Visual analytics and financial, non-geographical Time series: A review

Asra Khalid, Marc Kendal, Max Lensvelt, Karunakar Kotagaram

City University London, UK

Figure 1: A TreeMap representation of stock prices on the S&P500 index.

Abstract— This paper seeks to provide an introduction and overview of the ways in which visual analytics have been applied to non-geographical time series, in particular time series within the financial domain. In addition to an overview and discussion of the range of methods involved, the drivers of the growth of visual analytics within financial time series are explored and potential future developments are discussed.

1 INTRODUCTION

For good reason, the term ‘ubiquitous’ is used repeatedly in the papers reviewed herein to describe time series data. Time series data is encountered regularly in everyday life and found in a broad range of sectors, not limited to:

- health
- power grids
- finance
- telecommunications
- industrial processes

It consists of “sequences of numbers, usually representing measurements or observations of a real variables at equal time intervals” (Buono et al., 2005). This literature review is focused on ‘non-geographical time series’, but for brevity we will frequently refer to ‘time series’, with an implicit assumption that we are alluding to the non-geographic forms.

Within time series, the most common questions being asked are whether patterns exist within a certain time series or between a number of separate time series. These patterns are interesting to analysts for two interrelated reasons: they a) may allow for prediction of the future behaviour of the variable(s) in question or b) because they will allow for better understanding of why certain phenomenon acted in a certain way. For each of the sectors mentioned above we found one (or both) of these drivers at play:

- Analysing data produced by electrocardiograms to predict repeat heart attacks (El-Sherif et al., 1995)
- High volumes of energy consumption time series data are studied (along with weather and other forms of data) by energy traders to predict day-ahead power markets in order to optimise their purchasing and selling of power (Dagnely et al., 2015)
- Prediction of future stock price movements using historical behaviour (Lee and Jo, 1999)
- Analysis of Ohio State university internet traffic data to better understand application types and eventually to improve prediction and forecasting of network traffic statistics (You and Chandra, 1999)
- Research to better understand anomalous data and outliers within time series generated by industrial processes (e.g. chemical plants) (Brighenti and Sanz-Bobi, 2011)

This paper concerns the methodologies by which visual analytics can be used to aid analysts in achieving these aims. Due to the mentioned breadth of uses for time series data, we will focus particularly on one of the most frequently studied forms: financial data, and again more specifically on asset price data (e.g. equity markets). Given the vast amounts of money at stake in global financial markets, financial time series data is especially driven by the question of how to predict future price movements, although we also encountered analysis driven by desire to better understand important historical economic events like the ‘Dot-Com’ or 2008 global stock market crashes. For an analyst working with financial time series data a primary challenge is the volume of data that is available and relevant for an analyst. (Merino et al., 2006) express this challenge, and the resulting need

“Today’s financial community is being bombarded by massive amounts of information from real-time data sources. The market consists of thousands of companies quoting in different stock markets in many countries around the world, with a continuously increasing number of private and public investors selling and buying stocks. Every few seconds hundreds of trades characterized by stock price and trading volume are accomplished and observed through seconds, minutes, hours, days, months, and years. […] For data exploration to be more effective, it is important to include the human being in the data exploration process and combine their flexibility, creativity, and general knowledge to enable a better interpretation and evaluation of the probably plentiful results, and hence, take the most out of them.”

This ‘inclusion of the human’ through interactive and iterative visualisation and analysis feedback loops is a helpful picture of the ‘visual analytics’ that we are interested in. Before progressing, one additional challenge associated with financial time series that we should highlight is the character of time series as both temporally linear and cyclical, i.e. depending on the financial asset in question, to best address the underlying patterns within the data it may be necessary to treat the time series as either a linear stream of values or as a repeating sequence of values distributed over a day, week, month or year (or, as we will see in section ii), as both simultaneously.

The outline of this paper is as follows: in section ii) we provide an overview of the methodologies that are frequently used or emerging within the financial time series domain, in section iii) we discuss the respective merits and limitations of such methodologies, as well as why and how the practice of visual analytics for financial time series is evolving, and in section iv) we conclude our review.

2 Methods

This section provides an overview of methodologies that have been used for visual analytics within the financial time series domain, and also for select other methods that have been used with non-financial, non-geographic time series, but which we believe could be used to good effect within the financial domain.

We have divided the methodologies into four categories: Visualisation, clustering, interaction types and a short note on data preparation.

2.1. Visualization Techniques

2.11 Line chart

The line chart is somewhat synonymous with financial data: the falling (or rising) stock price, shown on a line chart, is one of the most frequently used data visualisations in mainstream media. This is a key strength of the medium: it is easily and intuitively understood and simple to create. Much of the research focus in recent decades has been on increasing the utility of the classical line chart through combining it with additional, different plot types or adding to the complexity of the line chart to increase utility (see ‘Candlestick chart’ below or (Kong and Agrawala, 2012) or (Ming Hao and Umeshwar Dayal, 2007) for further examples of such adjustments.

