Business Intelligence (BI)

Project Report

Project Members:

(i) Shufei TANG

(ii) Muhammad Adil FAYYAZ

Architecture of System:

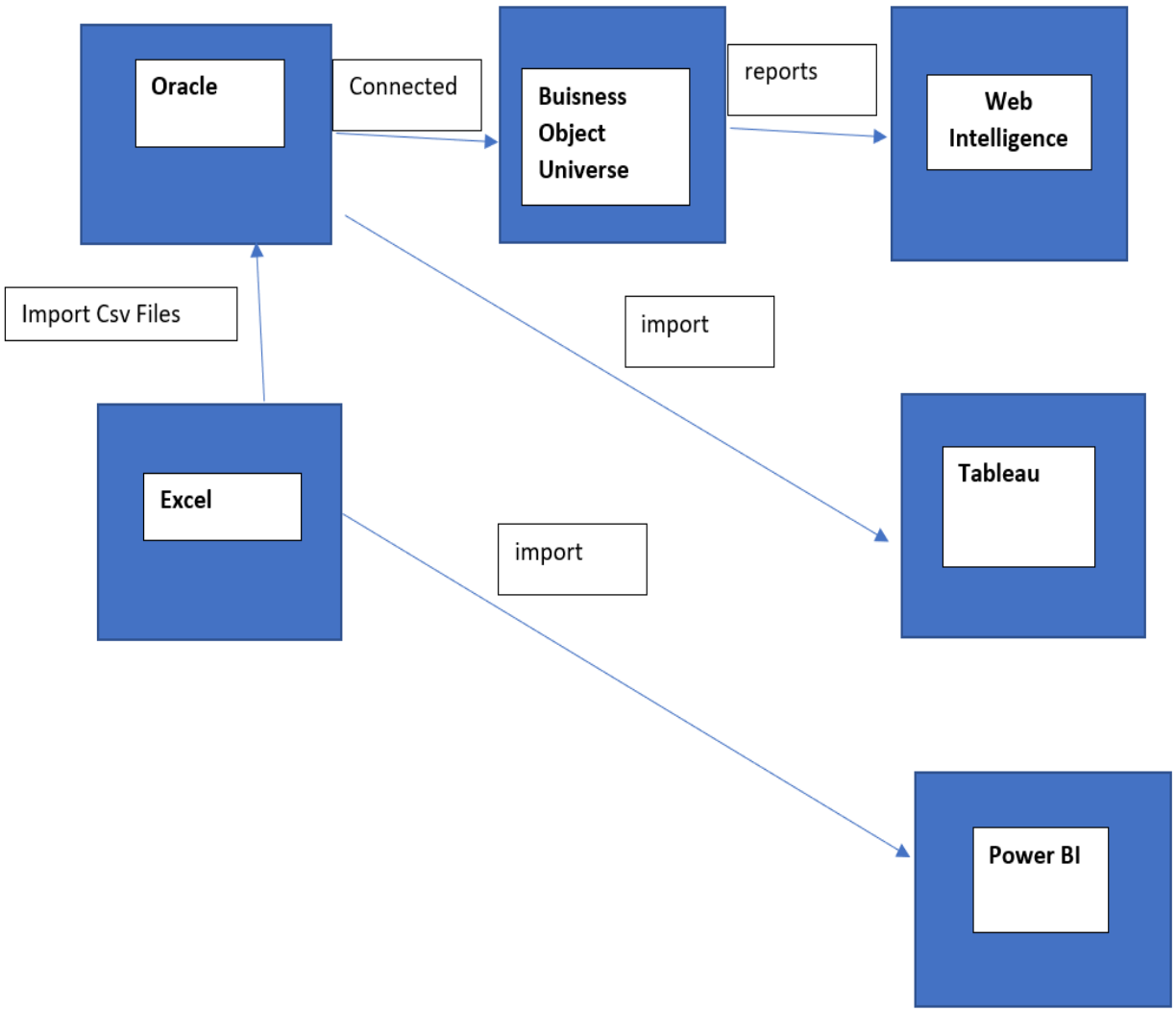


Figure 1: Architecture of System

Description of Data Warehouse:

