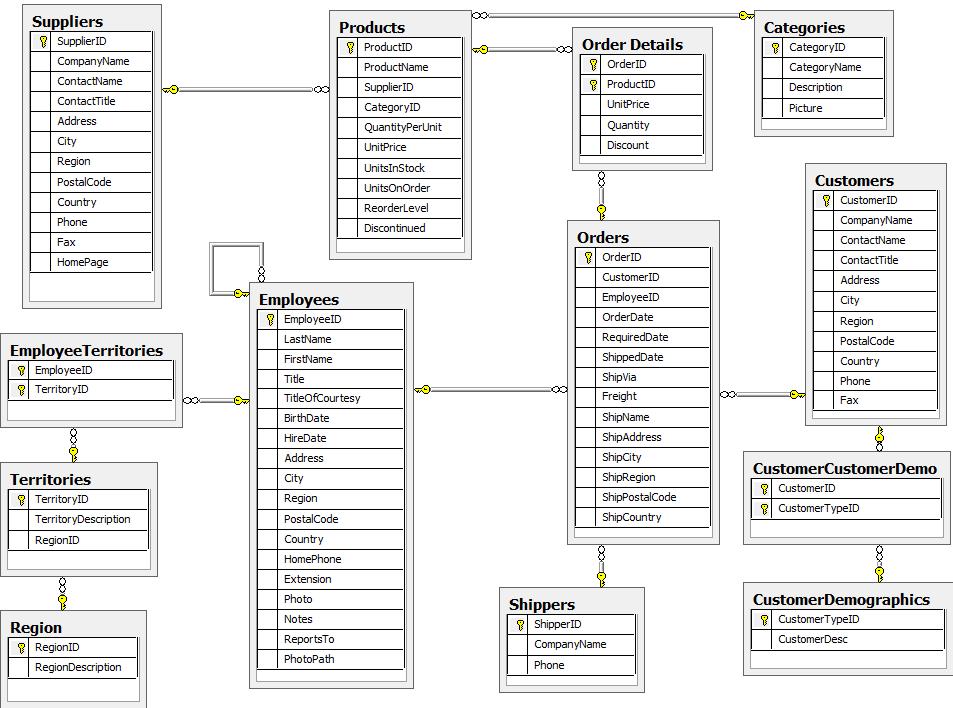
**Part 1 – Business Drivers**

Northwind is a fictitious food import-export company created by Microsoft.

Below is the complete Northwind’s database ER Diagram.



In order to have the star schema, from the ER Model shown above I identified the dimension tables first, then the fact table. Since we are focussing on sales, Sales is the fact table. The dimension tables contain each a primary key and the data by which we measure the fact table. The primary keys connect the dimension tables to the fact table. I identified the dimension tables from only a few tables: Customer, Supplier, Product, Employee. In addition I created a new dimension table, Time, which has been modified into DimDate.

**Part 2 – Data Modelling**

While the ER Diagram is normalised, the star schema is de-normalised. It is a dimensional model. Dimensional models are used to structure warehousing data. To de-normalise the model, some of the tables are combined.

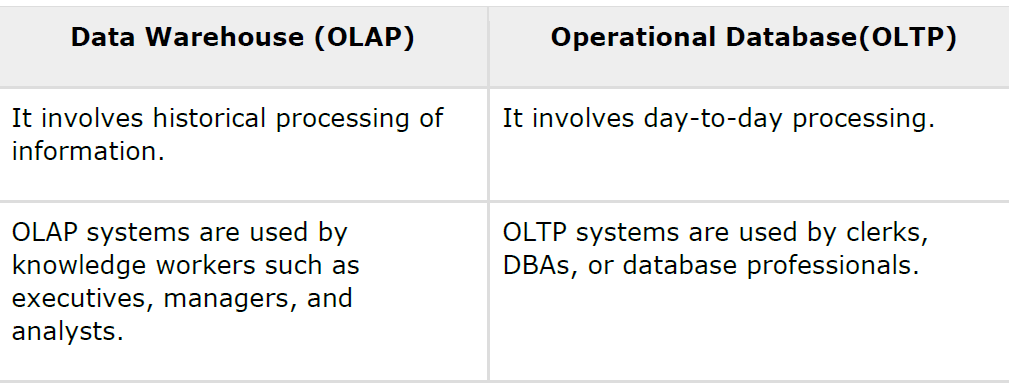
De-normalised models are easier to understand than normalised models.

