



DAX30 Price Risk Clustering

By:
Parker McClearn
 &
Spencer Embry

STEP #1

Business Objective

Clustering time series data with the DAX 30 stocks.

STEP #2

Background/Business Understanding

The DAX30 is a portfolio of thirty top companies within in the German market. Each of these companies are unique in their growth, but can also be key indicators in relation to how other stocks within this market index behave. In using cluster analysis, relationships can be made between individual stocks. These relationships provide insight for which companies within the market move similarly. This can be very beneficial to investors in order to create a diversified portfolio that provides maximal growth.

STEP #3

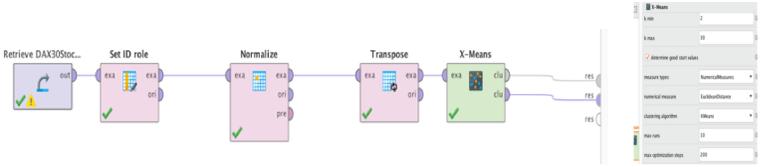
Data Description

Identify the proper trading station with the correct stock name and verify the closing stock price with the date given. The data contains 30 attributes (company stocks) along with stock prices ranging from 2004-2008. The date column's data type is polynomial and the 30 company stock attributes is real.

STEP #4

Data Preparation

Our first step in the data preparation process was to download the DAX 30 stock data. Then we set the date column to role ID. After setting the role ID, we standardized each price time series to perform a Z-transformation of the values so that they have a mean of 0 and a standard of 1 afterwards. For step 3, we transposed the data set (so that each time series is a row) and cluster the data so that each series falls into one cluster. The k min parameter= 2 and k max= 30, max optimization steps = 200 and max runs = 10.



STEP #5

Modeling

Clustering is the task of grouping a set of objects in such a way that objects in the same group are more similar to each other than those in other groups. After we ran our regression our stock data was compiled into 3 clusters. Cluster 0, 1, and 2. The DAX 30 contains 30 different stocks and after we adjusted our K-means and ran the regression, cluster 0 contained 21 stocks, cluster 1 = 8 stocks, and cluster 2= 1 stock. Each stock in each cluster performed in a similar way historically over time. Cluster 0 and 1 contains more established stocks. Therefore, these stocks had a larger negative impact on the stock market crash.

Cluster Model

- Cluster 0: 21 items
- Cluster 1: 8 items
- Cluster 2: 1 items
- Total number of items: 30

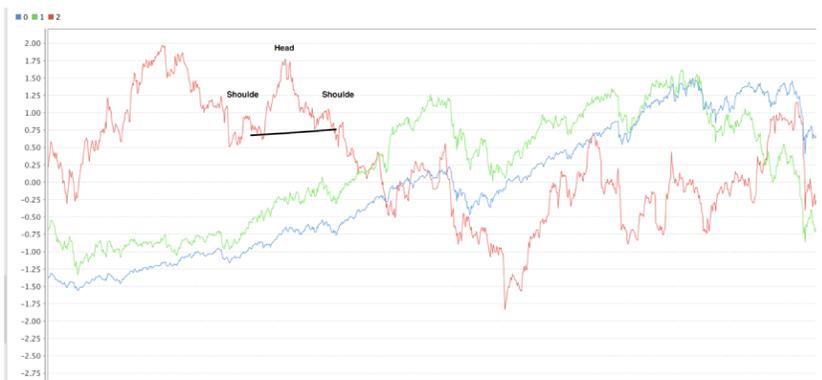
STEP #6

Evaluation

Once the items were clustered into three separate groups, it was essential to find a model in which best displayed the relationship between the clusters. In looking at various graphs and charts, we determined the plot of the X-means provided us with the most practical insight. In looking at Figure 2, the trends of each cluster over the time period are depicted. This graph is very similar to a typical chart you would see when looking at any stock market data, the only difference being the clusters. However, with the clusters added, other information can be derived in further analysis. In looking at clusters 0 and 1, it is clearly evident these clusters are similar and follow some of the same trends within the market. Cluster 2 is quite the opposite. This cluster only consists of 1 stock and is an apparent outlier in relation to the other 29 companies within the DAX 30. One other key observation to note is the fact that cluster 2 (Deutsche Telekom) seems to be inversely related to the other clusters at some point. In noticing this trend, one could use this information as safeguard for their investments.

Through the cluster 2 stock data we were able to use the head and shoulder method to confirm a bearish run in the future. The head & shoulder pattern is one of the most popular and reliable chart patterns in technical analysis. The head and shoulder pattern consist of a left shoulder, a head, and a right shoulder, and a line drawn as the neckline. The left shoulder is formed at the end of an extensive move. After the peak if the left shoulder is formed, there is a subsequent reaction and prices slide down to a certain extent. The prices rally up to form the head and subsequent reaction downward is formed with lesser volume. Next, the right shoulder is formed when prices move up again but remain below the central peak (head) and fall nearly equal to the first valley. The price then breaks below the neckline to complete the pattern. The price break indicates the increased urgency on the part of the sellers, which confirms that the stock is following a bearish (downward) trend.

Although this graphical depiction doesn't provide all the key figures for accurate predictions, it does provide great insight in regards to the trends of these clusters throughout the time frame. In grouping these clusters, the relationship between the pricing and risk of these stocks can be understood more thoroughly. A variety of other stock market models can be applied to this plot in order to find more predictive information. With our project however, the main objective was clustering the 30 stocks based on price risk.



STEP #7

Deployment

After analyzing the data and looking for ways to utilize this information, it becomes evident that the data can be used in a variety of ways. The main way clustering can be used in order to benefit someone with regards to investing is in diversifying/strengthening a portfolio. In knowing that two clusters operate based on similar trends, one could invest more between the two due to the fact they will likely produce the same levels of growth. If one knew that certain clusters will follow the same trends, they could "double down" and likely gain twice the profits. This information could be used in a variety of ways. For instance, both cluster 1 and 0 are nearly identical leading up to the 2008 market crash. Once the crash occurred, a significant difference can be noticed. If one were to only invest in cluster 1 companies, they would have experienced sharp declines, but if their investments were split between cluster 0 and 1, they would not have suffered as much loss. Cluster 2 would have helped to balance the portfolio out and reduce the amount lost. Diversifying a portfolio has always been acknowledged when investing, but without clustering different stocks, it can be difficult to determine what stocks need to purchase in order to create a balanced portfolio.

STEP #8

Conclusion

Clustering is a technique that can provide great information for a variety of data types. When dealing with finance and the stock market, data is extremely rich in order to proceed and make profits. Combining the power of clustering with the plethora of data that is available through various market archives can be extremely beneficial in order to reduce risk and operate more efficiently. Algorithmic data models are quickly becoming a cornerstone for individuals wanting to succeed in the financial sector. In knowing that, becoming familiar with clustering and k-means can be one of many techniques utilized in order to improve investment performance.

Reference

<https://www.schwab.com/active-trader/insights/content/how-use-head-and-shoulders-chart-patterns>
https://en.wikipedia.org/wiki/Cluster_analysis