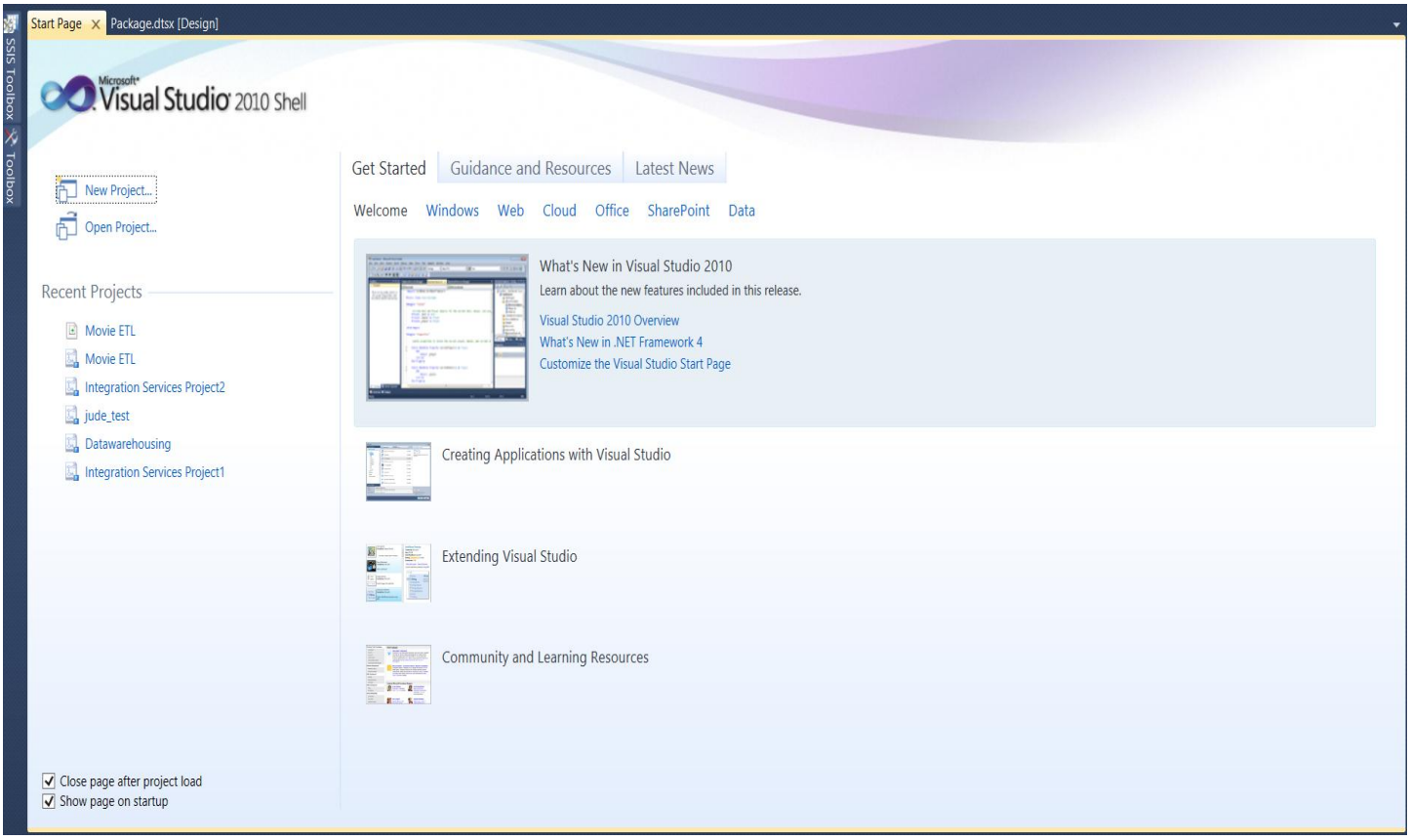
"Movie Management System"

