

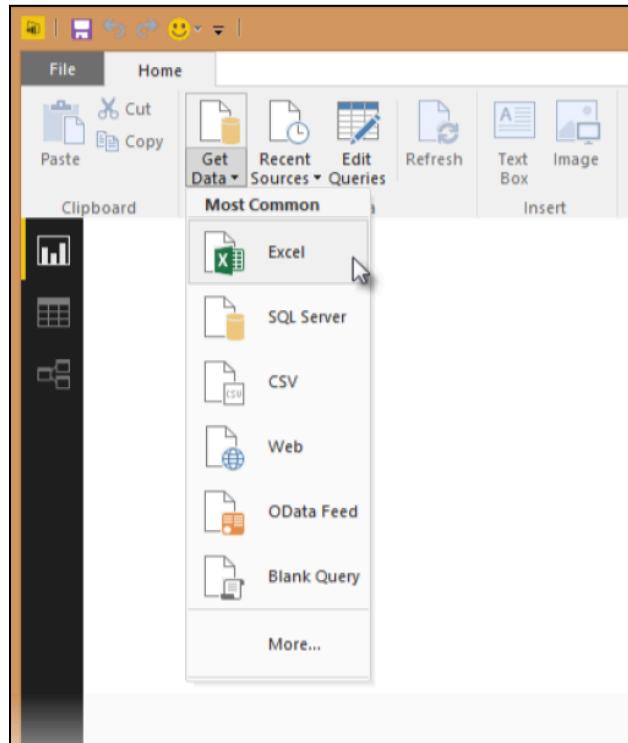
Practical 1. Import the legacy data from Excel and load in the target system

Steps 1: Create an excel sheet with data

	A	B	C	D	E	F	G
1	OrderDate	Region	Rep	Item	Units	Unit Cost	Total
2	9-1-14	Central	Smith	Desk	2	125.00	250.00
3	6-17-15	Central	Kivell	Desk	5	125.00	625.00
4	9-10-15	Central	Gill	Pencil	7	1.29	9.03
5	11-17-15	Central	Jardine	Binder	11	4.99	54.89
6	10-31-15	Central	Andrews	Pencil	14	1.29	18.06
7	2-26-14	Central	Gill	Pen	27	19.99	539.73
8	10-5-14	Central	Morgan	Binder	28	8.99	251.72
9	12-21-15	Central	Andrews	Binder	28	4.99	139.72
10	2-9-14	Central	Jardine	Pencil	36	4.99	179.64
11	8-7-15	Central	Kivell	Pen Set	42	23.95	1,005.90
12	1-15-15	Central	Gill	Binder	46	8.99	413.54
13	1-23-14	Central	Kivell	Binder	50	19.99	999.50
14	3-24-15	Central	Jardine	Pen Set	50	4.99	249.50
15	5-14-15	Central	Gill	Pencil	53	1.29	68.37
16	7-21-15	Central	Morgan	Pen Set	55	12.49	686.95
17	4-10-15	Central	Andrews	Pencil	66	1.99	131.34
18	12-12-14	Central	Smith	Pencil	67	1.29	86.43
19	4-18-14	Central	Andrews	Pencil	75	1.99	149.25
20	5-31-15	Central	Gill	Binder	80	8.99	719.20
21	2-1-15	Central	Smith	Binder	87	15.00	1,305.00
22	5-5-14	Central	Jardine	Pencil	90	4.99	449.10
23	6-25-14	Central	Morgan	Pencil	90	4.99	449.10
24	12-4-15	Central	Jardine	Binder	94	19.99	1,879.06

Step 2: Open Power BI desktop

Step 3: Go to Home Ribbon-> Get Data-> Excel and browse your excel file



Step 4: In the Navigator tab, select your table(Sheet1) from your dataset(Legacy dataset.xlsx)

OrderDate	Region	Rep	Item	Units	Unit Cost	Total
01-09-2014	Central	Smith	Desk	2	125	250
17-06-2015	Central	Kivell	Desk	5	125	625
10-09-2015	Central	Gill	Pencil	7	1.29	9.03
17-11-2015	Central	Jardine	Binder	11	4.99	54.89
31-10-2015	Central	Andrews	Pencil	14	1.29	18.06
26-02-2014	Central	Gill	Pen	27	19.99	539.73
05-10-2014	Central	Morgan	Binder	28	8.99	251.72
21-12-2015	Central	Andrews	Binder	28	4.99	139.72
09-02-2014	Central	Jardine	Pencil	36	4.99	179.64
07-08-2015	Central	Kivell	Pen Set	42	23.95	1005.9
15-01-2015	Central	Gill	Binder	46	8.99	413.54
23-01-2014	Central	Kivell	Binder	50	19.99	999.5
24-03-2015	Central	Jardine	Pen Set	50	4.99	249.5
14-05-2015	Central	Gill	Pencil	53	1.29	68.37
21-07-2015	Central	Morgan	Pen Set	55	12.49	686.95
10-04-2015	Central	Andrews	Pencil	66	1.29	83.34
12-12-2014	Central	Smith	Pencil	67	1.29	84.43
18-04-2014	Central	Andrews	Pencil	75	1.29	94.5
31-05-2015	Central	Gill	Binder	80	8.99	719.2
01-02-2015	Central	Smith	Binder	87	15	1305
05-05-2014	Central	Jardine	Pencil	90	4.99	449.1
25-06-2014	Central	Morgan	Pencil	90	4.99	449.1
04-12-2015	Central	Jardine	Binder	94	19.99	1879.06

Step 6: You will obtain this screen for queries

The screenshot shows the Power Query Editor interface with the following details:

- File**: Standard file navigation options.
- Home**: Active tab, showing various transformation tools like Close & Apply, New Source, Refresh, Enter Data, Data source settings, Manage Parameters, Refresh Preview, Advanced Editor, Properties, Choose Columns, Remove Columns, Keep Rows, Remove Rows, Sort, Split Column, Group By, Replace Values, Data Type: Date, Use First Row as Headers, Merge Queries, Append Queries, Combine Files, and Combine.
- Queries [1]**: Shows the query 'Sheet1' with 23 rows of data.
- Preview**: Displays the 23 rows of data from 'Sheet1'.
- Transform**: The 'Edit' button is highlighted with a blue oval.
- QUERY SETTINGS** pane (right):
 - Properties**: Name is set to 'Sheet1'.
 - Applied Steps**:
 - Source
 - Navigation
 - Promoted Headers
 - Changed Type (highlighted with a blue oval)

Practical 2: Perform the Extraction Transformation and Loading (ETL) process to construct the database in SQL server.

Step 1: Open Power BI.

Step 2: Get Data > Excel

Step 3: Choose your Excel data > Select your table > Edit

Step 4: Power Query Editor window will open.

The screenshot shows the Power Query Editor interface. The ribbon at the top has tabs like File, Home, Transform, Add Column, View, and Help. The 'Transform' tab is selected. The main area shows a table with columns labeled 'Column1' through 'Column8'. The 'Data_GA' query is selected in the left sidebar. The 'QUERY SETTINGS' pane on the right shows the 'Name' is set to 'Data_GA'. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 6:26 PM' and the date '3/5/2019'.

Step 5: Reduce > Remove rows > Remove top rows.

This screenshot is similar to the previous one, showing the Power Query Editor with the 'Data_GA' query. However, the 'Remove Top Rows' button in the ribbon is highlighted, indicating it has been used. The 'QUERY SETTINGS' pane remains the same, showing the name 'Data_GA'. The status bar at the bottom shows 'PREVIEW DOWNLOADED AT 6:26 PM' and the date '3/5/2019'.

Step 6: Enter the number of rows you want to reduce (here it is just one row from top)

Step 7: Transform > Use the first row as headers.

The screenshot shows the Power Query Editor interface with the 'Data_GA' query selected. In the ribbon, the 'Transform' tab is active. A context menu is open over the first row of the table, with the 'Use First Row as Headers' option highlighted. The 'QUERY SETTINGS' pane on the right shows the 'Name' as 'Data_GA'. The 'APPLIED STEPS' pane lists the 'Removed Top Rows' step. The table below has five columns labeled 'ntsb_no', 'aircraft_key', 'ev_date', 'latitude', and 'longitude'.

Your table has appropriate headers now.

Step 8: Right click on the columns which you don't want to use Remove Columns.

The screenshot shows the Power Query Editor interface with the 'Data_GA' query selected. In the ribbon, the 'Transform' tab is active. A context menu is open over the 'longitude' column, with the 'Remove Columns' option highlighted. The 'QUERY SETTINGS' pane on the right shows the 'Name' as 'Data_GA'. The 'APPLIED STEPS' pane lists the 'Removed Top Rows' and 'Changed Type1' steps. The table below has four columns labeled 'ntsb_no', 'aircraft_key', 'ev_date', and 'latitude'.

Step 9: Replace values by selecting that cell right click replace values...

The screenshot shows the Power Query Editor interface. On the left, the 'Queries [1]' pane lists a single query named 'Data_GA'. The main area displays a table with columns: ntsb_no, ev_date, ev_city, ev_state, ev_country, inj_tot_f, inj_tot_s, and ev_highest_injury. A context menu is open over the 'inj_tot_f' column, specifically at the cell for row 26 which contains 'null'. The menu path 'Number Filters' -> 'Replace Values...' is highlighted. The 'QUERY SETTINGS' pane on the right shows the query name as 'Data_GA' and the applied steps as 'Source', 'Navigation', 'Changed Type', 'Removed Top Rows', 'Promoted Headers', 'Changed Type1', and 'Removed Columns'. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 6:28 PM'.

Step 10: Provide the new value and click on OK.

This screenshot shows the same Power Query Editor session after the replacement. The 'inj_tot_f' column now contains '0' in the cell previously marked with 'null'. The rest of the data remains the same. The 'QUERY SETTINGS' pane still shows the query name as 'Data_GA' and the applied steps. The status bar at the bottom indicates 'PREVIEW DOWNLOADED AT 6:31 PM'.

All values having the values you wished to replace earlier will be replaced with new values.

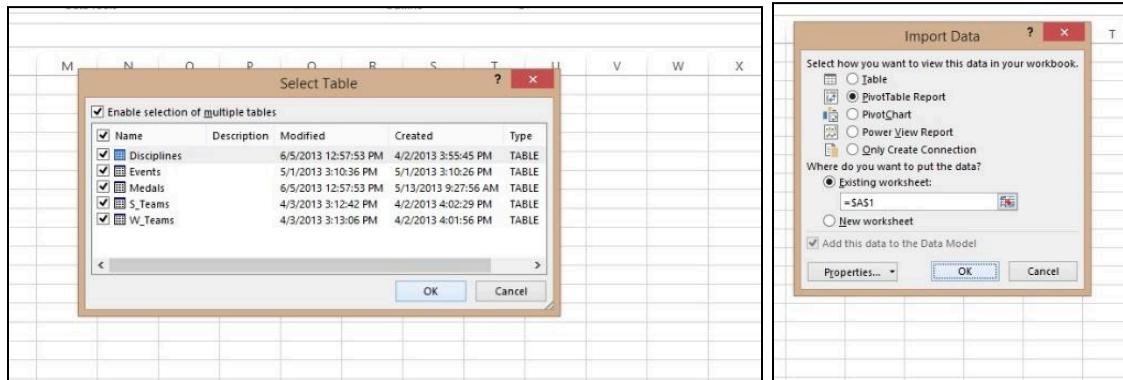
Practical 3: Import data warehouse data in Microsoft Excel and create Pivot table and Pivot Chart.

Step 1: Open a blank workbook.

Data -> Get External Data -> From Access.

Step 2: Select the OlympicMedals.accdb file and click Open.

Step 3: Check the Enable Selection of Multiple Tables box and select all the tables. Click OK. Click OK again.



Step 4: A pivot table is created using the imported tables.

Step 5: In PivotTable Fields, expand the Medals table. Find the NOC_CountryRegions and drag it to the Columns area.

Step 6: Find the Disciplines table and drag it to the Rows area.

The screenshot shows a Microsoft Excel window with a PivotTable named 'PivotTable1' in cell A1. The PivotTable Fields pane is open on the right, showing the 'Disciplines' table expanded. The 'Discipline' field is selected in the 'ROWS' area. The 'VALUES' area is set to 'NOC_Country...'. The PivotTable itself lists various sports from 1 to 30.

Step 7: filter disciplines to display only five sports: archery, diving, fencing, figure skating and speed skating.

Click anywhere in the PivotTable to ensure the excel PivotTable is selected. In the PivotTable Fields list, where the Disciplines table is expanded, hover over its Discipline field and a drop down arrow appears to the right of the field. Click the dropdown, click “Select All” to remove all selections, then scroll down and select archery, diving, fencing, figure skating and speed skating. Click OK.

The screenshot shows the 'Label Filters' dialog box open over the PivotTable Fields pane. The 'Value Filters' section is visible, showing a list of sports. The sports 'Archery', 'Diving', 'Fencing', 'Figure skating', and 'Speed skating' are checked. The 'OK' button is highlighted.

Step 8: In PivotTable Fields, from the Medals table, drag Medal to the VALUES area. Since Values must be numeric, Excel automatically changes Medal to Count of Medal.

