0	4	$300.00
Haileybury Hockey Club	**29**	**0**	**29**	**$400.00**
bawlfni01	10	0	10	$500.00
berlilo01	2	0	2	$300.00
currial01	14	0	14	$500.00
deyed01	3	0	3	$300.00
Les Canadiens	**0**	**0**	**0**	**$300.00**
cattajo01	0	0	0	$300.00
Montreal Wanderers	**252**	**62**	**314**	**$1,258.33**
atkinar01	1	2	3	$600.00
bakerbr01	21	4	25	$1,100.00
bawlfni01	1	2	3	$300.00
bellbi01	19	3	22	$1,100.00
camerha01	1	1	2	$300.00
chipcbi01	1	0	1	$300.00
cleghod01	114	23	137	$5,800.00
cleghsp01	79	27	106	$3,600.00
currial01	0	0	0	$300.00
doherfr01	0	0	0	$300.00
gardnji01	15	0	15	$800.00
haguebi01	0	0	0	$600.00
Ottawa Senators	**277**	**71**	**348**	**$1,135.71**
Renfrew Creamery Kings	**25**	**0**	**25**	**$400.00**
Grand Total	**587**	**133**	**720**	**$951.35**

Dim PlayersHierar...: Canada, England, Scotland, Austria, Belarus, Belgium, Brazil, Brunei Darussalam

Dim TeamsYear Pl...: 1909, 1910, 1911, 1912, 1913, 1914, 1915, 1916

Sheet2 | Sheet1

Afterward, dice data is presented as a 3d_stacked bar as well.



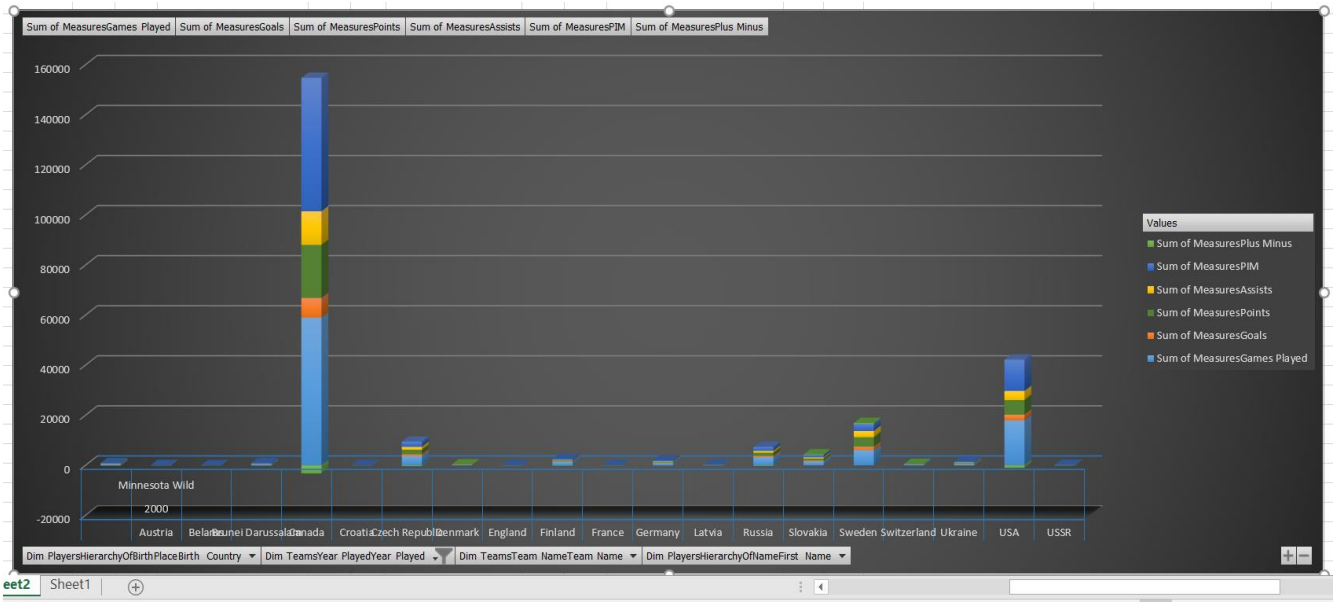
05. Pivot

The pivot table describes the country-wise players' data according to the teams' names which they were played under and the year match was held. Further, it contains scoring statistics such as games played, goals, points, assists, penalty minutes (PIM), and plus-minus. Since the pivot table gives calculated, summarize data a complex data set has been chosen between 1990 to 2011.