OLAP of the Application

Mohamed Niyaz



1. OLAP Introduction

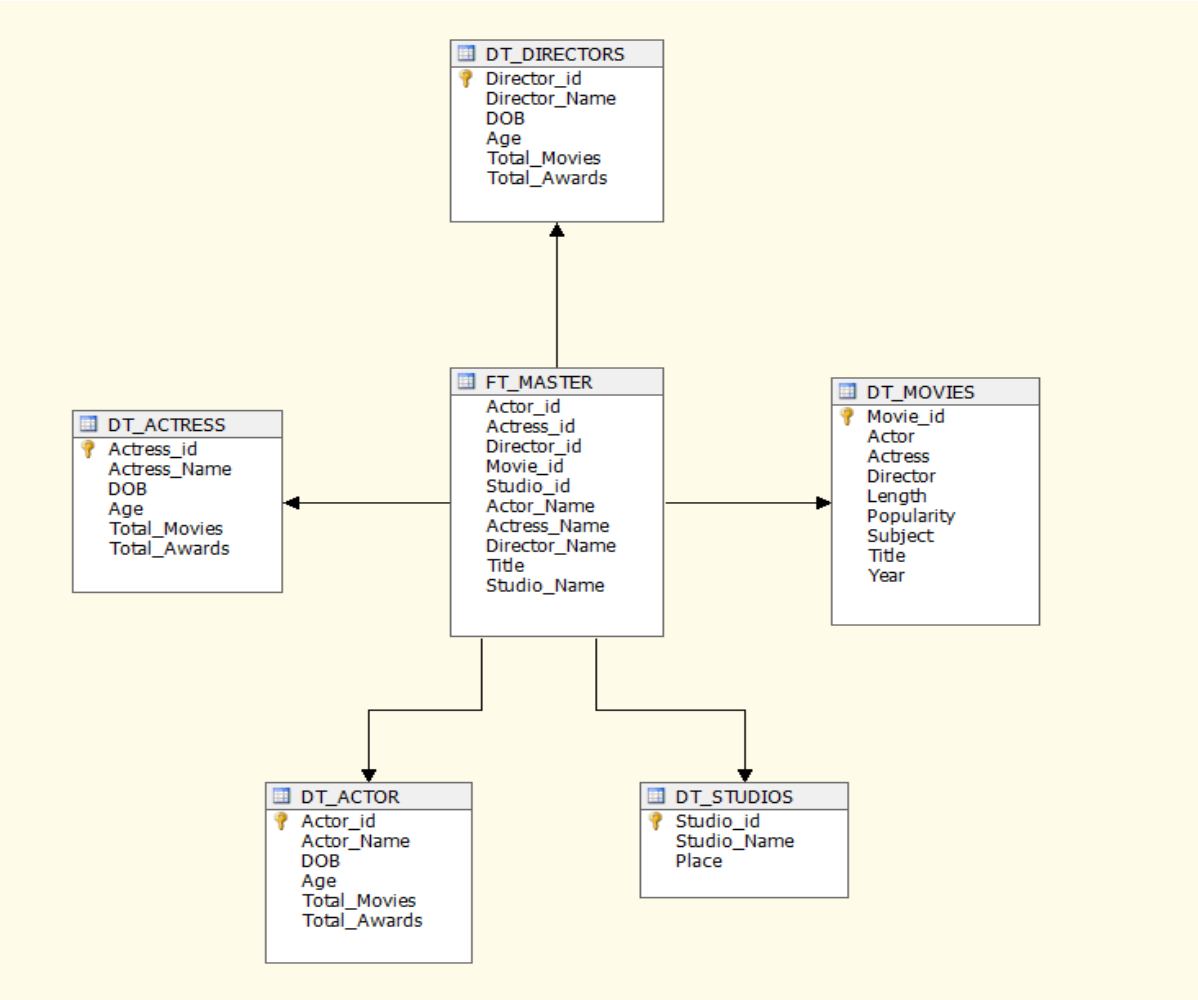
In computing, online analytical processing, or OLAP an approach to answering multi-dimensional analytical (MDA) queries swiftly. OLAP is part of the broader category of business intelligence, which also encompasses relational database report writing and data mining. Typical applications of OLAP include business reporting for sales, marketing, management reporting, business process management budgeting and forecasting, financial reporting and similar areas, with new applications coming up, such as agriculture. The term OLAP was created as a slight modification of the traditional database term OLTP (Online Transaction Processing).

OLAP tools enable users to analyze multidimensional data interactively from multiple perspectives. OLAP consists of three basic analytical operations: consolidation (roll-up), drill-down, and slicing and dicing. Consolidation involves the aggregation of data that can be accumulated and computed in one or more dimensions. Databases configured for OLAP use a multidimensional data model, allowing for complex analytical and ad-hoc queries with a rapid execution time. They borrow aspects of navigational databases, hierarchical databases and relational databases.