The screenshot shows a Microsoft Excel spreadsheet titled "Book1 - Excel (Product Activation Failed)". The PivotTable Fields pane on the right is open, showing the following fields:

- Choose fields to add to report:**
 - NOC_CountryRegion
 - Gender
 - Event,gender
 - Sport
 - DisciplineID
 - Event
 - Medal
 - MedalKey
 - DisciplineEvent
- S Teams:** A collapsed section.
- Drag fields between areas below:**
 - FILTERS:** Discipline
 - COLUMNS:** NOC_Country...
 - ROWS:** Discipline
 - VALUES:** Count of Medal

The main PivotTable area displays medal counts for various countries and sports. The data includes:

	AUS	AUT	BEL	BLR	BOH	BUL	CAN	CHN	CUB	DEN	EGY	ESP	EUA	EUN	FIN	FRA	FRG	GBR	GDR	GER	GRE	HUN	INA	ITA	JPN	KAZ	KOR	MEX	NED	NOR	POL	PRK	ROU	RUS	SLO	SWE	TUR	UKR	USA	VIE	YUG	ZAF
Archery	2	51								15		3	4	6	46	13	6					3	12	3	52	9	4															
Diving	17									11	60	2	2	4	3	1	7	7	24	2			9																			
Fencing	5	13	44							5	19	24	5	1	10	13	283	47	24	1	51		226	328	1	3	1	24	81	39												
Figure skating		25	3							28	7			5	9	4	18	3	19	13	11		12	2	2		3	7														
Speed skating	1	6	1	1						3	43	19		2	1	24	3	1	29	34		7	15	1	18	75	79	2	2													
Grand Total	5	20	44	99	1	5	3	82	120	24	7	2	4	21	30	34	348	53	64	50	126	2	238	3	358	21	1	73	12	111	86	87	2	39								

Step 9: From the Medals table, select Medal again and drag it into the FILTERS area.

Step 10: Let's filter the PivotTable to display only those countries or regions with more than 90 total medals.

In the PivotTable, click the dropdown to the right of Column Labels.

The screenshot shows a Microsoft Excel spreadsheet titled "Book1 - Excel (Product Activation Failed)". The PivotTable Fields pane on the right is open, showing the following filters:

- Choose fields to add to report:**
 - Gender
 - Event,gender
 - Sport
 - DisciplineID
 - Event
 - Medal
 - MedalKey
 - DisciplineEvent
- S Teams:** A collapsed section.
- Drag fields between areas below:**
 - FILTERS:** Medal
 - COLUMNS:** NOC_Country...
 - ROWS:** Discipline
 - VALUES:** Count of Medal

The main PivotTable area displays medal counts for countries with more than 90 total medals. The data includes:

	AUS	AUT	BEL	BLR	BOH	BUL	CAN	CHN	CUB	DEN	EGY	ESP	EUA	EUN	FIN	FRA	FRG	GBR	GDR	GER	GRE	HUN	INA	ITA	JPN	KAZ	KOR	MEX	NED	NOR	POL	PRK	ROU	RUS	SLO	SWE	TUR	UKR	USA	VIE	YUG	ZAF
Archery	2	51								15		3	4	6	46	13	6					3	12	3	52	9	4															
Diving	17									11	60	2	2	4	3	1	7	7	24	2			9																			
Fencing	5	13	44							5	19	24	5	1	10	13	283	47	24	1	51		226	328	1	3	1	24	81	39												
Figure skating		25	3							28	7			5	9	4	18	3	19	13	11		12	2	2		3	7														
Speed skating	1	6	1	1						3	43	19		2	1	24	3	1	29	34		7	15	1	18	75	79	2	2													
Grand Total	5	20	44	99	1	5	3	82	120	24	7	2	4	21	30	34	348	53	64	50	126	2	238	3	358	21	1	73	12	111	86	87	2	39								

Step 11: Select Value Filters and select Greater Than....

Step 12: Type 90 in the last field (on the right). Click OK.
Your PivotTable looks like the following screen.

The screenshot shows a Microsoft Excel spreadsheet titled "Book1 - Excel (Product Activation Failed)". The ribbon is visible with tabs like FILE, HOME, INSERT, etc. The PivotTable is named "Medal". The data shows medal counts for various sports across different countries. A tooltip for cell A10 (Speed skating) displays the value 277, with details: Row: Speed skating, Column: Grand Total.

	CHN	FRA	GER	HUN	ITA	NED	RUS	URS	USA	Grand Total
Archery	51	15	46	6	12	9	1	7	52	199
Diving	60	1	24	9	24	14	14	131		263
Fencing	44	19	283	51	226	328	24	41	145	48
Figure skating	3	7	18	11	12	2	3	29	42	51
Speed skating	1	19	34	7	75	8	60	73		178
Grand Total	99	120	348	126	238	358	111	103	268	355

Step 13: Let's start by creating a blank worksheet, then import data from an Excel workbook. Insert a new Excel worksheet, and name it Sports.

Step 14: Browse to the folder that contains the downloaded sample data file, and open OlympicSports.xlsx.

Step 15: Select and copy the data in Sheet1. If you select a cell with data, such as cell A1, you can press Ctrl + A to select all adjacent data. Close the OlympicSports.xlsx workbook.

On the Sports worksheet, place your cursor in cell A1 and paste the data.

The screenshot shows the "Sports" worksheet in Microsoft Excel. The data consists of a list of sports and their corresponding codes, starting with "Sport" in cell A1. The list includes Aquatics, Archery, Athletics, Badminton, Baseball, Basketball, Basque Pelota, Biathlon, Bobsleigh, Boxing, Canoe/Kayak, Cricket, Croquet, Curling, Cycling, Equestrian, Fencing, Football, Golf, Gymnastics, Handball, Hockey, Ice Hockey, Jeu de paume, Judo, Lacrosse, Luge, Modern Pentathlon, and Polo. The "Editing" tab of the ribbon is selected.

Sport	Code
Aquatics	S1
Archery	S2
Athletics	S3
Badminton	S4
Baseball	S5
Basketball	S6
Basque Pelota	S7
Biathlon	S8
Bobsleigh	S9
Boxing	S10
Canoe / Kayak	S11
Cricket	S12
Croquet	S13
Curling	S14
Cycling	S15
Equestrian	S16
Fencing	S17
Football	S18
Golf	S19
Gymnastics	S20
Handball	S21
Hockey	S22
Ice Hockey	S23
Jeu de paume	S24
Judo	S25
Lacrosse	S26
Luge	S27
Modern Pentathlon	S28
Polo	S29

Step 16: With the data still highlighted, press Ctrl + T to format the data as a table. You can also format the data as a table from the ribbon by selecting HOME > Format as Table. Since the data has headers, select My table has headers in the Create Table window that appears.

Step 17: Name the table. In TABLE TOOLS > DESIGN > Properties, locate the Table Name field and type Sports. The workbook looks like the following screen.
Save the workbook.

Step 18: Insert a new Excel worksheet, and name it Hosts.

Step 19: Select and copy the following table, including the table headers.

City	NOC_CountryRegion	Alpha-2 Code	Edition	Season
Melbourne / Stockholm	AUS	AS	1956	Summer
Sydney	AUS	AS	2000	Summer
Innsbruck	AUT	AT	1964	Winter
Innsbruck	AUT	AT	1976	Winter
Antwerp	BEL	BE	1920	Summer
Antwerp	BEL	BE	1920	Winter
Montreal	CAN	CA	1976	Summer
Lake Placid	CAN	CA	1980	Winter
Calgary	CAN	CA	1988	Winter
St. Moritz	SUI	SZ	1928	Winter
St. Moritz	SUI	SZ	1948	Winter

In Excel, place your cursor in cell A1 of the Hosts worksheet and paste the data.

Step 20: Format the data as a table. As described earlier in this tutorial, you press Ctrl + T to format the data as a table, or from HOME > Format as Table. Since the data has headers, select My table has headers in the Create Table window that appears.

A1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	City	NOC_CountryRegion	Alpha-2 Code	Edition	Season													
2	Melbourne	AUS	AS	1956	Summer													
3	Sydney	AUS	AS	2000	Summer													
4	Innsbruck	AUT	AT	1964	Winter													
5	Innsbruck	AUT	AT	1976	Winter													
6	Antwerp	BEL	BE	1920	Summer													

Step 21: Select the Edition column, and from the HOME tab, format it as Number with 0 decimal Places.

Step 22: Save the workbook. Your workbook looks like the following screen.

D1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	City	NOC_CountryRegion	Alpha-2 Code	Edition														
2	Melbourne	AUS	AS	1956														
3	Sydney	AUS	AS	2000														
4	Innsbruck	AUT	AT	1964														
5	Innsbruck	AUT	AT	1976	Winter													
6	Antwerp	BEL	BE	1920	Summer													

Step 23: On Sheet1, at the top of PivotTable Fields, click All to view the complete list of available tables.

A1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Medal	All																
2	Count of Medal	Column Labels																
3	Row Labels	T_BEL	CHN	FRA	GER	HUN	ITA	NED	RUS	URS	USA	Grand Total						
4	Archery	51	15	46	6	12	9	1	7	52		199						
5	Diving	60	1	24	9	24	14	131				263						
6	Fencing	44	19	283	51	226	328	24	41	145	48	1209						
7	Figure skating	3	7	18	11	12	2	3	29	42	51	178						
8	Speed skating	1	19	34	7	75	8	60	73			277						
9	Grand Total	99	120	348	126	238	358	111	103	268	355	2126						

PivotTable Fields

ACTIVE | ALL

Choose fields to add to report:

Disciplines

Medals

Events

Hosts

S_Teams

Step 24: Expand Sports and select Sport to add it to the PivotTable. Notice that Excel prompts you to create a relationship, as seen in the following screen.

The screenshot shows a Microsoft Excel spreadsheet titled 'Book1.xlsx'. The PivotTable Fields pane is open on the right side. In the 'Relationships between tables may be needed' section, there is a yellow box around the 'CREATE...' button. The 'Choose fields to add to report' area also has a yellow box around the 'CREATE...' button. The main table on the left shows medal counts for various sports across different countries.

Step 25: Click CREATE, in the highlighted PivotTable Fields area to open the Create Relationship dialog,

The screenshot shows the 'Create Relationship' dialog box open over the PivotTable Fields pane. The dialog box is titled 'Create Relationship' and contains fields for 'Table:' and 'Related Table:', both set to 'Sports'. The 'Related Column (Primary)' dropdown is set to 'SportID'. The 'Column (Foreign)' dropdown is set to 'SportID'. The 'OK' button is highlighted. The background shows the same PivotTable setup as the previous screenshot.

Step 26: In Table, choose Disciplines from the drop down list.

In Column (Foreign), choose SportID.

In Related Table, choose Sports.

In Related Column (Primary), choose SportID.

Click OK.

imported. Let's create the relationship.

4. Click **CREATE...** in the highlighted PivotTable Fields area to open the **Create Relationship** dialog, as shown in the following screen.

5. In **Table**, choose **Disciplines** from the drop-down menu.

6. In **Column (Foreign)**, choose **SportID**.

7. In **Related Table**, choose **Sports**.

8. In **Related Column (Primary)**, choose **SportID**.

9. Click **OK**.

The PivotTable changes to reflect the new relationship. But the PivotTable doesn't look right quite yet, because of the ordering of fields in the **ROWS** area. Discipline is a subcategory of a given sport, but since we arranged Discipline above Sport in the **ROWS** area, it's not organized properly. The following screen shows this unwanted ordering.

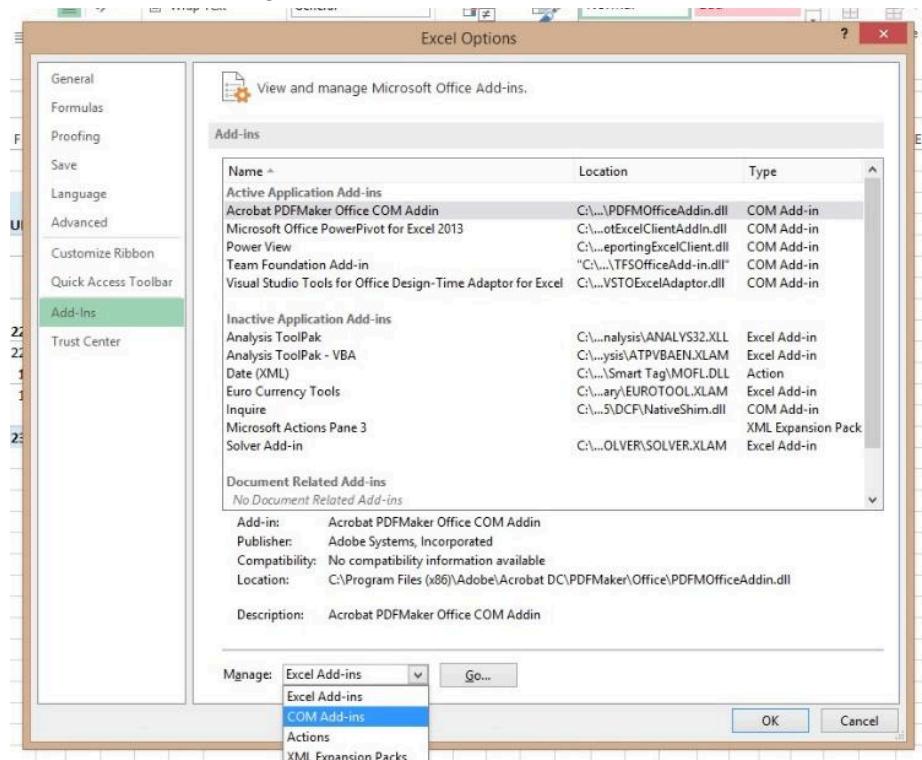
Step 27: In the ROWS area, move Sport above Discipline. That's much better, and the PivotTable displays the data how you want to see it, as shown in the following screen.