Figure 2: A line chart from the Financial Times showing the movements of the S&P500 stock price index.

2.12 Candlestick chart/Box plots

While ‘classical’ box plots are somewhat rare in the visualisation of financial data, one of their children – the ‘Candlestick chart’ is broadly used. The Candlestick chart can be thought of as a combination of a line and bar chart, with the ‘box’ generally showing the high, low, open and close elements of the daily price of the asset in question. The Candlestick chart is of 19th century Japanese origin, and today is broadly used in quantitative financial analysis within industry, and has been widely used for economic research within academia, particularly in relation to stock market prediction (see (Lee and Jo, 1999) or (Tsai and Quan, 2014))

Figure 3: A Candlestick chart compared to a bar chart to display stock price movements for Bank of America.

2.13 Scatter Plot and time-series scatter plot

A scatter plot is a common and useful way of exploring or displaying the correlation between two variables. An example could be to plot two stock prices (with aligned time index) on a single scatter plot, giving a clear and quick indication of how the two stock prices change with respect to one another (e.g. left to right decreasing slope would suggest negatively correlated prices).

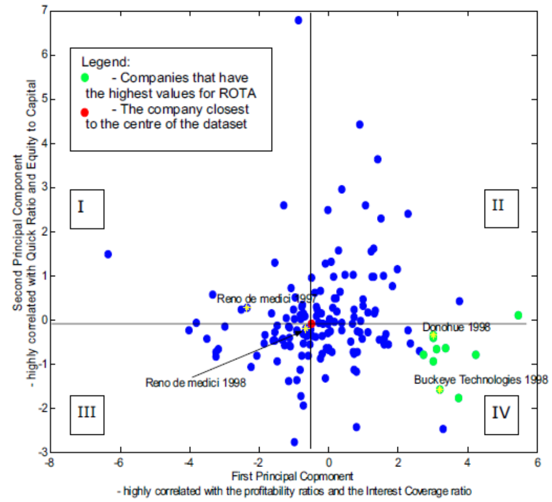

Figure 4: A scatter plot showing correlation between various stock price movements [reference ‘Multidimensional data visualisation a review’]

Another interesting usage is the time-series scatter plot, where frequently the price of the stock and the daily trade volume are shown together, with the price setting the circle height and the size of the circle defined by the volume. While this is used not infrequently within the financial media, we could not find examples of this method within the published finance literature but found examples of publications showing its use within the monitoring of Data Centres (Hao et al., 2010)

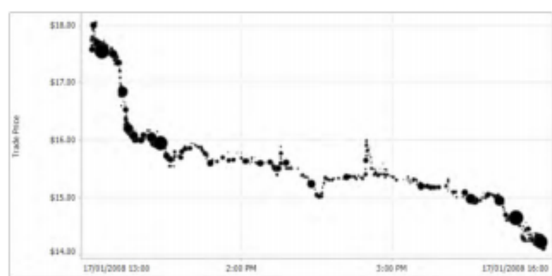

Figure 5: A time-series scatter plot with stock price on the y-axis and volume of stock trade as the size of the cirle

2.14 TreeMap and Heat Map

TreeMaps show hierarchical data sets. This technique (Shneiderman, 1992) is a powerful layout technique that has found frequent use in finance in recent years, in part due to the development of the Map of the Market tool, which is one of the very early internet financial visualizations, having found significant public use since its introduction on the SmartMoney website in 1998. A key asset for the TreeMap as commonly used is the ability to interact with the data (an analyst can generally zoom, filter, and view details on demand, as well as access the underlying tables of a certain cell if desired).

Figure 6: TreeMap visualisation of stock price as divided by asset sector.

Heat maps are similar to TreeMaps but represent each of the financial time series as an equally sized cells, whereas a Treemap uses the size of the box to represent a qualitative value and location to represent hierarchical relationships. For Heat Maps the square is coloured to show a value relative to the other boxes in the Heat map, while the location can represent the sorting of another quantitative or categorical value, for example the colour could show the traded volume for a large number of financial time series, and they could be ordered hierarchically by, say, company size, or categorically by sector.

