

Sri Lanka Institute of Information Technology



Data Warehousing and Business Intelligence

Assignment -02

Submitted by: - IT20268244

Kaleem F.A

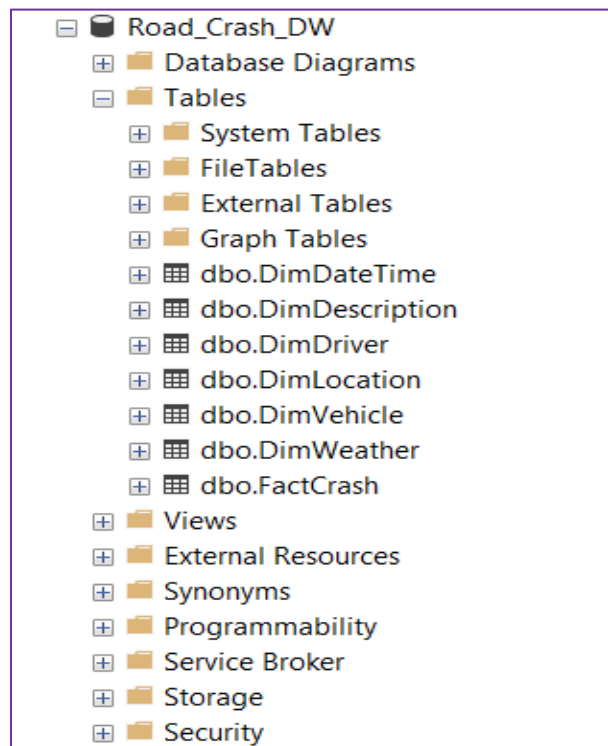
Table of Contents

Data Source Selection	3
ER Diagram	4
High Level Design Architecture	5
Relational Schema	6
SSAS Cube Implementation	7
Data Source View	8
Cube Structure	9
Hierarchies	13
KPI Creation.....	14
Role.	16
OLAP Operation Demonstration	17
Drill Down – OLAP Operation.....	19
Roll Up - OLAP Operation.....	20
Pivot - OLAP Operation	21
Slice - OLAP Operation	22
Dice - OLAP Operation	23
SSRS Reports	24
Report With A Matrix	25
Report with Multiple Parameter	26
Drill Down SSRS Report	27
Drill Through SSRS Report	28

Data Source Selection

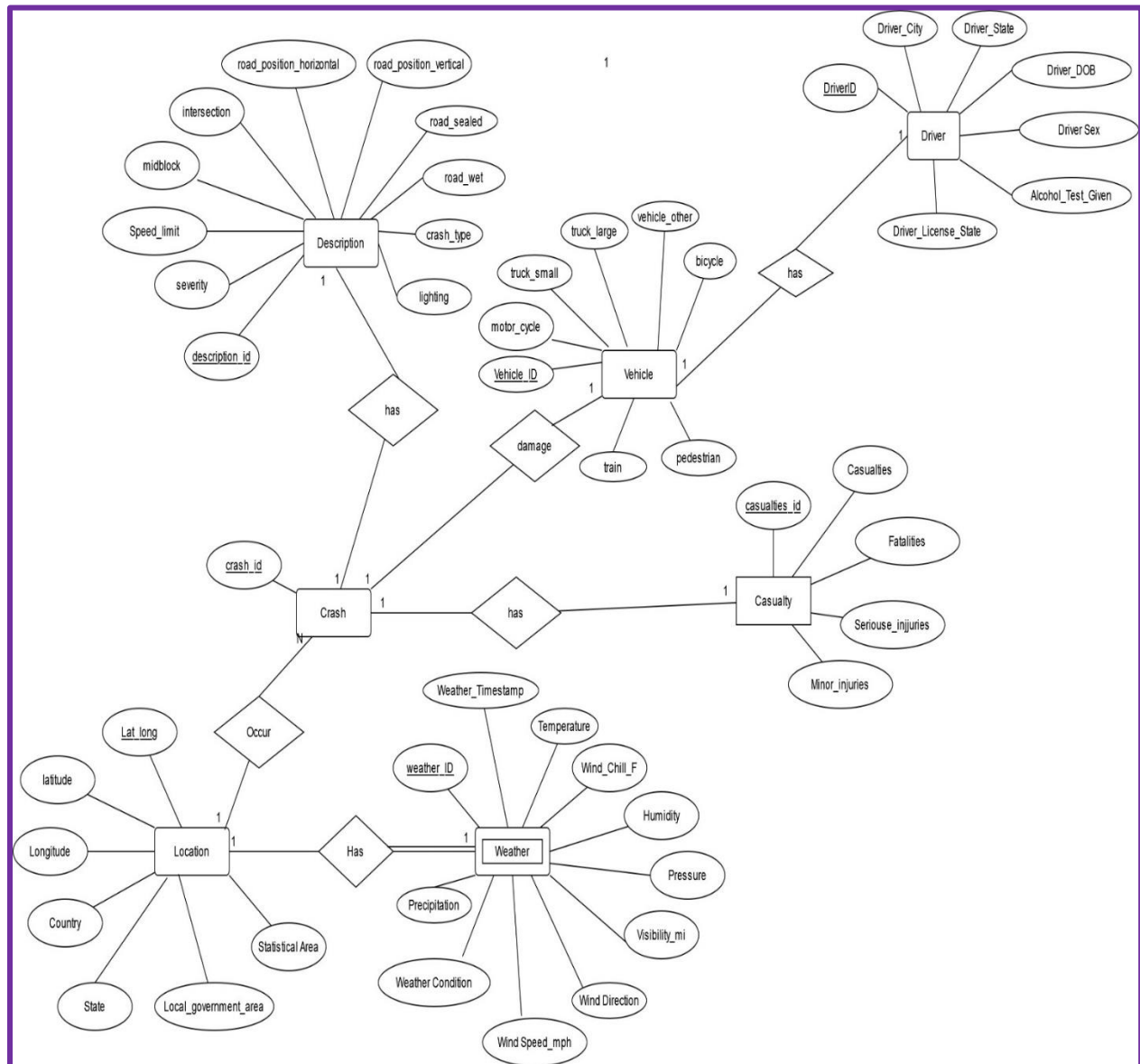
This is a dataset that contains data about on when, where and under what conditions car accidents occur in Australia and New Zealand. It records every information related to each accident which includes the vehicle type that got accident, the driver details who drove the vehicle, the location details including detailed weather condition of the specified location. Moreover, it also describes details such as the road condition, lightning in the road and what type of crash occurred. Finally, it records the most important detail which is the number of casualties, fatalities and injuries that occurred due to the accident.

Thereby to visually represent these data the Road_Crash_DW database has been used. Following is the Structure of the Datawarehouse table and the tables used to analyze and create charts and graphs for the selected dataset.

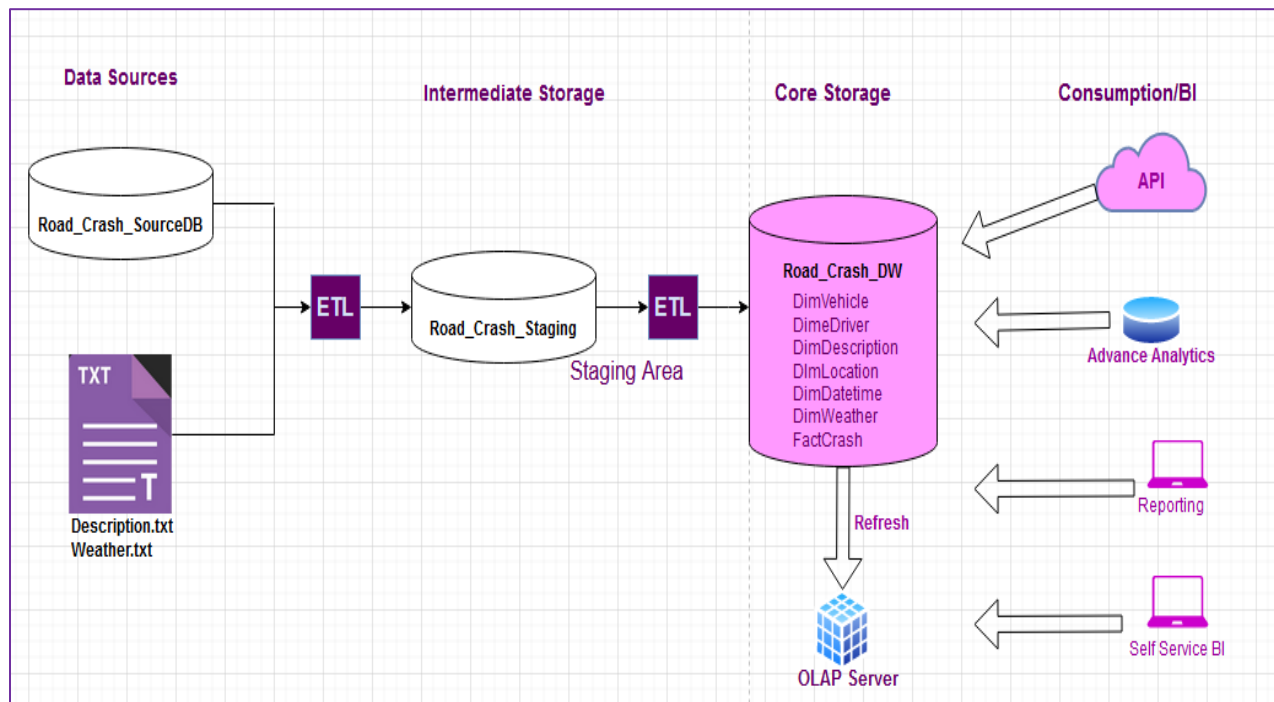


ER Diagram

Following is a ER diagram drawn to represent the relation between tables in the selected dataset in order to depict the relationship between each table.



High Level Design Architecture



The high-level architecture solution of the Australia & New Zealand Road Crash is as given above. The entire dataset is provided by two separate source formats: Source database and Text respectively.