Row Labels	Sum of MeasuresGames Played	Sum of MeasuresGoals	Sum of MeasuresPoints	Sum of MeasuresAssists	Sum of MeasuresPIM	Sum of MeasuresPlus Minus
⊞	**431**	**38**	**113**	**75**	**212**	**-43**
⊟ **Austria**	**35**	**4**	**9**	**5**	**8**	**-2**
⊟ **2000**	**35**	**4**	**9**	**5**	**8**	**-2**
⊟ Minnesota Wild	35	4	9	5	8	-2
Christoph	35	4	9	5	8	-2
⊟ **Belarus**	**1**	**0**	**0**	**0**	**2**	**-1**
⊟ **1991**	**1**	**0**	**0**	**0**	**2**	**-1**
⊟ San Jose Sharks	1	0	0	0	2	-1
Sergei	1	0	0	0	2	-1
⊟ **Brunei Darussalam**	**427**	**33**	**77**	**44**	**337**	**-37**
⊟ **1997**	**427**	**33**	**77**	**44**	**337**	**-37**
⊟ Carolina Hurricanes	427	33	77	44	337	-37
Craig	427	33	77	44	337	-37
⊟ **Canada**	**59029**	**7821**	**21210**	**13389**	**53324**	**-3200**
⊟ **1991**	**6266**	**1159**	**2864**	**1705**	**5381**	**-344**
⊟ San Jose Sharks	6266	1159	2864	1705	5381	-344
Adam	163	26	49	23	83	-3
Andrew	93	5	20	15	51	3
Ben	34	4	7	3	43	0
Bob	107	17	43	26	173	-9
Bobby	16	1	2	1	14	4
Brad	1	0	0	0	2	-2
Brent	81	11	37	26	34	8
Brian	20	3	19	16	8	9
Curtis	123	15	33	18	72	3
Dale	98	11	29	18	28	-20
Dan	310	49	213	164	246	24
Dave	37	3	4	1	46	4

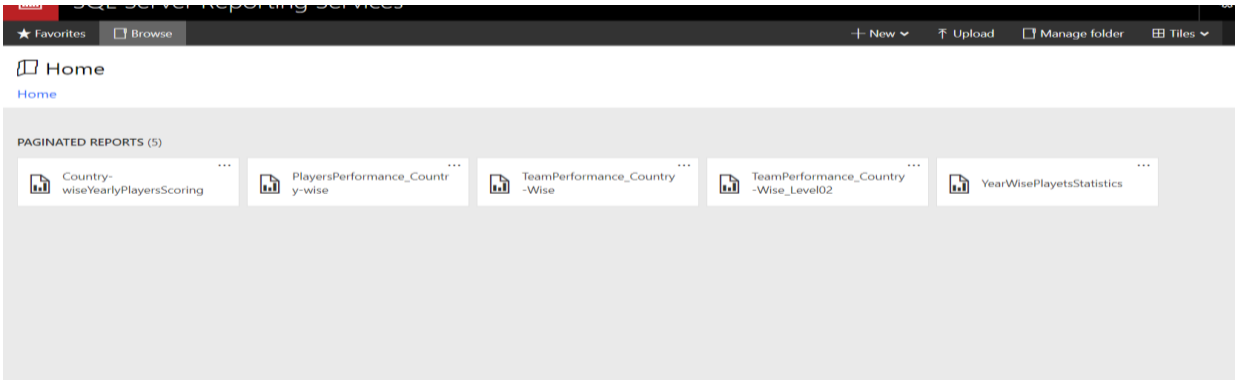
Sheet2 Sheet1

Further pivot table data is presented as a 3d stacked column chart as well.