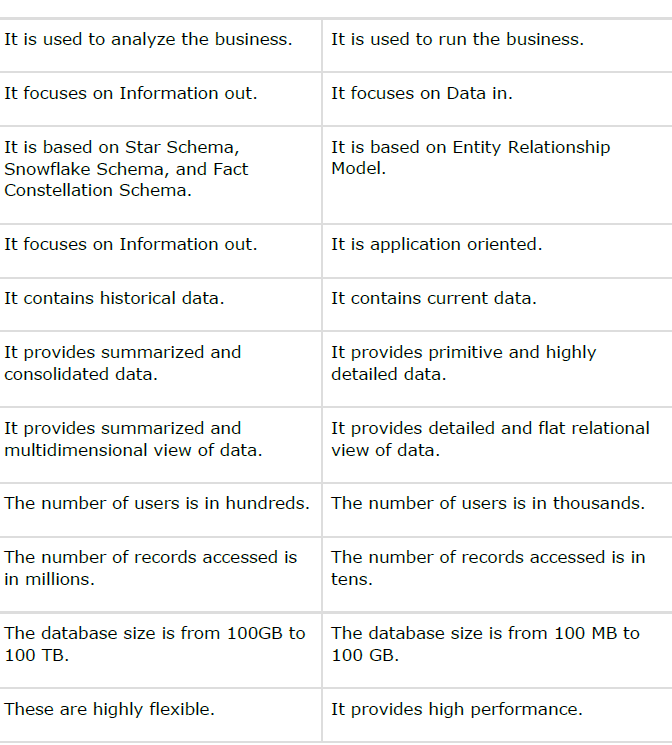
Warehousing’s models are based on Ralph Kimball’s dimensional modelling. Kimball suggested that the data loaded into the warehouse are divided into two types: Dimensions and Facts. Whereas dimensions are numeric and measurable, facts are simply pieces of information.

Fact tables are highly detailed to be as flexible as possible. By contrast, Dimension tables are static.

All the activities performed by an organisation are described as Business processes. These gather the performance metrics.

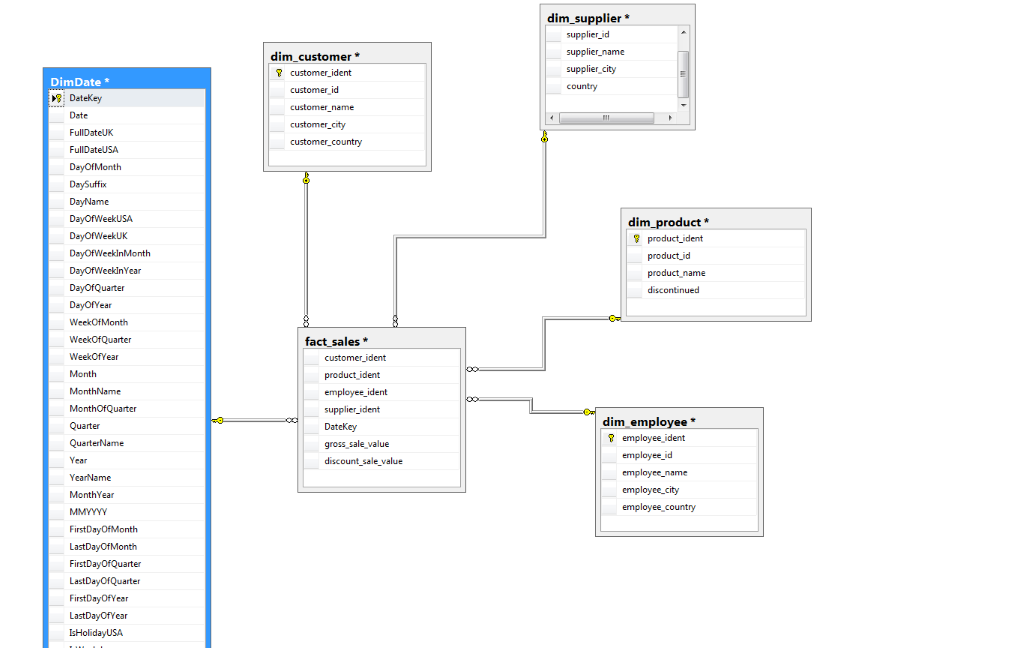
Business decisions are based on information, or data retrieval as opposed to data modification. Data retrieval models are referred to as OLAP (Online Analytical Processing). Data modification models are referred to as OLTP (Online Transaction Processing).