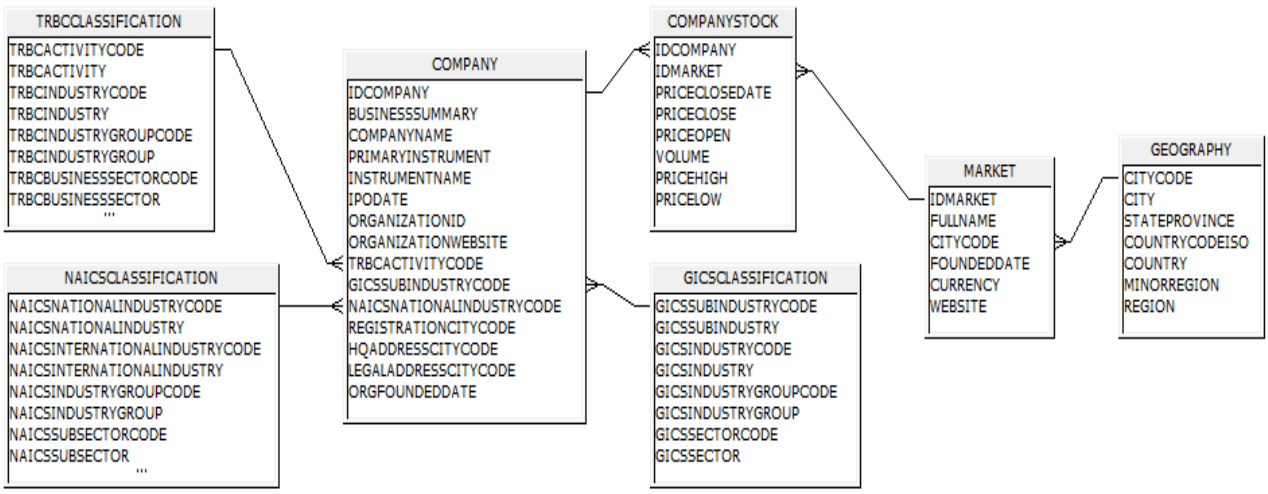
This database represents the stock market values of a series of companies. The data of this database is extracted from the Thomson-Reuters Eikon DataStream database (available at the University Library). Every market day (stock markets are not open every day and this depends on the market country) company shares values and some market indexes are recorded. Companies are classified according to administrative information (mainly relative to the stock market and the nationality of the headquarters of the company) as well as three business classifications (the Thomson Reuters Business Classification—TRBC, the Global Industry Classification Standard—GICS and the NorthAmerican Industry Classification System—NAICS). Companies put on sale on a stock market place financial instruments (usually shares, but this may be obligations, bonds, etc.). Stock market places monitor the exchanges through indexes (Dow Jones, Nikkei, CAC40, etc.) that represent the average activity of a certain number of financial instruments from some of the highest capitalized companies. After importing the data into Oracle, we modified the date format of priceClosedDate to the one as 'YYYY/MM/DD' in order to avoid confusion and error in following parts.

Sql Queries and Analytical Sql Queries:

A separate pdf named "queries report" is attached.

Business Universe:

Our fact was to study average volume and average price close. Then we made three dimensions: (1) Time (Time hierarchy only) (2) Market (Market geographical location hierarchy only) (3) Company (3 sub-hierarchies (TRBC, NAICS, GICS)) in order to study performance of company stocks based on different axes.



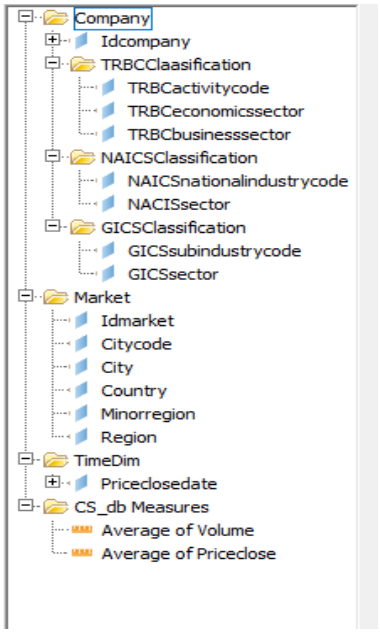


Figure 2: Business Universe Schema

A separate file named “Bo_Universe.unv” is attached.

Multidimensional Schema:

A separate pdf named “conceptual_Logical_Schema” is attached.

Dictionaries:

A separate pdf named “dictionaries” is attached.

Business Objects Reports:

(i) Top 3 Companies with highest average price close based on NACIS Sector during 2008 and 2009

A separate pdf named “IstReport” is attached.