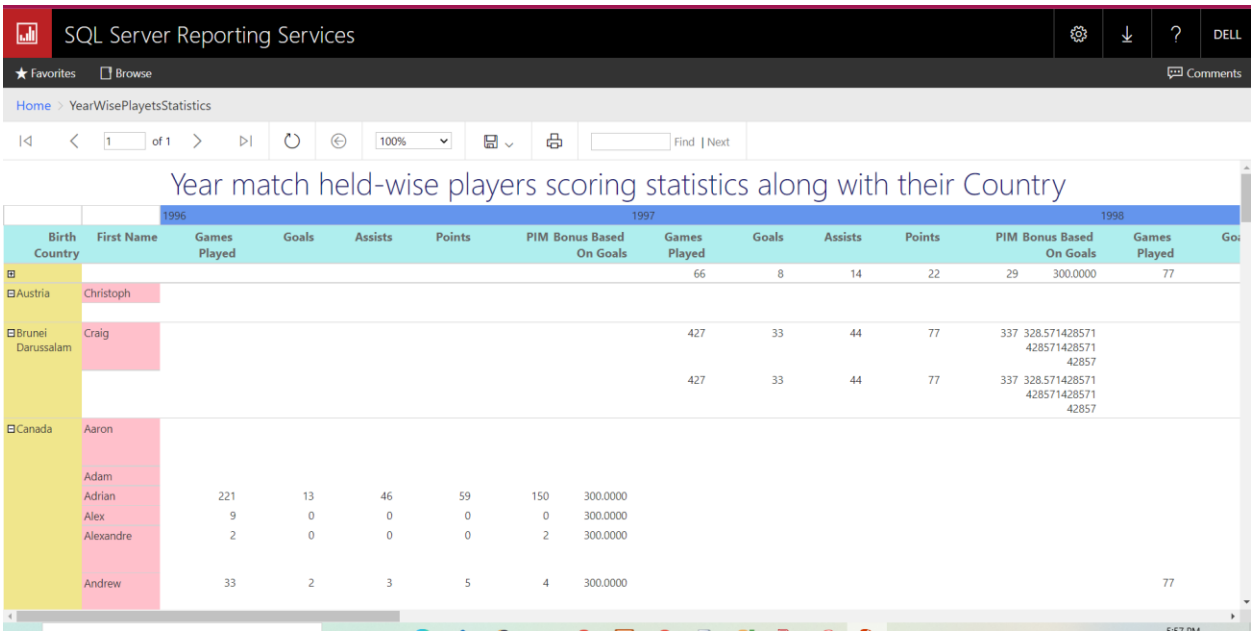
Step 04: SSRS Reports

First of all ,the connection with SSRS web portal has been made and and report builder was connected to that portal.



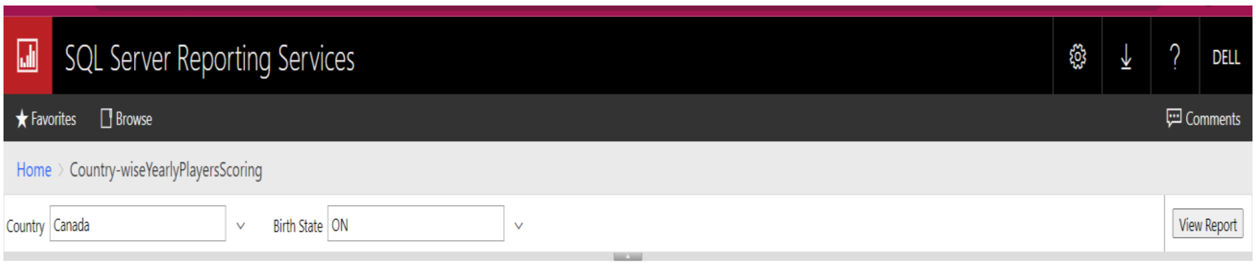
01. Report with a matrix

In here firstly, a separate report has been created in order to develop the matrix report which named as YearWisePlayersStatistics. It contains information about the year match held wise players' scoring statistics along with their countries. Then the developed Matrix report has been published in SSRS web portal. It executed correctly at there as shown in below.