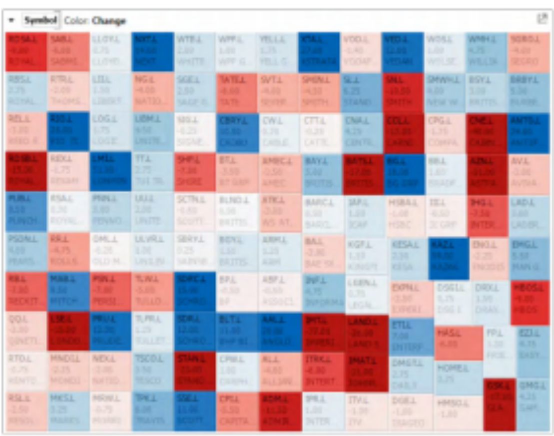
Figure 7: A heat map showing stock price movement (with red as loss and blue as gain)

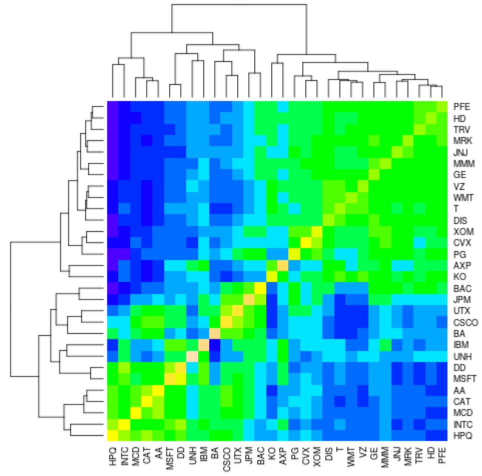

Figure 8: A dendrogram combined with a tree map using clustering to show which of the stock price movements for the companies listed on the Dow Jones Industrials index show correlation (green)

2.15 Horizon graph

One of the most common challenges is for reviewing many, perhaps hundreds, of financial time series at once. This is faced daily by, for example, financial portfolio managers. Horizon graphs are a very effective, relatively new method for achieving this in a limited space (Perin et al., 2013) A time-series is compressed to a thin line, with colour indicating normalized value. Many such lines are visualised juxtaposed, so that patterns and trends across multiple time series can be intuitively recognised.

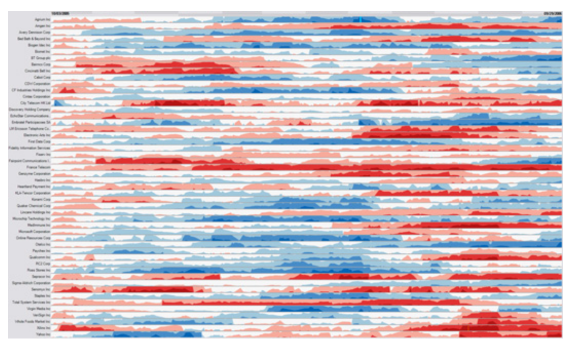
Figure 9: A horizon graph of UK stock prices

2.16 River Plot of Time Series Data

A river plot enables the user to visualize the distribution of several time series, without congestion of the image hampering readability. For example in the below river plot (Buono et al., 2007) several forecasted time series are displayed (the many outputs from a model) with a central forecast highlighted. It is possible to gain an intuition regarding how prices as well as volatility change with respect to time (based on the 'width' of the river). Often shading is used to demonstrate the 'density' of the time series underlying the graphic, at any point.

A recently developed tool that uses river plots is TimeSearcher 3, developed at the University of Maryland. Within industry TimeSearcher is used to help spot historical patterns in financial data by filtering data with similar shapes and characteristics. The river plot can be used to show hundreds of time series in parallel, to explore, for example, the way in which the stock prices and to give an idea of the 'spread' of movement within the group, as well as some summary line for the group aggregated trend.

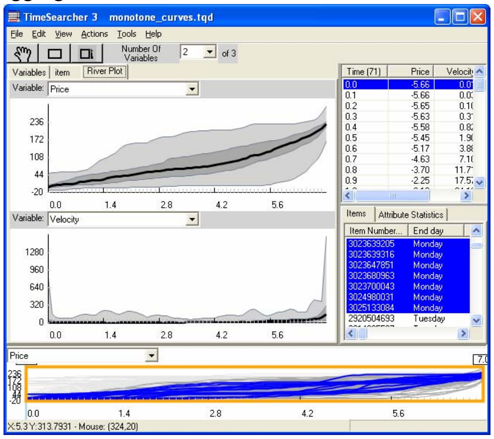

Figure 10: 4 River Plot of Share Prices using the TimeSearcher tool

2.17 Spiral

The next graphics were not found to be applied to financial time series in our review of published literature but are included because we feel that they could be illustrative or value adding within this domain.

Spiral graphs are efficient in identification of patterns and periodic structures in data. The time series 'y' axis is wrapped in a circular or spiraled shape, and segments pertaining to the period over which a pattern is expected (or being investigated) are shown on concentrically increasing shells. Sometimes the magnitude of the time series is represented through a graduated colour scale (Weber et al., 2001). In the figure below, two visualizations of sunshine intensity: In the spiral visualization it is much easier to compare days, to spot cloudy time periods, or to see events like sunrise and sunset.