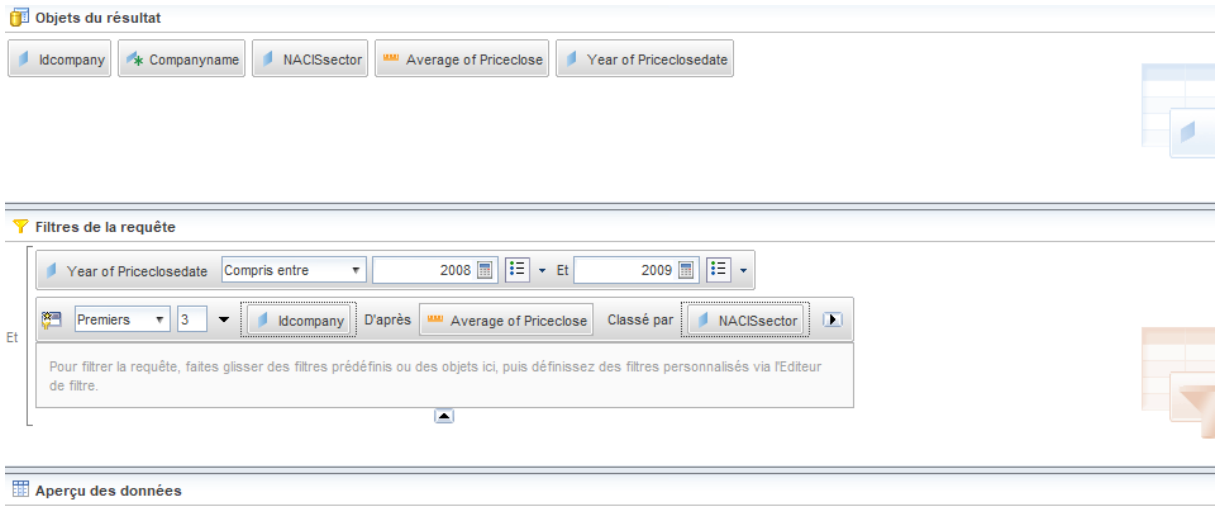


Figure 3: Filters Applied in first report

Using this filter allows us to get the three companies that had highest average price value recorded when the stock market closed by NACIS sector. Then from the result table of this query we calculate how many companies have appeared in this table per sector to know whether the top 3 positions of a sector changed a lot, as an indicator of how competitive the sector is. We found out that the top 3 posi

tions in each sector barely changed which is understandable because we only have data of 2 years. Then after making sure that the competition in each sector is stable, we tried to use the top 3 companies as the representatives of their respective sector, to see the performance of each sector in year 2008 and 2009.