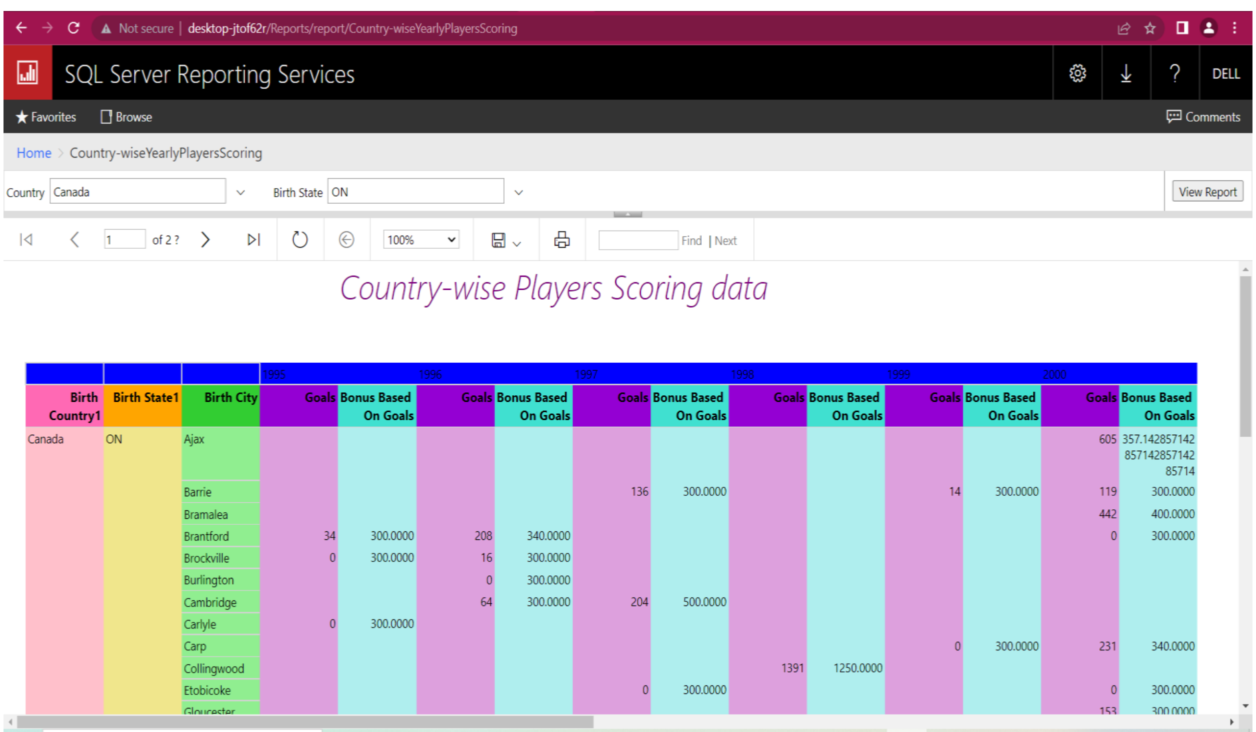


02. <u>Report with more than one parameter</u>

To develop an SSRS report firstly, the details which are going to take as parameters were decided. So the Country and state were chosen. In that case, a report was developed named as Country-wiseYearlyPlayersScoring. The first parameter is the Birth Country with a list of values for the user to select. The second parameter is the Birth state, where the values for the Birth state should be populated based on the selected value(s) for the Birth Country. And finally, the report populated country-wise players' information based on the Birth state selected for the second parameter.



Then the developed report with more than one parameter has been published in the SSRS web portal. It was executed correctly there as shown below.

