

CMSC838S: Information Visualization

Application Report

Analyzing the Forbes 2000 List using Spotfire

V. Shiv Naga Prasad - shivnaga@cs.umd.edu

Introduction

We analyzed the Forbes 2000 List for the year 2004. Forbes 2000 is a ranking of the world's 2000 biggest companies in terms of their sales, profits, assets and market value. It includes companies from 51 countries, spanning 27 industries, and represents a sizeable chunk of the world's commerce.

The list consists of 2000 records, each having 7 fields: ranking (1 to 2000), company name, country, sales, profits, assets and market value (in billions of U.S. dollars). We thought the list would be interesting as there is ample scope for comparison of statistics of different industries, searching for companies that are exceptions to industry statistics etc. Spotfire is suitable for visualizing the Forbes 2000 data as the list does not contain temporal information and is relatively low dimensional; otherwise TimeSearcher or HCE would have been more suitable.

There are several interesting insights that could be obtained by exploring a rich data set like Forbes 2000 – especially with some domain knowledge. We will highlight only a few of them, emphasizing the role played by Spotfire in their discovery.

Correlation between Assets and Market Value

In general market value of a company is correlated with its assets. However, in the scatter plot of the market value versus the assets for all companies, we can see two regions of exceptions – one for the cases in which a company has large amounts of assets but low market value, and the opposite. Selecting the records within these regions and using the labels to give details about their industries showed the reason for the exceptions.

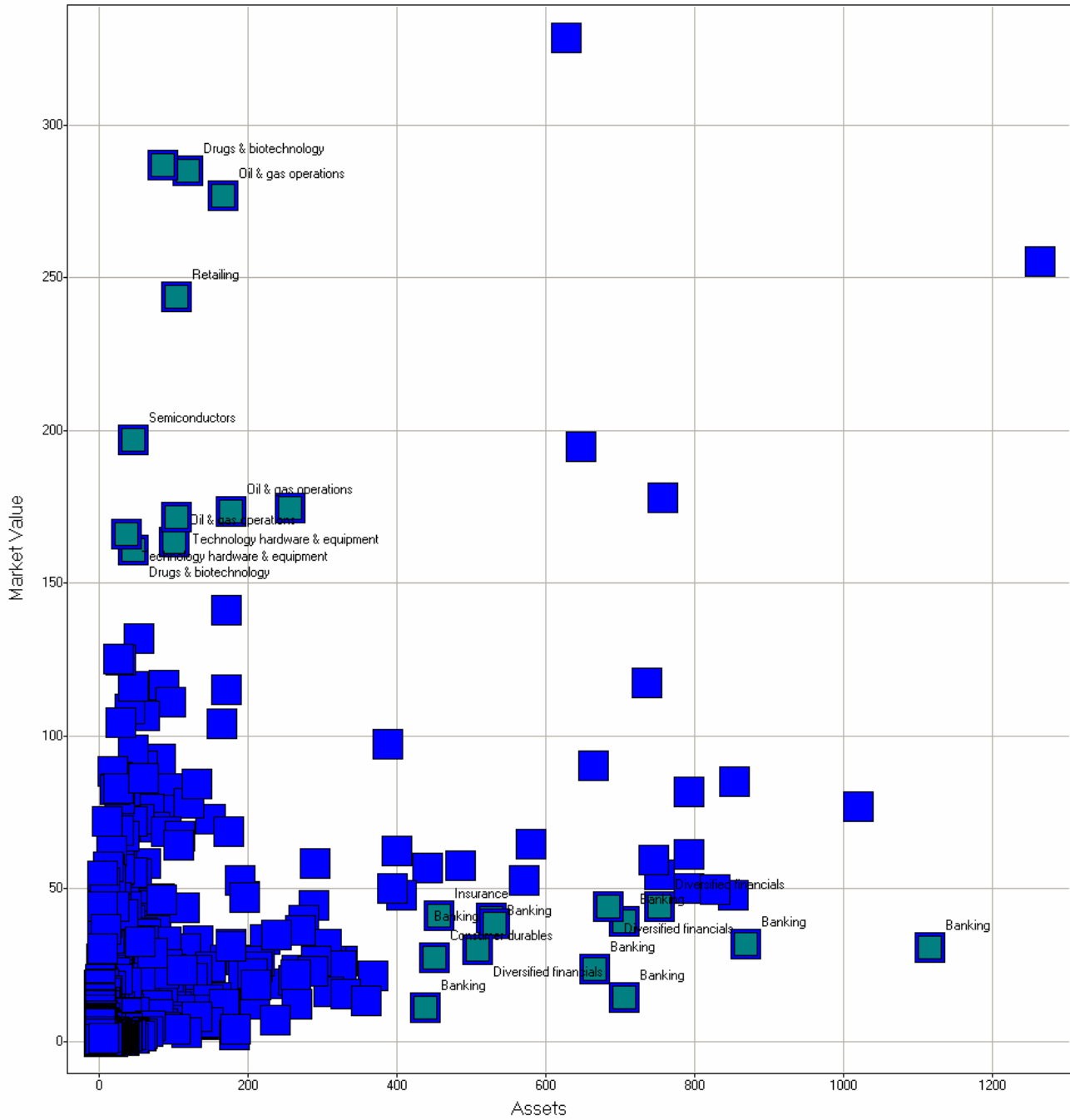
The companies having large assets but low market value usually belonged banking, finance or insurance companies. Companies in these industries are characterized by the risky nature of their assets. Therefore, the face value of their assets does not completely determine their market value – several other factors like risk assessment and macroeconomic behavior come into play too.