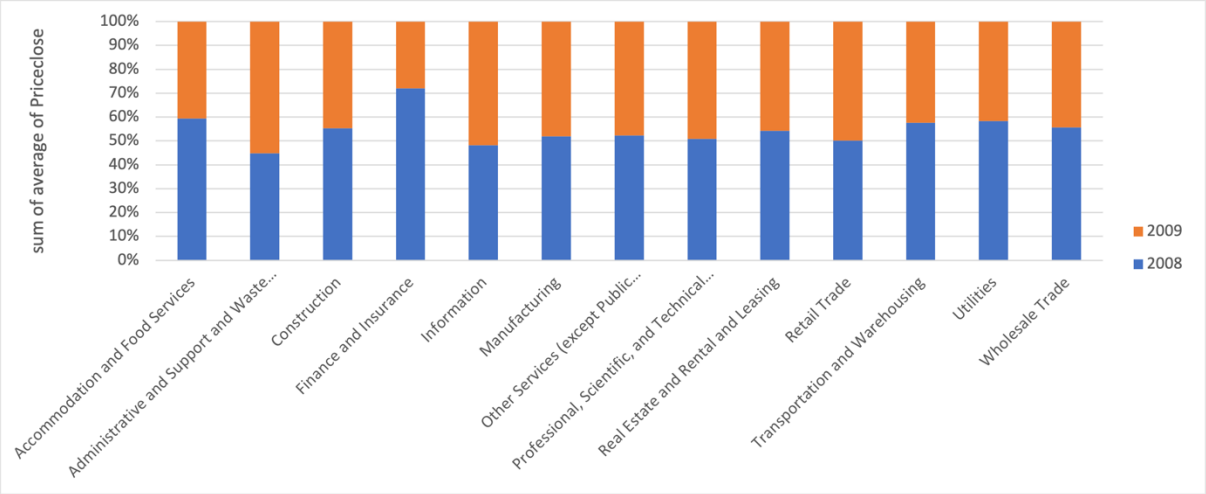


Figure 4: Top 3 companies with highest average price close

From this chart we can see the distribution of the sectors' market value in 2008 and 2009. For example, we can see that Administration and Support sector had higher value through 2009 while Finance and Insurance sector faced a huge decrease. This can help us know better the performance by sector and with help of more data from other years, we may also tell the tendency and even make prediction of the sectors' market value.

(ii) Comparison between volume and price close of TRBC Technology Sector

A separate pdf named "2ndReport" is attached.

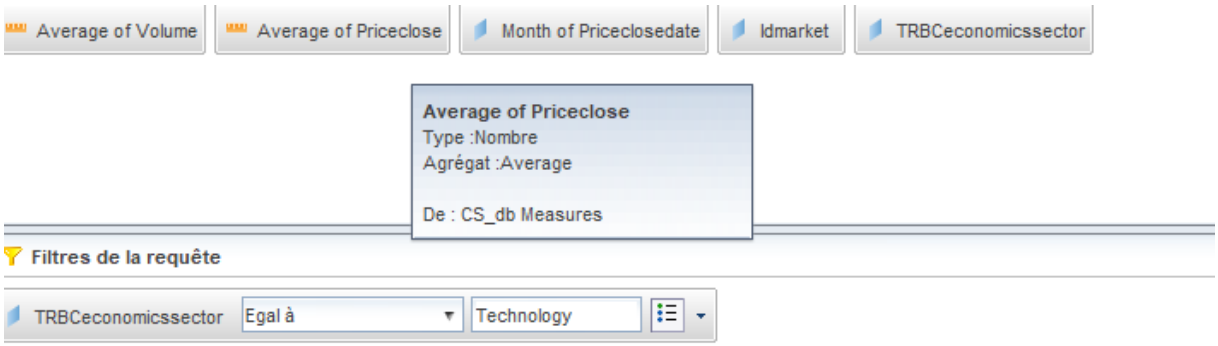


Figure 5: Filters Applied in 2nd report

For this report we compare the average volume and average price close of TRBC technology sector to see if the number of financial instruments exchanged and the price value have correlation between them. We picked technology because compared to other sectors this is the one that had highest price and volume — far beyond the other sectors thus makes it unreasonable to compare between sectors.

