FAKULTÄT FÜR !INFORMATIK

Evaluation of Interactive Visualization Methods to Compare Multivariate Heterogeneous Time Series

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Wirtschaftsinformatik

eingereicht von

Christian Kainz
Matrikelnummer 0425423

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung:
Betreuerin: Univ.-Prof. Mag. Dr. Silvia Miksch
Mitwirkung: Dipl.-Ing. Dr. Wolfgang Aigner

Wien, 20.11.2009

(Unterschrift Verfasser) (Unterschrift Betreuerin)

Eidesstattliche Erklärung

Ich erkläre an Eides statt, daß ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benützt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 20. November 2009

Abstract

This master thesis explores various comparison methods for multivariate time series in the application area of stock markets. The data is usually compared by juxtaposition or by superimposition. But even a large enough difference between the price ranges can make a simple comparison of the data difficult.

Stock market data is not limited to stocks. It also includes stock (market) indices such as the Dow Jones and NASDAQ. Stock indices have their own proprietary unit, which is different to the unit of a stock. Time series which have different units are also called heterogeneous data.

To compare heterogeneous data by superimposition multiple y-axes are often used. But in most cases the arrangement of the axes makes comparisons between different variables completely arbitrary and comparisons are often misleading. The visualization pioneer Jacques Bertin has studied this problem. One suggestion is to index the values, which transforms all data into values which reflect the percent change compared to an indexing point.

The first part of the research is concerned with the identification of relevant comparison methods for multivariate time series. The major part of the research deals with the evaluation of an advanced comparison method based on indexing. Another important part is the investigation of the importance of the used axis scale. Differences between linear and logarithmic scale are analyzed for effects on user performance.

The major part of this work is a comparative study about three visual comparison methods (visualization types) for multivariate time series. The three tested comparison methods are juxtaposition, superimposition and indexing. 24 test persons participated in the study. Each participant had to complete 14 tasks for each one of the three visualization types. The task completion time and the task correctness for every task were measured and later used for statistical analyses.

This work further presents state of the art research about other visualizations suited for visual comparison tasks. The prototype application incorporates several common stock market visualizations such as line charts, OHLC charts and candlestick charts. Basic comparison methods like juxtaposition and superimposition are available. A more advanced comparison method based on indexing was implemented.

The usability test results support the assumption that the indexing method enables the user to perform comparison tasks with much less estimation errors. The task completion time is not significantly different. The free selection of the indexing point makes comparisons for a certain time period more effective and delivers more precise results. A post-test survey showed that the majority of the participants favor the indexing method over the two other visualization types. The test results for the usage of different scales indicate that tasks were faster completed when using logarithmic scales. The task correctness rate was not significantly different between linear and logarithmic scales.

Zusammenfassung

Diese Diplomarbeit untersucht verschiedene Methoden zum Vergleich multivariater Zeitreihen für den Anwendungsbereich Aktienmärkte. Die Daten werden in der Regel durch Nebeneinanderlegung oder durch Überlagerung verglichen. Aber schon eine ausreichend große Differenz zwischen den Preisen kann selbst einen recht einfachen Vergleich erschweren.

Börsendaten sind nicht nur auf Aktien beschränkt. Sie umfassen auch Aktien (Markt-) Indizes wie den Dow Jones und NASDAQ. Aktienindizes besitzen eine eigene proprietäre Einheit, die der Einheit einer Aktie verschieden ist. Zeitreihen, die verschiedenen Einheiten haben werden auch als heterogen bezeichnet.

Heterogene Daten werden oft durch Überlagerung mittels mehreren y-Achsen verglichen. Aber in den meisten Fällen ist die Anordnung der Achsen völlig willkürlich und Vergleiche sind oft irreführend. Jacques Bertin, ein Pionier im Bereich der Visualisierung, hat dieses Problem studiert. Ein Vorschlag ist, die Werte zu indizieren. Alle Werte werden entsprechend der prozentualen Veränderung bezogen auf den Basispunkt umgewandelt.