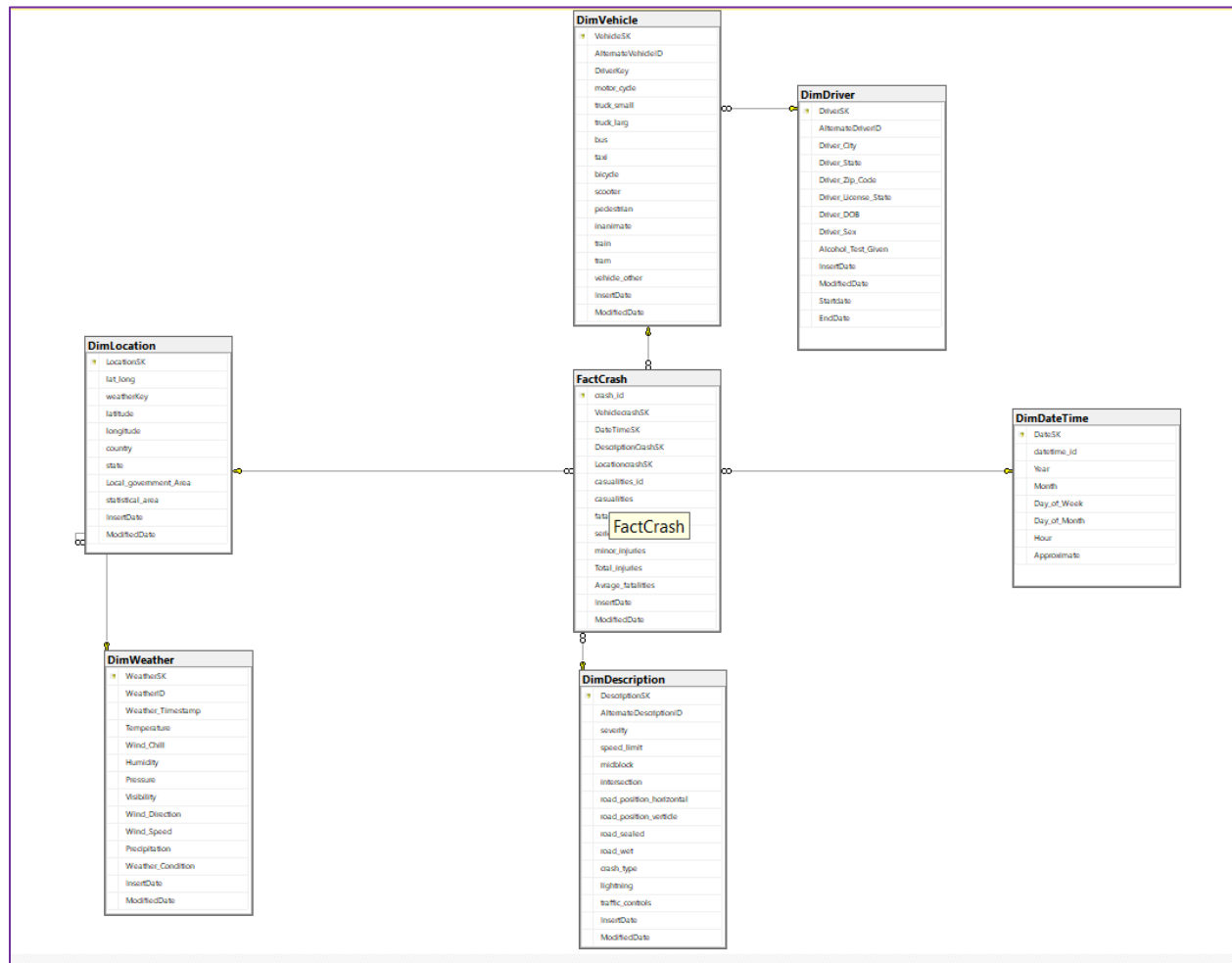
As the first step the entire source data *Road_crash_SourceDB* , *Description.txt* and *Weather.txt* is loaded to the staging database named *Road_Crash_Staging*. This database ultimately serves as a single database including all source data as one database.

AS the next step once the data is loaded to the staging area the data is cleaned, validated and then necessary transformations are done in order to load data from the staging area to the Datawarehouse(*Road_Crash_DW*).

The data from data warehouse is then refreshed to create OLAP cubes which can be used by end users in order to carry out Analysis on the data set. Here data visualization can be either done through OLAP server or directly from Data warehouse which is specifically called Self-service BI.

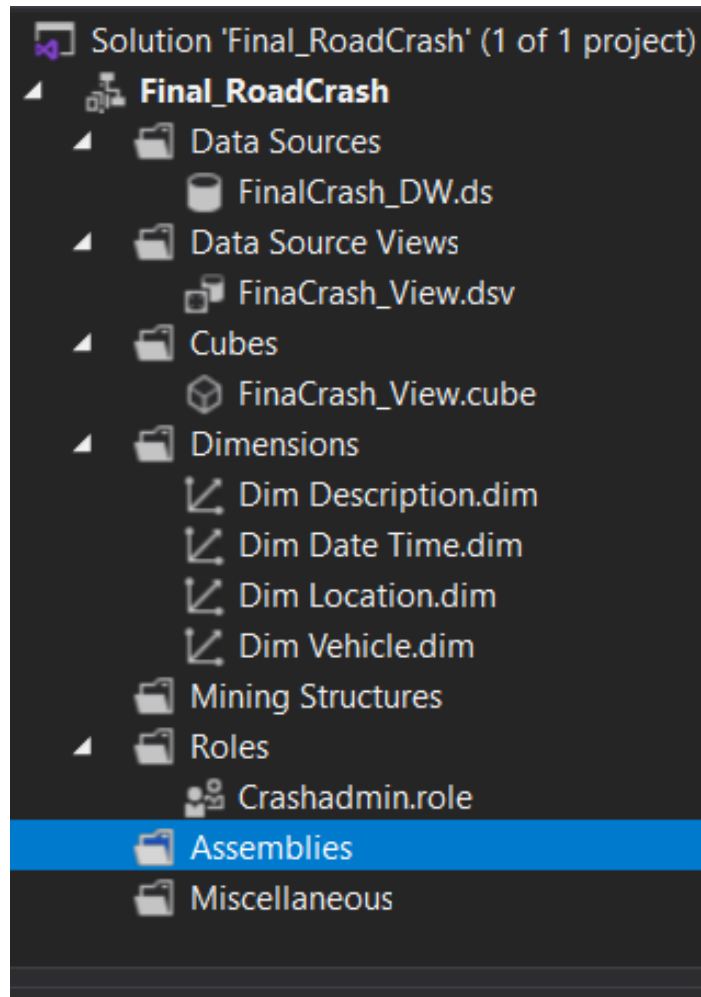
Relational Schema

The Schema designed for the Australia & New Zealand Road Crash Dataset is a snowflakes schema with six-dimension tables and one Road_Crash fact table. Thereby the entities are normalized. Following is the snowflake schema for the model,



SSAS Cube Implementation

SQL server data tools 2019 was used to create Multidimensional analysis service and Data Mining Project to create the cube structure for the above dataset. (Road_Crash_DW - Datawarehouse)

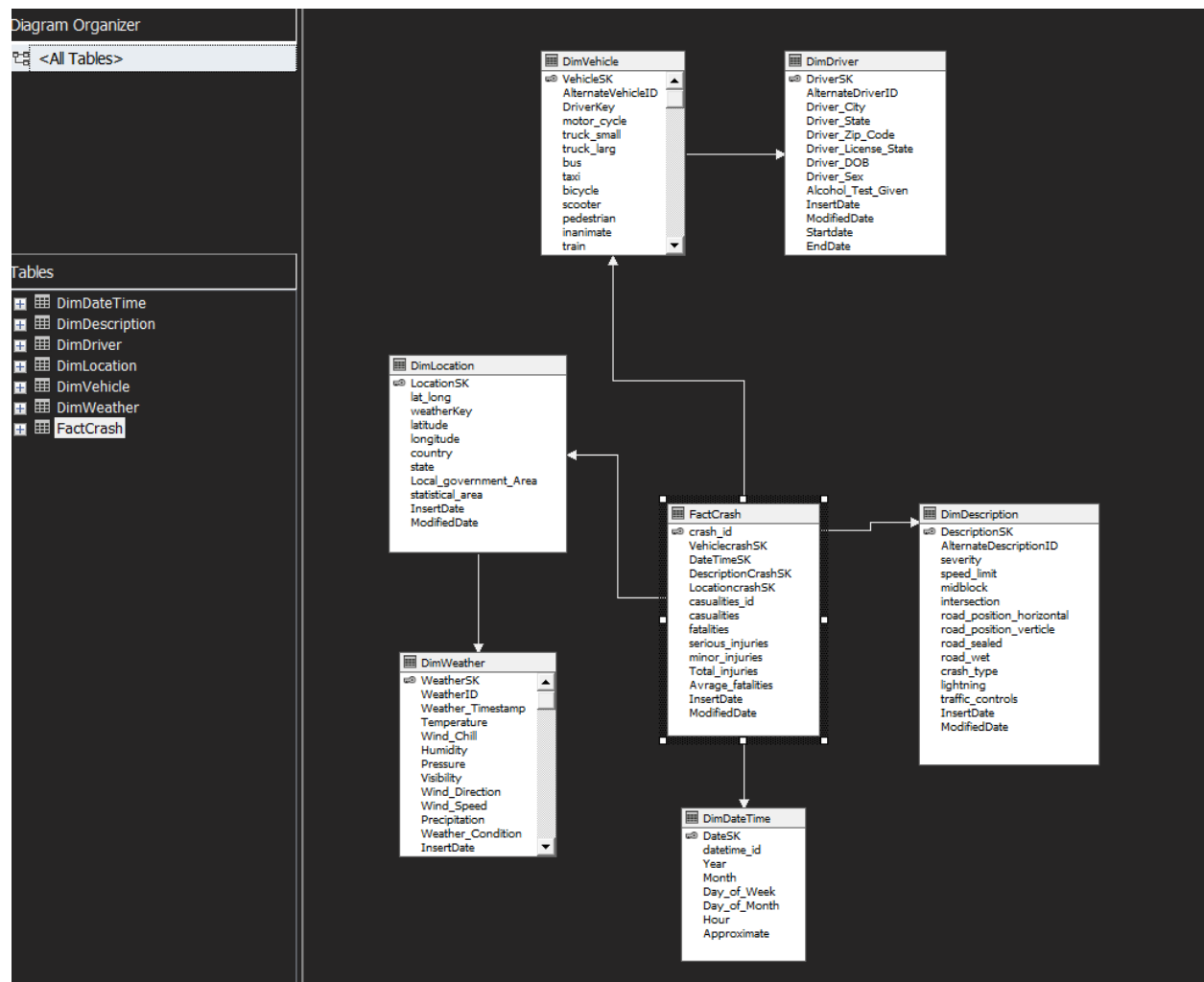


Accordingly to create the above cube a , following steps were carried out,

1. Data source Creation. (FinalCrash_DW)
2. Data Source View Creation. (FinalCrash_View)
3. Cube creation. (FinalCrash_View.Cube)

Data Source View

Once a data source was created the data source view was created by making use of the source Datawarehouse Road_Crash_DW.