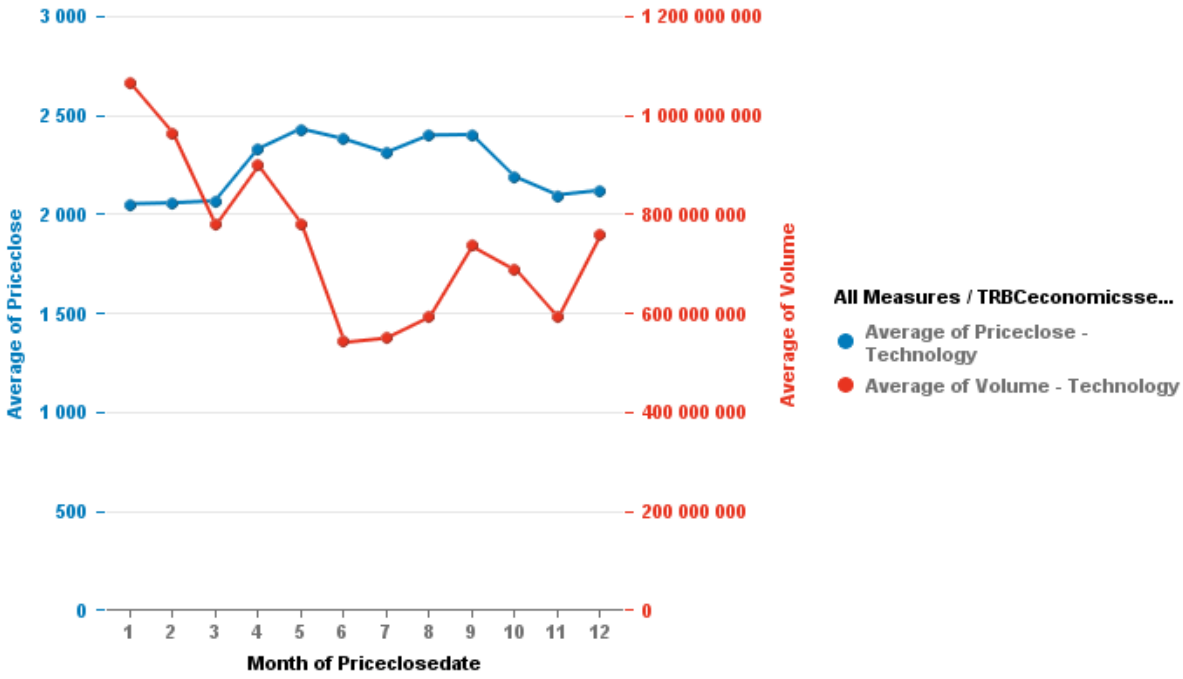


Figure 6: Comparison between avg priceClose and avg Volume in Trbc Technology

After executing the query, we got the result table then generated this line chart. The reason why we chose month instead of year is that we only have 2 years' data, so it is hard to find pattern in term of year in this case. In this chart we can see a huge drop of volume exchanged during summer and the rise of price close in this period. This could be a seasonal pattern and the reason behind it could be useful for the companies in this sector and investors. Besides the respective rise and fall of volume and price close, we can see that when price value increased the volume would decrease, we can thus conclude that these two have a negative correlation though the reason behind it will need more study.

(iii) Comparison of average volume by region

A separate pdf named "3rdReport" is attached.

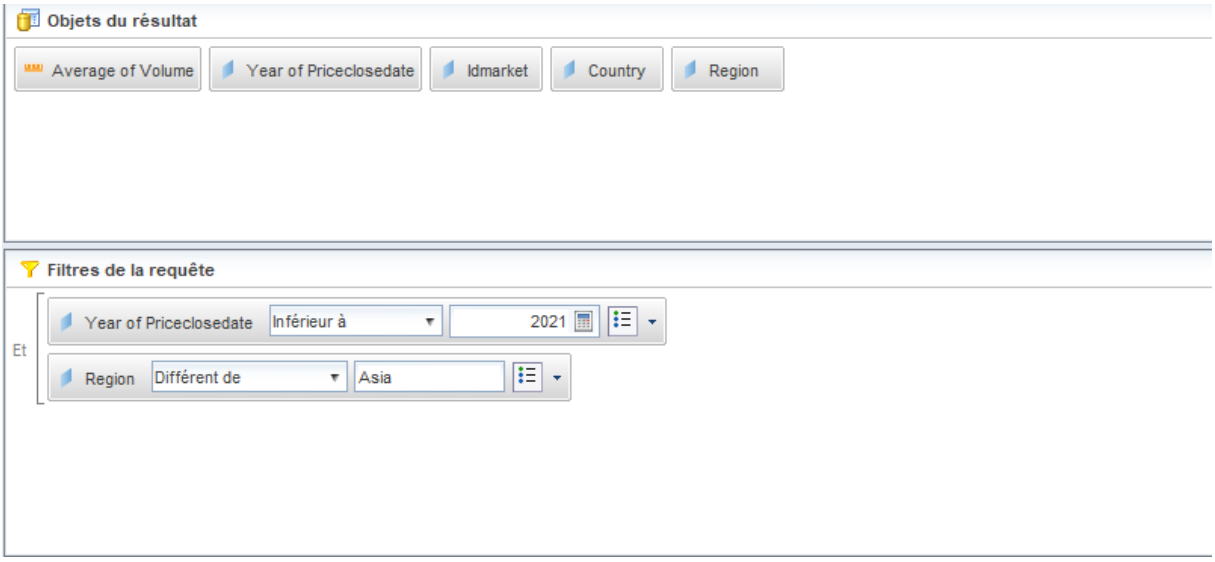


Figure 7: Filters Applied in 3rd report

For this query we would like to see the how busy the markets in different regions are, in term of the volume exchanged in these markets. We originally wanted to see the comparison between all three regions in the database, but it turned out that the volume exchanged in Asia was so big that it made the comparison meaningless — we could barely see the other two regions in the chart. It explains why we added a filter to exclude Asia from the regions.