The star schema, or data mart, is particularly good to meet the needs of a specific department, in this case the sales department, as it includes only the data relevant to that department.

Similarly to a normalised model, I needed to link the fact table to the dimension tables. Each dimension table has a new ID, called ‘ident’ and set it as primary key. I connected all the primary keys to the fact table, populating it as foreign keys.



**Part 3 – Implement Tables**

I uploaded the data from the operation database onto SQL Server, using SQL queries.

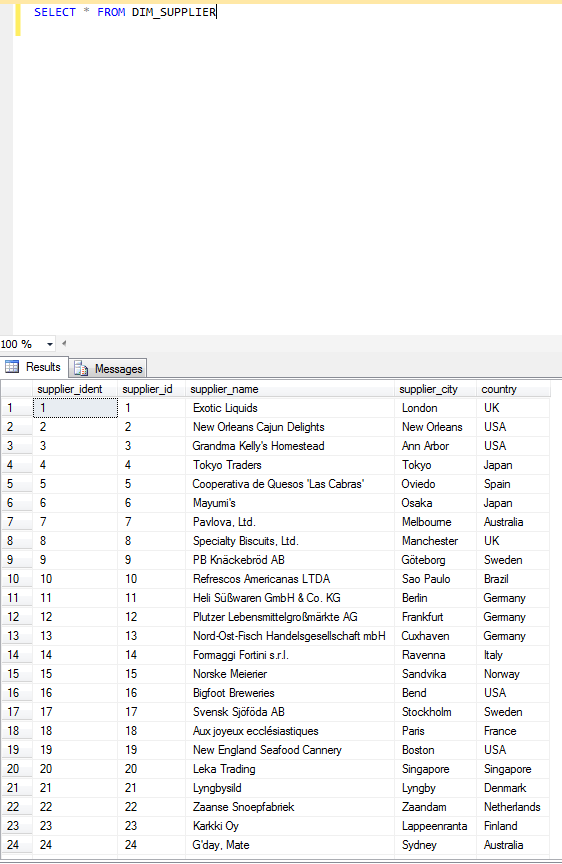
From the ER model, I created both the dimension tables and the fact table. Ultimately I populated the tables and designed the diagram.

As an example, below is the dimension Supplier populated.



Microsoft SQL Server is a powerful and relatively easy tool to create and populate tables by using SQL queries.

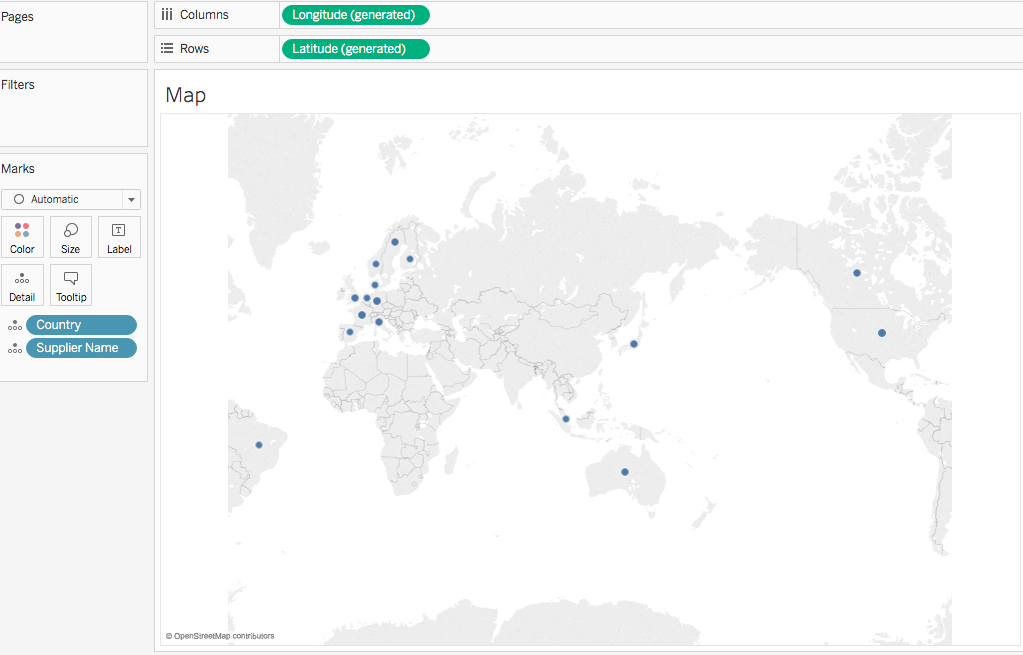
As shown below it is also easy to make queries. I queried the Suppliers with one simple line of code.