On the other end of the spectrum, companies able to attain high market value without hoarding up too many assets usually belong to software or technology companies. In their

case, the value of the basic infrastructure is much smaller; the market value depends more upon their future prospects and current profits.

This highlights the utility of Spotfire for not only visualizing global trends, but also for discovering and assessing exceptions to these trends.

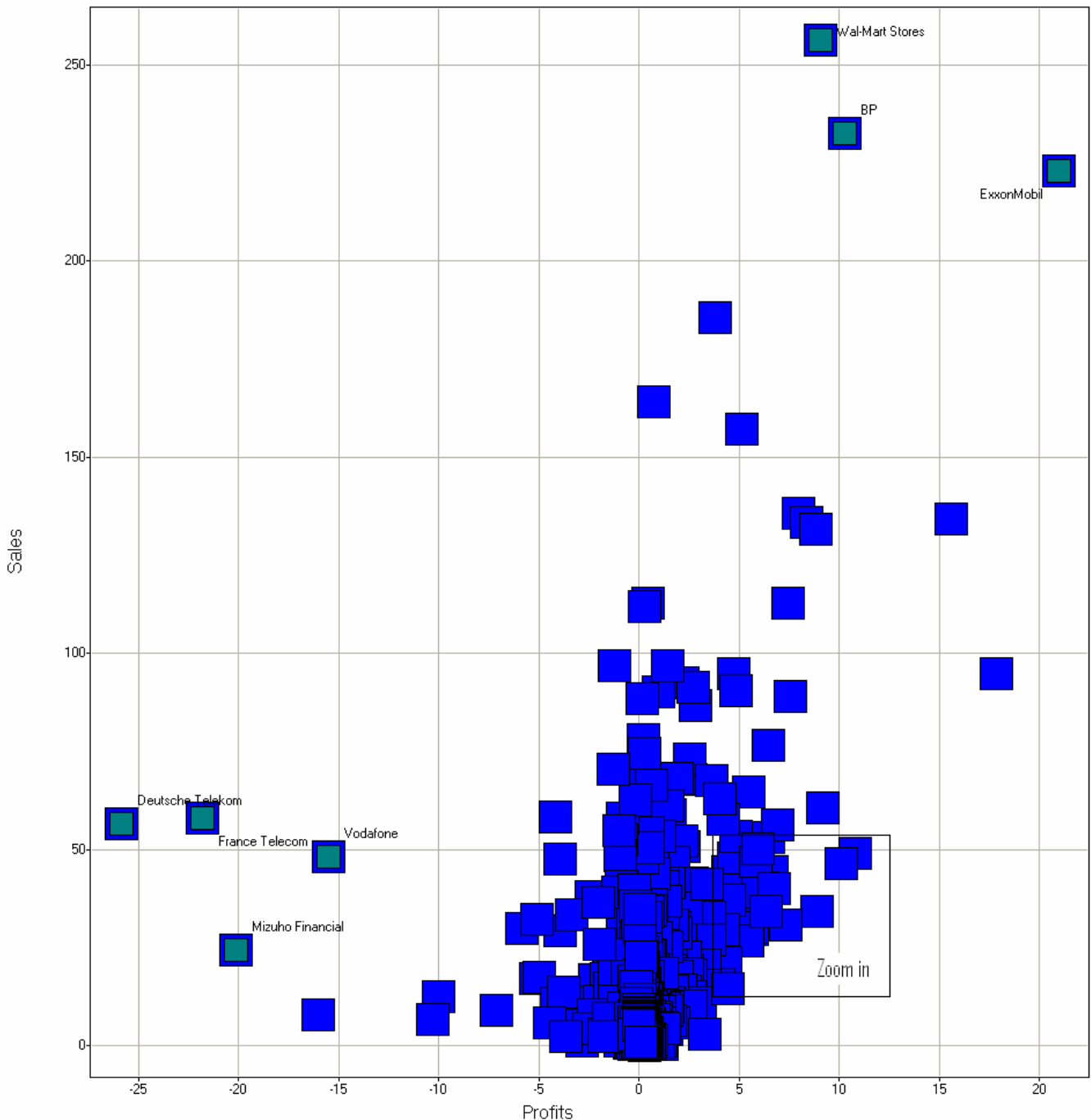
Scatter Plot of Market Value vs. Assets. For All Companies