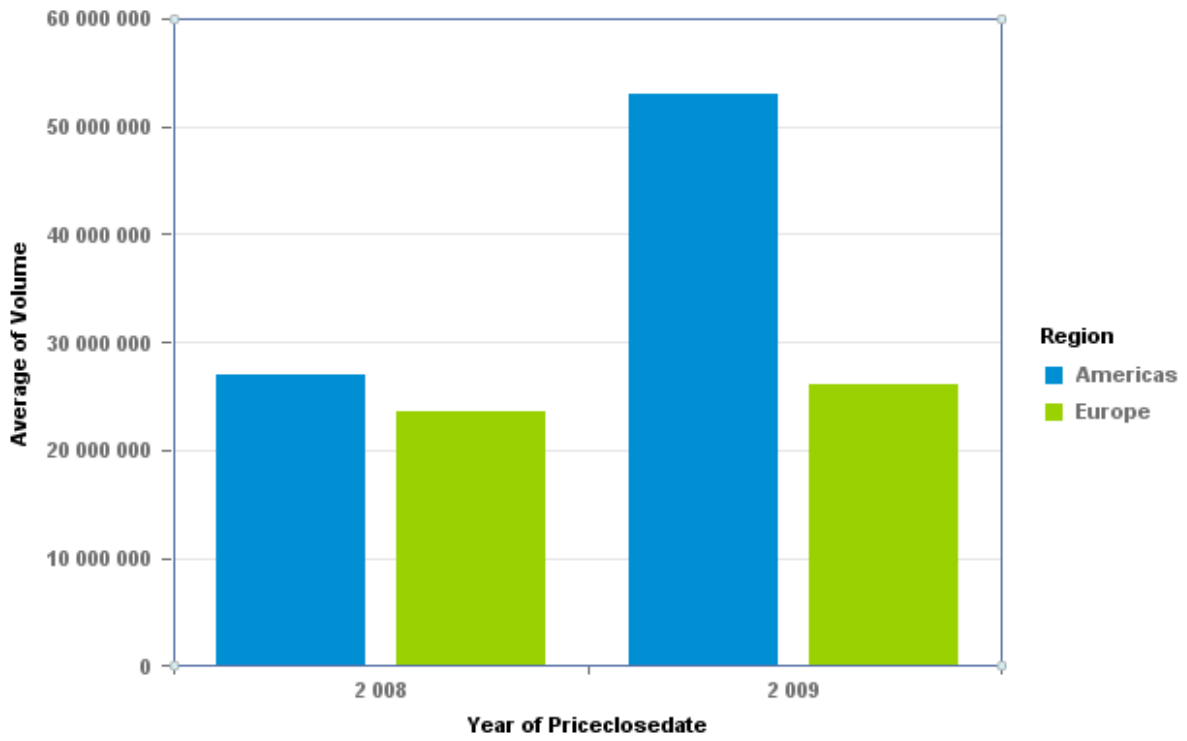


Figure 8: Comparison of Average Volume by Region

This chart clearly demonstrates the difference and respective growth of volume in America and Europe. For both years, the American markets had more volumes exchanged than the European ones and we can see that the American markets also had a larger growth in year 2009. It could be that the American countries recovered more quickly from the financial crisis than European countries, but it is just a hypothesis before studying more data.

Comparative study of the reporting tools (Tableau/PowerBI/Zoho Analytics):

	TABLEAU	POWERBI	ZOHO ANALYTICS
ADVANTAGES	compatible with OLAP, plenty choices for data import, drag&drop, can merge several reports into one	Integration among other Microsoft Products, easy to start with for Excel users	can import data from csv and other sheet form, drag&drop, automate data refreshes
DISADVANTAGES	Slow when processing large files, no auto-save (risk of losing work when it crashes)	Cannot use multiple datasets as sources, relatively slow, cannot connect to google sheets	not compatible with OLAP, expensive for small business
FUNCTIONALITIES	Data Import/Export, Ad hoc Query, Annotations, Customizable Dashboard, Collaboration Tools……	Data Import/Export, Activity Dashboard, Ad hoc Query, Collaboration Tools, Data Visualization, Data Blending, Data Cleansing……	Data Import/Export, Ad hoc Query, Data Connectors, Dashboard Creation……
PRICE	70$/month for one user or 145$ for one user+5 viewers	9.9$ or 20$/month for individual users, 4995$/month for organizations	On-premise: free for personal use, 30€/user for professional use. Cloud: 24€ - 455€/month based on the number of user and extra features.