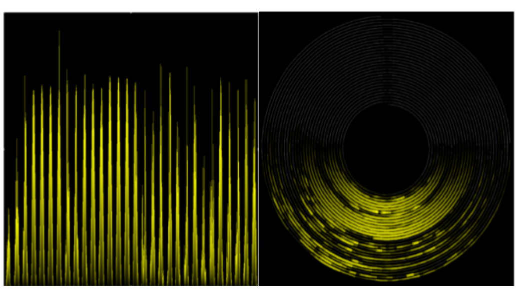

Figure 11: Sunshine intensity as shown on a bar plot and a spiral graphic.

2.18 Dendogram

A dendogram is not usually associated with time-series, but as shown in Speigel et al and in the graphic below, it can be useful for demonstrating the separate clusters within a time series. This has not been used within financial time series, to the best of our knowledge, but could be an interesting new method for clarifying clustering of similar events (e.g. days where the stock price rose or fell by more than a certain amount or where volume traded was unusually high) within a financial time series.

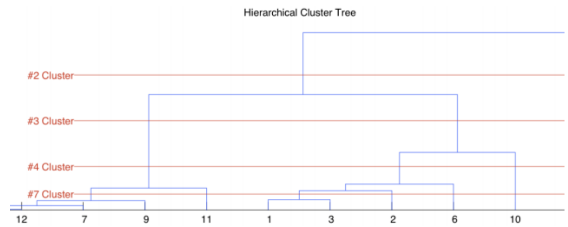

Figure 12: A dendogram used to show the clustering within a time series

Hierarchical clustering looks at a distance matrix, also known as a dissimilarity matrix, and groups observations together into a "hierarchy" of groupings. This can be done agglomeratively (group the closest observations together over and over until there is only one group) or divisively (start with one massive cluster of observations and split the groups over and over until you only have individuals). After the hierarchy formation is complete, the distances between the centers of groups can be plotted.

2.19 Calendar based visualization

As with spiral and dendograms, calendar based visualisations were not seen in the financial time-series literature, but we believe could be very effective for meeting the common needs within financial industry. Van Wijk et al used a coloured calendar with a line plot to show both the annual clustering of their data, as well as the intra-day behaviour of each cluster (on the line plot) (Van Wijk and Van Selow, 1999). This was achieved by first clustering similar daily data patterns, and visualizing the average patterns as subsequently visualizing the data on the line graph and the corresponding days on a calendar. Colors indicate corresponding clusters and patterns.

This presentation provides a quick insight into both standard patterns and days that were outliers. Furthermore, it is well suited to interactive exploration (e.g. choosing a certain day which seems to be an outlier and exploring further). As such it could be extremely useful for exploring large volumes of financial time series data, particularly if you were interested in how certain clusters of the stocks or assets in question behave on an intraday basis.

The figure below shows an analysis by Van Wijk et al of Dutch power consumption. The five main clusters are shown here. During week-ends power consumption is fairly constant. The correlation with the seasons is clearly visible. On February 4th an outlier can clearly be seen: on further examination it is understood that this is due to a public holiday in the Netherlands. Such exploratory analysis could be helpful for exploring financial time series and outliers within groups of time series.

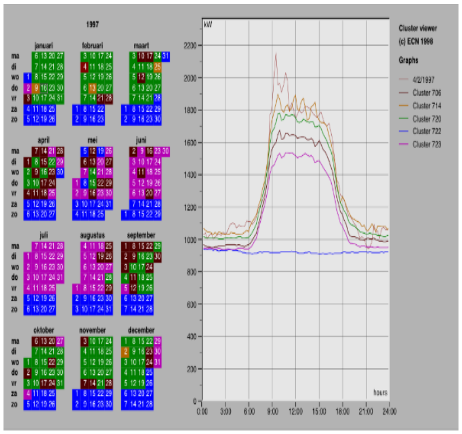

Figure 13: Cluster analysis of power demand by ECN

2.2 Clustering

Clustering is a key method for increasing the insight that can be gleaned from a visualisation and is often an intermediary and iterative step within visual analytics. Clustered data objects are similar within a cluster and dissimilar to the objects outside the cluster, though the measures for determining similarity and difference are numerous. Clustering is also referred as un-supervised classification. The popular clustering methodologies are distance-based clustering methods (partitioning algorithms like K-means,K-mediods,K-Medians), hierarchical clustering (agglomerative,divisive), density-based clustering and probabilistic and generative model-based clustering. Clustering of financial time series has been the focus of a large volume of research both within academia and industry for at least 50 years (Saha et al., 2009) and (Press, 2016) and is widely used in quantitative financial analysis within industry.