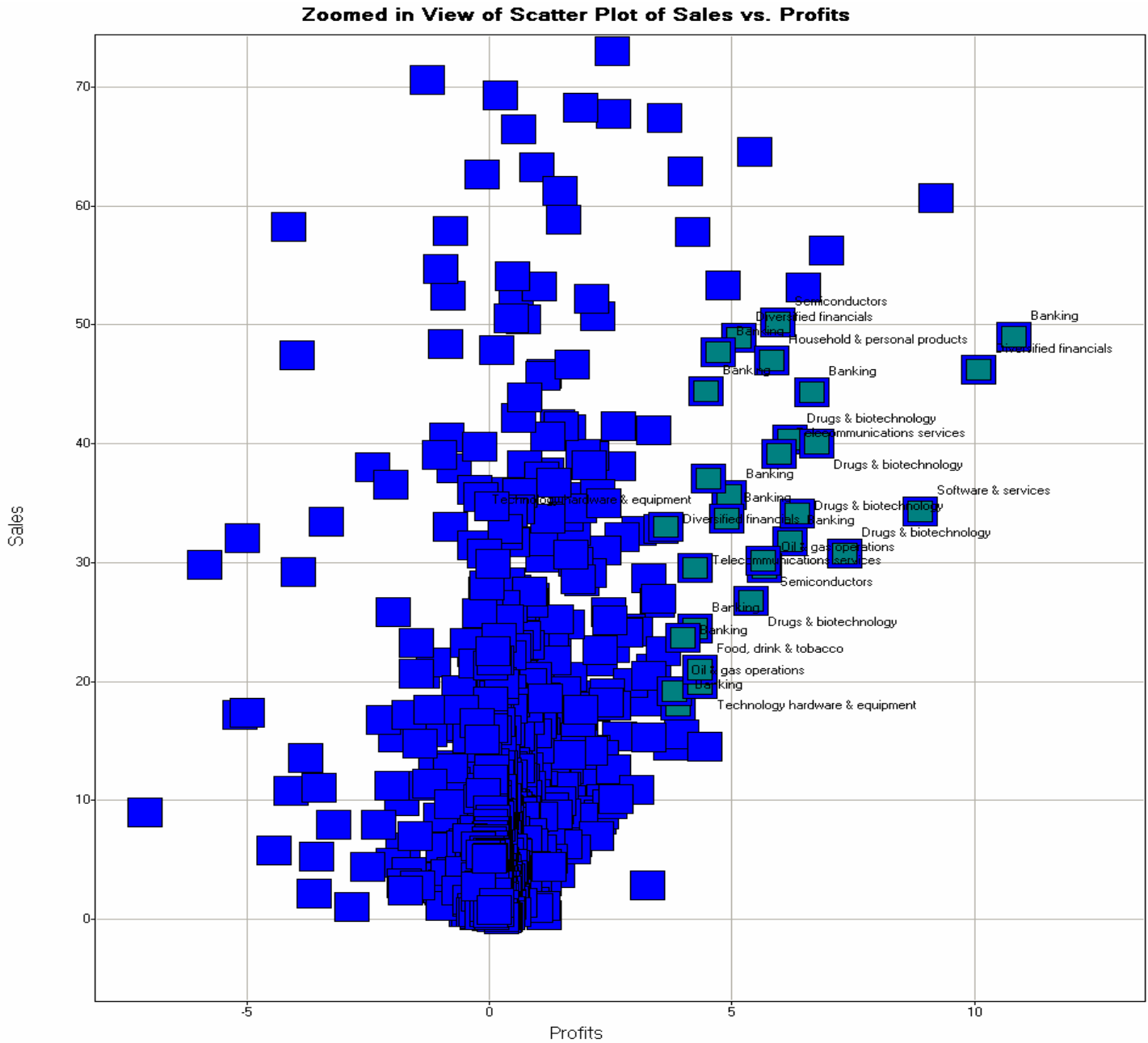
Analysis of the Sales vs. Profits

The dependence of profits on sales varies with the industry – for high volume sectors, large sales are need for high profits. On the other hand technology firms usually have high profits in spite of relatively low sales volumes. We drew a scatter plot of the sales with respect to the profits for the listed companies. For most companies, high sales usually implied profits – the exceptions being three telecom companies and one finance company. Checking recent business articles showed that these companies were indeed in the “risky stock” category. The other exceptions were Wal-Mart, Exxon and B.P., which had exceptionally high sales volumes – Wal-Mart of course leads the pack.

Scatter Plot of Sales vs. Profits



The plot also shows a cluster in the region corresponding to high profits with relatively low sales volumes. The second plot in this section shows a zoom-in on this region. Most of the companies turned out to be belonging to technology and biotech sectors. This highlights the zoom-filter feature of Spotfire.



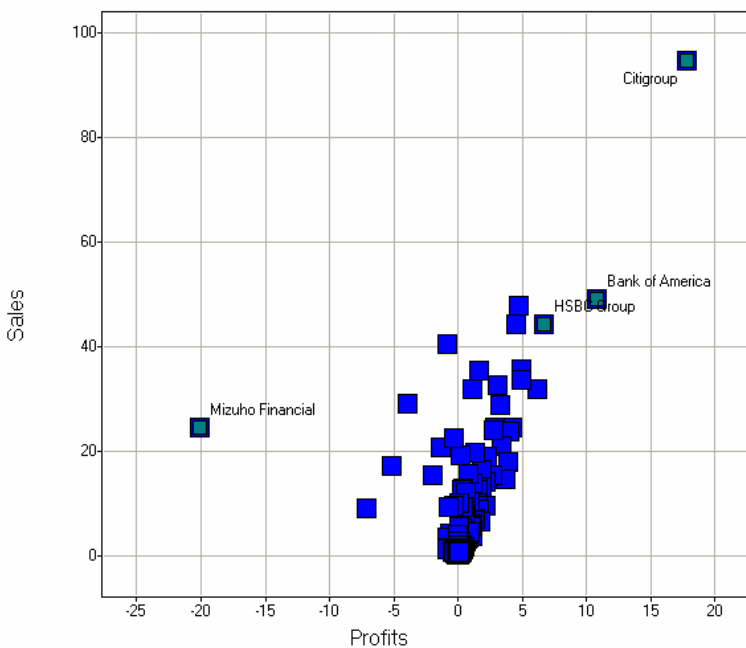
Interesting Observations about Particular Industries