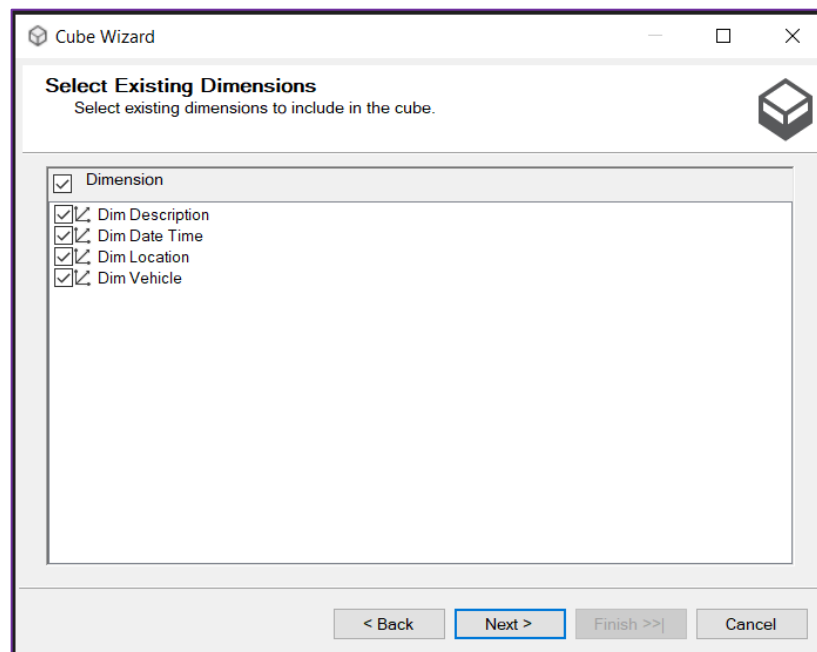
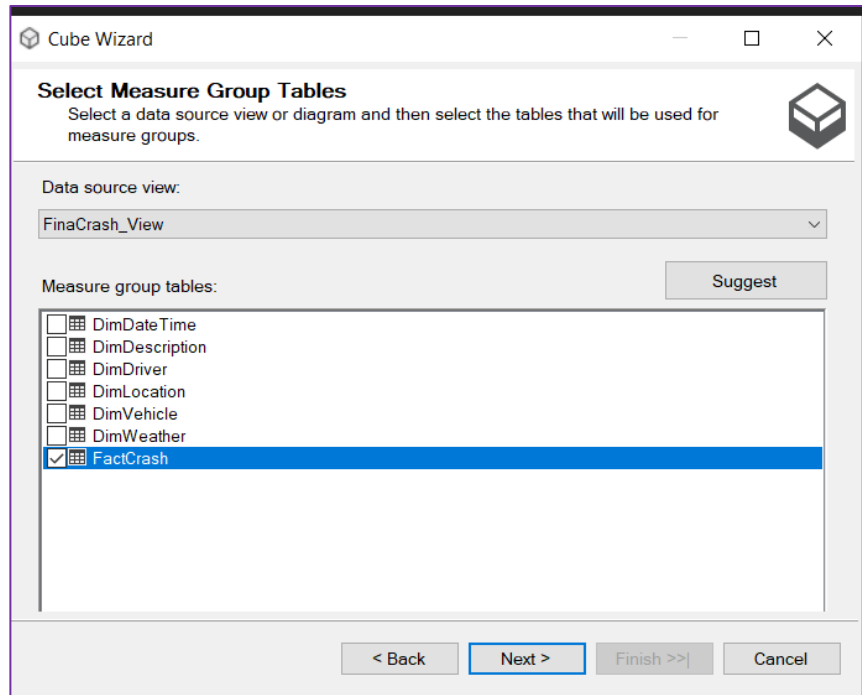


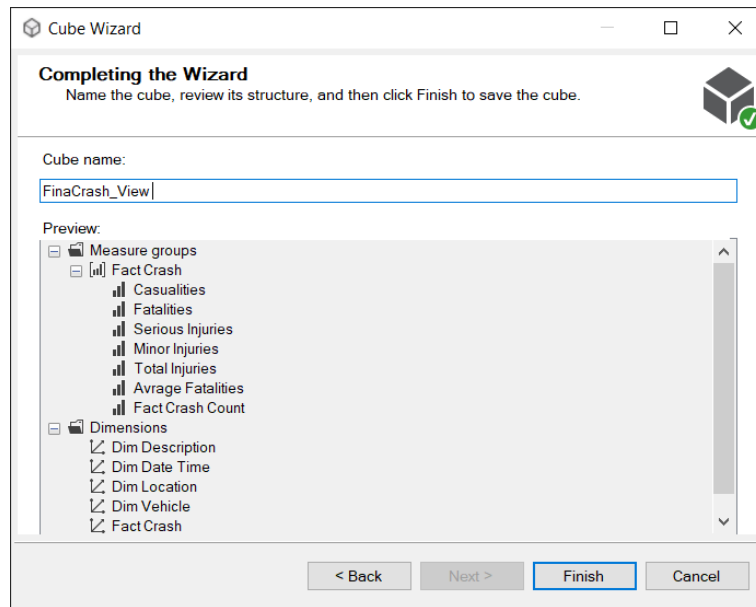
The relationship among the tables were created using the Surrogate key of each table and foreign key of the fact table.

Cube Structure

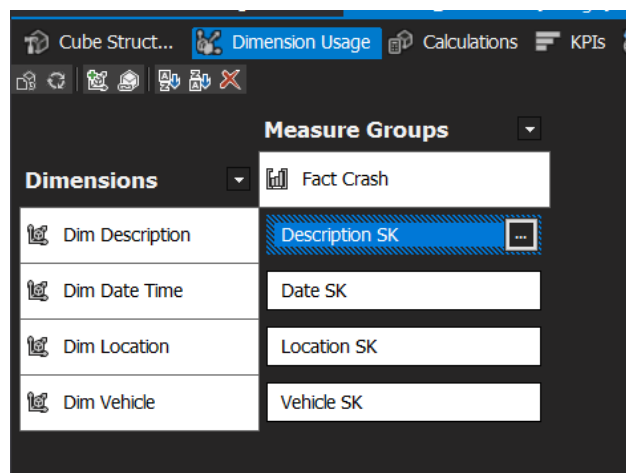
As the final step the cube structure (FinalCrash_view_cube) was designed by using the data source view created in the above step.

Next following steps were proceeded in order to deploy the cube,





- The following snap shows how the measure table is connected with the dimension tables.

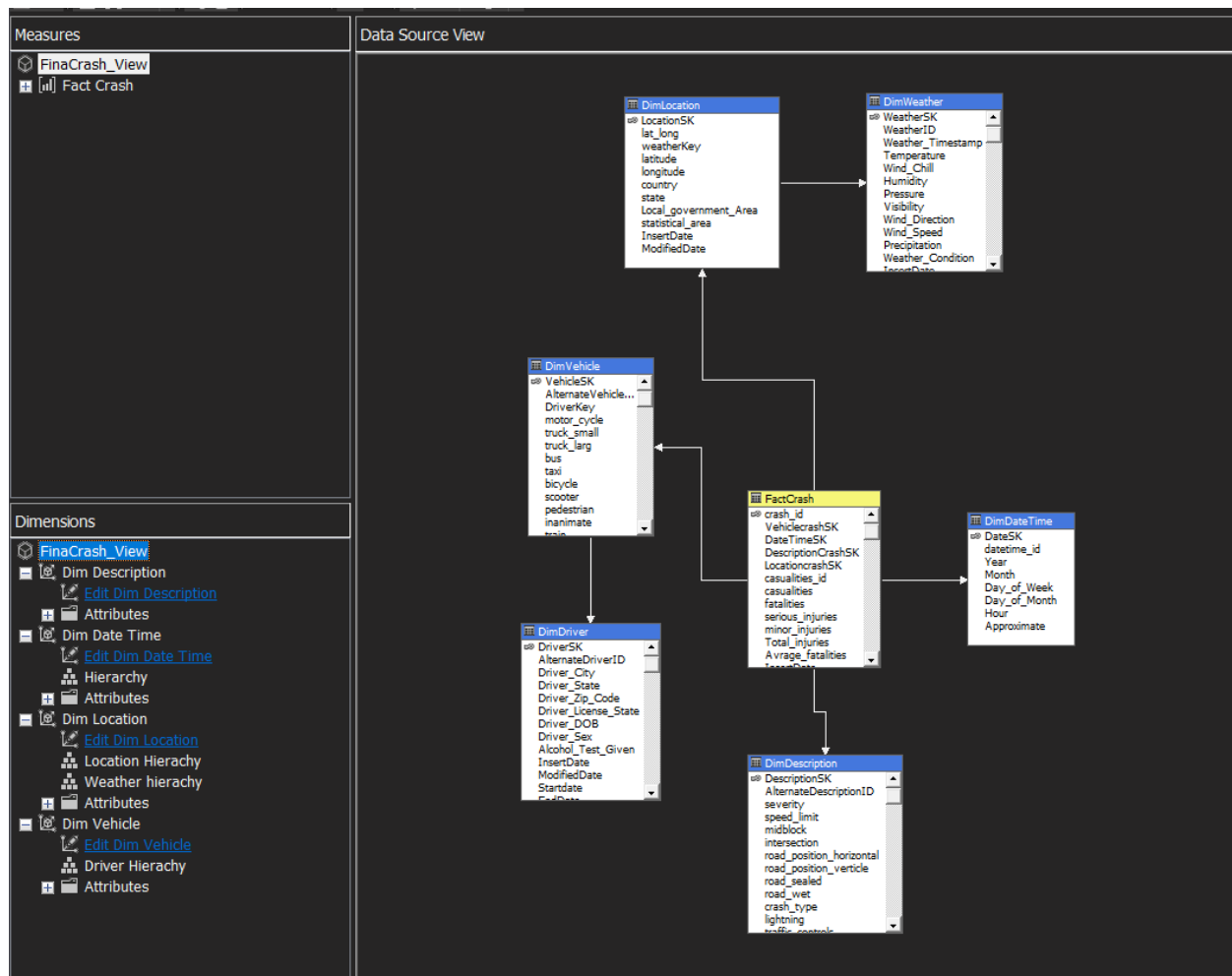


The final cube structure created is as shown below.

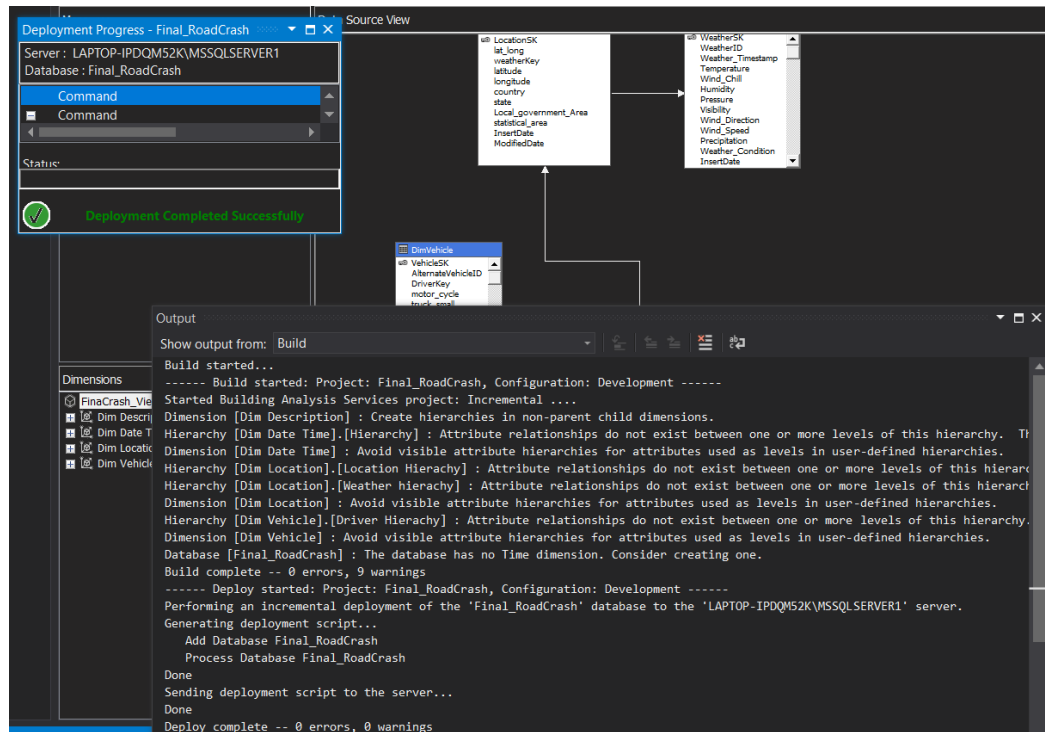
Accordingly, 7 tables are created including both dimension and fact measure table.

The main 4 dimension tables connected to the fact table is , DimLocation, DimDescription, DimVehicle and DimDatetime.

Finally the DimDriver and DimWeather relate to DimVehicle and DimLocation as hierarchical implementations respectively.