Good quality clusters have high intra class similarity and low inter class similarity. A key component of clustering is the function used to measure the similarity between data being compared. For a variety of time series analysis similarity measure is of fundamental importance. In traditional databases, similarity measure is exact match based. However, in time series data, similarity measure is generally carried out in an approximate manner. It is a real valued function that quantifies the similarity between two objects. There are many methods to find similarity measure of time series data such as Euclidean distance, Pearson's correlation coefficient, dynamic time warping, probability-based distance. Dynamic Time Warping Distance(DTW) is an algorithmic technique used to compare discrete sequences to sequences of continuous values, this technique is commonly used in speech recognition. Dynamic time warping deals well with temporal drift and has superiority over Euclidean distance. However, it has high computation cost and has some limitations on huge data sets.

Here we discuss K-means clustering and Density-based Clustering techniques with a focus on financial domain.

2.21 K-means clustering

K-means clustering (Hartigan, 1975) defines a prototype in terms of a centroid, which is the mean of group of points. This method aims to partition n observed objects into k clusters. Each object belongs to one cluster. All objects are treated with the equal importance and thus a cluster. With the predetermined k, the algorithm proceeds by alternating between two steps: assignment step and update step. Assignment step assigns each example to its closest cluster (centroid). Update step uses the result of assignment step to calculate the new means (centroids) of newly formed clusters. (Cai et al., 2012). The convergence speed of the k-means algorithm is fast in practice but the optimal k value is not known in advance. When clustering is used for data compression, every point must be clustered. In financial domain outliers are of much importance and cannot be eliminated.

2.22 Density based clustering

Another clustering approach is density based (Tan, Steinbach and Kumar, 2006) which does not partition the sample space by mean centroid, but instead density based information is used, by which tangled, irregular contoured but well distributed dataset can be clustered correctly (Cai et al., 2012). OPTICS (Ankerst et al., 1999) is a density based clustering technique to get insight into the density distribution of a dataset. It makes up for the weakness of the K-means algorithm for lack of knowledge of how to choose the value k. OPTICS creates an augmented ordering of examples based on the density distribution (Cai et al., 2012).

2.3 Interaction with visualisations

Once a visualisation method has been selected, it can often be enhanced by one or more of a number of interactive techniques, (see Weber et al., 2001) :

- Scrolling extends the display area and allows for the representation of larger data sets. However, a comparison of data elements is only possible in the currently visible subset.
- Zooming is another approach to the visualization of large data sets. Initially a low resolution view is presented and the user can decide to zoom into interesting regions. Again, comparisons are only possible across the visible subset and important detail might not be visible in the overview.
- Focusing & linking [2] extends the idea of zooming by providing not only zoomed versions of the detail data, applying also different, more effective visualization techniques for the selected frame.
- Brushing provides such additional information as popups which are automatically displayed as a roll-over effect.

These are now very widely used by stock price viewing platforms, for example free services Yahoo Finance and Google Finance both provide zooming and scrolling of financial line graphics, with pop-up additional information, as basic functionality of stock price visualisations.

2.4 Data preparation

Commonly data is prepared to allow for user understanding, interaction and further analysis. While this paper is not focused on such methods, due to their interrelation with the visual analytics methods discussed, we provide a brief note on two of the most frequent: sampling and aggregation.

In sampling (Astrom ,1969) a selection of the data is taken to provide a lower-volume representation. For example, with a time-series, a rate of m/n samples can be

dimension after dimensionality reduction. (Fu, 2011). However, a low sampling rate results in a distorted, compressed time series.
Another method is to aggregate the data, for example by taking the average value of a number of sub-sections of the data. Examples of this include Piecewise Aggregate Approximation(PAA) (Keogh et al., 2000)] and the Segmented Sum of Variation(SSV) (Lee et al., 2003).

3 Discussion
This section will focus on three main questions: i) why and how is the role of visual analytics for financial time series changing? ii) which visualisations and visual analytics methods are most appropriate for which data types or analytical questions and iii) what tools are there for interactive visualisaiton and analysis of financial data?

3.1 Why and how is the role of visual analytics for financial time series changing?

The traditional line graph, with price at the vertical axis and time running horizontally has for a long time been the most widely used visual tool for representing financial time series data both during analysis and as visual communication. Innovations in prior decades were generally oriented toward improving the efficacy of the existing chart types, as opposed to the creation of entirely new graphics and tools. These 'evolutions' on the traditional line graph have included the growth in use of moving averages, adding relative percentage gains/losses, combining with other known graphics (e.g. to form the candlestick chart) or by overlaying with separate graphics. (Marghescu, 2007; Weber et al., 2001; Ziegler et al., 2010)
However in recent years we have seen significant innovation outside of the classical visual representations. The large and rich financial sector provides a strong incentive for commercialization of new tools from academia or research domains. An example of this is the introduction of the 'horizon' graphic (see figure 14) from academic research project into a product marketed by Panopticon Software is a good example. The horizon methodology allows for a very quick digestion of financial time-series data, either as a mid-point in analyses or as a final visual product for communication.