Next, we explored the statistics of each industry using the scatter plots tool. Outliers turn out to be the key source of information in these plots as they correspond to companies showing exceptions to the industry trends. This could be used to instantly spot “winner” and “loser” stocks. The visualization was enhanced by Spotfire’s ability to link selections across different scatter plots. Using this feature, it was easy to verify whether certain companies consistently out/under performed the norm.

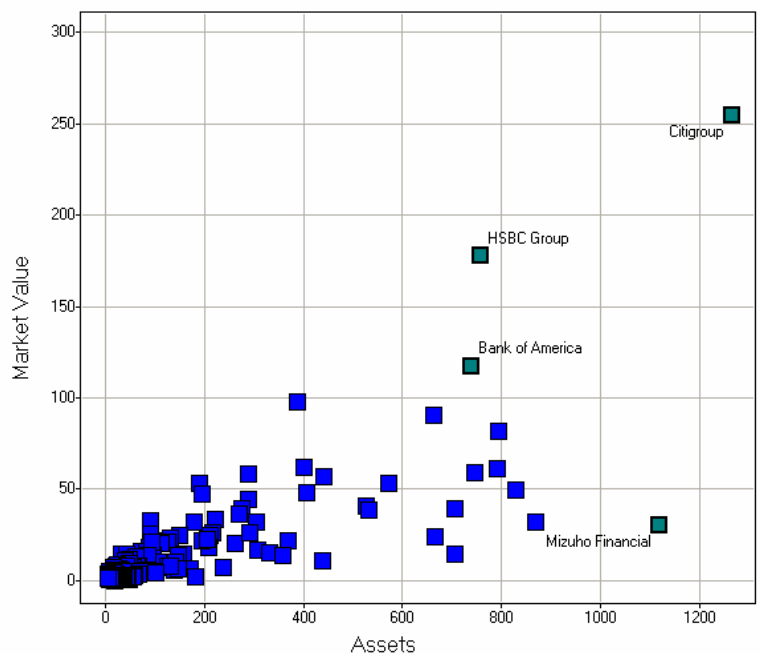
Example 1: Banking Industry

The following scatter plots show variation of sales vs. profits, and market value vs. assets for companies belonging to the banking industry. We can see that three companies - Citigroup, Bank of America and HSBC group - out perform others by big margins – they are the “good” outliers. This is reflected in their relatively high market values (Citigroup is ranked 1 in Forbes 2000). On the other hand, the company named “Mizuho financial” has exceptional losses even though it has large volumes of sales and assets. This is a basic signature for risky stock and is indicated by its low market value and rank.