2. OLAP Design

Our OLAP Process is carried through **Microsoft BI** tool and below screen shot explains the high level overview of the **Schema**

The below Screenshot depicts **data flow task** which is part of ETL process



3. OLAP Process

- ***Roll-up*** **:** A roll-up involves summarizing the data along a dimension. The summarization rule might be computing totals along a hierarchy or applying a set of formulas such as movie rating = Total Rating/Number of Votes.
- ***Drill-Down*** **:** Drill Down allows the user to navigate among levels of data ranging from the most summarized (up) to the most detailed (down).
- ***Slicing*** **:** Slice is the act of picking a rectangular subset of a cube by choosing a single value for one of its dimensions, creating a new cube with one fewer dimension. Finding out the movie that released in the year of 2004 are "sliced" out of the data cube.
- ***Dicing*** : The dice operation produces a sub cube by allowing the analyst to pick specific values of multiple dimensions. The new cube will show limited number of movies, the time and region dimensions cover the same range as before**.**

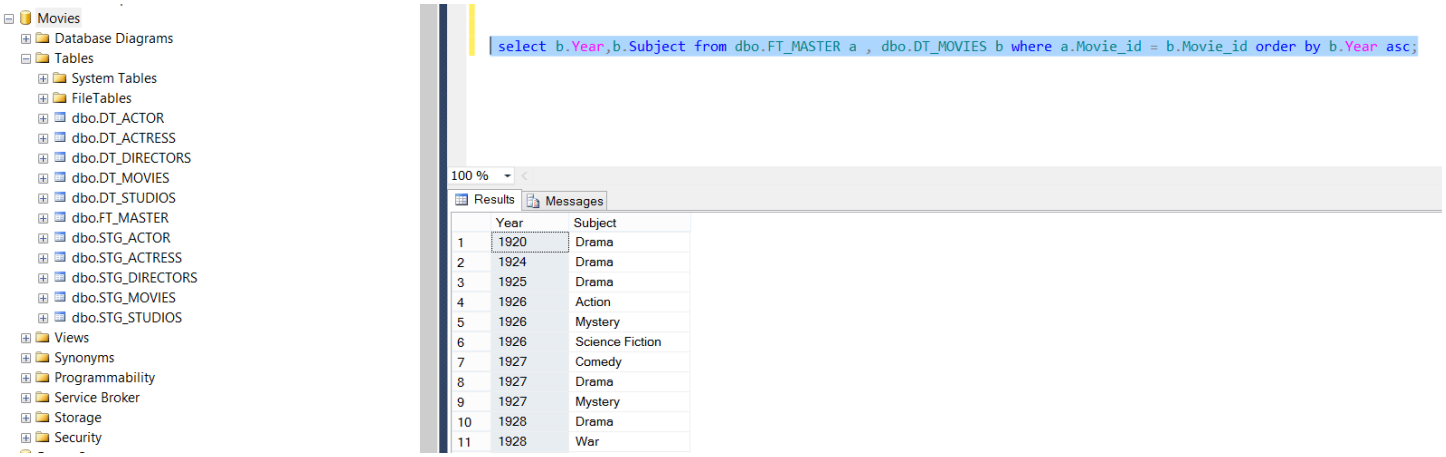
4. Table Hierarchies

- Actor table : Actor id < Actor Name < DOB < Age
- Actress table : Actress id < Actress Name < DOB < Age
- Director table : Director id < Director Name < DOB < Age
- Movies table : Movie id < Title < Genre < Popularity < Year
- Studio table : Studio id < Studio Name < Place
- Master table : Actor Name < Actress Name < Director Name < Title < Studio Name

5. OLAP Operations (Includes Screenshot and Sql implementation)

Consider a base cuboid Master table and Movies table [Master, Movies]

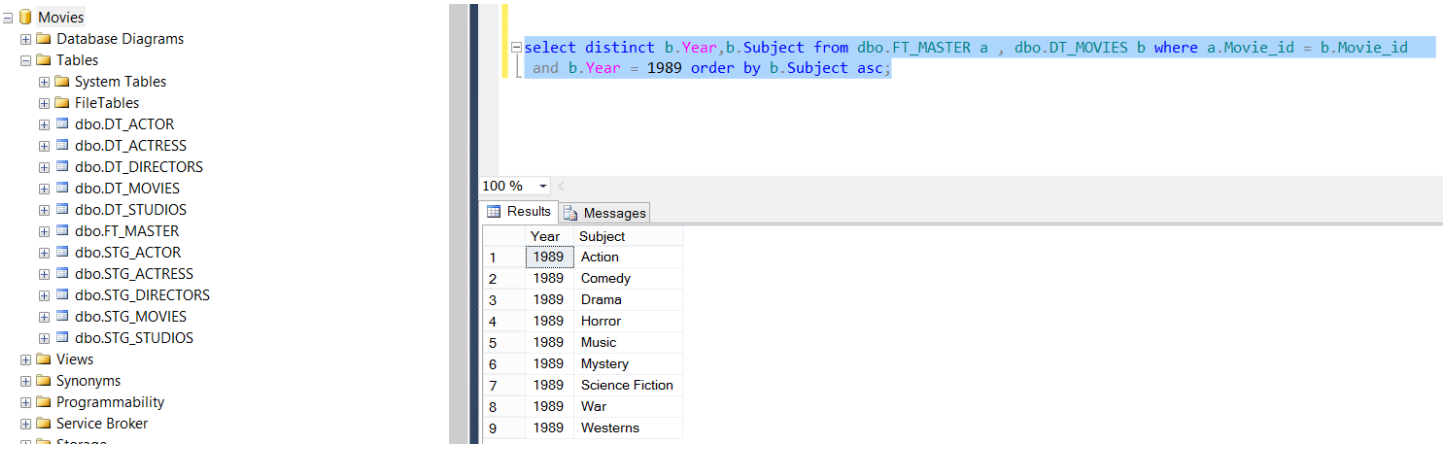
Step1: Create a Base cuboid with Year and Subject = [Year, Subject]