Finally the cube was deployed after following the above steps

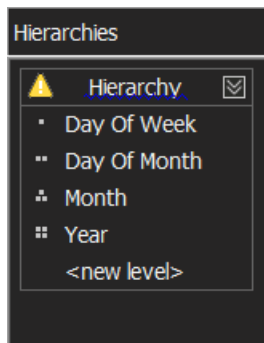


Hierarchies

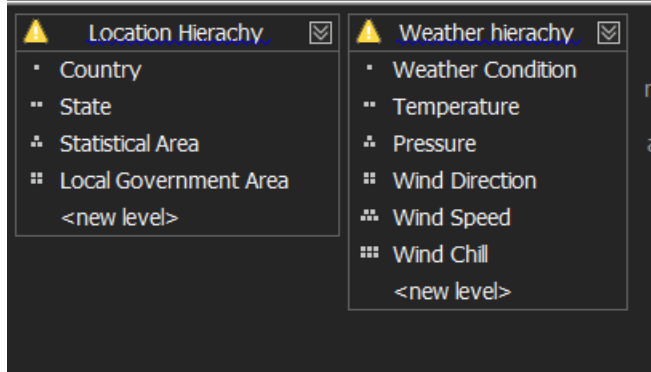
According to the data loaded in data warehouse, the built snowflake schema reflects many hierarchical relationships . Hierarchies are created for Driver, Date, Location and Weather. In addition, hierarchical relationships between Vehicle and driver, Location and weather also exists. The main purpose of hierarchy creation is to provide the ability to drill down to detailed level data and to roll up to create aggregated data so that it will make the business analysis purposes efficient.

Following Hierarchies were created while creating the cube,

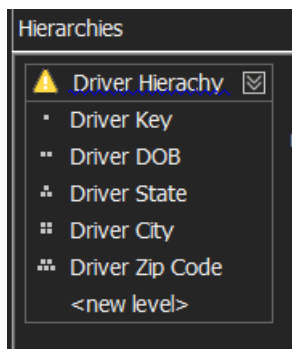
Date Hierarchy derived from the Dim DateTime Table



Location and Weather Hierarchy Derived from the Location table,



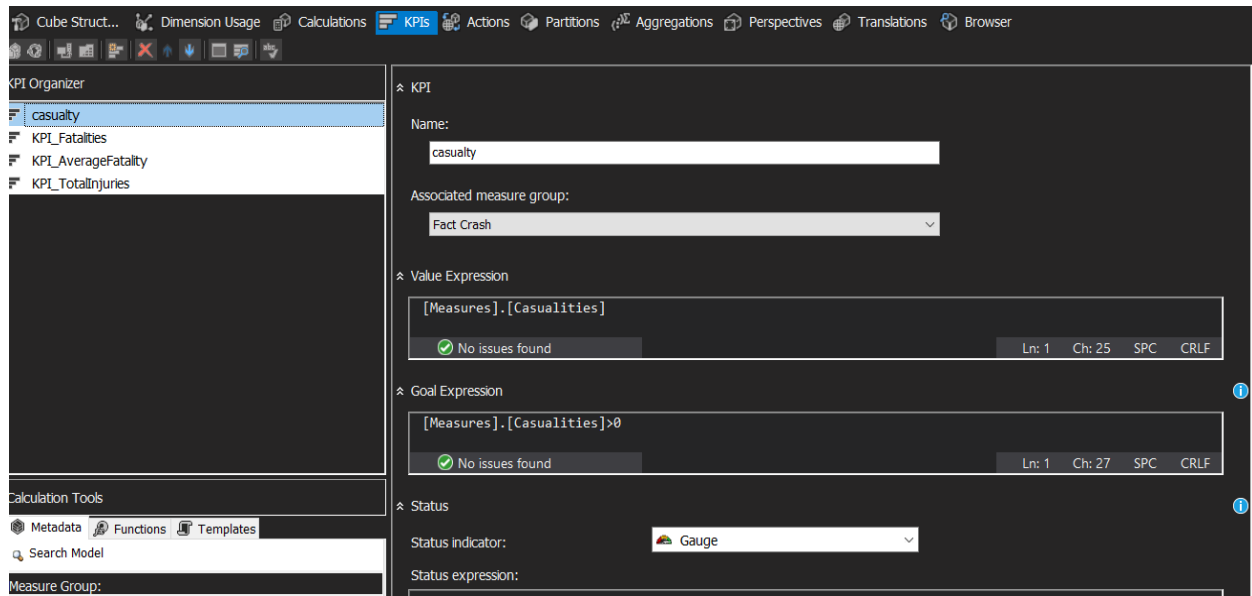
Driver Hierarchy Derived from the Drive table



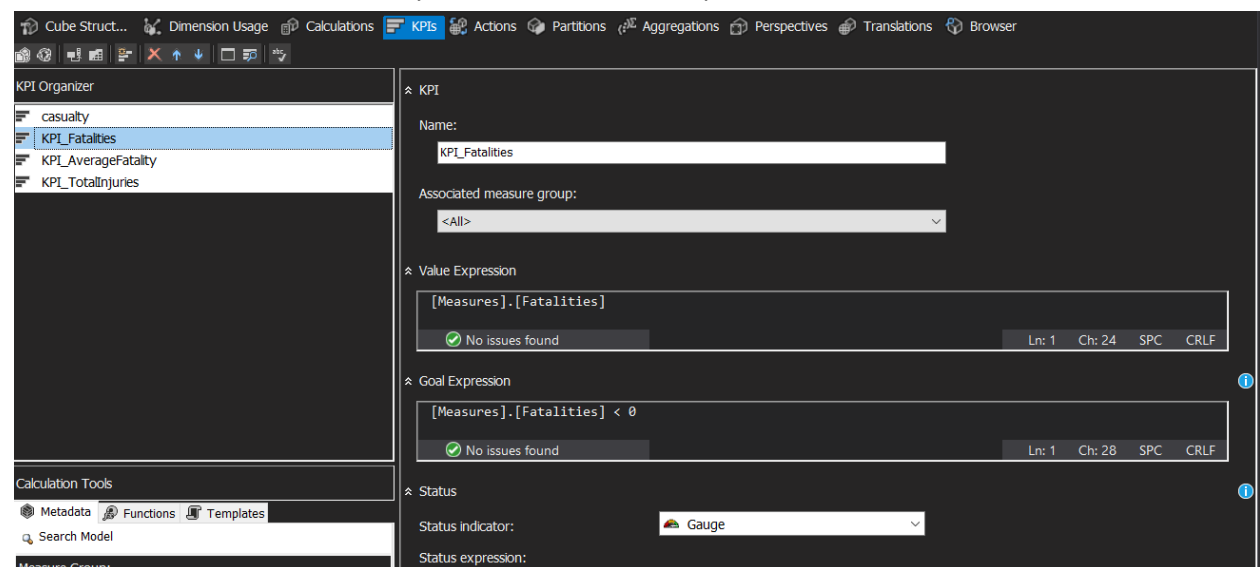
KPI Creation

KPI (Key Performance Indicator) shows the amount of progress that is done to achieve a business goal. Thereby the measure table FactCrash is used to create the following KPI's.

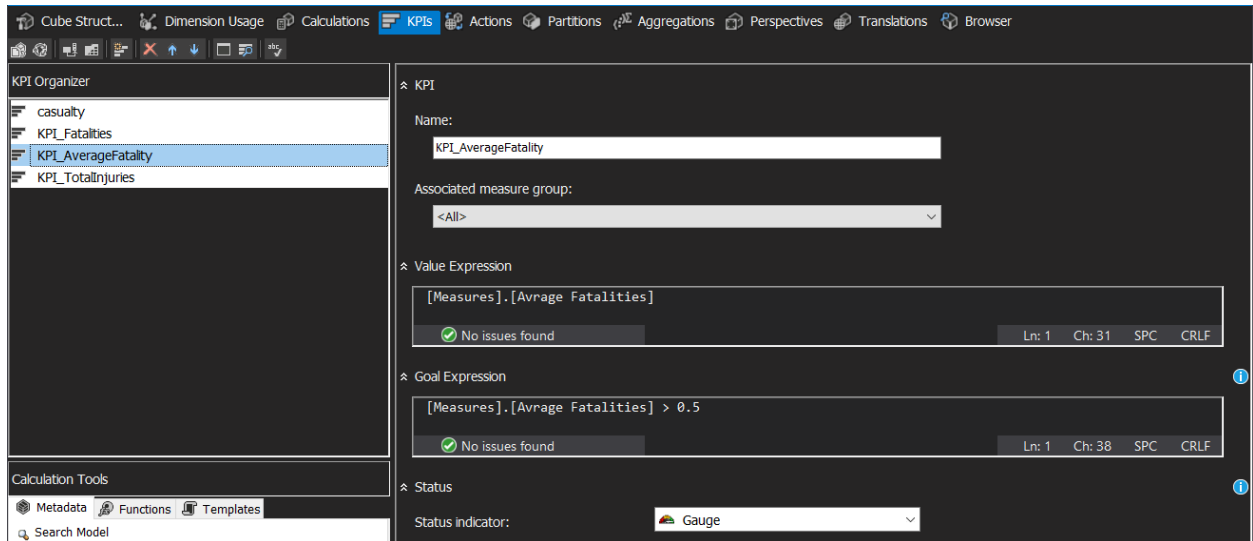
- This is KPI created to check if casualty is greater than 0 if it is greater goal is true else false. Thereby the accidents that was caused due to faults of the third parties can be identified



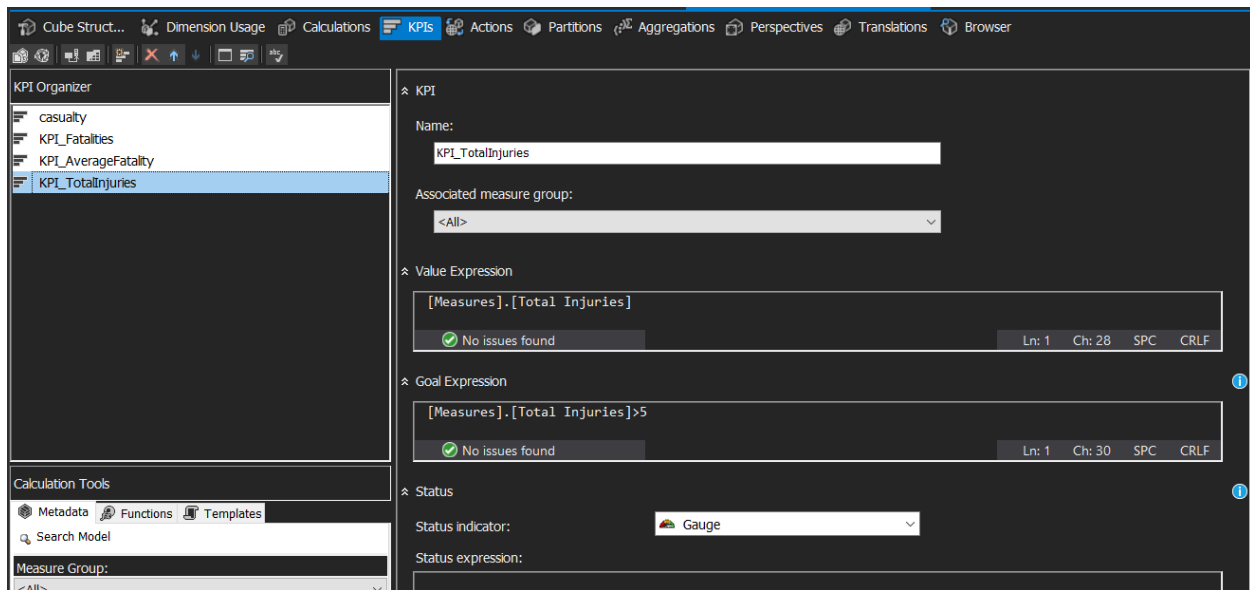
- Here a KPI is created to check if the fatality count is less than 0. Thereby if it is less then the measure is true else false. Accordingly, we could identify if the accident is due to the drivers mistake or not and identify the reasons to further predictions.