	CHIN	FRA	GER	HUN	ITA	NED	RUS	URS	USA	Grand Total	
Aquatics	60	1	24	9	24	14	131			263	
Diving	60	1	24	9	24	14	131			263	
Archery	51	15	46	6	12	9	1	7	52	199	
Archery	51	15	46	6	12	9	1	7	52	199	
Fencing	44	19	283	51	226	328	24	41	145	48	1209
Fencing	44	19	283	51	226	328	24	41	145	48	1209
Skating	4	26	18	45	12	9	78	37	102	124	455
Figure skating	3	7	18	11	12	2	3	29	42	51	178
Speed skating	1	19	34		7	75	8	60	73	277	
Grand Total	99	120	348	126	238	358	111	103	268	355	2126

Practical 4: Import the cube in Microsoft Excel and create the Pivot table and Pivot Chart to perform data analysis.

Step 1: Go to FILE > Options > Add-Ins.

Step 2: In the Manage box near the bottom, click COM Add-ins > Go



Step 3: Check the Microsoft Office Power Pivot in Microsoft Excel 2013 box, and then click OK.

Step 4: The following data is displayed.

Row Labels	Column Labels	CHN	FRA	GER	HUN	ITA	NED	RUS	URS	USA	Grand Total
Aquatics	All	60	1	24	9	24	14	131		263	
Diving	All	60	1	24	9	24	14	131		263	
Archery	All	51	15	46	6	12	9	1	7	52	199
Fencing	All	44	19	283	51	226	328	24	41	145	48
Fencing	All	44	19	283	51	226	328	24	41	145	48
Skating	All	4	26	18	45	12	9	78	37	102	124
Figure skating	All	3	7	18	11	12	2	3	29	42	51
Speed skating	All	1	19	34	7	75	8	60	73		277
Grand Total	All	99	120	348	126	238	358	111	103	268	355
											2126

Step 5: The Excel workbook includes a table called Hosts. We imported Hosts by copying it and pasting it into Excel, then formatted the data as a table.

Step 6: In Excel, click the Hosts tab to make it the active sheet.

On the ribbon, select POWER PIVOT > Tables > Add to Data Model. This step adds the Hosts table to the Data Model.

The screenshot shows the Microsoft Excel interface with the 'PowerPivot' ribbon tab selected. A tooltip for 'Add to Data Model' is displayed over a cell in the table. The table data includes:

	City	NOC_CountryRegion	Alpha-2 Code	Edition	Season
1	Melbourn	AUS	AS	956	Summer
2	Sydney	AUS	AS	2000	Summer
3	Innsbruck	AUT	AT	1964	Winter
4	Innsbruck	AUT	AT	1976	Winter
5	Antwerp	BEL	BE	1920	Summer
6	Antwerp	BEL	BE	1920	Winter

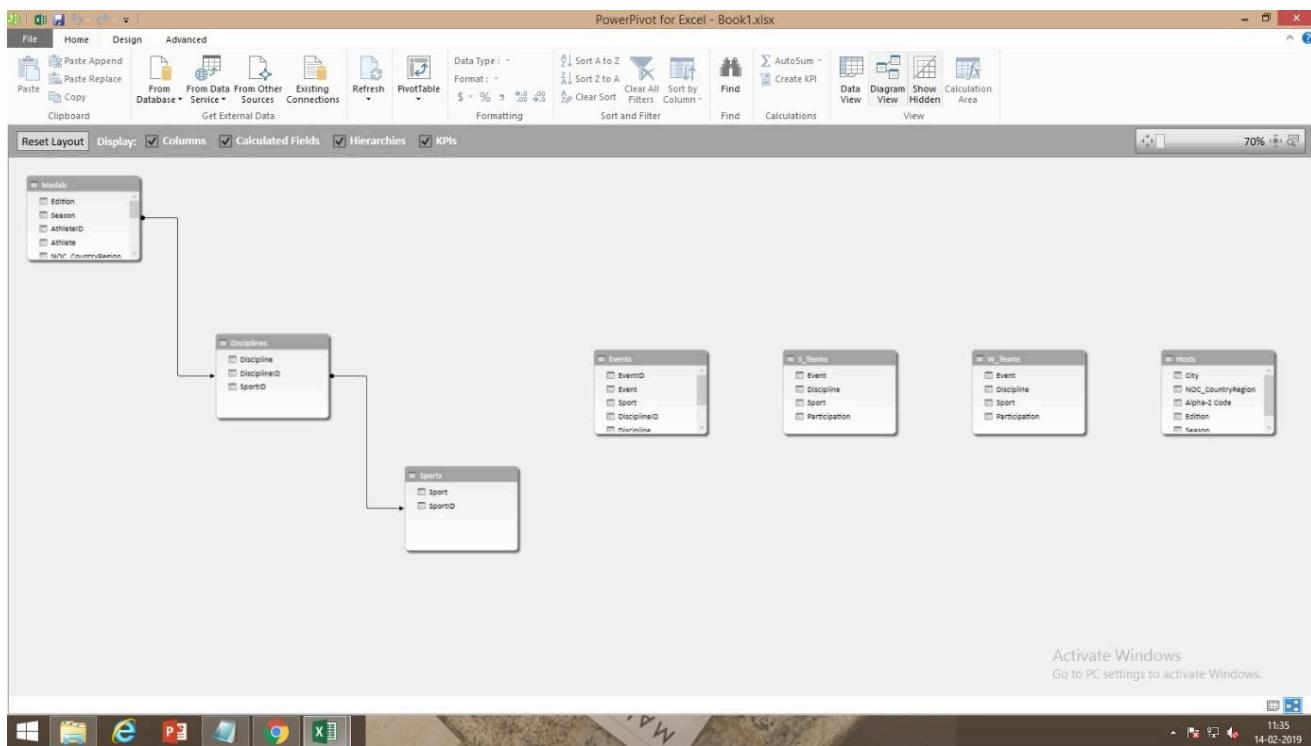
Step 7: the Power Pivot window shows all the tables in the model, including Hosts. Click through a couple of tables. In Power Pivot you can view all of the data that your model contains, even if they aren't displayed in any worksheets in Excel, such as the Disciplines, Events, and Medals data below, as well as S_Teams,W_Teams, and Sports.

The screenshot shows the 'PowerPivot for Excel - Book1.xlsx' window. The ribbon is set to 'Table Tools'. The main area displays the 'Hosts' table with columns: City, NOC_CountryRegion, Alpha-2 Code, Edition, and Season. Below the table, tabs for Disciplines, Events, Medals, S_Teams, W_Teams, and Sports are visible.

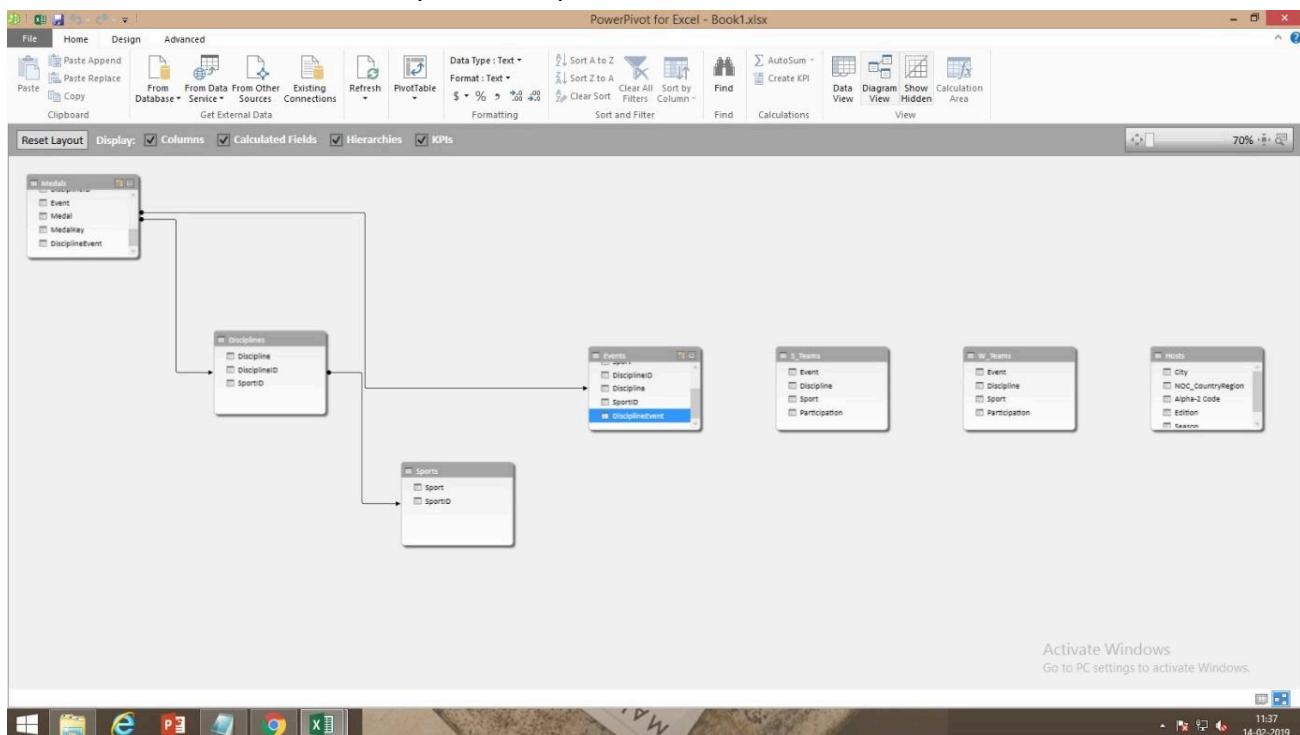
City	NOC_CountryRegion	Alpha-2 Code	Edition	Season
Melbourn	AUS	AS	1956	Summer
Sydney	AUS	AS	2000	Summer
Innsbruck	AUT	AT	1964	Winter
Innsbruck	AUT	AT	1976	Winter
Antwerp	BEL	BE	1920	Summer
Antwerp	BEL	BE	1920	Winter

Step 8: In the Power Pivot window, in the View section, click Diagram View.

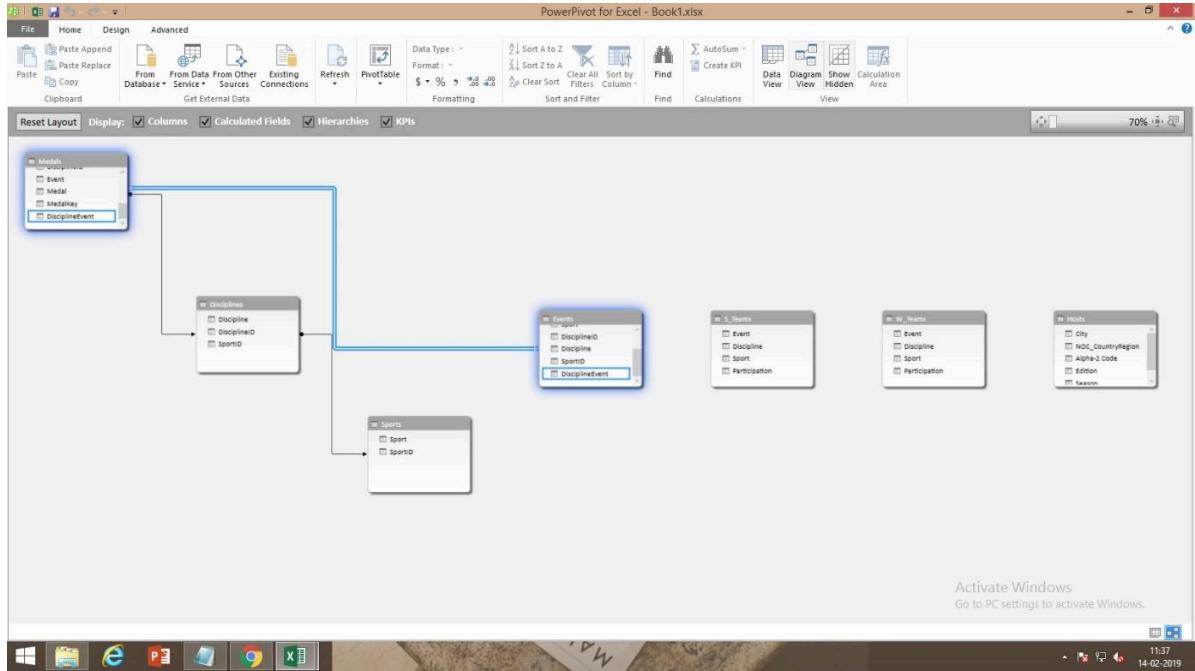
Step 9: Use the slide bar to resize the diagram so that you can see all objects in the diagram. Rearrange the tables by dragging their title bar, so they're visible and positioned next to one another. Notice that four tables are unrelated to the rest of the tables: Hosts, Events, W_Teams, and S_Teams.