Sql implementation:

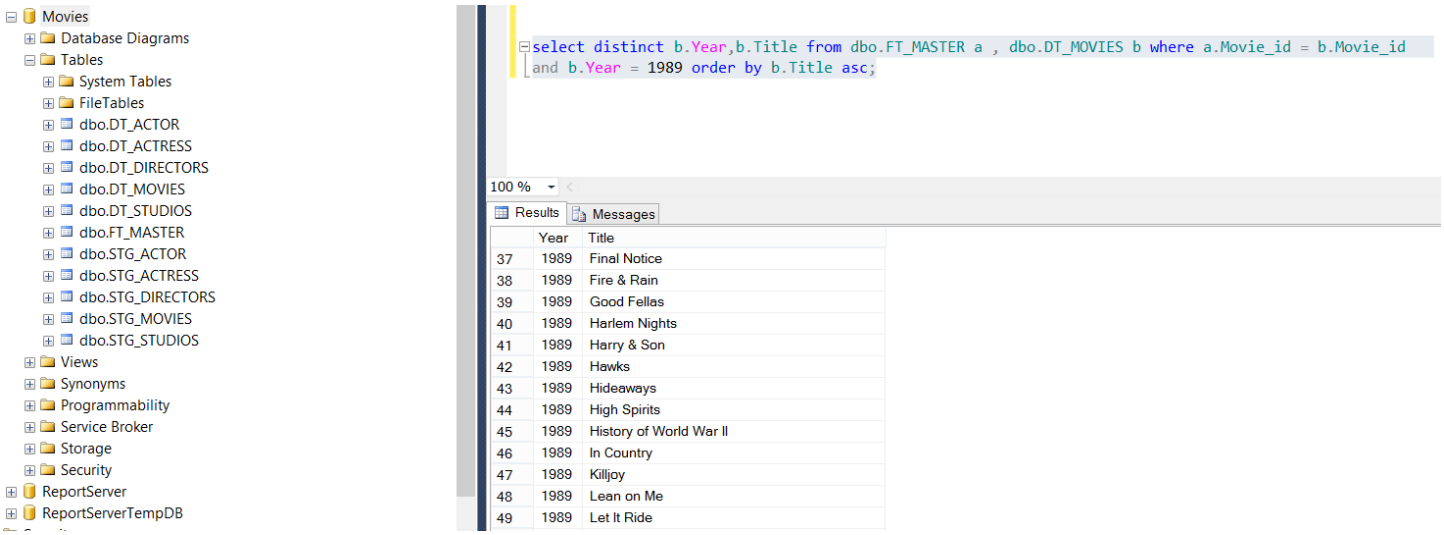
```
select b.Year,b.Subject from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id = b.Movie_id order by
b.Year asc;
```

Step2: Slice on the Year by 1989




```
select distinct b.Year,b.Subject from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id = b.Movie_id
and b.Year = 1989 order by b.Subject asc;
```