- This is a KPI created to check if the average fatality rate is greater than 0.5 or not in order to identify the severity of the accidents caused.

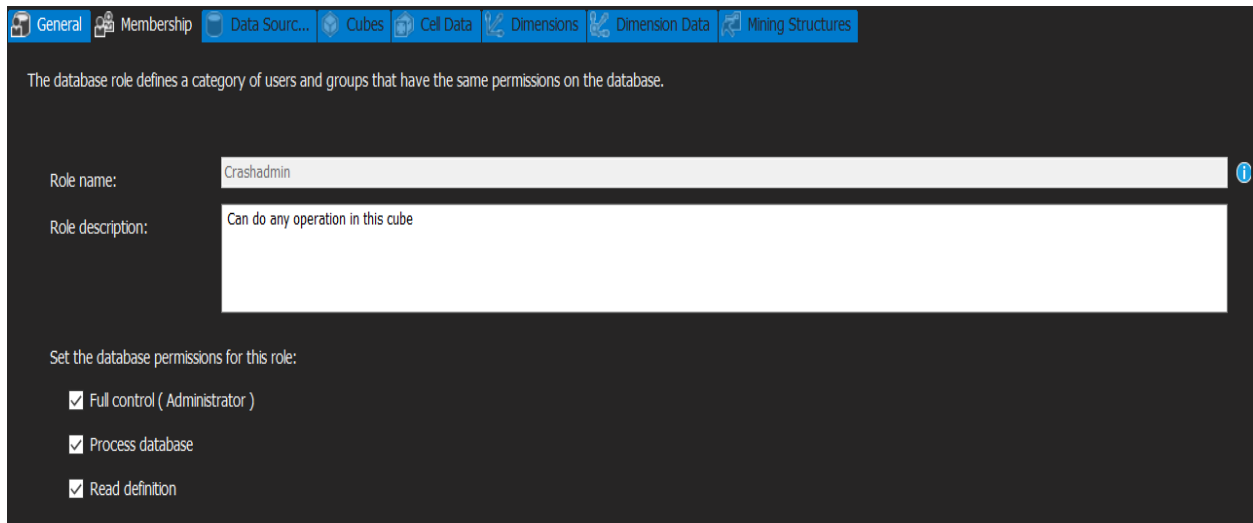


- This is a KPI created to check if the Total injury count is greater than 5 or not , which also can be used to identify the severity of the accident.



Role.

A role is created in order to provide access to users to ensure who can do what. A role can be created by allowing only specific permissions. In this step a fully control role (Administrator) with all permissions is created.



The screenshot shows a software interface with a dark theme. At the top, there is a horizontal menu bar with several tabs: 'General' (selected), 'Membership', 'Data Sourc...', 'Cubes', 'Cel Data', 'Dimensions', 'Dimension Data', and 'Mining Structures'. Below the menu bar, a text box states: 'The database role defines a category of users and groups that have the same permissions on the database.' Below this, there are two input fields. The first is labeled 'Role name:' and contains the text 'Crashadmin'. The second is labeled 'Role description:' and contains the text 'Can do any operation in this cube'. Below these fields, there is a section titled 'Set the database permissions for this role:' followed by three checked checkboxes: 'Full control (Administrator)', 'Process database', and 'Read definition'.

General Membership Data Sourc... Cubes Cel Data Dimensions Dimension Data Mining Structures

The database role defines a category of users and groups that have the same permissions on the database.

Role name: Crashadmin

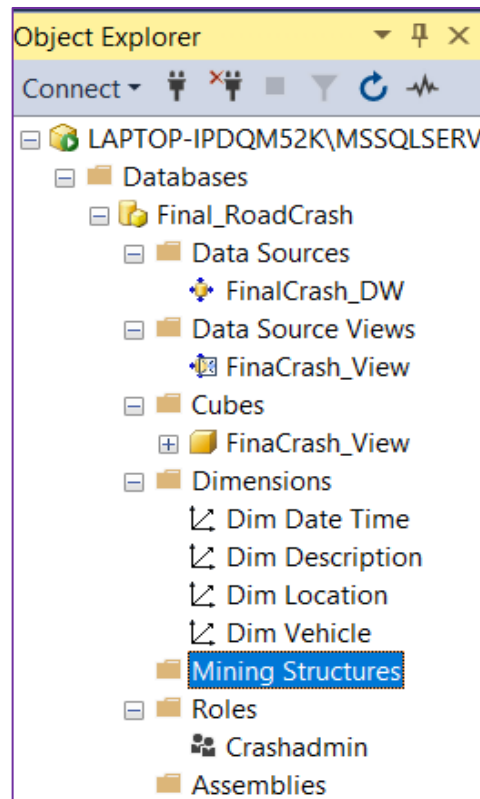
Role description: Can do any operation in this cube

Set the database permissions for this role:

- ☒ Full control (Administrator)
- ☒ Process database
- ☒ Read definition

OLAP Operation Demonstration

Once the cube was deployed successfully, it was able to be viewed in the SQL Server Management Studio under Analysis service.



Browsing the cube

Using the Created cube, we can browse through it and create Queries to be passed as MDX queries so that it can be used for visualizing and analyzing services.

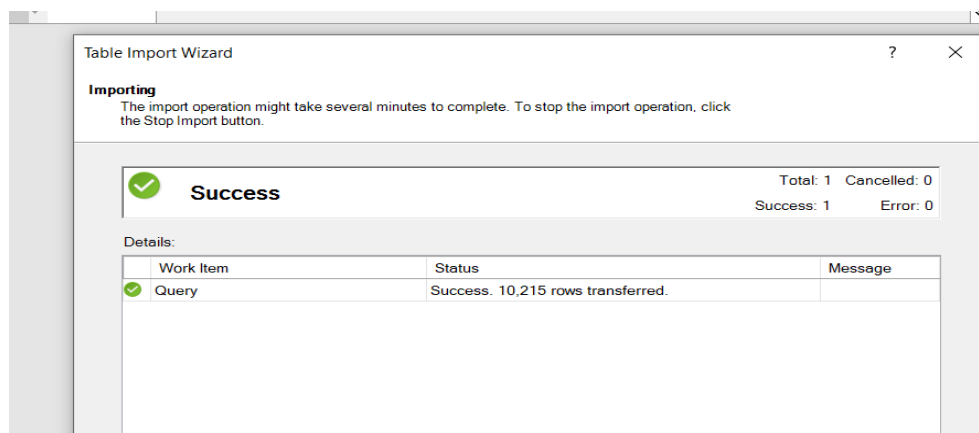
The screenshot shows the 'FinaCrash_View' cube browser in SQL Server Analysis Services. The 'Measure Group' is set to '<All>'. The table displays data for the year 2006, with columns for Year, Month, State, Casualties, Fatalities, casualty Goal, KPI_Fatalities Goal, KPI_AverageFatality Goal, and KPI_TotalInjuries Goal.

Year	Month	State	Casualties	Fatalities	casualty Goal	KPI_Fatalities Goal	KPI_AverageFatality Goal	KPI_TotalInjuries Goal
2006	1	SA	180	5	True	False	True	True
2006	1	VIC	57	1	True	False	True	True
2006	1	Unk...	(null)	(null)	False	False	False	False
2006	10	SA	200	1	True	False	True	True
2006	10	VIC	57	2	True	False	True	True
2006	10	Unk...	(null)	(null)	False	False	False	False
2006	11	SA	174	0	True	False	True	True
2006	11	VIC	53	0	True	False	True	True
2006	11	Unk...	(null)	(null)	False	False	False	False
2006	12	SA	174	4	True	False	True	True

The MXD Query made by extracting the date hierarchy, from date table , casualties and fatalities from the fact table and All KPI'S created Is showed Below,

```
SELECT NON EMPTY { [Measures].[Casualties], [Measures].[Fatalities], KPIGoal("casualty"), KPIGoal("KPI_Fatalities"),
KPIGoal("KPI_AverageFatality"), KPIGoal("KPI_TotalInjuries") }
ON COLUMNS, NON EMPTY { ([Dim Date Time].[Year].[Year].ALLMEMBERS *
[Dim Date Time].[Month].[Month].ALLMEMBERS * [Dim Location].[State].[State].ALLMEMBERS ) }
DIMENSION PROPERTIES MEMBER_CAPTION, MEMBER_UNIQUE_NAME ON ROWS
FROM ( SELECT ( { [Dim Date Time].[Hierarchy].[All] } ) ON COLUMNS
FROM [FinaCrash_View]) WHERE ( [Dim Date Time].[Hierarchy].[All] )
CELL PROPERTIES VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME, FONT_SIZE, FONT_FLAGS
```

Next By providing server name and the MXD query I connected the Cube with Excel,



Finally, the **Power Pivot table** was created below in Excel,

[Dim Date...]		fx									
	[Dim Date TimeYe...]	[Dim Date TimeMont...]	[Dim Date Time...]	[Dim Locat...]	[MeasuresFatalities]	[MeasuresCasualties]	[MeasuresKPI_Averag...]	[MeasuresKPI_Totalln...]	[MeasuresKPI_Fatalities Goal]	[Measures]	
1	2006	1	22	SA	1	8	True	True	False	True	
2	2006	1	24	SA	1	5	True	False	False	True	
3	2006	10	31	SA	1	5	True	False	False	True	
4	2006	12	24	SA	1	4	True	False	False	True	
5	2006	12	8	SA	1	8	True	True	False	True	
6	2006	2	25	SA	1	8	True	True	False	True	
7	2006	2	26	SA	1	9	True	True	False	True	
8	2006	2	28	SA	1	5	True	False	False	True	
9	2006	5	11	SA	1	4	True	False	False	True	
10	2006	5	30	SA	1	10	True	True	False	True	
11	2006	6	15	SA	1	6	True	False	False	True	
12	2006	6	18	SA	1	6	True	False	False	True	
13	2006	7	18	SA	1	3	True	False	False	True	
14	2006	8	10	SA	1	7	True	True	False	True	
15	2006	8	21	SA	1	6	True	False	False	True	
16	2006	9	3	SA	1	5	True	False	False	True	
17	2006	9	7	SA	1	10	True	True	False	True	
18	2007	10	24	SA	1	8	True	True	False	True	
19	2007	10	30	SA	1	9	True	True	False	True	
20	2007	11	1	SA	1	6	True	False	False	True	
21	2007	11	10	SA	1	9	True	True	False	True	
22	2007	11	16	SA	1	7	True	True	False	True	