Der erste Teil der Forschung beschäftigt sich mit der Identifikation von relevanten Methoden zum Vergleich von multivariaten Zeitreihen. Der Forschungsschwerpunkt ist die Evaluierung der Methode zum Vergleichen, welche auf einer Indizierung basiert. Ein weiterer wichtiger Forschungsteil ist die Untersuchung der verwendeten Skala einer Achse. Es werden die Unterschiede zwischen linearen und logarithmischen Skalen und deren Auswirkungen auf die Performance der Anwendungen analysiert.

Der wichtigste Teil dieser Arbeit ist die vergleichende Studie über drei Methoden (Visualisierungs-Typen) zum visuellen Vergleich multivariater Zeitreihen. Die drei getesteten Methoden sind Nebeneinanderlegung, Überlagerung und Indexierung. 24 Probanden nahmen an der Studie teil. Jeder Teilnehmer mussten 14 Aufgaben für jede der drei Visualisierung-Typen ausführen. Die benötigte Zeit und die Richtigkeit wurden für jede Aufgabe gemessen und später für statistische Analysen verwendet.

Die vorliegende Arbeit präsentiert aktuelle Forschungsthemen zu weiteren geeigneten Visualisierungen für visuelle Vergleiche. Der begleitende Prototyp zu dieser Arbeit umfasst die Darstellung von üblichen Aktienmarkt Visualisierungen wie Liniendiagramme, OHLC-Diagrammen und Candlestick-Diagrammen. Simple Vergleichsmethoden wie Nebeneinanderstellung und Überlagerung sind verfügbar. Eine fortgeschrittene Methode zum Vergleichen basierend auf Indizierung wurde ebenfalls implementiert.

Die Usability-Test Ergebnisse stützen die Annahme, dass die Indizierungs-Methode es dem Benutzer ermöglicht Vergleiche mit sehr viel weniger Fehlern bei der Schätzung von Werten durchzuführen. Die benötigte Zeit der Aufgabe wird dadurch nicht signifikant beeinflusst. Die freie Wahl des Indexierungspunktes macht Vergleiche für einen bestimmten Zeitraum effizienter und liefert genauere Ergebnisse. Eine Post-Test-Umfrage ergab, dass die Mehrheit der Teilnehmer die Indizierungs-Methode gegenüber den beiden anderen Vi-

sualisierungs-Typen bevorzugt. Die Testergebnisse für die Nutzung der verschiedenen Skalen zeigen, dass die Aufgaben schneller abgeschlossen wurden wenn logarithmischen Skalen verwendet wurden. Die Richtigkeit war nicht signifikant unterschiedlich zwischen linearen und logarithmischen Skalen.

Contents

1 Introduction

“A picture is worth a thousand words”

This famous phrase is summarizing a phenomenon of human cognition that is the reason for the success of data and information visualization.

Measuring of data is important to gain insight and help to make better analysis. The measuring process can create a lot of data. The easiest way to store data is to write the values in tabular form. With increasing amount of values the tabular form needs more time to read and is harder to analyze.

In the last 200 years people started to visualize the measured data. In this process the measured data points are transformed into abstract visual items. These visual items form a graphical representation of the measured data. A human can process such a visualization very fast compared to the large table of numbers referencing the measured values of the data points.

For example the determination of the gradient of a line should be easier most of the time easier than to compare numbers. Through visualization the identification and analysis of time series can thus be simplified. Although there are some disadvantages like the reduced preciseness.

Another important point of visualizations is the ability of a clearer communication about the visualized data. It should be not surprising that data is often visualized to increase the understanding of the viewer about the data.

The British scientist William Playfair was one of the first to enhance the cognition of data by using visualizations to encode data. In the late 18th century this approach was certainly pioneer work. At that time the most common visualizations were maps. Maps are a good example for the phrase above. They successfully encode geographic coordinates onto a two-axial grid. Textual descriptions are not very useful for this task.

Data visualization provides the unique ability to compress a huge amount of information into a graph. The human brain can process visual encoded data faster than any numerical values in a table. Visualizations reveal the shape of the data, which is not to be found in numeric tables. For example, it shows how the data has developed over a certain time period or how some variables are influencing some other variables.

The strength of visualizations lies on the unique possibility to display relationships between the used data sets. Trends and patterns are easily spotted by using graphs. It should be no surprise that visualizations help the viewer to understand the data. Also the characteristics like peak values of the underlying data can be easier communicated.