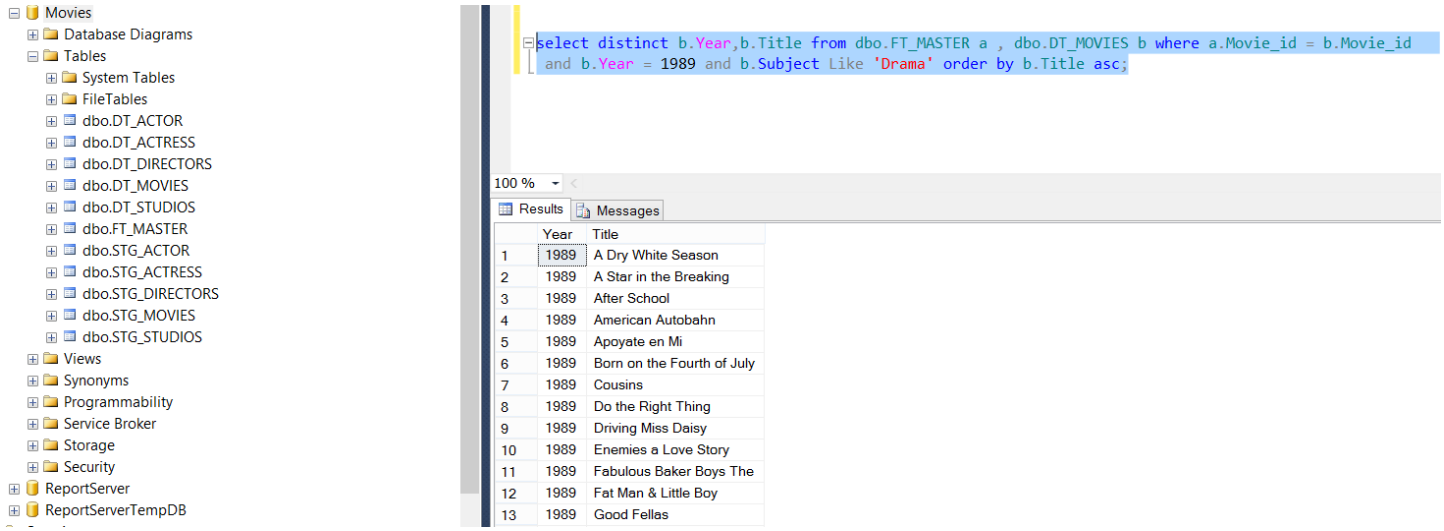
Step3: Rollup on the Subject to obtain the list of titles for the year 1989



Sql implementation:

```
select distinct b.Year,b.Title from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id = b.Movie_id
and b.Year = 1989 order by b.Title asc;
```

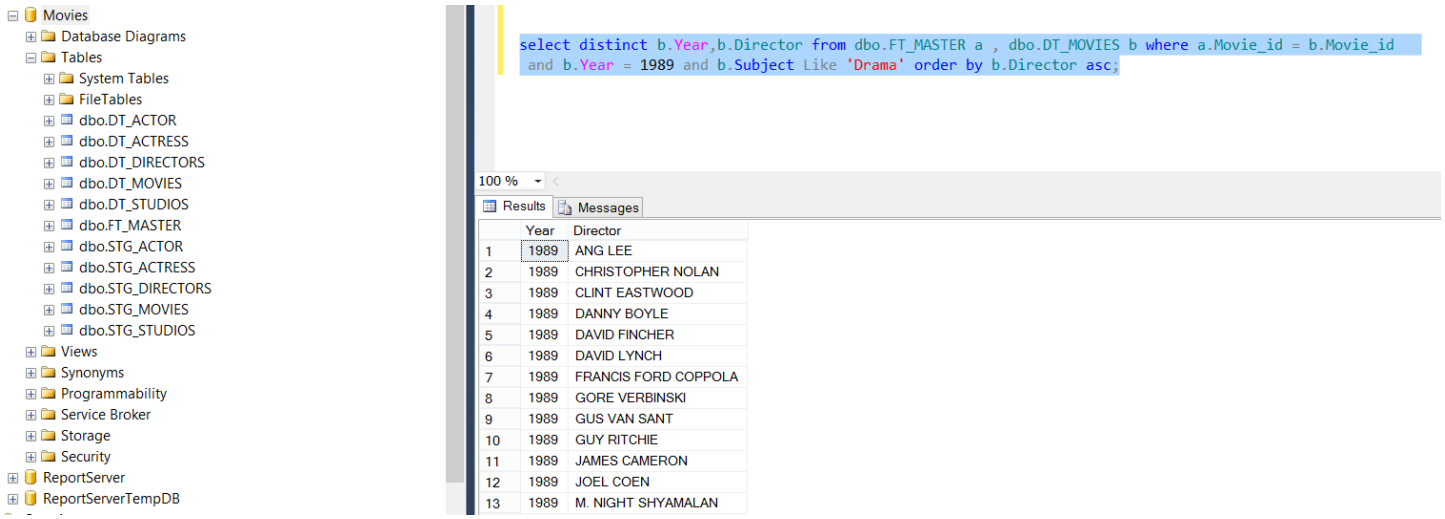
Step4: Slice on the Subject to Drama to obtain the list of movies for 1989 with genre Drama



Sql implementation:

```
select distinct b.Year,b.Title from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id = b.Movie_id
and b.Year = 1989 and b.Subject Like 'Drama' order by b.Title asc;
```