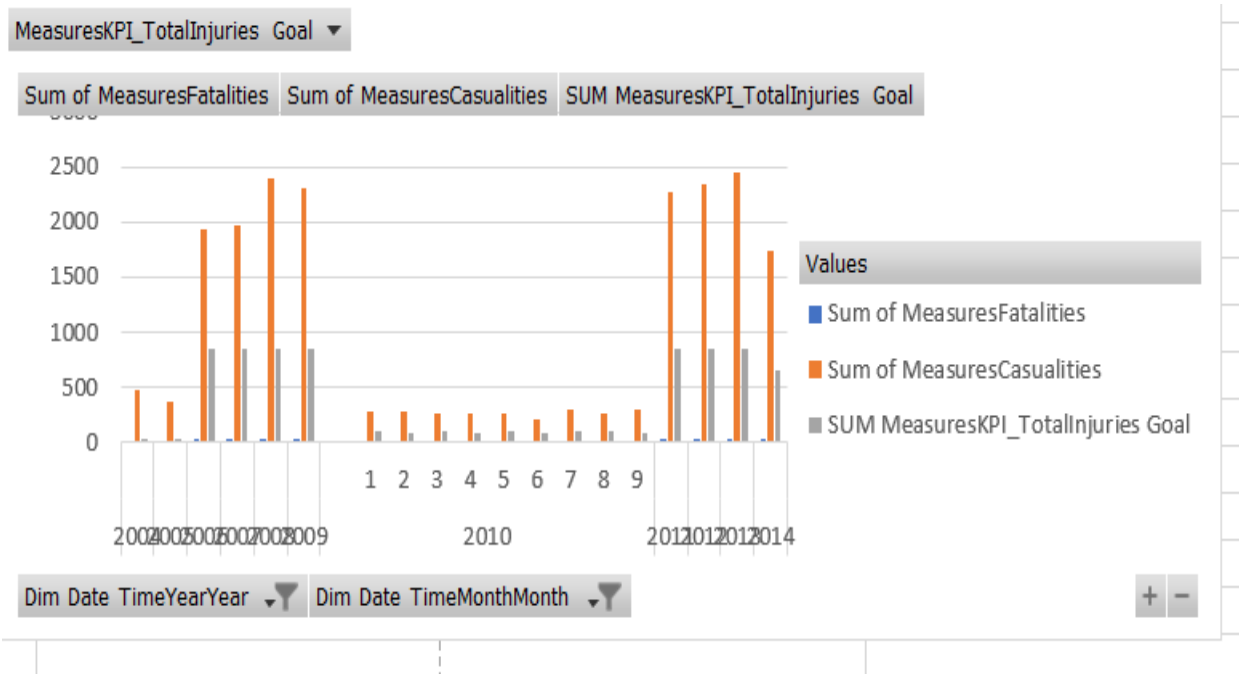
- Accordingly using the above exported columns OLAP operations were demonstrated in EXCEL

Drill Down – OLAP Operation

Drill down operation is performed by stepping down a concept hierarchy for a dimension Which means going from the Top most level to lower levels in the hierarchy.

Below I have created a Drill down visualization where the date hierarchy is stepping down from year to month and demonstrating the sum of casualty and sum of fatalities and the count of total injuries for each month in a particular year.

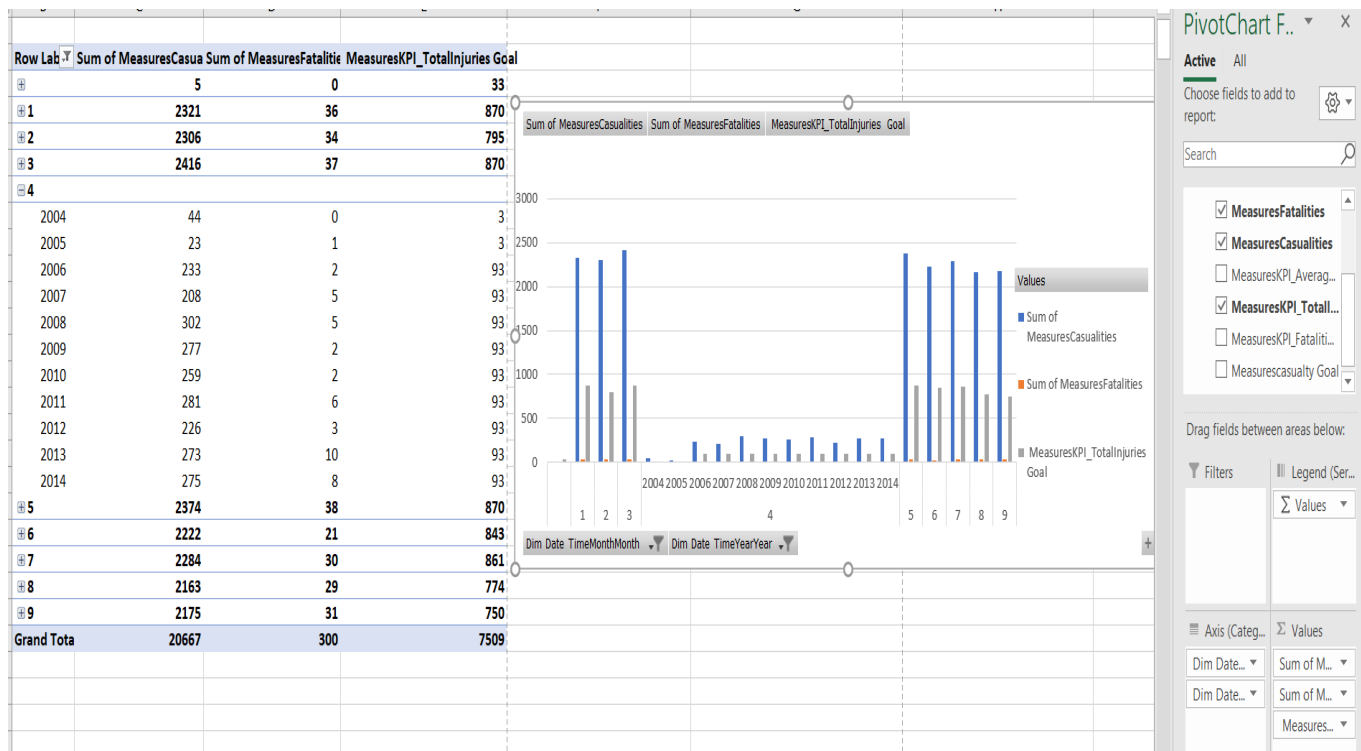
MeasuresKPI_TotalInjuries Goal All			
Row Labels	Sum of MeasuresFatalities	Sum of MeasuresCasualties	SUM MeasuresKPI_TotalInjuries Goal
2004	9	470	30
2005	3	361	30
2006	23	1927	849
2007	29	1979	849
2008	39	2406	852
2009	31	2308	849
2010			
	0	1	3
1	6	284	96
2	4	272	87
3	5	266	96
4	2	259	93
5	6	271	96
6	1	207	93
7	2	291	96
8	6	262	96
9	4	291	93
2011	28	2274	849
2012	37	2340	852
2013	35	2459	849
2014	30	1739	651
Grand Total	300	20667	7509

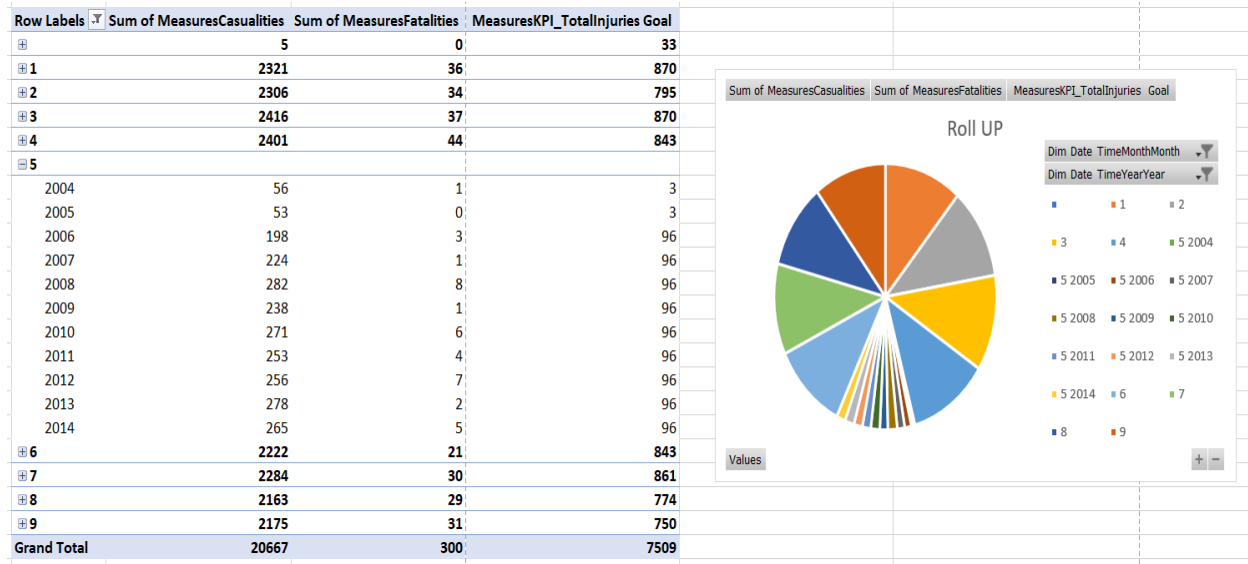


Roll Up - OLAP Operation

Roll up is used to demonstrate show the flow of hierarchies stepping up from a lower level to a higher level in order to visualize the data in a different view.

Here I have used the date hierarchy to move up from month to year. along with the sum of casualties, sum of fatalities and count of KPI_Totalinjuries goal.

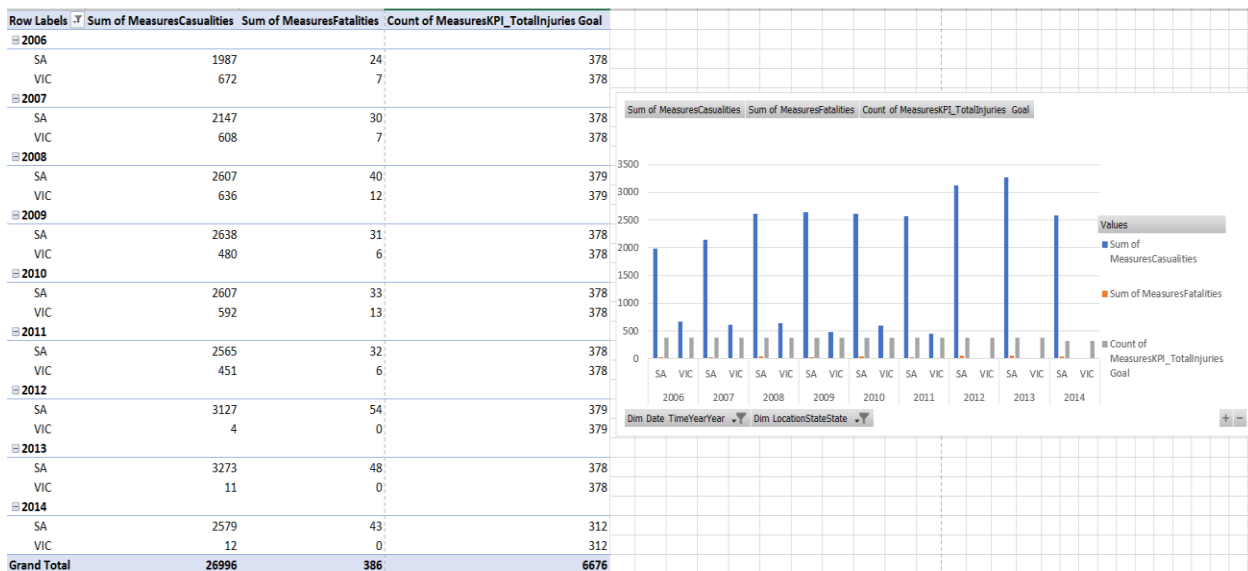




Pivot - OLAP Operation

The pivot operation is another way called for rotation.. It rotates the data axes in order to provide a different view of presenting the data.