Scatter Plot of Sales vs. Profits for Banking Industry

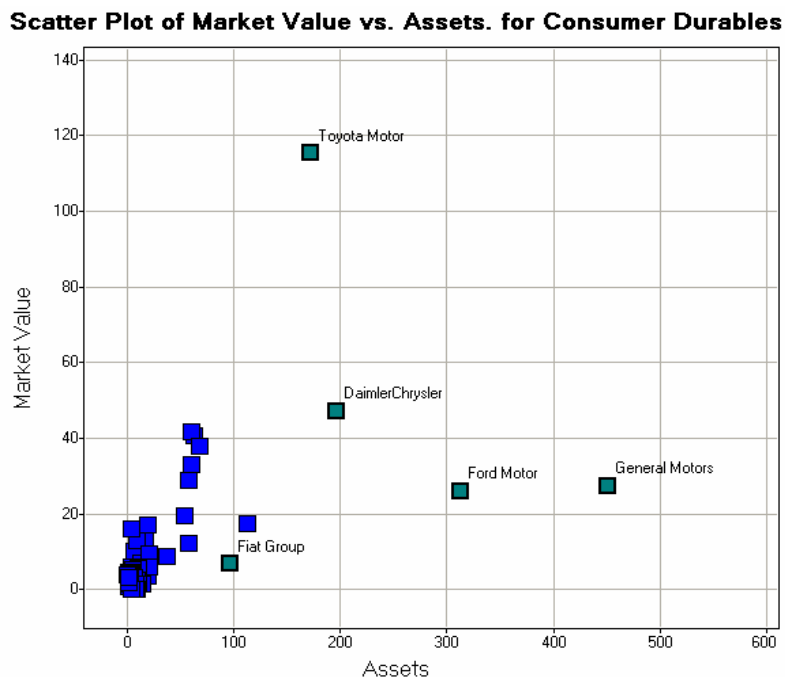
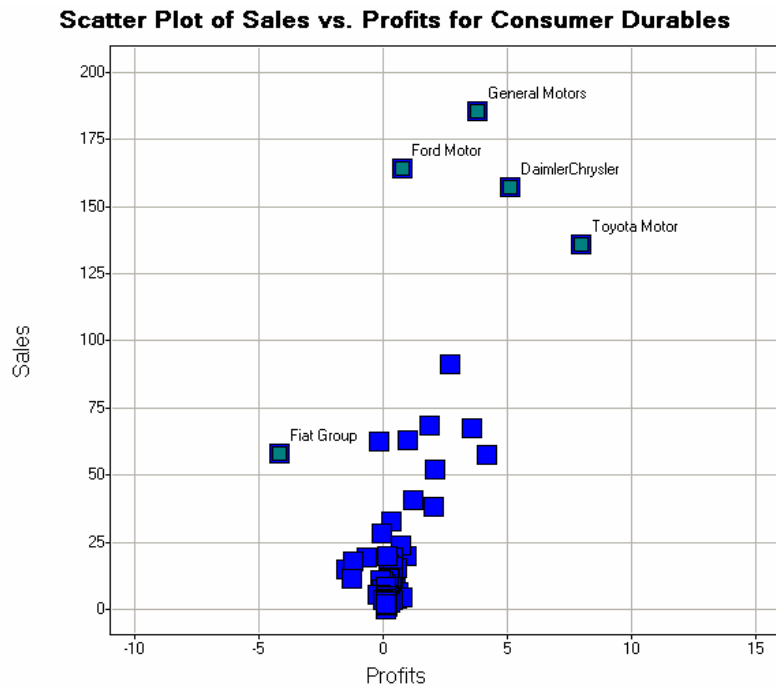


Scatter Plot of Market Value vs. Assets. for Banking Industry



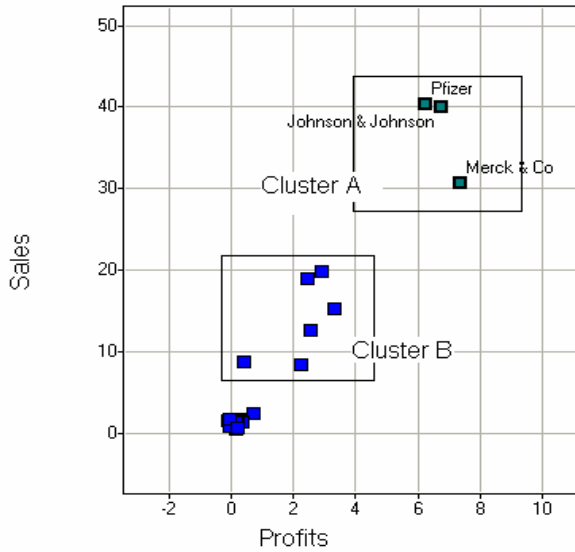
Example 2: Consumer Durables (Automobile) Industry

The plots for the consumer durables or automobile industry show similar characteristics. This industry is characterized by the need for large investment in infrastructure and emphasis on sales volumes. We can see from the sales vs. profits plot that high profits require large sales volumes. Moreover, keeping the assets low leads to low risk and is considered a good management practice. This is clearly reflected in the outliers – Toyota with its high sales and low assets, has the highest market value and has the top rank in this industry.