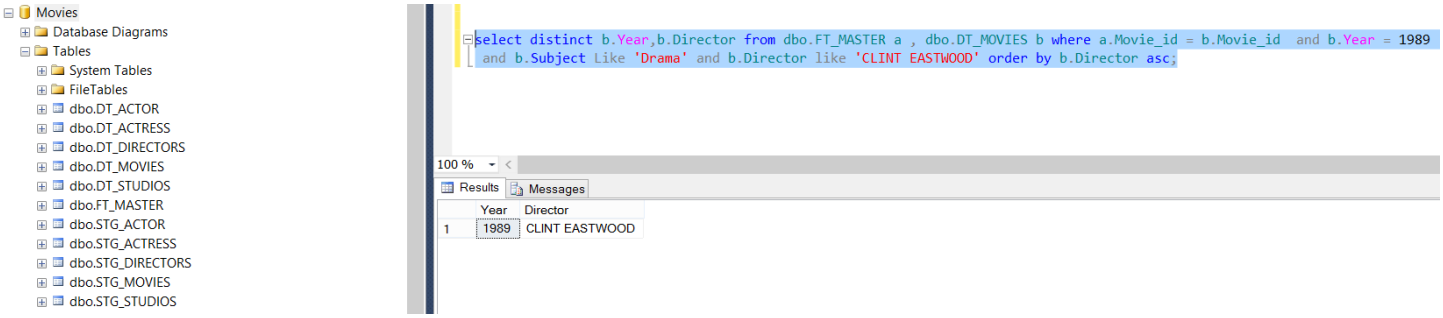
Step5: Rollup on the Title to obtain list of director who has directed in the year 1989 with genre Drama



Sql implementation:

```
select distinct b.Year,b.Director from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id = b.Movie_id
and b.Year = 1989 and b.Subject Like 'Drama' order by b.Director asc;
```

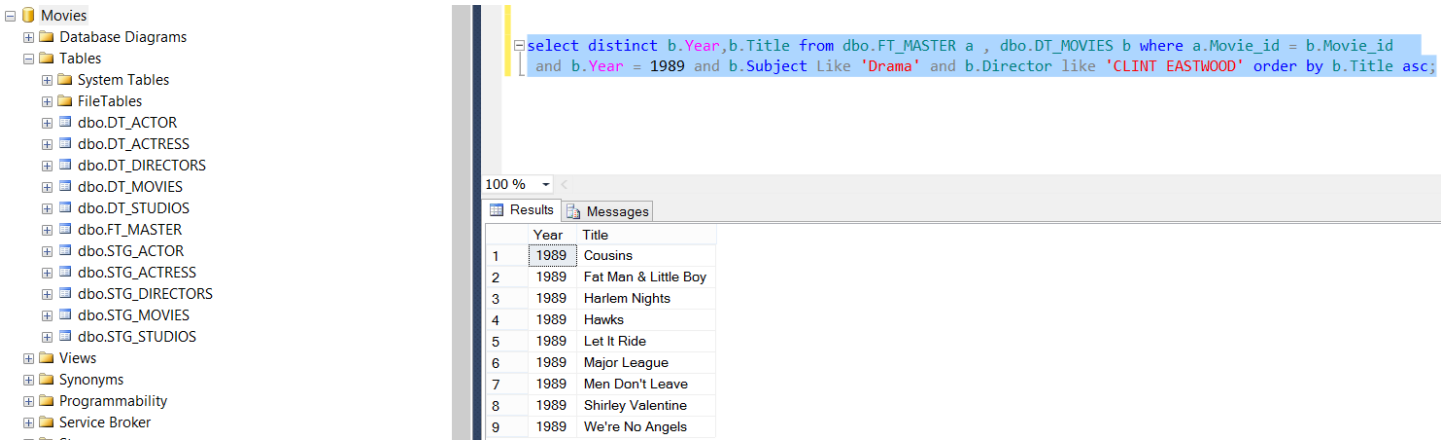
Step6: Slice on the director to Clint Eastwood to obtain the list of movies for 1989 with genre Drama



Sql implementation:

```
select distinct b.Year,b.Director from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id = b.Movie_id
and b.Year = 1989 and b.Subject Like 'Drama' and b.Director like 'CLINT EASTWOOD' order by b.Director
asc;
```