Below I have created a pivot chat and table to show the sum of casualties and sum of fatalities and the count KPI_total injuries goal in each state in each year.

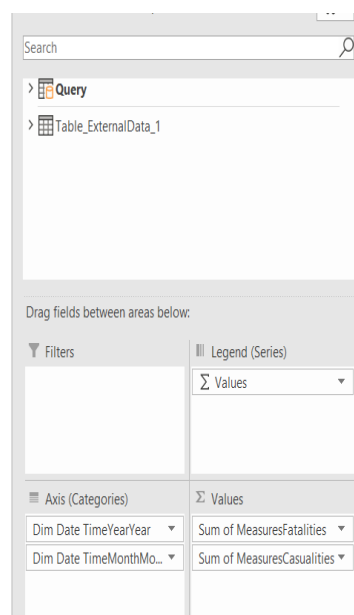
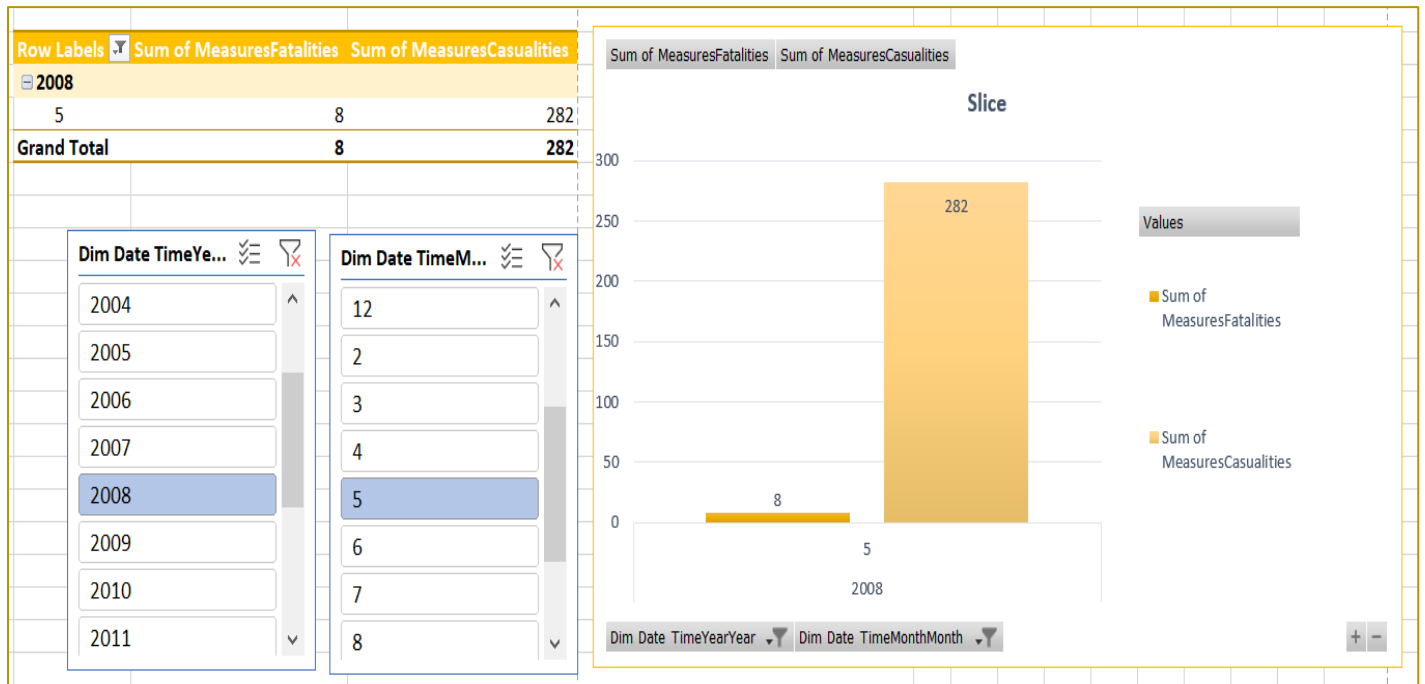


Slice - OLAP Operation

The slice operation selects one dimension from a given cube and provides a new sub-cube.

Thereby to demonstrate the slice operation I have used the DimDate dimension in order to represent the sum of casualties and fatalities in a specific year for each month along with the usage of a slicer.

Here it demonstrates the 5th months sum of casualty and fatalities for the year 2008.

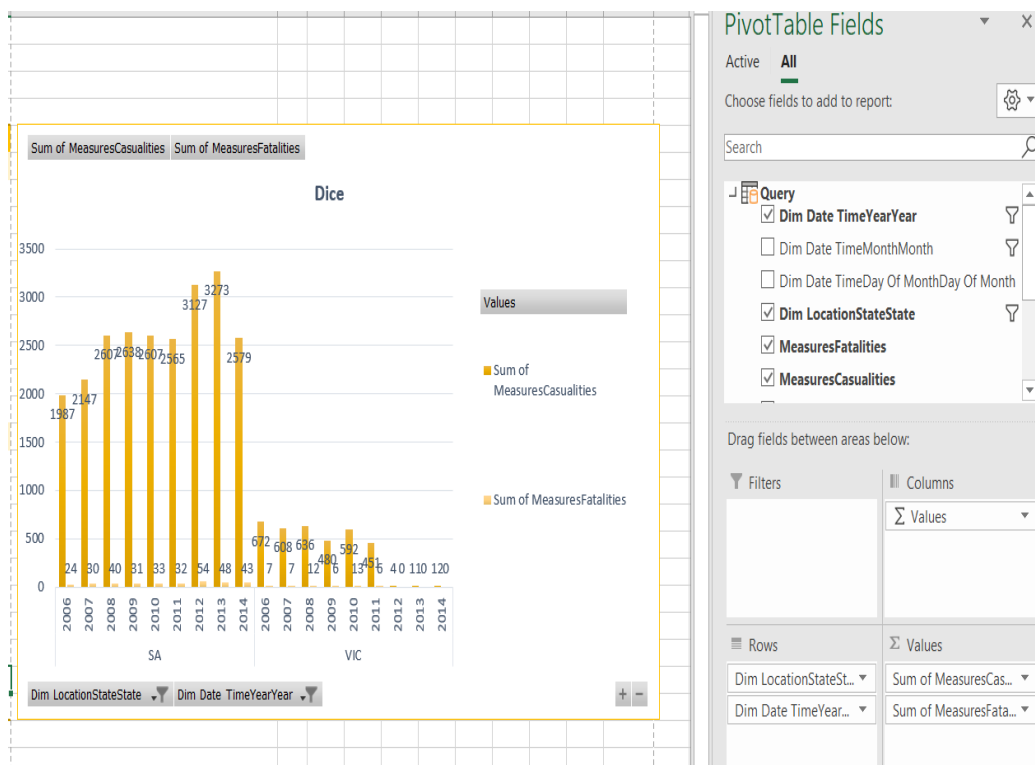


Dice - OLAP Operation

Dice selects two or more dimensions from a given cube and provides a new sub-cube to show how data is connected in a different perspective.

Below I Have created the Dice operation for the sum of casualties and fatalities in each state for a specific year.

Row Labels	Sum of MeasuresCasualties	Sum of MeasuresFatalities
SA		
2006	1987	24
2007	2147	30
2008	2607	40
2009	2638	31
2010	2607	33
2011	2565	32
2012	3127	54
2013	3273	48
2014	2579	43
VIC		
2006	672	7
2007	608	7
2008	636	12
2009	480	6
2010	592	13
2011	451	6
2012	4	0
2013	11	0
2014	12	0
Grand Total	26996	386

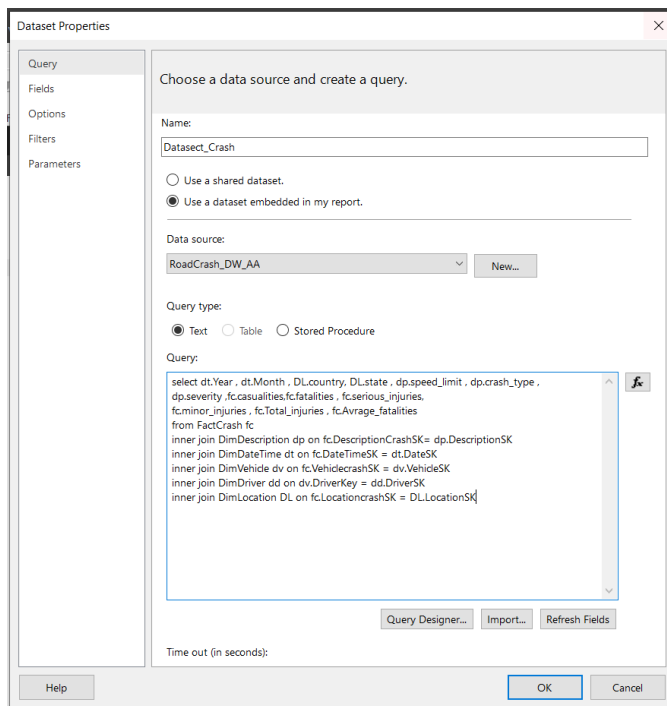


SSRS Reports

In order to report the data using SQL Server Reporting Service and Report Builder the reports were created in order to give a detail idea for the users

And then it was connected with SSRS reporting web portal to display the reports.

- **As the initial stage a data set was called to create a tabular report before going in detail.**



- **A representation of a tabular report using SSRS web portal.**

SQL Server Reporting Services

Home > SSRS report > TablereportFinal

1 of 2? Find | Next

Road Crash IN Australia and New Zealand

Year	Month	state	casualties	fatalities
2000		VIC	0	0
2001		VIC	0	0
	1	VIC	59	0
	2	VIC	66	0
	3	VIC	49	0
	4	VIC	42	0
	5	VIC	37	0
	6	VIC	36	1
	7	VIC	57	0
	8	VIC	52	1

Report With A Matrix

Matrix is created in order to visualize data Row and column wise along with aggregate functions. The below matrix report is created to show the casualties, fatalities, minor injuries and major injuries in an accident for a specific state for each year and month

To create this I have filtered and taken the years between 2006 to 2014 only.