Example 3: Pharmaceutical Industry

Scatter Plot of Sales vs. Profits for Drugs

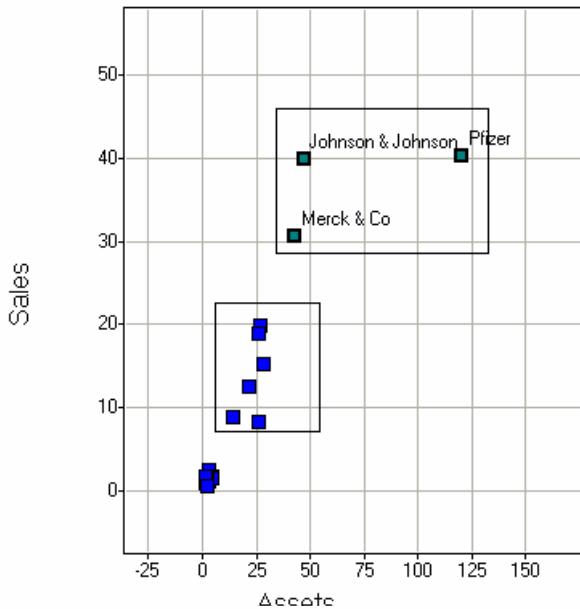


The sales vs. profits scatter plot of the pharmaceutical industry showed an interesting clustered distribution. We refined the view by plotting only US based drug companies – shown on the left.

We can see that the companies form three distinct clusters. These clusters remain consistent across the three scatter plots. The clusters are indicated with rectangles in the different plots.

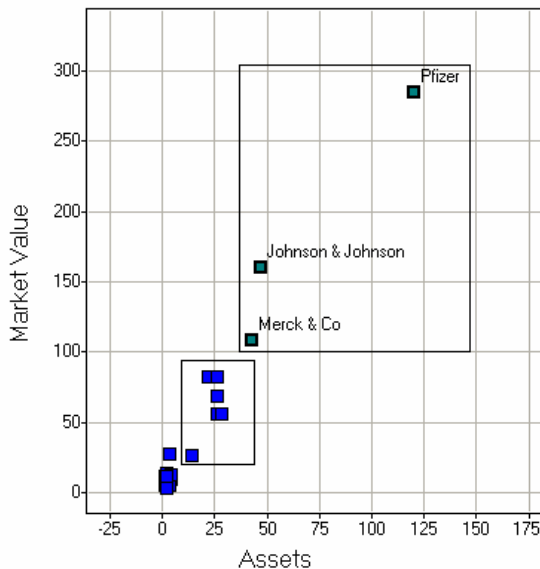
The drugs industry is characterized by the emphasis on volumes of sales. The profits and hence the ranking is determined to large extent by the volume of sales. In fact, the three top companies (Cluster A) together constitute more than a quarter of the total sales of all the 21 companies put together.

Scatter Plot of Sales vs. Assets for Drugs



We can see that the ability of linking the selection of records across different visualizations helps in spotting trends.

Scatter Plot of Market Value vs. Assets



Summary

We presented some observations about the statistics of companies listed in Forbes 2000 made using Spotfire. The exercise brought up the following aspects of Spotfire:

- The process of dynamic and interactive exploration with Spotfire is much more rewarding and efficient than by using a programming-based data analysis tool like Matlab etc.
- The ability to link selections across visualizations is crucial.
- Even simple tool like a 2D scatter plot can be effective, provided the user is allowed to easily modify it. The quality of the insights depend more upon domain knowledge and familiarity with the interface.
- We found it difficult to get the pie chart tool to work. The menu structure for this tool was not intuitive and the help was not very useful.