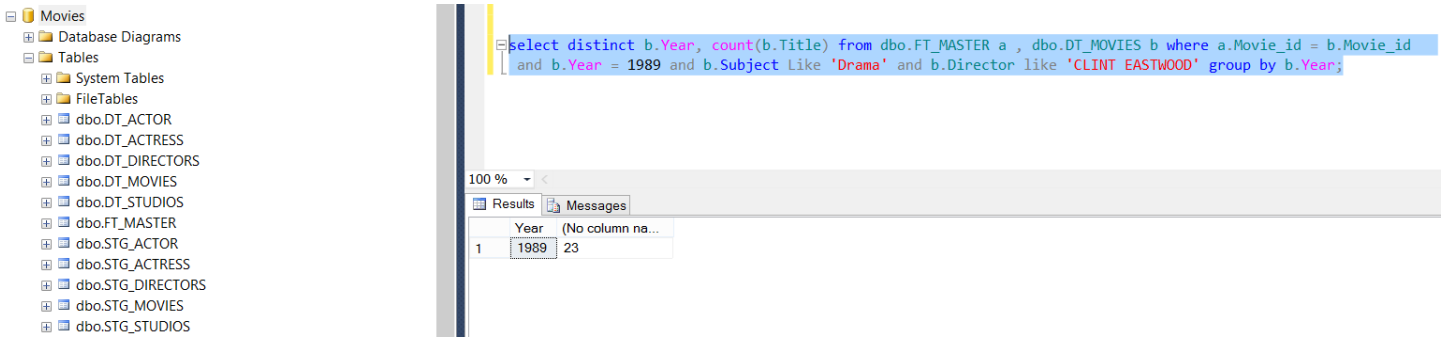
Step7: Rollup on the title to obtain list of movies for 1989 with genre Drama directed by Clint Eastwood



Sql implementation:

```
select distinct b.Year,b.Title from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id = b.Movie_id
and b.Year = 1989 and b.Subject Like 'Drama' and b.Director like 'CLINT EASTWOOD' order by b.Title asc;
```

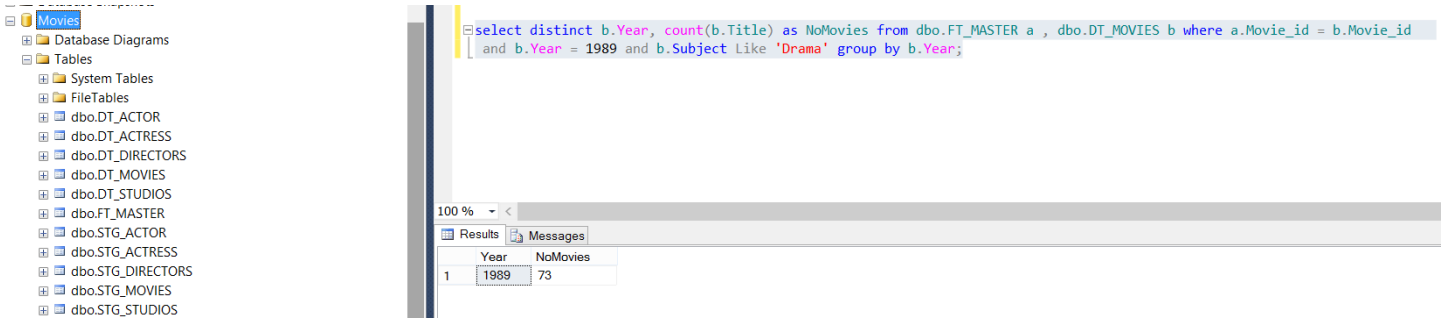
Step8: Dice on the title to obtain the number of movies for 1989 with genre Drama directed by Clint Eastwood



Sql implementation:

```
select distinct b.Year, count(b.Title) from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id =
b.Movie_id and b.Year = 1989 and b.Subject Like 'Drama' and b.Director like 'CLINT EASTWOOD' group by
b.Year;
```

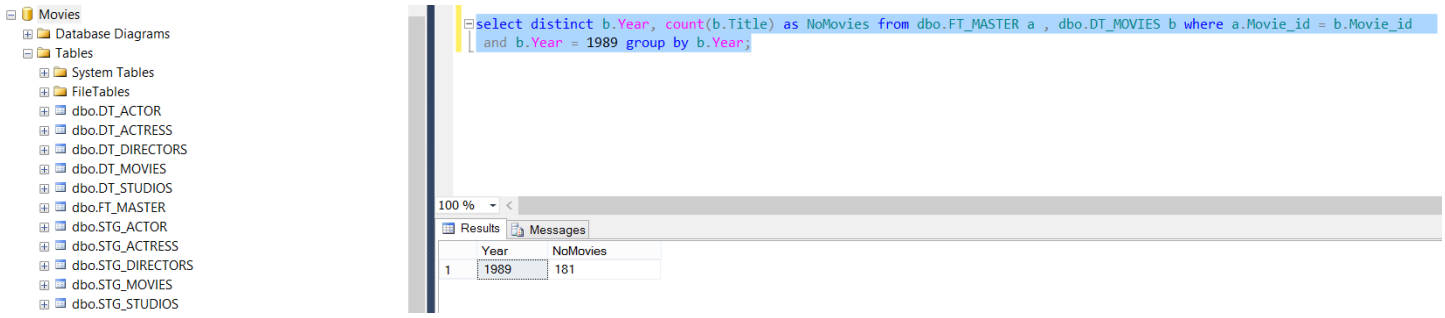
Step9: Drill down on director to find the number of movies directed for the year 1989 with genre drama