**Part 4 – Reporting and Analysis**

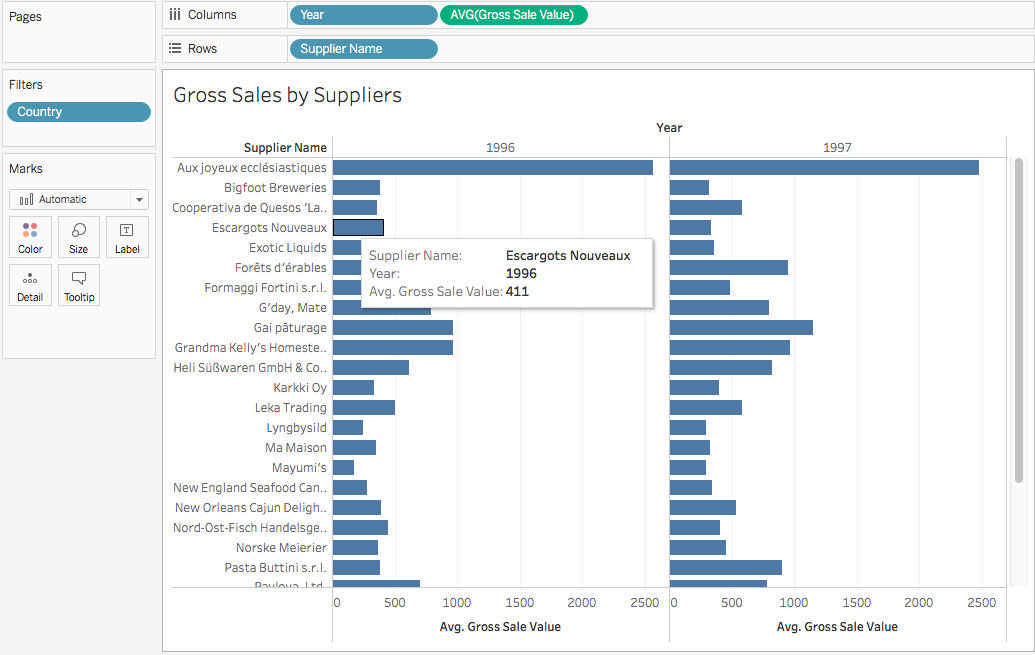
I connected Tableau to Microsoft SQL Server, selected the relevant Database and extracted the data. I was able to create some visualisation from the data.

First I created a map table to have a geographic representation of Suppliers all over the world. It is already clear that most suppliers are from Europe and that there are no Suppliers from Africa and most of Asia.

