

Data Analytics & Visualization using Qlik

Sumedha Ashish Mhatre

Lecturer, Department of Computer Engineering, V.P.M's Polytechnic, Thane, Maharashtra, India

Abstract – Data Analytics is a process of analysis, cleansing, transforming and mining useful information, suggesting conclusions and supporting decision-making. Qlik is one of those BI enterprises who have come up with a user friendly and intuitive tools like QlikView and QlikSense to enable Data Analytics through its powerful engine and advanced visualization interfaces which together help provide Business Intelligence to enterprises and decision makers. This document reviews the power of one such tool called Qlikview

Key Words: Data Analytics, Data Visualization, Qlikview

1. INTRODUCTION

Data Analytics has come a long way since Mainframes stored Megabytes of data in them to now when we have petabytes of data generated every minute. Back then every byte stored costed enterprises quite a lot of money and still there were analysts who decided what to store and what to clean. Today, there are data analysts who use several mathematical models to co-relate data and get insights from it to enable decision makers to make informed critical decisions about the businesses and processes. Data exists in structured as well as non-structured form. We shall discuss structured data in this document and how Qlikview can be used to get powerful insights from the data.

1.1 Data Analytics

Data Analytics can be referred to as a process where you first identify the source of the data, cleanse it, transform it and then mine it to come to conclusions about it. For example, a sales agent analyzes his sales data, compares it with his peers and comes to a conclusion how is he performing with respect to his peers. Such kind of analysis is usually done on the data which is stored in Enterprise Data Warehouses or Data Marts. Data can be structured or unstructured. To store structured data, enterprises usually house Relational databases like Teradata, Oracle, IBM DB2, MS SQL Server, MySQL amongst others.

The process of Data Analysis includes the following:

Data Requirements:

The data is necessary as input to the analysis one is undertaking which is usually based upon the requirements from the customer who is directing the analysis.

Data Collection:

Data is collected from variety of sources which could include a flat file, a hadoop file system or even a table from a relational database. Sometimes data may even come from sensors in the environment such as recording systems, temperature gauges in the oceans across the globe or may be even from a generator in a hydro electric power station. All this data is collected and stored in an agreeable format on a single computer or on a cluster of computers in a distributed environment.

Data Cleansing:

Data is usually cleansed before it is even transformed and analyzed. The need to cleanse data arises due to the way data is stored and formatted in different ways. Data cleansing involves getting rid of all the unwanted data, standardizing data coming in from all the different sources and at times may be even deleting unwanted characters from the data which might not be readable in the tool that would be used to analyze it.

Data Transformation:

Once the source of the data is identified, requirements are known and the data is cleansed, the time comes to transform it in a way which can lead to mining powerful insights out of the data. This would involve joining tables, files, referring smaller dimensions which would eventually lead to creating a data model which would act as an interface for an analyst to analyze the data.

1.2 Data Visualization

Data Analytics if facilitated using visualizations can be more powerful, intuitive and convincing for the end user or customer for whom we analyze. It involves visual representation of plain or aggregated data. The goal of a visualization is to provide powerful message to the user about the data using statistical graphics, plots and information graphics. Effective visualization helps users analyze and reason about the data and evidence. It makes complex data more accessible, understandable and usable.

Data Visualization is both an art and science. It is viewed as a branch of descriptive statistics. Qlikview is one such power tool which has best of the statistical graphics and plots which help effectively depict the insights into the data.

2. Qlikview

Qlikview allows users to rapidly build and develop analytic applications without the need for professional development skills, driving faster response to changing business requirements, shorter time to value, and drive greater insights across the organization. Qlikview has all the inbuilt tools to connect to variety of sources like an excel file, flat file or even a database. Qlikview provides connectors for sourcing data, scripting language to cleanse, transform and model this data and a powerful user interface to plot out all this data.

The capability which Qlikview distinguishes itself from other tools in the market is its ability to process all the data in-memory. This makes the tool more responsive and user friendly. So more the memory to process the data, faster will be the response time of every click on the tool.