A report done in Tableau:

A file named “TableauReport.twb” is attached.

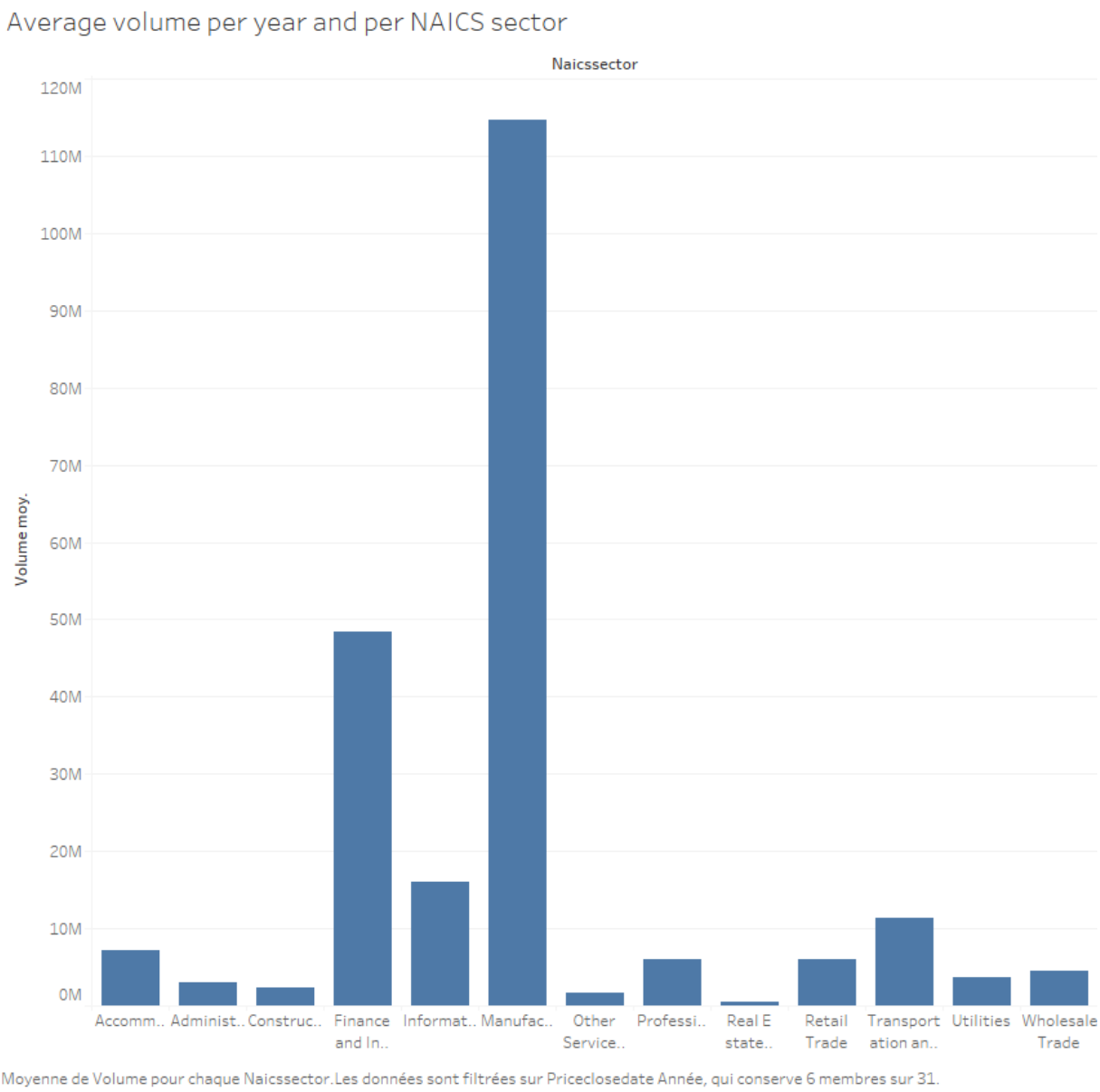


Figure 9: a report in Tableau

During the use we found that its drag&drop functionality is very user friendly, and it offers more flexible chart choices than SAP BO. The best part is that it can merge several dashboards into one report thus we do not have to export one by one. However the query filters are not as clear as the one in SAP BO it also has higher crash rate, for example we lost our work by the end of the day but it never happened with BO.