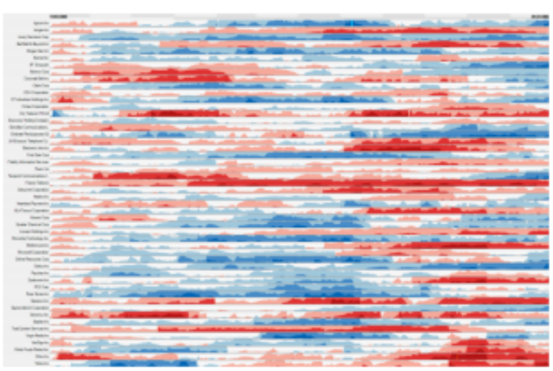

Figure 14: Example of a horizon graphic

We see three primary drivers of this trend: increasing volumes of data, new analytic techniques and renewed focus on financial regulation. As touched upon in section 1, the vastly higher volume of data is increasing the need analyses and humans. In parallel the ever increasing computational power available, as well as the rise in availability of big data and machine learning techniques, have increased the potential range and power of quantitative analysis and accordingly boosted interest in fields like visual analytics that can enhance the efficacy of such analysis. Finally, non-technological drivers also provide supportive drive for the growth of the field. Following the 2008 financial crisis there has been renewed vigour from North American and European financial regulators. A key challenge that such regulators face is the sheer volume of data over which they are expected to understand and to provide a degree of surveillance on.
Faced with such a task there is desire for techniques that can help regulators to understand financial data, as well as to guide the attention of analysts to the more interesting sections. Visual analytics is a good candidate for this. For these reasons the Office of Financial Research in the United States and the European Commission have both written papers on the role of visual analytics specifically in aiding the efforts towards improved financial regulation. Collaborative efforts between financial regulators and academia have driven the development of new tools, such as the VisRisk platform. While VisRisk is not specific to time-series data, this data type is certainly included in its scope. A key functionality is that it allows users to easily interactively see various groups of companies or assets with similar risk profiles, and to see how these change over time. It is an example of how innovation and regulation might work together to drive the development of visual analytical techniques that are established in academic circles, and with other data domains, into the financial sector.

3.2 Advantages of respective methods
Selection of an appropriate visualisation type is essential if the broader visual analytics process will be effective. Section 2 provided an overview of some of the methods that are available for visualisation, and the below table provides a summary of which of the more commonly used methods may be preferable for a specific situation:

Line chart	Useful for exploring low numbers of time series, and particularly for situations where broad trends (and not detailed short intervals) are of primary interest but strongly limited for comparison of several parallel time series (e.g. monitoring hundreds of stock prices)
Candlestick chart	Similar to a line chart but more oriented to behaviour over small time frames (e.g. past several time periods) as the additional detail of the 'candlestick' box plot will be lost over very long time spans.
Horizon graph	A relatively new visualisation type but which can be extremely effective way for seeing correlations between large numbers of separate time series in a single view, a frequent problem faced by financial users of time series data.

scatter plot	variable values for a single entity over a certain period, (for example, this method is frequently used to show a stock's trade volume and price, where price is represented by height on y-axis and trade volume is represented by size).
Spiral Plot	Particularly useful for large volumes of data and for highlighting periodic patterns, i.e. fluctuations that occur monthly.
Tree map	Not frequently used within financial time series but could be highly effective for visualisation of hierarchical time series data (see section 3.3 for more, with an example)
Heat map	Not frequently used with financial time series but can be an effective way for displaying large volumes of data items (see section 3.3 for more, with example)

A comprehensive comparison of the merits of individual approaches is provided by Rajanen, as summarised in the below (Marghescu, 2007).

	Task – See Section 2.1					
Visualization technique	Outliers detection	Dependency analysis	Clustering	Cluster description	Class description	Comparison
Line graphs	✓	✓			✓	✓
Permutation matrix	✓	✓				✓
Survey plot	✓	✓			✓	✓
Scatter plot matrix	✓	✓				✓
Parallel coordinates	✓	✓				✓
Star glyphs	✓	✓				
Treemaps	✓				✓	✓
PCA	✓	✓	✓	✓		✓
Sammon's mapping					✓*	
Self organizing maps – all views combined	✓	✓	✓	✓	✓	✓
Dendrogram			✓			

Table.1: Appropriate uses of visualisation methods by Rajanen [reference 'Multidimensional data visualisation a review]

3.3 Products for visual interaction with financial time series

This section will cover some of the most commonly used visual analytical products suitable for analysing financial data. Often, latest techniques and tools for visualising financial data are confidential and not openly shared, in order for large trading houses, banks and hedge funds to maintain competitive advantage. However, several tools are commonly used, and we give an overview of two in this section.