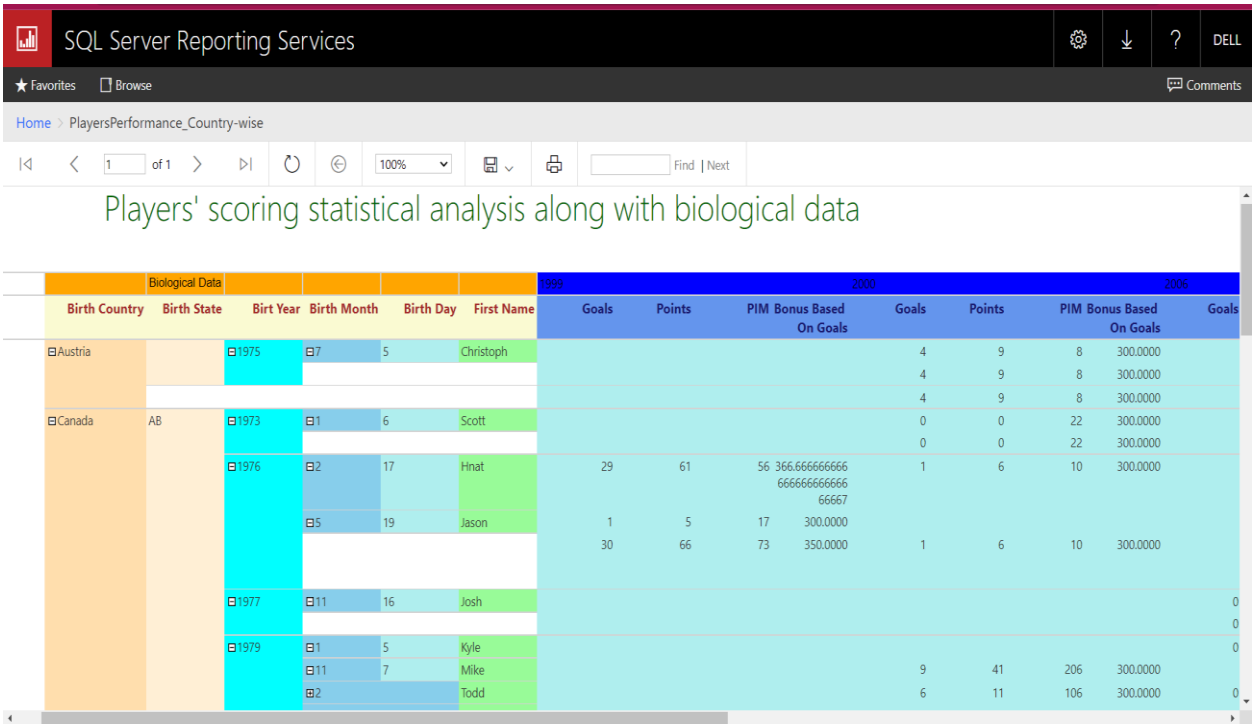


			1995		1996		1997		1998		1999		2000	
Birth Country1	Birth State1	Birth City	Goals	Bonus Based On Goals	Goals	Bonus Based On Goals	Goals	Bonus Based On Goals	Goals	Bonus Based On Goals	Goals	Bonus Based On Goals	Goals	Bonus Based On Goals
Canada	ON	Ajax											605	357.142857142857142857142857142 85714
		Barrie					136	300.0000			14	300.0000	119	300.0000
		Bramalea											442	400.0000
		Brantford	34	300.0000	208	340.0000							0	300.0000
		Brockville	0	300.0000	16	300.0000								
		Burlington			0	300.0000								
		Cambridge			64	300.0000	204	500.0000						
		Carlyle	0	300.0000										
		Carp									0	300.0000	231	340.0000
		Collingwood							1391	1250.0000				
		Etobicoke					0	300.0000					0	300.0000
		Gloucester											153	300.0000

03. SSRS drill-down report

In this task firstly, another SSRS report was created and it was named PlayersPerformance_Country-wise. Then the report was developed to present information about players' scoring statistics along with their biological information. Further Birth Country, birth year, and month columns were developed to drill –down and view information.

Then the developed drill_down report has been published in the SSRS web portal. It was executed correctly there as shown below.