Step 10: Notice that both the Medals table and the Events table have a field called DisciplineEvent. Upon further inspection, you determine that the DisciplineEvent field in the Events table consists of unique, non-repeated values.



Step 11: Create a relationship between the Medals table and the Events table. While in Diagram View, drag the DisciplineEvent field from the Events table to the DisciplineEvent field in Medals. A line appears between them, indicating a relationship has been established. Click the line that connects Events and Medals. The highlighted fields define the relationship, as shown in the following screen.



Step 12: To connect Hosts to the Data Model, we need a field with values that uniquely identify each row in the Hosts table. Then we can search our Data Model to see if that same data exists in another table. Looking in Diagram View doesn't allow us to do this. With Hosts selected, switch back to Data View.

Step 13: The following screen appears.

	City	NOC_CountryRegion	Alpha-2 Code	Edition	Season	Add Column
1	Melb...	AUS	AS	1956	Summer	
2	Sydney	AUS	AS	2000	Summer	
3	Innsbr...	AUT	AT	1964	Winter	
4	Innsbr...	AUT	AT	1976	Winter	
5	Antw...	BEL	BE	1920	Summer	
6	Antw...	BEL	BE	1924	Winter	
7	Montre...	CAN	CA	1976	Summer	
8	Lake Pl...	CAN	CA	1980	Winter	
9	Calgary	CAN	CA	1988	Winter	
10	St. Mo...	SUI	SZ	1928	Winter	
11	St. Mo...	SUI	SZ	1948	Winter	
12	Beijing	CHN	CH	2008	Summer	
13	Berlin	GER	GM	1936	Summer	
14	Garmi...	GER	GM	1936	Winter	
15	Barcel...	ESP	SP	1992	Summer	
16	Helsinki	FIN	FI	1952	Summer	
17	Paris	FRA	FR	1900	Summer	
18	Paris	FRA	FR	1924	Summer	
19	Cham...	FRA	FR	1924	Winter	
20	Greno...	FRA	FR	1968	Winter	
21	Albert...	FRA	FR	1992	Winter	
22	London	GBR	UK	1908	Summer	
23	London	GBR	UK	1908	Winter	
24	London	GBR	UK	1948	Summer	

Step 14: Select the Hosts table in Power Pivot. Adjacent to the existing columns is an empty column titled Add Column. Power Pivot provides that column as a placeholder. There are many ways to add a new column to a table in Power Pivot, one of which is to simply select the empty column that has the title Add Column.

Step 15: In the formula bar, type the following DAX formula.

=CONCATENATE([Edition],[Season])

Step 17: Let's rename the calculated column to EditionID. You can rename any column by double-clicking it, or by right-clicking the column and choosing Rename Column. When completed, the Hosts table in Power Pivot looks like the following screen.

	City	NOC	CountryRegion	Alpha-2 Code	Edition	Season	EditionID	Add Column
1	Melb...	AUS	AS	1956	Summer	1956Summer		
2	Sydney	AUS	AS	2000	Summer	2000Summer		
3	Innsbr...	AUT	AT	1964	Winter	1964Winter		
4	Innsbr...	AUT	AT	1976	Winter	1976Winter		
5	Antw...	BEL	BE	1920	Summer	1920Summer		
6	Antw...	BEL	BE	1920	Winter	1920Winter		
7	Montr...	CAN	CA	1972	Summer	1972Summer		
8	Lake P...	CAN	CA	1980	Winter	1980Winter		
9	Calgary	CAN	CA	1988	Winter	1988Winter		
10	St. Mo...	SUI	SZ	1928	Winter	1928Winter		
11	St. Mo...	SUI	SZ	1948	Winter	1948Winter		
12	Beijing	CHN	CH	2008	Summer	2008Summer		
13	Berlin	GER	GM	1936	Summer	1936Summer		
14	Garmi...	GER	GM	1936	Winter	1936Winter		
15	Barcel...	ESP	SP	1992	Summer	1992Summer		
16	Helsinki	FIN	FI	1952	Summer	1952Summer		
17	Paris	FRA	FR	1900	Summer	1900Summer		
18	Paris	FRA	FR	1924	Summer	1924Summer		
19	Cham...	FRA	FR	1924	Winter	1924Winter		
20	Greno...	FRA	FR	1968	Winter	1968Winter		
21	Albert...	FRA	FR	1992	Winter	1992Winter		
22	London	GBR	UK	1908	Summer	1908Summer		
23	London	GBR	UK	1908	Winter	1908Winter		
24	London	GBR	UK	1948	Summer	1948Summer		

Step 18: Create a new column in the Medals table, like we did for Hosts. In Power Pivot select the Medals table, and click Design > Columns > Add. Notice that Add Column is selected.

	Edition	Season	AthleteID	Athlete	NOC	CountryRegion	Gender	Event_gender	Sport	Discipline	#	Event	Medal	MedalKey	DisciplineEvent	Add Column
1	01-01-19...	Winter	A29666	URS	Men	M	Skiing	D18	4x10km ...	Gold	M10187	D184x10km relay				
2	01-01-19...	Winter	A29667	SWE	Men	M	Skiing	D18	4x10km ...	Bronze	M10188	D184x10km relay				
3	01-01-19...	Winter	A29668	FIN	Men	M	Skiing	D18	4x10km ...	Silver	M10189	D184x10km relay				
4	01-01-19...	Winter	A29729	URS	Men	M	Skiing	D18	4x10km ...	Bronze	M10319	D184x10km relay				
5	01-01-19...	Winter	A29730	NOR	Men	M	Skiing	D18	4x10km ...	Silver	M10320	D184x10km relay				
6	01-01-19...	Winter	A29731	FIN	Men	M	Skiing	D18	4x10km ...	Gold	M10321	D184x10km relay				
7	01-01-19...	Winter	A29732	URS	Men	M	Skiing	D18	4x10km ...	Bronze	M10322	D184x10km relay				
8	01-01-19...	Winter	A29733	NOR	Men	M	Skiing	D18	4x10km ...	Silver	M10323	D184x10km relay				
9	01-01-19...	Winter	A29734	FIN	Men	M	Skiing	D18	4x10km ...	Gold	M10324	D184x10km relay				
10	01-01-19...	Winter	A29736	URS	Men	M	Skiing	D18	4x10km ...	Bronze	M10326	D184x10km relay				
11	01-01-19...	Winter	A29737	NOR	Men	M	Skiing	D18	4x10km ...	Silver	M10327	D184x10km relay				
12	01-01-19...	Winter	A29738	FIN	Men	M	Skiing	D18	4x10km ...	Gold	M10328	D184x10km relay				
13	01-01-19...	Winter	A29790	URS	Men	M	Skiing	D18	4x10km ...	Bronze	M11192	D184x10km relay				
14	01-01-19...	Winter	A29791	NOR	Men	M	Skiing	D18	4x10km ...	Silver	M11193	D184x10km relay				
15	01-01-19...	Winter	A29792	FIN	Men	M	Skiing	D18	4x10km ...	Gold	M11194	D184x10km relay				
16	01-01-19...	Winter	A29855	URS	Men	M	Skiing	D18	4x10km ...	Bronze	M11326	D184x10km relay				
17	01-01-19...	Winter	A29856	SWE	Men	M	Skiing	D18	4x10km ...	Gold	M11327	D184x10km relay				
18	01-01-19...	Winter	A29857	FIN	Men	M	Skiing	D18	4x10km ...	Silver	M11328	D184x10km relay				
19	01-01-19...	Winter	A29858	URS	Men	M	Skiing	D18	4x10km ...	Bronze	M11329	D184x10km relay				
20	01-01-19...	Winter	A29859	SWE	Men	M	Skiing	D18	4x10km ...	Gold	M11330	D184x10km relay				
21	01-01-19...	Winter	A29860	FIN	Men	M	Skiing	D18	4x10km ...	Silver	M11331	D184x10km relay				
22	01-01-19...	Winter	A29865	URS	Men	M	Skiing	D18	4x10km ...	Bronze	M11336	D184x10km relay				
23	01-01-19...	Winter	A29866	SWE	Men	M	Skiing	D18	4x10km ...	Gold	M11337	D184x10km relay				
24	01-01-19...	Winter	A29867	FIN	Men	M	Skiing	D18	4x10km ...	Silver	M11338	D184x10km relay				

Step 19: In the formula bar above the table, type the following DAX formula. “=YEAR([Edition])”

Step 20: When you finish building the formula, press Enter. Values are populated for all the rows in the calculated column, based on the formula you entered. Rename the column by right-clicking CalculatedColumn1 and selecting Rename Column. Type Year, and then press Enter.

The screenshot shows the PowerPivot ribbon with the 'Home' tab selected. A calculated column named 'Year' has been created, containing the value '1956' for every row. The table includes columns for EditionID, Season, AthleteID, Athlete, NOC_CountryRegion, Gender, Event_gender, Sport, Discipline, Event, Medal, MedalKey, DisciplineEvent, and Year. The 'DisciplineEvent' column contains the formula '=YEAR([Edition])'. The status bar at the bottom indicates 'Record: 1 of 32,551' and the date '14-02-2019'.

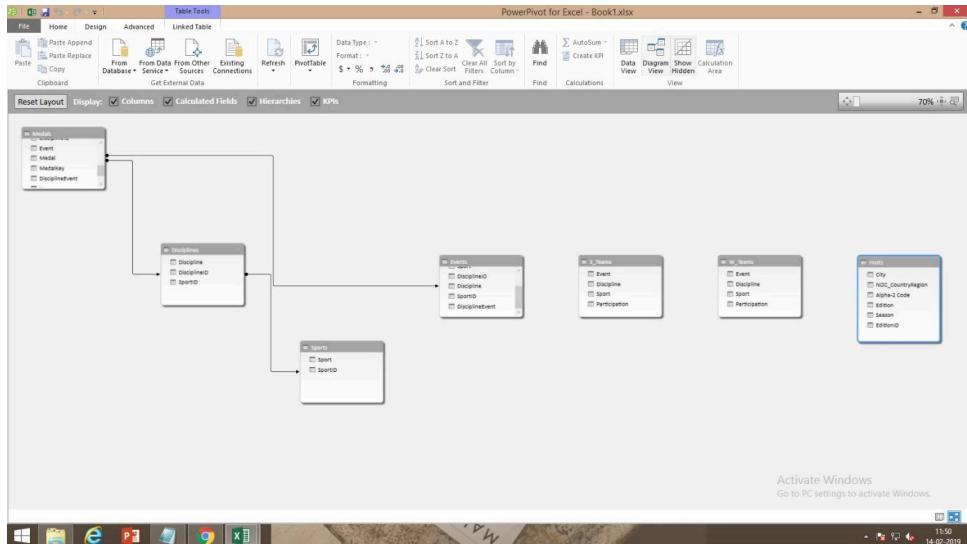
Step 21: Create the EditionID calculated column, so select Add Column. In the formula bar, type the following DAX formula and press Enter. “=CONCATENATE([Year],[Season])”

The screenshot shows the PowerPivot ribbon with the 'Home' tab selected. A calculated column named 'EditionID' has been created, containing values like '1956Winter' and '1964Winter' for each row. The table includes columns for EditionID, AthleteID, Athlete, NOC_CountryRegion, Gender, Event_gender, Sport, Discipline, Event, Medal, MedalKey, DisciplineEvent, Year, and EditionID. The 'EditionID' column contains the formula '=CONCATENATE([Year],[Season])'. The status bar at the bottom indicates 'Record: 1 of 32,551' and the date '14-02-2019'.

Step 22: Sort the column in ascending order.

Step 23 In the Power Pivot window, select Home > View > Diagram View from the ribbon.

Step 24: Expand Hosts so you can view all of its fields.



Step 26: Position the Hosts table so that it is next to Medals.

Drag the EditionID column in Medals to the EditionID column in Hosts. Power Pivot creates a relationship between the tables based on the EditionID column, and draws a line between the two columns, indicating the relationship.

Step 27: Expand the Events table so that you can more easily see all of its fields.

Press and hold Ctrl, and click the Sport, Discipline, and Event fields. With those three fields selected, right-click and select Create Hierarchy.