Associated Query Language:

Association between the data is what analysts always tries to find when they deal with data coming in from different sources. Traditional ways to find association is by referring a data model or a relational diagram or digging into the data and finding relations using structured query language.

What if all all these relations are built within the tool and the tool does your work to identify these relations and shows it to you in a digram? This is what Associated Query Language of Qlikview does with your data. Every matching column between multiple sources is automatically associated as soon as the data is loaded inside the memory. This feature of Qlikview is so powerful that it reflects the usage of multiple where clauses in a Structured query language but without having any knowledge of common columns for the user who loads the data inside memory using the script editor.

Scripting Ability:

Qlikview provides a scripting editor which allows user or a developer to source data from multiple sources, manipulate, transform and even store it in Qlikview document format which can be reused further by another qlikview document. This editor provides structured query like features like joining the two files, manipulating the columns, transposing the data, aggregating the data amongst others. It provides the ability to create a data model which would be further used to design a user interface on top of it as a final layer for presentation.

Below figure Fig – 1, provides a snapshot of how the script editor looks like in Qlikview:

```

1 | QuartersMap:
2 | MAPPING LOAD
3 | rowno() as Month,
4 | 'Q' & Cell (rowno()/3) as Quarter
5 | AUTOGENERATE (12);
6 |
7 | Temp:
8 | Load
9 | min(OrderDate) as minDate,
10 | max(OrderDate) as maxDate
11 | Resident Orders;
12 |
13 | Let varMinDate = Num(Peek('minDate', 0, 'Temp'));
14 | Let varMaxDate = Num(Peek('maxDate', 0, 'Temp'));
15 | DROP Table Temp; |
16 |
17 | TempCalendar:
18 | LOAD
19 | $(varMinDate) + IterNo()-1 As Num,
20 | Date($(varMinDate) + IterNo() - 1) as TempDate
21 | AutoGenerate 1 While $(varMinDate) + IterNo() - 1 <= $(varMaxDate);
22 |
23 | MasterCalendar:
24 | Load
25 | TempDate AS OrderDate,
26 | week(TempDate) As Week,
27 | Year(TempDate) As Year,
28 | Month(TempDate) As Month,
29 | Day(TempDate) As Day,
30 | ApplyMap('QuartersMap', month(TempDate), Null()) as Quarter,
31 | Week(weekstart(TempDate)) & '-' & WeekYear(TempDate) as WeekYear,
32 | WeekDay(TempDate) as WeekDay
33 | Resident TempCalendar
34 | Order By TempDate ASC;
35 | Drop Table TempCalendar;
    
```

Fig -1: Qlikview Script Editor

User Interface:

Data visualization in Qlikview is provided by the dashboard which is an interface between the user and Qlik engine. The tool provides powerful visualizations from a basic table structure showing most granular data to the most appealing statistical graphics providing insights into the aggregated data. Some of the inbuilt available charts include pie chart, bar chart, combo chart, scatter chart, line chart, radar chart, grid chart, block chart, funnel chart, pivot table, straight table amongst others. The dashboard provided enables user to move around the objects with drag and drop convenience. You need not know any kind of pixel information to plot any object on the dashboard. Below figure Fig - 2 shows one such dashboard.



Fig -2: Qlikview Dashboard

3. Why Qlikview is different?

Structured Query Language was not designed for modern analytics, yet all other Business Intelligence tools are based on query based approach to analysis. This limits you to predefined, linear exploration within subsets of data. Only Qlik's Associated engine brings together all your data so you can explore without limits. Qlikview powerful, on-the-fly calculation and aggregation instantly updates all analytics with each click. Critical thinking is no longer derailed by slow queries or ongoing data preparation needs. This lets analysis move at the speed of thought.

4. CONCLUSION

Data Visualizations have now become need of the hour and there are certainly plethora of tools out in the market which would help enterprises get powerful insights from the data they already have. Qlikview visualization tool has very much kept itself alive in the market by constantly innovating the way visualizations work, adding different new sources of data which get added everyday while maintaining speed to market. Qlik Inc continues to invest in Qlikview and also added more products to its inventory to enable embedded analytics, Geo Analytics and printing capabilities.

REFERENCE

[1] www.qlik.com