	Biological Data					1999				2000				2006
Birth Country	Birth State	Birt Year	Birth Month	Birth Day	First Name	Goals	Points	PIM	Bonus Based On Goals	Goals	Points	PIM	Bonus Based On Goals	Goals
Austria		1975	7	5	Christoph					4	9	8	300.0000	
										4	9	8	300.0000	
										4	9	8	300.0000	
Canada	AB	1973	1	6	Scott					0	0	22	300.0000	
										0	0	22	300.0000	
		1976	2	17	Hnat	29	61	56	366.6666666666666666666666666667	1	6	10	300.0000	
			5	19	Jason	1	5	17	300.0000					
						30	66	73	350.0000	1	6	10	300.0000	
		1977	11	16	Josh									0
														0
		1979	1	5	Kyle									0
			11	7	Mike					9	41	206	300.0000	
			2		Todd					6	11	106	300.0000	0

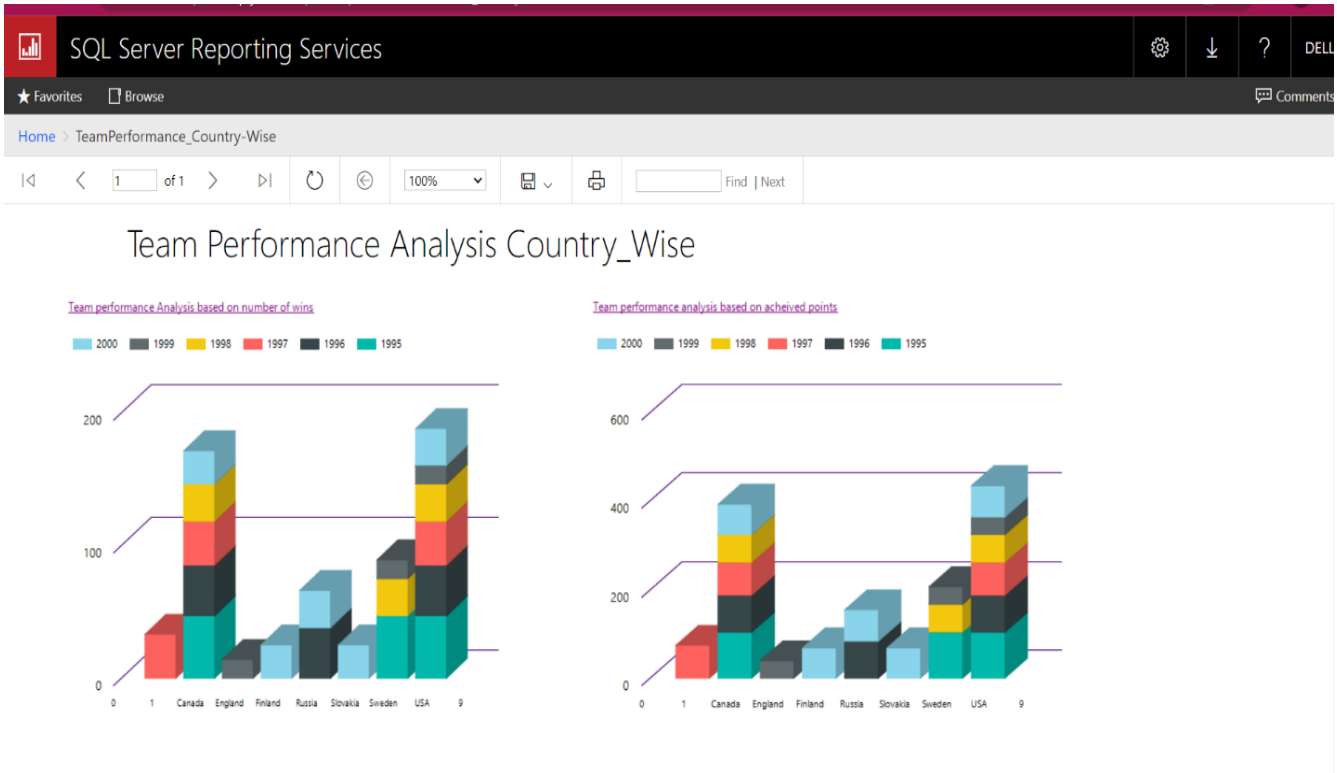
04. SSRS drill-through report

Here, firstly two source details were chosen to take as the two-level of drill-through report namely country and state. So that two SSRS reports were created as TeamPerformance_Country-Wise and TeamPerformance_Country-Wise_Level02.
In that case, each of these reports presented information about teams' performance based on the year