One of the most well-known 'modern' visualisation tools (Ziegler et al., 2010)for analysing financial data such as share prices fluctuations and movements of market sectors is the Map of the Market – which is used to produce a TreeMap of stock price movements within the market.

Figure 15: TreeMap of stock price behaviour across the S&P500. For a current example see *https://finviz.com/map.ashx.*

The Map of the Market allows for interactivity to enable mouseover or clicking, which reveals more in-depth information such as historical data and comparisons with other share price movements. Additionally, composition can be used such as the aligning to the left the stocks with the largest movement over the time period under investigation, (usually a day). There are limitations of the tool however (Ziegler et al., 2010). First, only a snapshot of the changes over fixed time interval can be shown, which precludes the potential to investigate the market sectors or assets over a longer period of time.

Additionally, rectangles represent companies with large market capitalizations and consequently companies with small market capitalization are less visiable represented on the visualisation, but may be of more interest to the analyst.
Another example of a successful tool is Impactoria (see http://www.market-topology.com/) that looks to unravel hidden relationships between financial time-series through developing correlation coefficients and mapping the relationships between the time series in a ‘tree’ in order to give traders valuable insight.

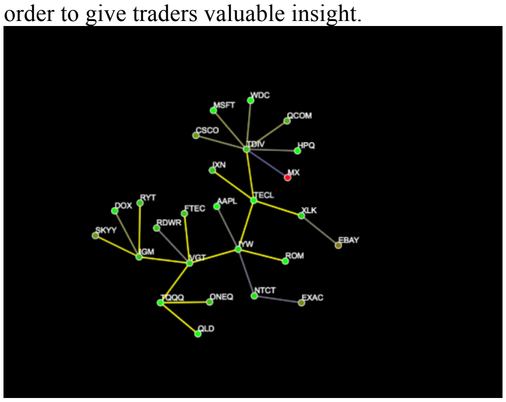

Figure 16: Examples of visualisations of relationships between large numbers of time series representing financial assets. [reference http://www.market-topology.com]

5. Conclusion

References

Ankerst, M., Breunig, M.M., Kriegel, H.-P., Sander, J., 1999. OPTICS: ordering points to identify the clustering structure, in: ACM Sigmod Record. ACM, pp. 49–60.

Brighenti, C., Sanz-Bobi, M.A., 2011. Auto-Regressive Processes Explained by Self-Organized Maps. Application to the Detection of Abnormal Behavior in Industrial Processes. IEEE Transactions on Neural Networks 22, 2078–2090. doi:10.1109/TNN.2011.2169810