Sql implementation:

```
select distinct b.Year, count(b.Title) from dbo.FT_MASTER a , dbo.DT_MOVIES b where a.Movie_id =
b.Movie_id  and b.Year = 1989 and b.Subject Like 'Drama' group by b.Year;
```

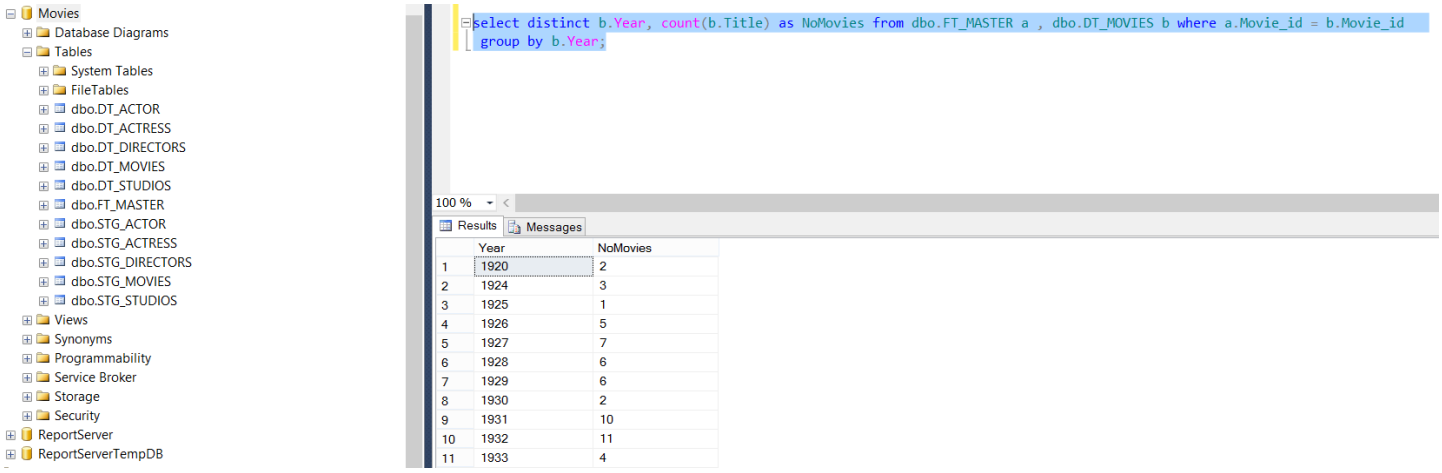
Step10: Drill down on subject to find the number of movies directed for the year 1989



Sql implementation:

```
select distinct b.Year, count(b.Title) as NoMovies from dbo.FT_MASTER a , dbo.DT_MOVIES b where
a.Movie_id = b.Movie_id  and b.Year = 1989 group by b.Year;
```

Step11: Drill down on Year to find the number of movies directed so far.



Sql implementation:

```
select distinct b.Year, count(b.Title) as NoMovies from dbo.FT_MASTER a , dbo.DT_MOVIES b where
a.Movie_id = b.Movie_id  group by b.Year;
```

6. Tools Used

- **Microsoft SQL Enterprise Edition-** The tool has inbuilt SQL server, Repository service, Integration Service, Analytical Services , SQL Database Management Studio. It's a tool which has both the ETL and the OLAP functionality with it.

- **Pl-Sql –** This is an Oracle commands which will be used to provide a flexible database activity in order to promote best achievements in obtaining an elegant working system.

- **SSIS –** Used for ETL (Extract transform Load)

- **SSAS –** Used for OLAP(Roll up , drill down , slicing, dicing , cube creation)

- **Hardware's-** Windows 8 64 bit operating , 16 GB DDR3 Ram, i7 Intel Processor, 2 GB DDR5

7. Challenges

- OLAP operations can be performed only the cleanse data, so getting the cleanse data was one among the challenges.
- OLAP operations have to be performed on fact and dimension tables, forming multiple relations and hierarchy was really challenging.
- Forming the cuboid was the biggest challenge among all these, we have to make sure the entity relation is perfectly right and data are fully cleansed.

No refinement to the warehouse design.