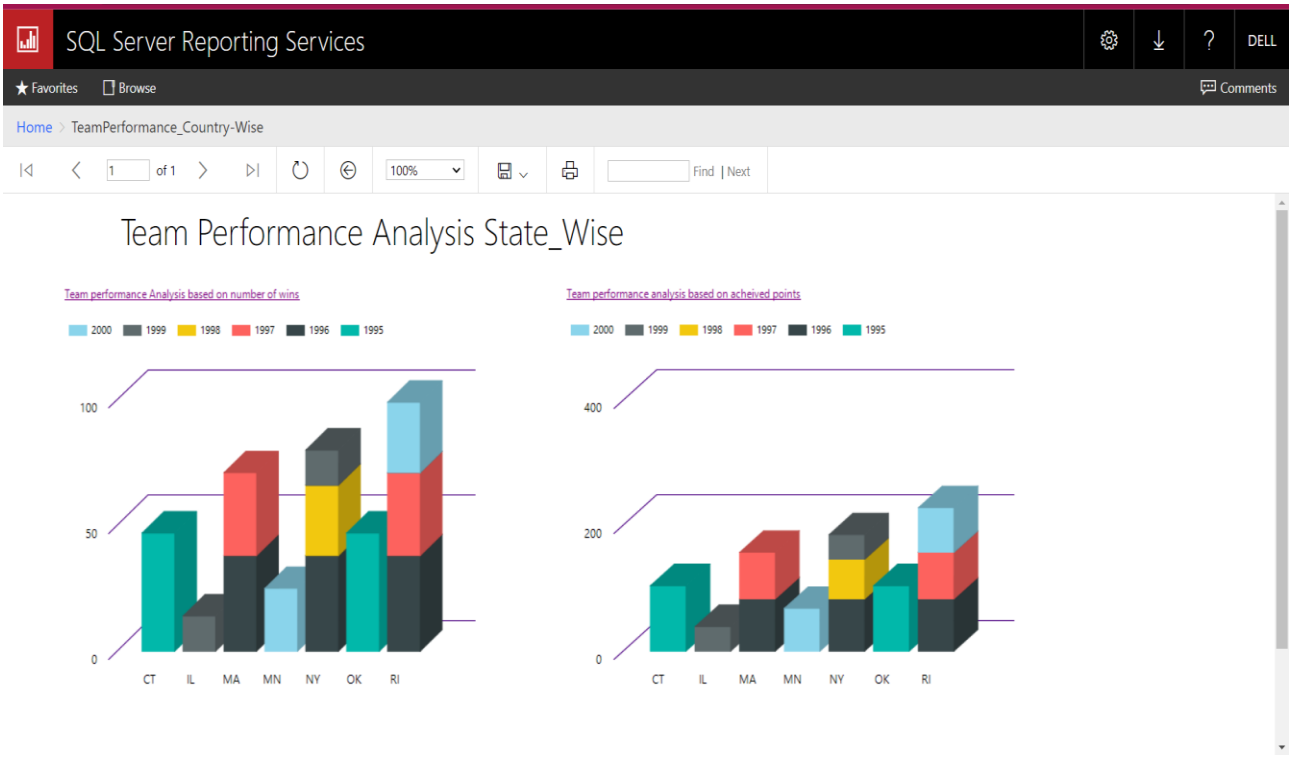
- TeamPerformance_Country-Wise Report - About Team performance analysis country-wise.
- TeamPerformance_Country-Wise_Level02 Report – About Team performance analysis state-wise.

Thereafter, the developed drill_through report has been published in the SSRS web portal. It was executed correctly there as shown below. Further according to the drill_through concept via the Team performance analysis country-wise, can correctly move to Team performance analysis state-wise report. So if the user wants can view Team performance information country-wise or state-wise as a 3D stacked chart.

❖ Level 01



❖ Level 02


SQL Server Reporting Services
DELL
Favorites
Browse
Comments
Home
TeamPerformance_Country-Wise
1
of 1
100%
Find | Next
Team Performance Analysis State_Wise
Team performance Analysis based on number of wins
2000 1999 1998 1997 1996 1995
100
50
0
CT IL MA MN NY OK RI
Team performance analysis based on acheived points
2000 1999 1998 1997 1996 1995
400
200
0
CT IL MA MN NY OK RI