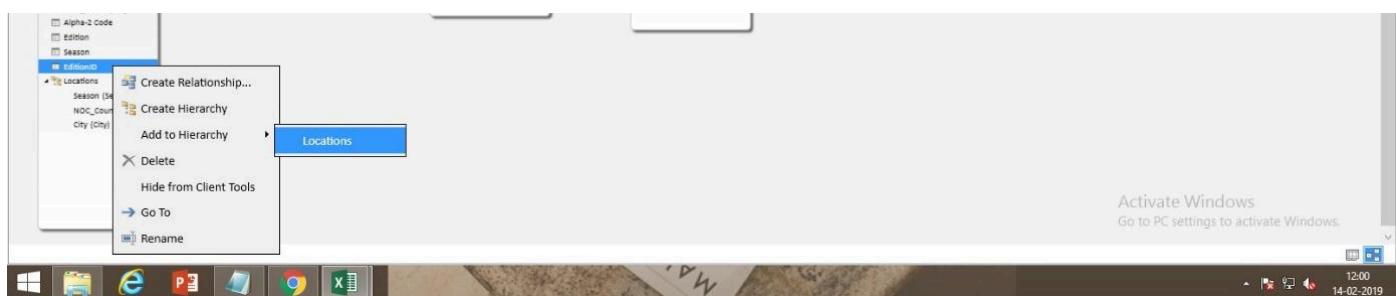
Step 28: Still in Diagram View in Power Pivot, select the Hosts table and click the Create Hierarchy button in the table header.

Step 29: An empty hierarchy parent node appears at the bottom of the table.

Step 30: Type Locations as the name for your new hierarchy.

Step 31: There are many ways to add columns to a hierarchy. Drag the Season, City and NOC_CountryRegion fields onto the hierarchy name (in this case, Locations) until the hierarchy name is highlighted, then release to add them.

Step 32: Right-click EditionID and select Add to Hierarchy. Choose Locations



Step 33: Go back to Excel. In Sheet1, remove the fields from the ROWS area of PivotTable Fields.

Step 34: Remove all the fields from the COLUMNS area.

Step 35: The only remaining fields in the PivotTable fields are Medal in the FILTERS area, and Count of Medal in the VALUES area.

Step 36: From the PivotTable Fields area, drag SDE from the Events table to the ROWS area.

Sport	Count of Medal
Aquatics	3545
Archery	305
Athletics	3411
Badminton	120
Baseball	335
Basketball	940
Basque Pelota	4
Bathtub	290
Bobsleigh	344
Boxing	842
Canoe / Kayak	1002
Cricket	24
Croquet	8
Curling	21
Cycling	1003
Equestrian	675
Fencing	1539
Football	1387
Golf	20
Gymnastics	2060
Handball	886
Hockey	1325
Ice Hockey	596
Judo	435
Lacrosse	59
Luge	139
Modern Pentathlon	174

Step 37: Then drag Locations from the Hosts table into the COLUMNS area. Just by dragging those two hierarchies, your PivotTable is populated with a lot of data, all of which is arranged in the hierarchy you defined in the previous steps.

Step 38: Let's filter that data a bit, and just see the first ten rows of events. In the PivotTable, click the arrow in Row Labels, click (Select All) to remove all selections, then click the boxes beside the first ten Sports.

Sport	Count of Medal
Aquatics	3545
Archery	305
Athletics	3411
Badminton	120
Baseball	335
Basketball	940
Basque Pelota	4
Bathtub	290
Bobsleigh	344
Boxing	842
Canoe / Kayak	1002
Cricket	24
Croquet	8
Curling	21
Cycling	1003
Equestrian	675
Fencing	1539
Football	1387
Golf	20
Gymnastics	2060
Handball	886
Hockey	1325
Ice Hockey	596
Judo	435
Lacrosse	59
Luge	139
Modern Pentathlon	174

Step 39: You can expand any of those Sports in the PivotTable, when we expand the Aquatics sport, we see all of its child discipline elements and their data. When we expand the Diving discipline under Aquatics, we see its child events too, as shown in the following screen. We can do the same for Water Polo, and see that it has only one event.

Step 40: In the PivotTable Fields area, remove Locations from the COLUMNS area.

Step 41: Your PivotTable will have the following screen.

	Medal	All
Row Labels	Count of Medal	
Aquatics	3545	
Diving	84	
plain high diving	9	
plunge for distance	3	
synchronized diving 10m platform	36	
synchronized diving 3m springboard	36	
Swimming	2428	
Synchronized S.	153	
Water Polo	880	
water polo	880	
Archery	305	
Athletics	3411	
Badminton	120	
Baseball	335	
Basketball	940	
Basque Pelota	4	
Biathlon	290	
Bobsleigh	344	
Boxing	842	
Grand Total	10136	

Step 42: Then remove SDE from the ROWS area. You're back to a basic PivotTable.

Step 43: From the Hosts table, drag Season, City, NOC_CountryRegion, and EditionID into the COLUMNS area, and arrange them in that order, from top to bottom.

	Column Labels	All	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Medal	All													
2	Count of Medal	709	709	709	1259	1259	1259	1998	1998	1998	1998	1859	1859	1859	1705
3	Column Labels	Summer	Amsterdam	Amsterdam Total	Antwerp	Antwerp Total	Athens	Athens Total	Atlanta	Atlanta Total	Barcelona	Barcelona Total	Barcelona Total	Barcelona Total	Barcelona Total
4		Amsterdam	NED Total	BEL	BEL Total	GRC	GRC Total	USA	USA Total	ESP	ESP Total	ESP Total	ESP Total	ESP Total	ESP Total
5		1928Summer	1920Summer	1920Summer	2004Summer	2004Summer	1996Summer	1996Summer	1996Summer	1992Summer	1992Summer	1992Summer	1992Summer	1992Summer	1992Summer
6															
7															
8															

Step 44: From the Events table, drag Sport, Discipline, and Event into the ROWS area, and arrange them in that order, from top to bottom.

Step 45: In the PivotTable, filter Row Labels to the top ten Sports.

Step 46: Collapse all the rows and columns

Step 50: Expand Aquatics, then Diving and Water Polo . Your workbook looks like the following screen.

Book1.xlsx - Excel (Product Activation Failed)

PIVOTTABLE TOOLS

FILE HOME INSERT PAGE LAYOUT FORMULAS DATA REVIEW ACROBAT POWERPIVOT Team ANALYZE DESIGN

Manage Calculated Fields KPIs Align Vertically Align Horizontally Add to Data Model All Detect Settings

Data Model Calculations Slicer Alignment Tables Relationships

A15 : Synchronized S.

Count of Medal

Row Labels	Column Labels											
	All	Amsterdam Total	Antwerp Total	Athens Total	Atlanta Total							
1 Medal	75	75	75	320	320	250	250	250	250	250	250	
2												
3 Count of Medal	All	Amsterdam Total	Antwerp Total	Athens Total	Atlanta Total							
4												
5	Summer	NED Total	BEL Total	GRC Total	USA Total							
6	Amsterdam	1920Summer	1920Summer	2004Summer	1996Summer							
7	NED											
8	Aquatics	75	75	75	320	320	250	250	250	250	250	
9	Diving		3	3	24	24	30	30	30	30	30	
10	plain high diving		3	3	24	24	39	39	39	39	39	
11	plunge for distance											
12	synchronized diving 10m platform											
13	synchronized diving 3m springboard											
14	Swimming	51	51	51	185	185	181	181	181	181	181	
15	Synchronized S.	24	24	24	33	33	30	30	30	30	30	
16	Water Polo	24	24	24	78	78	39	39	39	39	39	
17	water polo	24	24	24	78	78	39	39	39	39	39	
18	Archery		80	80	24	24	24	24	24	24	24	
19	Athletics	108	108	108	183	183	180	180	180	180	180	
20	Badminton				24	24	24	24	24	24	24	
21	Baseball				71	71	71	60	60	60	60	
22	Basketball				70	70	70	72	72	72	72	
23	Basque Pelota											
24	Biathlon											
25	Bobsleigh											
26	Boxing	24	24	24	44	44	44	48	48	48	48	
27	Grand Total	207	207	207	736	736	658	658	658	658	658	
28												
29												
30												

Sheet1 Sports Hosts

Activate Windows Go to settings to activate Windows UPDATE

12:19 14-02-2019

Practical 5: Apply the what – if Analysis for data visualization. Design and generate necessary reports based on the data warehouse data.

A book store has 100 books in storage. You sell a certain % for the highest price of \$50

C8				
	X	✓	f _x	=B4*(1-C4)
1	A	B	C	D
2				E
3		total number of books	% sold for the highest price	
4		100	60%	
5				
6			number of books	unit profit
7		highest price	60	\$50
8		lower price	40	\$20
9				
10			total profit	\$3,800
11				

and a certain % for the lower price of \$20.

If you sell 60% for the highest price, cell D10 calculates a total profit of
 $60 * \$50 + 40 * \20
= \$3800.

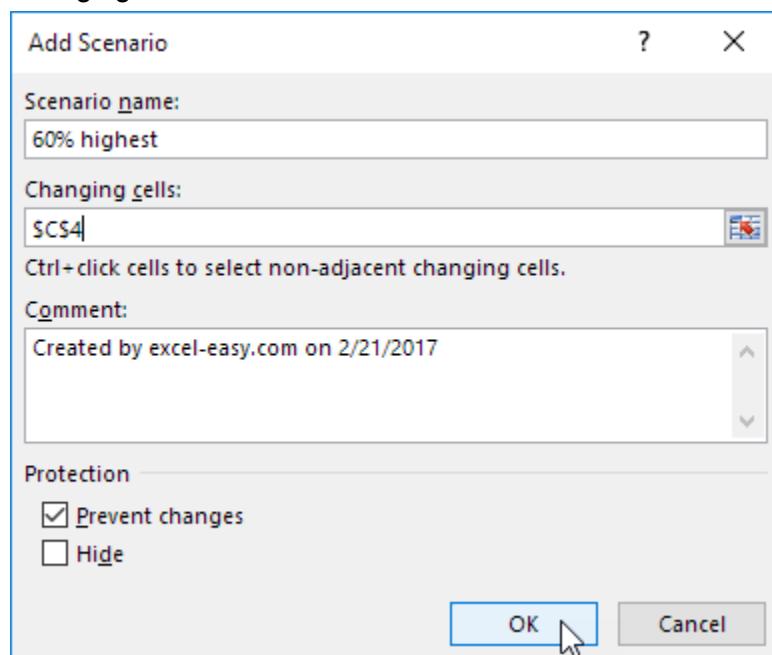
Create Different Scenarios

But what if you sell 70% for the highest price? And what if you sell 80% for the highest price? Or 90%, or even 100%? Each different percentage is a different scenario. You can use the

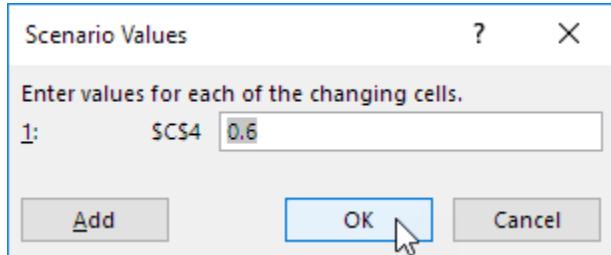
Scenario Manager to create these scenarios.

Note: You can simply type in a different percentage into cell C4 to see the corresponding result of a scenario in cell D10. However, what-if analysis enables you to easily compare the results of different scenarios. Read on

1. On the Data tab, in the Forecast group, click What-If Analysis.
 2. Click Scenario Manager. The Scenario Manager dialog box appears.
 3. Add a scenario by clicking on Add.
 4. Type a name (60% highest), select cell C4 (% sold for the highest price) for the Changing cells and click on OK.

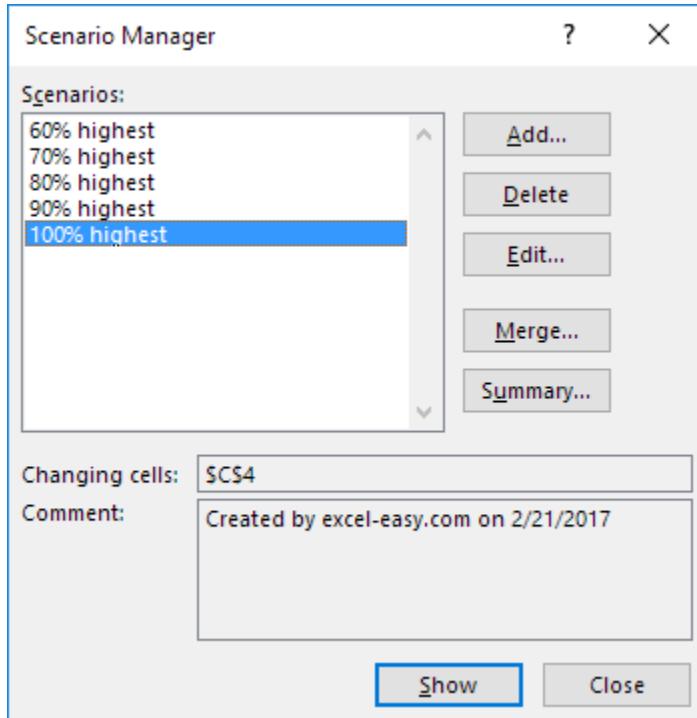


5. Enter the corresponding value 0.6 and click on OK again.



6. Next, add 4 other scenarios (70%, 80%, 90% and 100%).

Finally, your Scenario Manager should be consistent with the picture below:



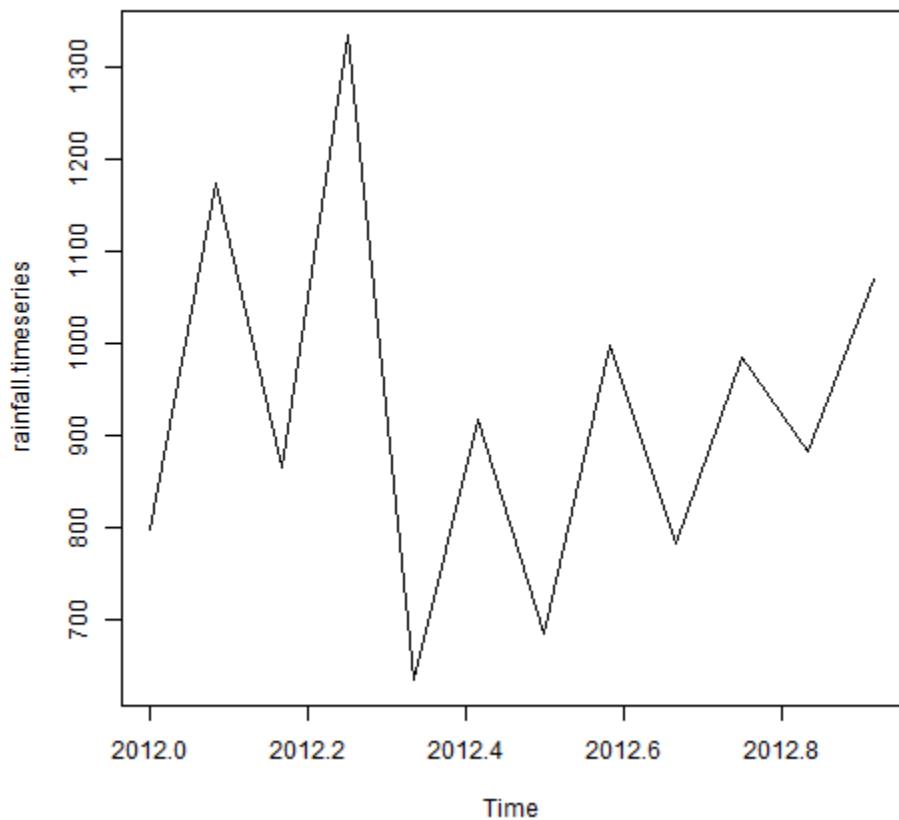
Practical 6: Implementation of Classification algorithm in R Programming.

Consider the annual rainfall details at a place starting from January 2012. We create an R time series object for a period of 12 months and plot it.

```
# Get the data points in form of a R vector.  
rainfall <-  
c(799,1174.8,865.1,1334.6,635.4,918.5,685.5,998.6,784.2,985,882.8,1071)  
# Convert it to a time series object.  
rainfall.timeseries <- ts(rainfall,start = c(2012,1),frequency = 12)  
# Print the timeseries data.  
print(rainfall.timeseries)  
# Give the chart file a name.  
png(file = "rainfall.png")  
# Plot a graph of the time series.  
plot(rainfall.timeseries)  
# Save the file.  
dev.off()
```

Output: When we execute the above code, it produces the following result and chart –

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	
2012	799.0	1174.8	865.1	1334.6	635.4	918.5	685.5	998.6	784.2	
	Oct	Nov	Dec							
2012	985.0	882.8	1071.0							



Practical 7: Practical Implementation of Decision Tree using R Tool

```
install.packages("party")
```

The package "party" has the function ctree() which is used to create and analyze decision tree.

Syntax: The basic syntax for creating a decision tree in R is –

```
ctree(formula, data)
```

Input Data: We will use the R in-built data set named readingSkills to create a decision tree. It describes the score of someone's readingSkills if we know the variables "age", "shoeSize", "score" and whether the person is a native speaker or not.

Here is the sample data.

```
# Load the party package. It will automatically load other
# dependent packages.
library(party)

# Print some records from data set readingSkills.
print(head(readingSkills))
```

When we execute the above code, it produces the following result and chart –

```
nativeSpeaker  age  shoeSize      score
1       yes     5  24.83189  32.29385
2       yes     6  25.95238  36.63105
3       no      11  30.42170  49.60593
4       yes     7  28.66450  40.28456
5       yes     11  31.88207  55.46085
6       yes     10  30.07843  52.83124
```

Loading required package: methods
Loading required package: grid

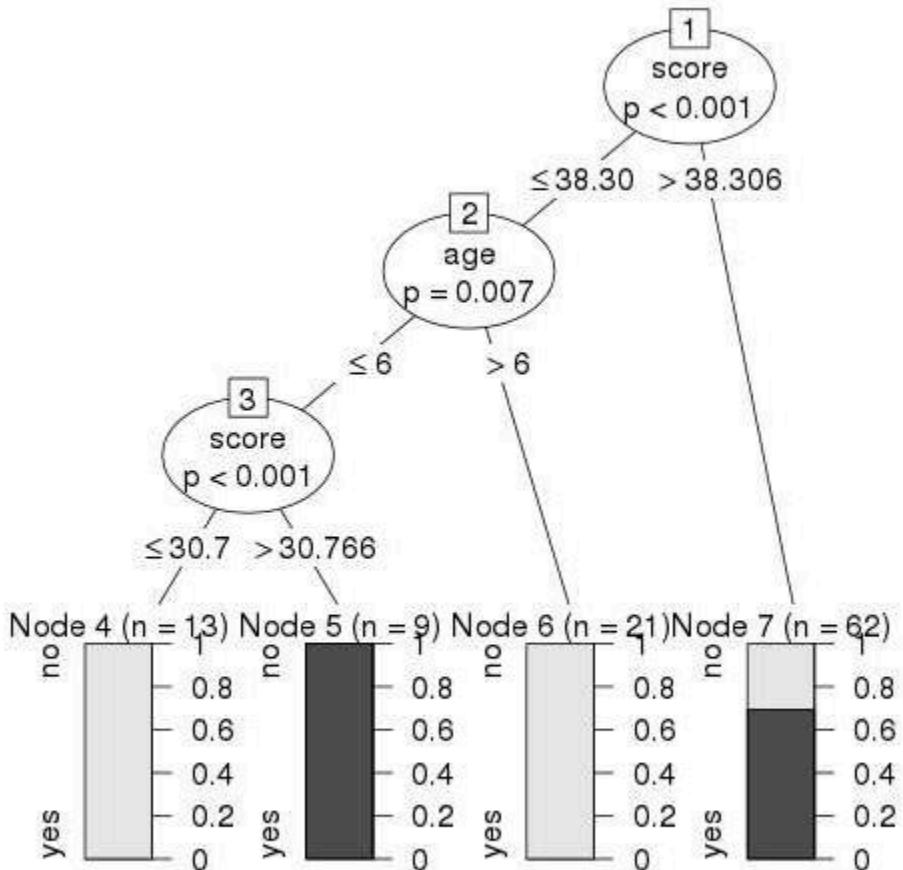
We will use the ctree() function to create the decision tree and see its graph.

```
# Load the party package. It will automatically load other
# dependent packages.
library(party)
# Create the input data frame.
input.dat <- readingSkills[c(1:105),]
# Give the chart file a name.
png(file = "decision_tree.png")
# Create the tree.
output.tree <- ctree(
  nativeSpeaker ~ age + shoeSize + score,
  data = input.dat)
# Plot the tree.
plot(output.tree)
```

```
# Save the file.  
dev.off()
```

Output:

```
null device  
1  
Loading required package: methods  
Loading required package: grid  
Loading required package: mvtnorm  
Loading required package: modeltools  
Loading required package: stats4  
Loading required package: strucchange  
Loading required package: zoo  
  
Attaching package: 'zoo'  
  
The following objects are masked from 'package:base':  
  
as.Date, as.Date.numeric  
  
Loading required package: sandwich
```



Practical 8: Perform the data clustering using clustering algorithm.

Step 1 : Install Library > `install.packages('caret',dependencies=TRUE)`

Step 2 : Install Library > `install.packages('ggplot2',dependencies=TRUE)`

Step 3 : Install Library > `install.packages('lattice',dependencies=TRUE)`

Step 4 : `library('caret')`

Step 5 : Choose file(.csv) > `data<-read.csv(file.choose(),header=FALSE)`

```
> data<-read.csv(file.choose(),header=FALSE)
> require("datasets")
> data("iris")
> str(iris)
'data.frame': 150 obs. of 5 variables:
 $ Sepal.Length: num 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num 3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 $
```

Step 7 : Create a new class having only the data that we want.

```
> iris.new<- iris[,c(1,2,3,4)]
> iris.class<- iris[,"Species"]
> head(iris.new)
  Sepal.Length Sepal.Width Petal.Length Petal.Width
1          5.1        3.5       1.4        0.2
2          4.9        3.0       1.4        0.2
3          4.7        3.2       1.3        0.2
4          4.6        3.1       1.5        0.2
5          5.0        3.6       1.4        0.2
6          5.4        3.9       1.7        0.4
> |
```

Step 8: Normalize the data.

```
> normalize <- function(x){
+   return ((x-min(x))/(max(x)-min(x)))
+ }
> iris.new$Sepal.Length<- normalize(iris.new$Sepal.Length)
> iris.new$Sepal.Width<- normalize(iris.new$Sepal.Width)
> iris.new$Petal.Length<- normalize(iris.new$Petal.Length)
> iris.new$Petal.Width<- normalize(iris.new$Petal.Width)
> head(iris.new)
  Sepal.Length Sepal.Width Petal.Length Petal.Width
1  0.22222222  0.6250000  0.06779661  0.04166667
2  0.16666667  0.4166667  0.06779661  0.04166667
3  0.11111111  0.5000000  0.05084746  0.04166667
4  0.08333333  0.4583333  0.08474576  0.04166667
5  0.19444444  0.6666667  0.06779661  0.04166667
6  0.30555556  0.7916667  0.11864407  0.12500000
> |
```

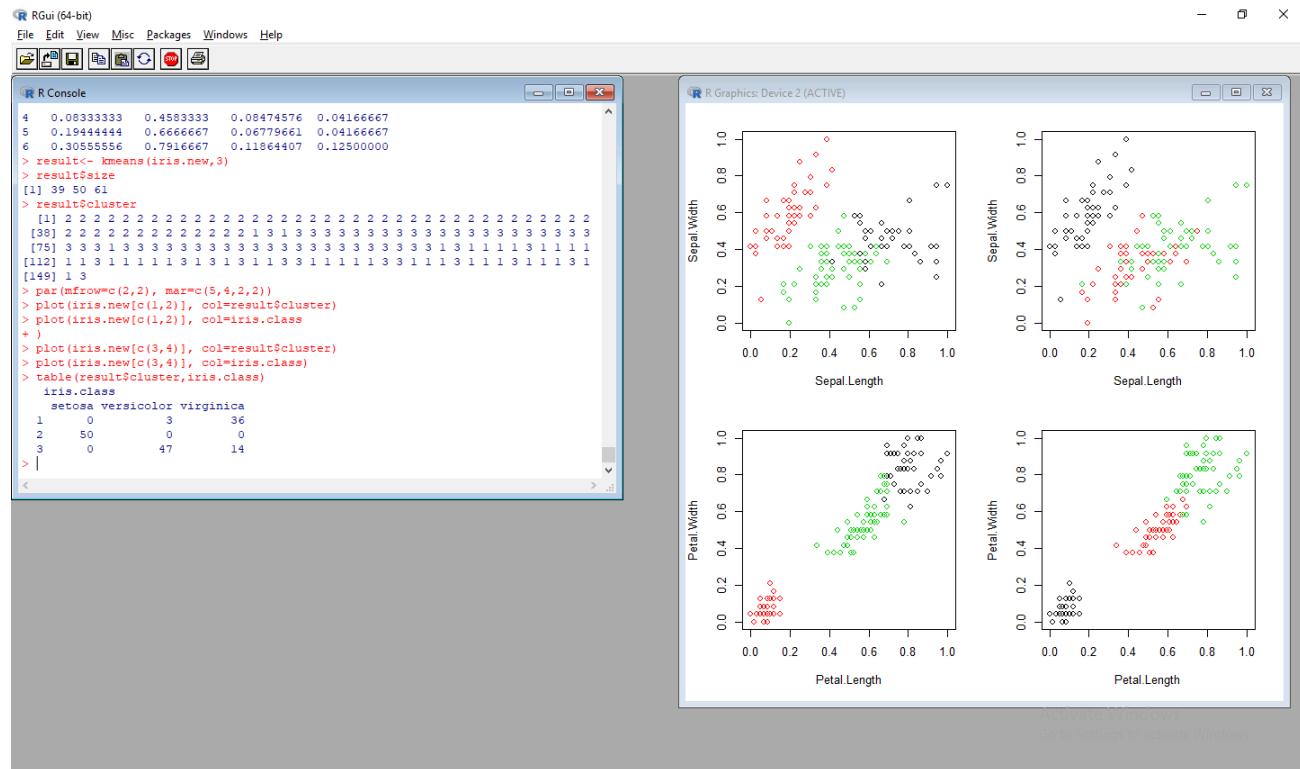
Step 9 : Apply K-Means Clustering with 3 clusters using kmeans function.

```
> result<- kmeans(iris.new,3)
> result$size
[1] 39 50 61
> |
```

Step 10: View Results.