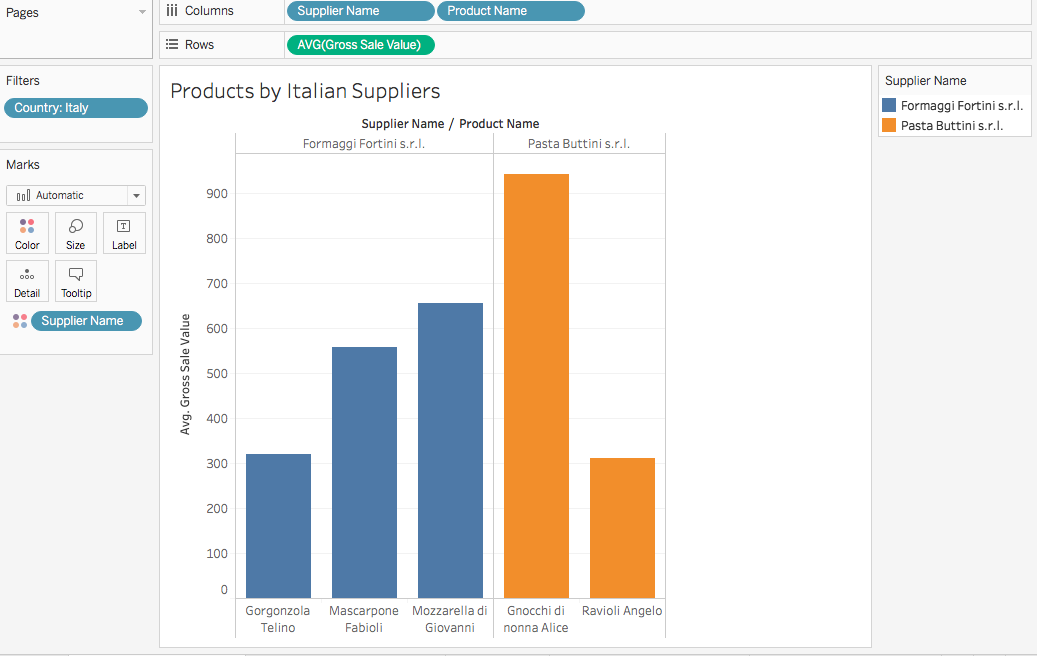


In the second table I chose to have a list of all the Suppliers, their names, and their Gross Sales by year.

In the table I can distinctly see that the Supplier with the highest Gross Sale Value is ‘Aux joyeux ecclestistiques’.



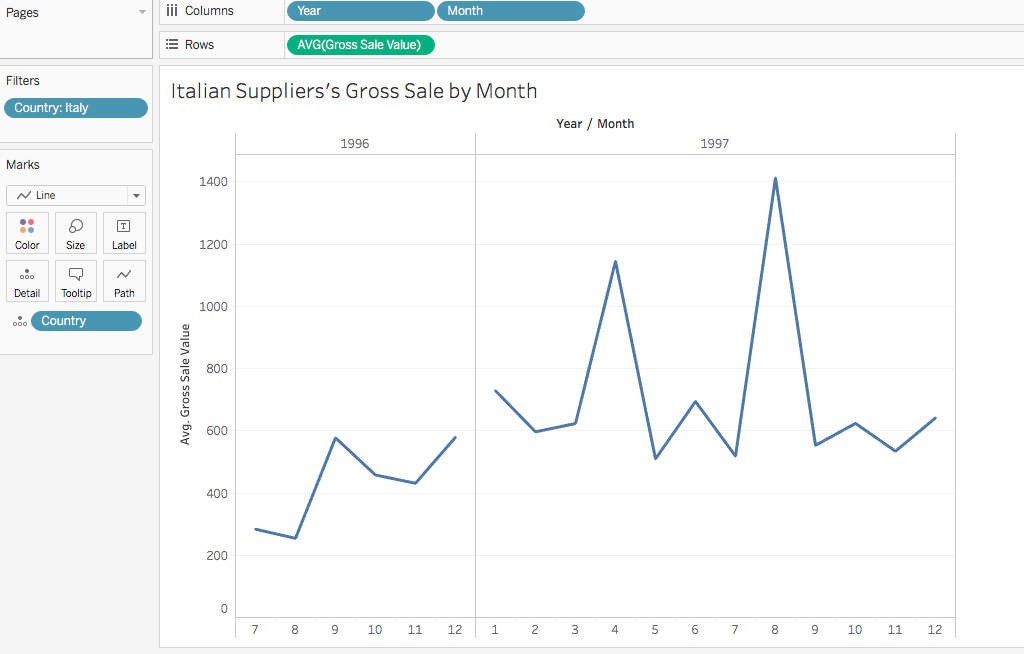
In the third table I focussed on the Gross Sale of Italian suppliers by product.



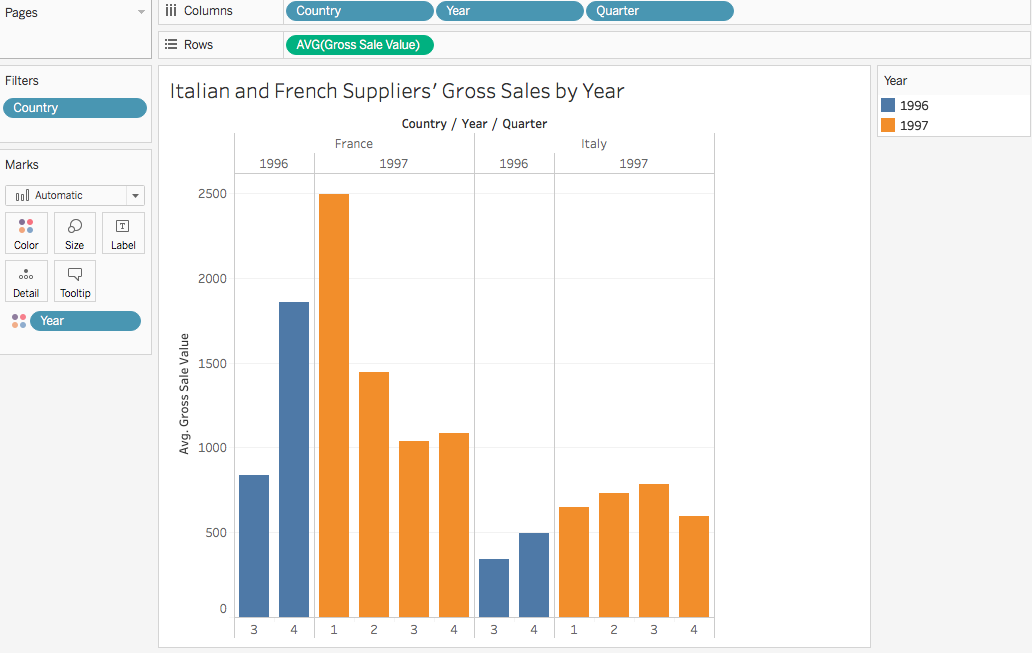
The table shows that Northwest has only two Italian Suppliers, ‘Formaggi Fortini s.r.l.’ and ‘Pasta Buttini s.r.l.’.