SQL Server Reporting Services													
★ Favorites Browse													
Home > SSRS report > Finalmatrix													
< < 1 of 1 > > ↺ ↻ 100% ⌵ ⌲ Find Next													
Road Crash Occured Per Year and Month													
	SA	VIC								Total			
Year	Month	casualties	fatalities	minor injuries	serious injuries	casualties	fatalities	minor injuries	serious injuries	casualties	fatalities	minor injuries	serious injuries
2006 Total		1987	24	1764	199	672	7	592	73	2659	31	2356	272
2007 Total		2147	30	1889	228	608	7	539	62	2755	37	2428	290
2008 Total		2607	40	2288	279	636	12	558	66	3243	52	2846	345
2009 Total		2638	31	2333	274	480	6	431	43	3118	37	2764	317
2010 Total		2607	33	2293	281	592	13	539	40	3199	46	2832	321
2011 Total		2565	32	2254	279	451	6	405	40	3016	38	2659	319
2012 Total		3127	54	2775	298	4	0	3	1	3131	54	2778	299
2013 Total		3273	48	2867	358	11	0	10	1	3284	48	2877	359
2014 Total		2579	43	2236	300	12	0	12	0	2591	43	2248	300
Total		23530	335	20699	2496	3466	51	3089	326	26996	386	23788	2822

A row is expanded from the above displayed report and shown here,

SQL Server Reporting Services													
★ Favorites Browse													
Home > SSRS report > Finalmatrix													
< < 1 of 1 > > ↺ ↻ 100% ⌵ ⌲ Find Next													
Year	Month	casualties	fatalities	minor injuries	serious injuries	casualties	fatalities	minor injuries	serious injuries	casualties	fatalities	minor injuries	serious injuries
2006 Total		1987	24	1764	199	672	7	592	73	2659	31	2356	272
2007 Total		2147	30	1889	228	608	7	539	62	2755	37	2428	290
2008 Total		2607	40	2288	279	636	12	558	66	3243	52	2846	345
2009 Total		2638	31	2333	274	480	6	431	43	3118	37	2764	317
2010 Total		2607	33	2293	281	592	13	539	40	3199	46	2832	321
2011 Total		2565	32	2254	279	451	6	405	40	3016	38	2659	319
2012 Total		3127	54	2775	298	4	0	3	1	3131	54	2778	299
2013 Total		3273	48	2867	358	11	0	10	1	3284	48	2877	359
2014						0	0	0	0	0	0	0	0
	1	212	2	191	19	0	0	0	0	212	2	191	19
	2	227	4	199	24	1	0	1	0	228	4	200	24
	3	275	5	223	47	3	0	3	0	278	5	226	47
	4	274	8	230	36	1	0	1	0	275	8	231	36
	5	264	5	235	24	1	0	1	0	265	5	236	24
	6	283	2	239	42	0	0	0	0	283	2	239	42
	7	198	4	178	16	0	0	0	0	198	4	178	16
	10	293	7	250	36	2	0	2	0	295	7	252	36
	11	272	2	239	31	3	0	3	0	275	2	242	31
	12	281	4	252	25	1	0	1	0	282	4	253	25
Total		2579	43	2236	300	12	0	12	0	2591	43	2248	300
Total		23530	335	20699	2496	3466	51	3089	326	26996	386	23788	2822

Report with Multiple Parameter

To create a report with multiple parameters the initial data set was modified using a where clause. Accordingly, the State and Local government area from DimLocation has been taken as the values listed before viewing the report. Once the parameters are passed the annual fatalities casualties and their totals are displayed with relevant to the state and local government areas provided.

Moreover, multiple values in both state and local government area can be added as parameters before in order to view the report.

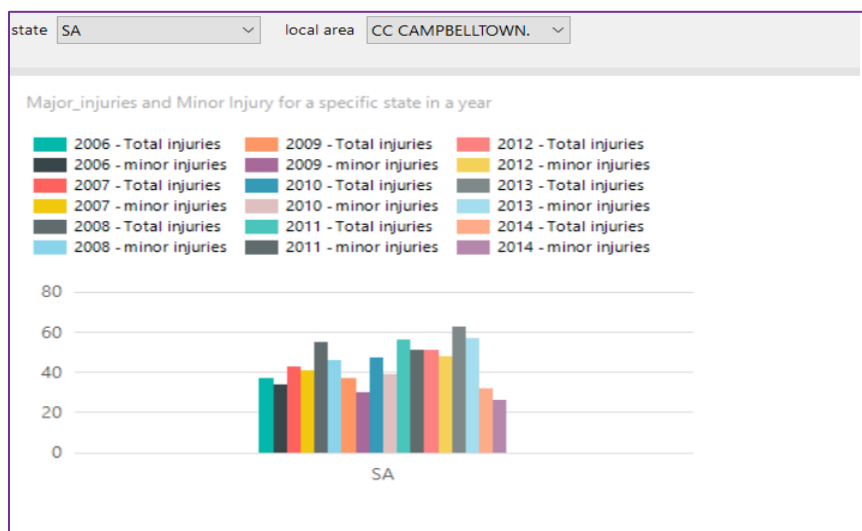
Home > SSRS report > Finalmultiparameter

state SA local area CC MARION,CC MITCHAM. View P

1 of 1 100% Find | Next

Road Crash Report for selected state and government area annually.

		2006		2007		2008		2009		2010		2011	
state	Local government Area	casualties	fatalities	casualties	fatalities	casualties	fatalities	casualties	fatalities	casualties	fatalities	casualties	fatalities
SA	CC MARION.	85	0	84	1	125	0	95	1	121	1	126	3
	CC MITCHAM.	84	0	106	1	132	0	119	1	105	0	129	4



The above is a graphycal representation of the multi parameter report

Drill Down SSRS Report

A SSRS drill down report enables to go down in a hierarchy to analyze data. Accordingly date hierarchy is used in order to display the drilling down from year to month in each state in order to visualize the casualties, fatalities , serious and minor injuries.

Home > SSRS report > FinalDrill down													
<div> <div> <div>1 of 1</div> <div>100%</div> <div>Find Next</div> </div> </div>													
Road Crash Drill Down Report													
		SA				VIC				Total			
Year	Month	casualties	fatalities	serious injuries	minor injuries	casualties	fatalities	serious injuries	minor injuries	casualties	fatalities	serious injuries	minor injuries
2006		1987	24	199	1764	672	7	73	592	2659	31	272	2356
2007		2147	30	228	1889	608	7	62	539	2755	37	290	2428
2008		2607	40	279	2288	636	12	66	558	3243	52	345	2846
2009		2638	31	274	2333	480	6	43	431	3118	37	317	2764
2010		2607	33	281	2293	592	13	40	539	3199	46	321	2832
2011		2565	32	279	2254	451	6	40	405	3016	38	319	2659
2012		3127	54	298	2775	4	0	1	3	3131	54	299	2778
2013		3273	48	358	2867	11	0	1	10	3284	48	359	2877
2014		2579	43	300	2236	12	0	0	12	2591	43	300	2248
Total		23530	335	2496	20699	3466	51	326	3089	26996	386	2822	23788

Road Crash Drill Down Report														
		SA				VIC				Total				
	Year	Month	casualties	fatalities	serious injuries	minor injuries	casualties	fatalities	serious injuries	minor injuries	casualties	fatalities	serious injuries	minor injuries
2006	2006		1987	24	199	1764	672	7	73	592	2659	31	272	2356
2007	2007		2147	30	228	1889	608	7	62	539	2755	37	290	2428
2008	2008						2	0	0	2	2	0	0	2
		1	208	4	15	189	33	0	2	31	241	4	17	220
		2	247	4	21	222	38	0	3	35	285	4	24	257
		3	228	5	18	205	35	1	2	32	263	6	20	237
		4	226	5	29	192	76	0	12	64	302	5	41	256
		5	224	6	25	193	58	2	10	46	282	8	35	239
		6	191	0	34	157	43	0	3	40	234	0	37	197
		7	208	3	22	183	51	1	8	42	259	4	30	225
		8	247	1	34	212	51	1	0	50	298	2	34	262
		9	187	2	21	164	53	4	9	40	240	6	30	204
		10	204	0	20	184	79	1	6	72	283	1	26	256
		11	210	6	16	188	54	1	5	48	264	7	21	236
		12	227	4	24	199	63	1	6	56	290	5	30	255
		2607	40	279	2288	636	12	66	558	3243	52	345	2846	
2009	2009		2638	31	274	2333	480	6	43	431	3118	37	317	2764
2010	2010		2607	33	281	2293	592	13	40	539	3199	46	321	2832

Drill Through SSRS Report

Drill through report is used to go through a information deeply. There are levels in this graph where from one specific level we can drill in to another level to get more information about the specific column we drilled through in the first level.

- Following is the query for level 01 to Visualize the sum of casualty in a specific year for a specific state

The screenshot shows the 'Query' tab in the SSRS Query Designer. The 'Name' field is set to 'DataSet1'. Under 'Data source', 'Road_Crash_DW' is selected. Under 'Query type', 'Text' is selected. The query text in the bottom pane is as follows:

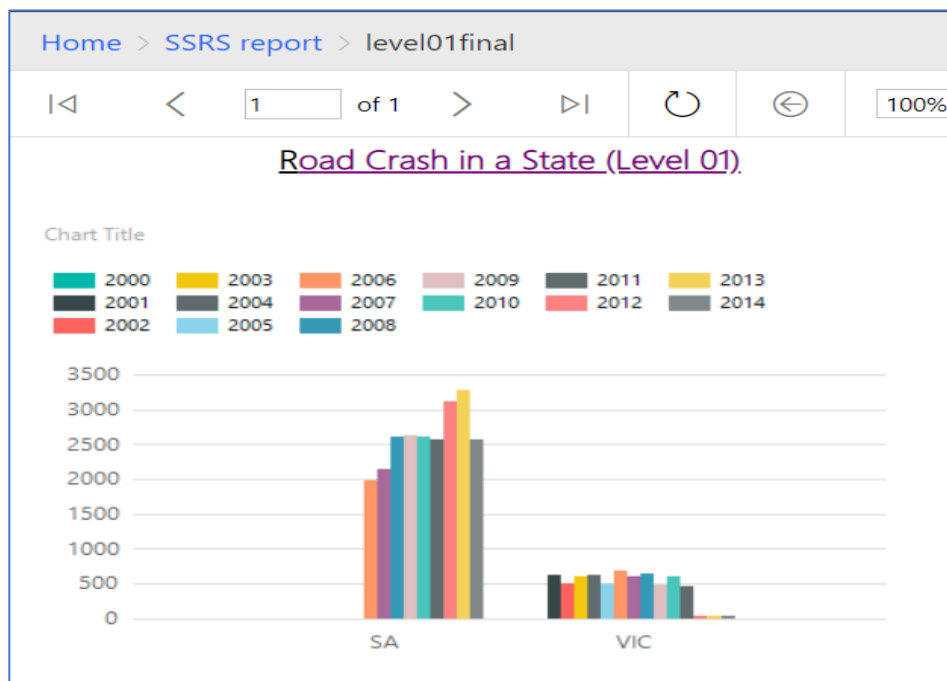
```
select dt.Year , DL.state , fc.casualties
from FactCrash fc
inner join DimDescription dp on fc.DescriptionCrashSK= dp.DescriptionSK
inner join DimDateTime dt on fc.DateTimeSK = dt.DateSK
inner join DimLocation DL on fc.LocationcrashSK = DL.LocationSK
order by dt.Year
```

- Following is the query for level 02, which drill through the level one casualty Column in order to retrieve more information of the accident like local government area , casualty, fatality , minor injuries and major injuries.

The screenshot shows the 'Query' tab in the SSRS Query Designer. The 'Name' field is set to 'description'. Under 'Data source', 'Road_Crash_DW' is selected. Under 'Query type', 'Text' is selected. The query text in the bottom pane is as follows:

```
select dt.Year , DL.state , DL.Local_government_Area,
fc.casualties,fc.fatalities , fc.serious_injuries,
fc.minor_injuries
from FactCrash fc
inner join DimDescription dp on fc.DescriptionCrashSK= dp.DescriptionSK
inner join DimDateTime dt on fc.DateTimeSK = dt.DateSK
inner join DimLocation DL on fc.LocationcrashSK = DL.LocationSK
where DL.state IN ( @state)
order by dt.Year
```

- Following is the first graph retrieved in the drill through process which shows the year, state and total casualty in the accidents.



This is the level 02 graph that will be displayed after clicking on the casualty column in the above displayed level 01 graph.

- Over here the data for year 2013 in state 'SA' is displayed once clicked on that particular column in the above graph.