Step 11: Initialize row and column area to draw the clusters graphically.

Step 12 : Draw the clusters with respect to sepal length and width and petal length and width.



Practical 9: Perform Linear Regression on given data warehouse data.

Step 1: Open data in Power BI. (Get Data > Excel > Edit).

Step 2 : Add more columns as per formula of XY, XSquare, YSquare. For performing linear regression.

Step 3 : Right click on FieldsSheet and add a new measure. Which may act as some constant values

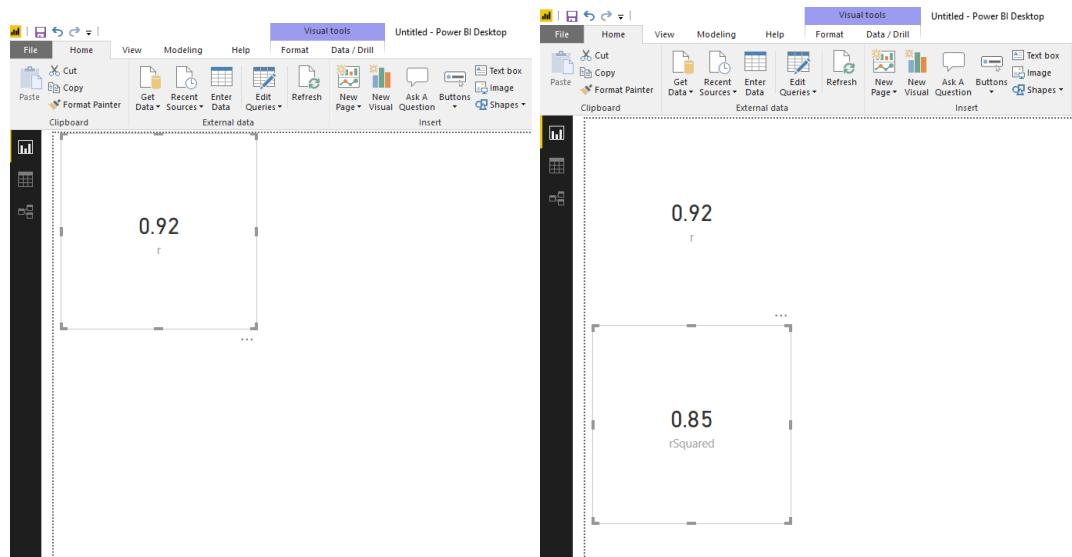
Step 4 : Calculate XSum.

Step 5 : Calculate XYSum.

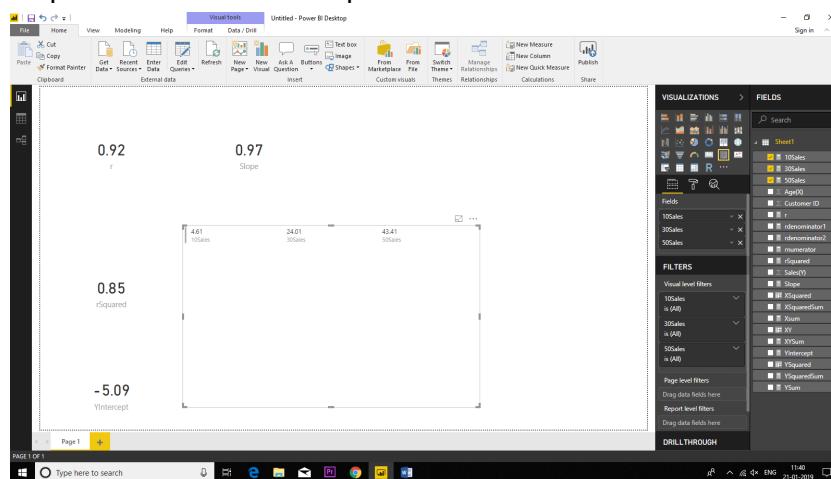
Step 6 : Calculate rnumerator.

Step 7 : Calculate r.

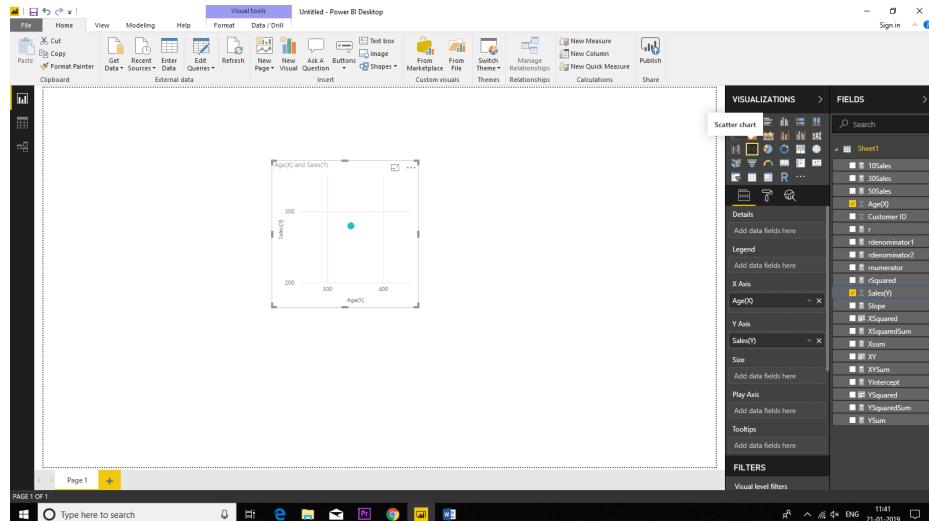
Step 8 : Select the Card Visualizations to view calculated results., Calculate Slope



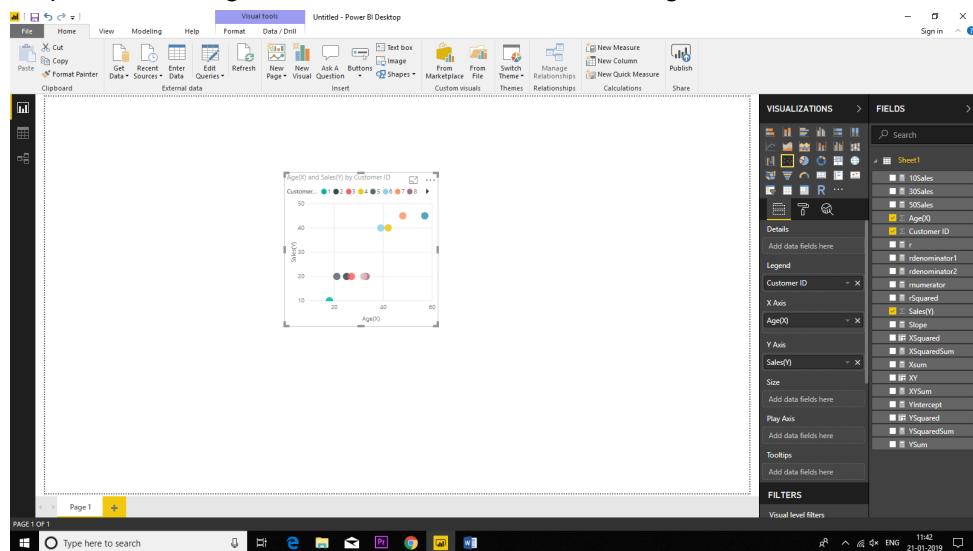
Step 9: Calculate YIntercept



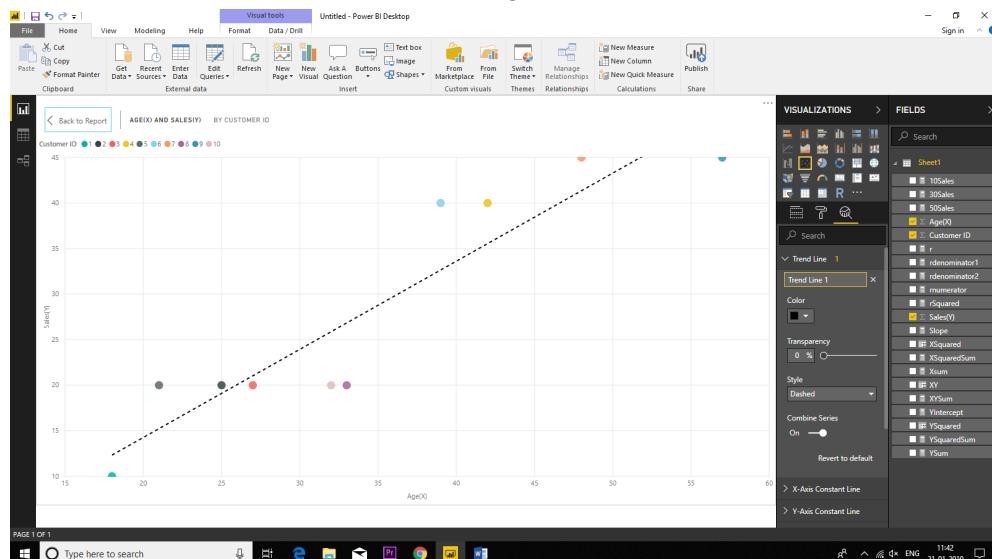
Step 10 : Select Scatter Chart from Visualizations.



Step 13 : Add Legend as Customer ID, X axis as Age and Y axis as Sales.



Step 14 : Select Trend Line to see the graph.



Practical 10: Perform the logistic regression on the given data warehouse data. Data Analysis and Visualization using Excel

Power View is a feature of Microsoft Excel 2013 that enables interactive data exploration, visualization, and presentation encouraging intuitive ad-hoc reporting.

- Create Power View Sheet: Make sure your Power View Add-In Is enabled in Excel 2013
(Note: Install Ms Office 2013 Professional. If Power View Add-In Is not enabled then download and install Silverlight.)

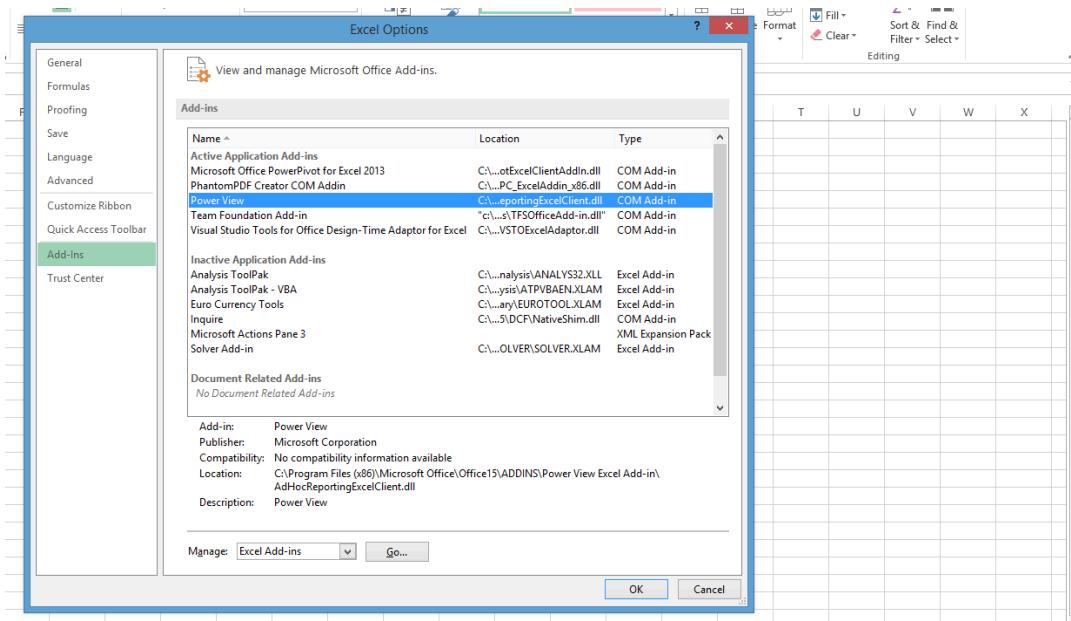
Step 1: Open Excel 2013 (Professional version) → Open blank Workbook

Step 2: Go to File → Options

Step 3 – Click on Add-Ins.

Step 4 – In the Manage box, click the drop-down arrow and select Excel Add-ins.

Step 5 – All the available Add-ins will be displayed. If Power View Add-ins is enabled, it appears in Active Application Add-ins.



Step 6: If Power View does not appear in Active Add-Ins, follow these steps:

Step 1 – In the Excel Options Window, Click on Add-Ins.

Step 2 – In the Manage box, click the drop-down arrow and select COM Add-ins

Step 3 – Click on the Go button. A COM Add-Ins Dialog Box appears.

Step 4 – Check the Power View Check Box.

Step 5 – Click OK.