Buono, P., Aris, A., Plaisant, C., Khella, A., Shneiderman, B., 2005. Interactive pattern search in time series, in: Electronic Imaging 2005. International Society for Optics and Photonics, pp. 175–186.
Buono, P., Plaisant, C., Simeone, A., Aris, A., Shmueli, G., Jank, W., 2007. Similarity-based forecasting with simultaneous previews: A river plot interface for time series forecasting, in: Information Visualization, 2007. IV'07. 11th International Conference. IEEE, pp. 191–196.
Cai, F., Le-Khac, N.-A., Kechadi, M.-T., 2012. Clustering approaches for financial data analysis: a survey, in: Proceedings of the International Conference on Data Mining (DMIN). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), p. 1.
Dagnely, P., Ruette, T., Tourwé, T., Tsiporkova, E., Verhelst, C., 2015. Predicting hourly energy consumption. Can you beat an autoregressive model, in: Proceeding of the 24th Annual Machine Learning Conference of Belgium and the Netherlands, Benelearn, Delft, The Netherlands.
El-Sherif, N., Denes, P., Katz, R., Capone, R., Mitchell, L.B., Carlson, M., Reynolds-Haertle, R., Investigators, C.A.S.T.-A.E. (CAST/SAECG) S., 1995. Definition of the best prediction criteria of the time domain signal-averaged electrocardiogram for serious arryhthmic events in the postinfarction period. Journal of the American College of Cardiology 25, 908–914. doi:10.1016/0735-1097(94)00504-J
Fu, T., 2011. A review on time series data mining. Engineering Applications of Artificial Intelligence 24, 164–181. doi:10.1016/j.engappai.2010.09.007
Hao, M.C., Dayal, U., Keim, D.A., Schreck, T., 2007. Multi-resolution techniques for visual exploration of large time-series data, in: EUROVIS 2007. pp. 27–34.
Hao, M.C., Dayal, U., Sharma, R.K., Keim, D.A., Janetzko, H., 2010. Variable binned scatter plots. Information Visualization 9, 194–203. doi:10.1057/ivs.2010.4
Hartigan, J.A., 1975. Clustering algorithms, Wiley series in probability and mathematical statistics. Wiley, New York.
Keim, D.A., Nietzschmann, T., Schelwies, N., Schneidewind, J., Schreck, T., Ziegler, H., 2006. A spectral visualization system for analyzing financial time series data, in: Eurographics/IEEE TCVG Symposium on Visualization. pp. 195–202.
Kong, N., Agrawala, M., 2012. Graphical overlays: Using layered elements to aid chart reading. IEEE transactions on visualization and computer graphics 18, 2631–2638.
Lee, K.H., Jo, G.S., 1999. Expert system for predicting stock market timing using a candlestick chart. Expert Systems with Applications 16, 357–364.
Levy-Carciente, S., Jaffé, K., Londoño, F., Palm, T., Pérez, M., Piñango, M., Reyes, P., 2006. Econophysics to unravel the hidden dynamics of commodity markets, in: Practical Fruits of Econophysics. Springer, pp. 77–81.
Marghescu, D., 2007. „Multidimensional Data Visualization Techniques for Financial Performance Data: A Review". Turku Centre for Computer Science.
Merino, C.S., Sips, M., Keim, D.A., Panse, C., Spence, R., 2006. Task-at-hand interface for change detection in stock market data, in: Proceedings of the Working Conference on Advanced Visual Interfaces. ACM, pp. 420–427.
Perin, C., Vernier, F., Fekete, J.-D., 2013. Interactive horizon graphs: improving the compact visualization of multiple time series, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 3217–3226.
Press, C., 2016. A New Look at Clustering of Stock Prices Author (s): Victor Niederhoffer Source : The Journal of Business , Vol . 39 , No . 2 (Apr ., 1966), pp . 309-313
Saha, A., Malkiel, B.G., Grecu, A., 2009. The clustering of extreme movements: stock prices and the weather. Journal of investment management 20.

Schaefer, M., Wanner, F., Kahl, R., Zhang, L., Schreck, T., Keim, D., 2011. A novel explorative visualization tool for financial time series data analysis. Citeseer.
Schreck, T., Tekušová, T., Kohlhammer, J., Fellner, D., 2007. Trajectory-based visual analysis of large financial time series data. ACM SIGKDD Explorations Newsletter 9, 30–37.
Shneiderman, B., 1992. Tree visualization with tree-maps: 2-d space-filling approach. ACM Transactions on graphics (TOG) 11, 92–99.
Silva, B., Cavique, L., Marques, N., 2013. Simulating price interactions by mining multivariate financial time series, in: Proceedings of the 3rd International Conference on Ubiquitous Data Mining-Volume 1088. CEUR-WS. org, pp. 44–48.
Tan, Steinbach and Kumar, 2006. Introduction to Data Mining. Addison Wesley.
Tsai, C.-F., Quan, Z.-Y., 2014. Stock Prediction by Searching for Similarities in Candlestick Charts. ACM Transactions on Management Information Systems 5, 1–21. doi:10.1145/2591672
Van Wijk, J.J., Van Selow, E.R., 1999. Cluster and calendar based visualization of time series data, in: Information Visualization, 1999.(Info Vis' 99) Proceedings. 1999 IEEE Symposium on. IEEE, pp. 4–9.
Weber, M., Alexa, M., Müller, W., 2001. Visualizing Time-Series on Spirals., in: Infovis. pp. 7–14.
Xie, Z., 2011. Exploratory Visualization of Data Pattern Changes in Multivariate Data Streams. Google.
You, C., Chandra, K., 1999. Time series models for internet data traffic, in: Local Computer Networks, 1999. LCN'99. Conference on. IEEE, pp. 164–171.
Zhang, Y., Glass, J.R., 2011. A Piecewise Aggregate Approximation Lower-Bound Estimate for Posteriorgram-Based Dynamic Time Warping., in: Interspeech. pp. 1909–1912.
Ziegler, H., Jenny, M., Gruse, T., Keim, D.A., 2010. Visual market sector analysis for financial time series data, in: Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium on. IEEE, pp. 83–90.
Ziegler, H., Nietzschmann, T., Keim, D.A., 2008. Visual analytics on the financial market: Pixel-based analysis and comparison of long-term investments, in: 2008 12th International Conference Information Visualisation. IEEE, pp. 287–295.