The former supplies three types of cheese: ‘Gorgonzola Telino’, ‘Mascarpone Fabioli’ and ‘Mozzarella di Giovanni’. The latter supplies two types of pasta: ‘Gnocchi di nonna Alice’ and ‘Ravioli Angelo’.

The product with the highest Gross Sale Value is ‘Gnocchi di nonna Alice’ supplied by ‘Pasta Buttini s.r.l.’.

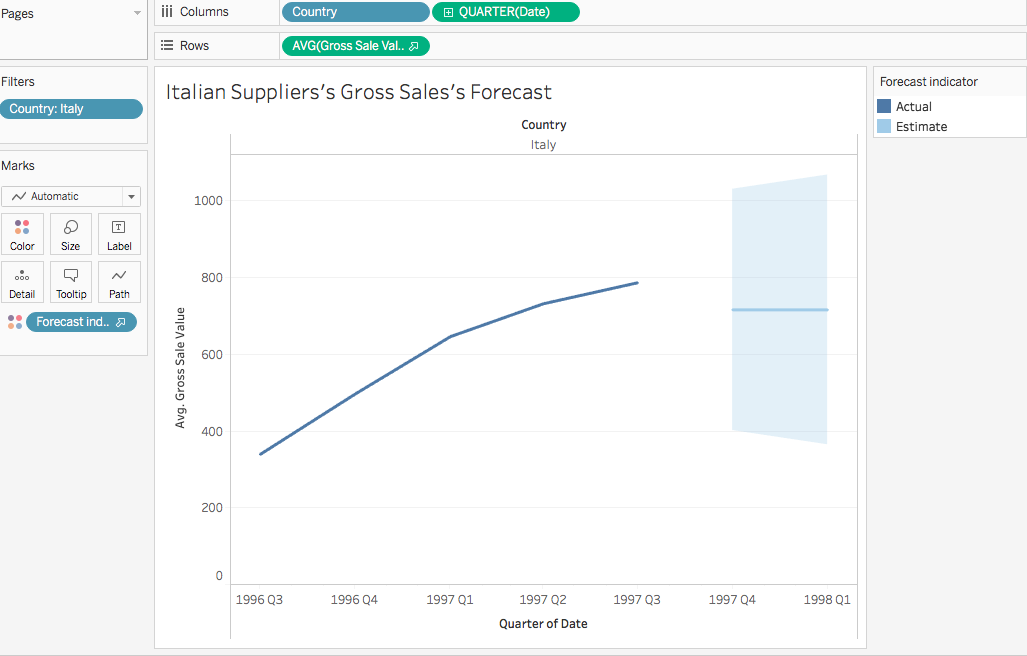


Here is a graph of Italian suppliers’ Gross Sale Value by month. In this graph we can see that 1997 was more profitable than 1996 and that the most profitable months were April and August. The least profitable month was February. The graph also shows that the Gross Sale Value changes dramatically month by month.



Above is a graph chart of the comparison between Italian and French suppliers, grouped by quarters.

French suppliers’ Gross Salue Value was significantly higher than Italian Suppliers’ both in 1996 and in 1997.



Above is the forecast for Italian Suppliers’ Gross Sales’ Value

Dashboard



The strength of Microsoft’s SQL Server is in creating tables, making queries and designing database’s diagrams.

In contrast, Tableau is particularly good at visualising the data with charts and graphs, especially at how they change over time and in specific geographic regions.