(Note: Above step is required only if Power View is not present in active Add-In)

Now, you are ready to create the Power View sheet

Step 6: Go to data tab →get external data→ from access → olympicmedals.accdb

The screenshot shows the Microsoft Excel ribbon with the 'DESIGN' tab selected. A table named 'Table_W_Teams' is displayed in the worksheet. The table structure is as follows:

	A	B	C	D
1	Event	Discipline	Sport	Participation
2	Combined 15 + 15km mass start	Cross Country S	Skating	duo
3	Individual	Figure skating	Skating	individual
4	giant slalom	Alpine Skiing	Skating	individual
5	500m	Speed Skating	Skating	individual
6	3000m	Speed Skating	Skating	individual
7	10km pursuit	Biathlon	Biathlon	individual
8	15km	Biathlon	Biathlon	individual
9	Half-pipe	Snowboard	Skating	individual
10	1500m	Short Track S.	Skating	individual
11	curling	Curling	Curling	team
12	Team	Nordic Combined	Skating	team
13	15km	Cross Country S	Skating	individual
14	ice dancing	Figure skating	Skating	duo
15	4x7.5km relay	Biathlon	Biathlon	team
16	aerials	Freestyle Ski.	Skating	individual
17	four-man	Bob sleigh	Bob sleigh	team
18	K90 individual (70m)	Ski jumping	Skating	individual
19	downhill	Alpine Skiing	Skating	individual
20	Alpine combined	Alpine Skiing	Skating	individual
21	K120 team (90m)	Ski jumping	Skating	team
22	10km	Biathlon	Biathlon	individual
23	Snowboard Cross	Snowboard	Skating	individual
24	2000m	Short Track S.	Skating	individual
25	4x5km relay	Cross Country S	Skating	team
26	50km	Cross Country S	Skating	individual
27	3000m	Speed Skating	Skating	individual
28	moguls	Freestyle Ski.	Skating	individual
29	5000m relay	Short Track S.	Skating	team
30	Individual	Skeleton	Bob sleigh	individual

Step 7 – Click on Insert tab.

Step 8 – Click on Power View in Reports Group.

The screenshot shows the Microsoft Excel ribbon with the 'POWERVIEW' tab selected. A Power View sheet is displayed in the worksheet. The left side of the screen shows a table visualization of the data. The right side shows the Power View Fields pane, which includes sections for Filters, Active, and Fields. The Active section shows the fields selected for the visualization: Discipline, Event, Participation, and Sport. The Fields section lists the available fields: Event, Discipline, Sport, and Participation.

The Power View sheet is created for you and added to your Workbook with the Power View. On the Right-side of the Power View, you find the Power View Fields. Under the Power View Fields, you will find Areas.

In the Ribbon, if you click on Design tab, you will find various Visualization options.

- Create Charts and other Visualizations

For every visualization you want to create, you start on a Power View sheet by creating a table, which you then easily convert to other visualizations, to find one that best illustrates your Data.

Step 1 – Under the Power View Fields, select the fields you want to visualize.

Step 2 – By default, the Table View will be displayed. As you move across the Table, on the top-right corner, you find two symbols – Filters and Pop out.

Step 3 – Click on the Filters symbol. The filters will be displayed on the right side. Filters has two tabs. View tab to filter all visualizations in this View and Table tab to filter the specific values in this table only.

- Visualization – Matrix

A Matrix is made up of rows and columns like a Table. However, a Matrix has the following capabilities that a Table does not have –

1. Display data without repeating values.
2. Display totals and subtotals by row and column.
3. With a hierarchy, you can drill up/drill down.

Collapse and Expand the Display

The Matrix Visualization appears.

You can select required fields for matrix and add required row to ΣValues area to get matrix as output

The screenshot shows a Microsoft Excel spreadsheet titled "Book1 - Excel". On the left, there's a table visualization with the following data:

Discipline	Sport	Count of Event
Luge	Luge	2
	Total	2
Nordic Combined	Skiing	3
	Total	3
Short Track S.	Skating	5
	Total	5
Snowboard	Skiing	3
	Total	3
Speed Skating	Skating	7
	Total	7
	Total	20

To the right of the table is a "Filters" pane and a "Power View Fields" pane. The "Filters" pane shows a single filter for "Count of Event" (All) with a value of 2. The "Power View Fields" pane shows fields from "Table_W_Teams" selected: Discipline, Event, Participation, and Sport. The "Σ VALUES" section contains the formula "# Count of Event". The "ROWS" section has Discipline and Sport selected. The "COLUMNS" section is empty.

- Visualization – Charts

Step 1 – Create a Table Visualization from Medals data.

TYBScIT Semester - VI

53003160083

Business Intelligence

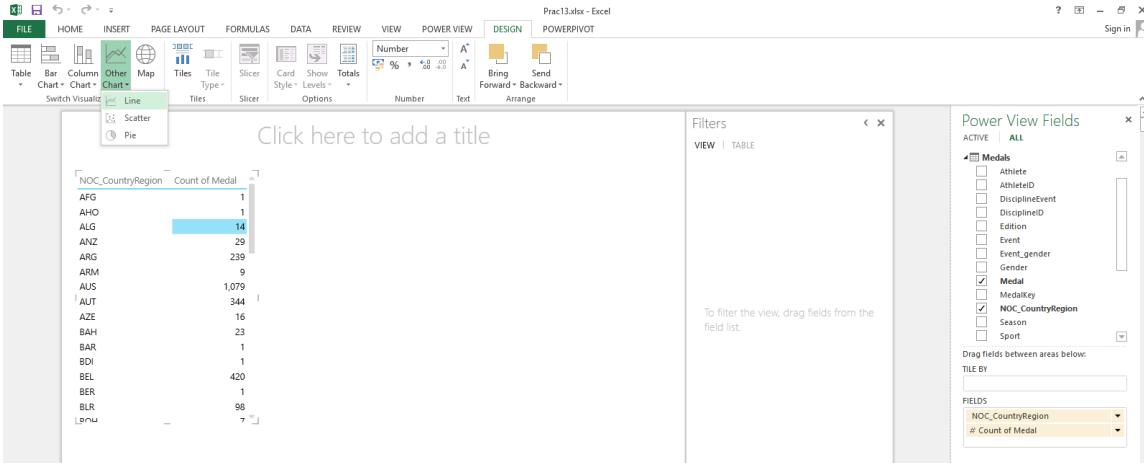
You can use Line, Bar and Column Charts for comparing data points in one or more data series. In these Charts, the x-axis displays one field and the y-axis displays another, making it easy to see the relationship between the two values for all the items in the Chart.

Line Charts distribute category data evenly along a horizontal (category) axis, and all numerical value data along a vertical (value) axis.

Step 2 – Create a Table Visualization for two Columns, NOC_Country Region and Count of Medal.

Step 3 – Create the same Table Visualization below.

Step 4 – Click on Other Chart in the Switch Visualization group.

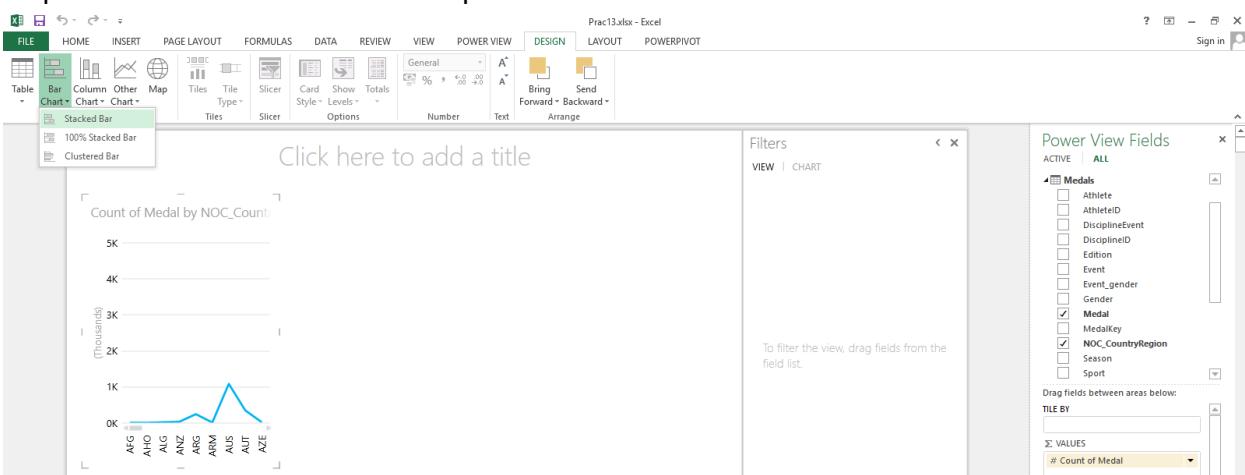


In a Bar Chart, categories are organized along the vertical axis and values along the horizontal axis. In Power View, there are three subtypes of the Bar Chart: Stacked, 100% stacked, and Clustered.

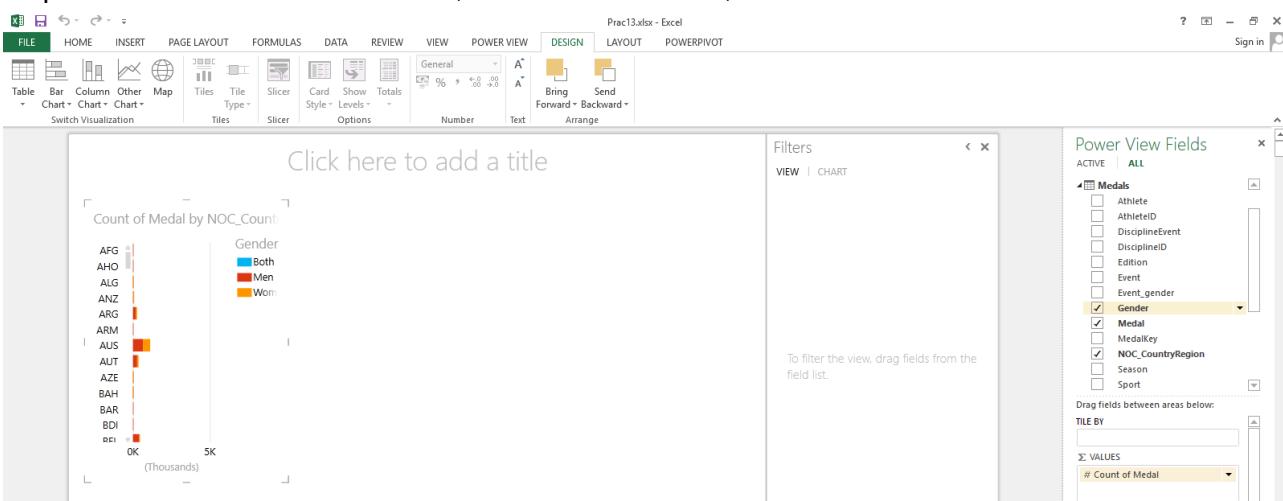
Step 7 – Click on the Line Chart Visualization.

Step 8 – Click on Bar Chart in the Switch Visualization Group.

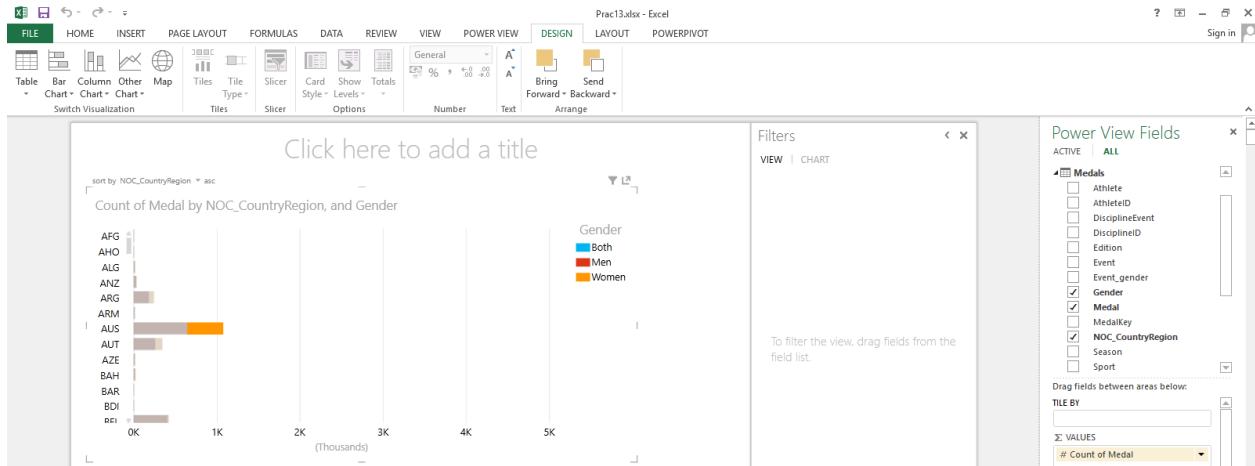
Step 9 – Click on the Stacked Bar option.



Step 10 – In the Power View Fields, in the Medals Table, select the Field Gender also.



Step 11 – Click on one of the bars. That portion of the bar is highlighted. Only the row containing the Data specific to the selected bar is displayed in the table above.



You can use the column charts for showing data changes over a period of time or for illustrating comparison among different items. In a Column Chart, the categories are along the horizontal axis and values are along the vertical axis.

In Power View, there are three Column Chart subtypes: Stacked, 100% stacked, and Clustered.

Step 12 – Click on the Stacked Bar Chart Visualization.

Step 13 – Click on Column Chart in the Switch Visualization group.