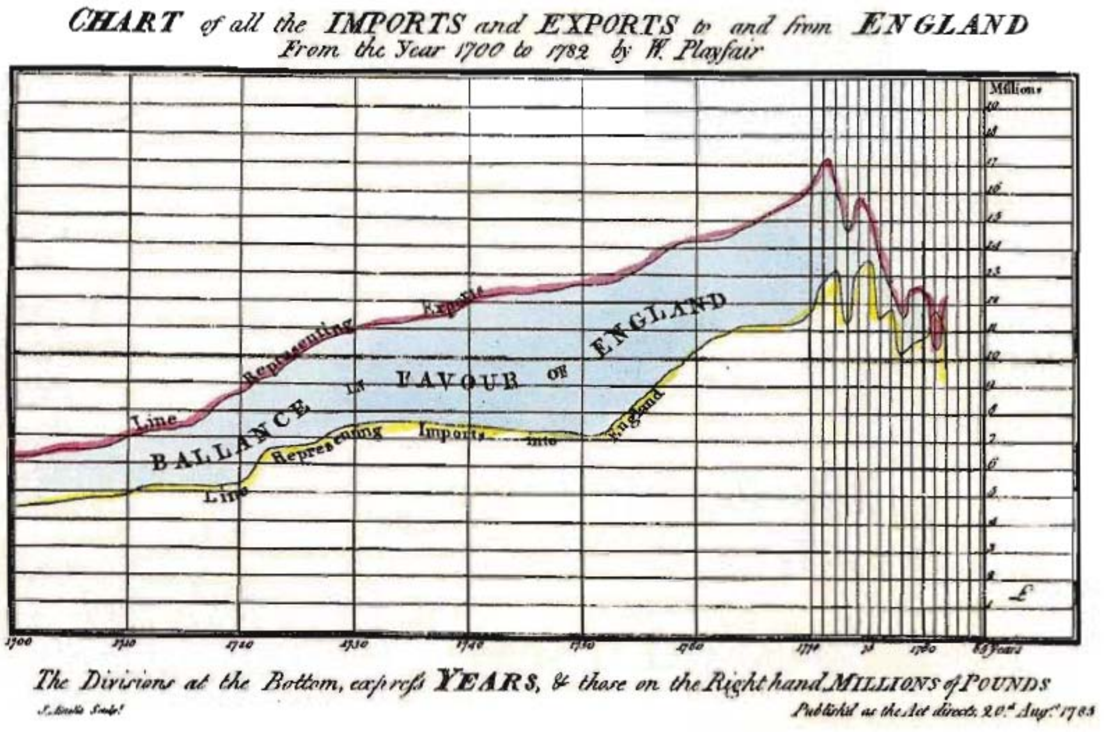


Figure 1: Import and Export by W. Playfair [Tufte, 2001]

1.1 Motivation

Time series is a special type of data series. Each data item refers to a specific point in time. The data items, ordered along the time dimension, represent the historical development. Another important aspect of time series is the potential to compare multiple time series against each other by temporally aligning the time series points.

Comparison of time series can be very important to analysis. Correlations, common trends and delays can help to understand dependencies and influences between time series. Certainly the field of statistics provides calculation methods for these and other characteristics between time series data. But visualizations should not be seen as a replacement. They are rather complementary sources of information. Visualizations form a shape of the data, which helps to quickly identify the time series.

Time series data is mostly visualized by line graphs. This graph enables a very efficient display of the historic values.

An important application of time series comparison is visual analytics of stock market data. The development of stock prices and volume are measured continuously and accurately. Depending on the needed accuracy of the stock prices a line chart, OHLC chart or candlestick chart is used to encode the data into a graph.

The comparison of multiple stocks is an important task for every investor. By analyzing the historic values and deriving a future trend one can infer a decision whether to buy more stocks, hold the stocks or sell some stocks. Visualization can help with this task by showing the form of the data.

1.2 Problem Description

The present master thesis and its research topics are originated from the problem description in [Aigner, 2009]. The following paragraph is an English translation of the original text in German.

Line graphs are extremely well for displaying development of time series and trends in the data. This is one of the reasons why line graphs are very popular for time series data. The display of univariate data is easily done. But the display of multivariate data has some restrictions. When using multiple variables it is important that all variables share the same unit and possess a similar value range. If the values of two variables are too different, the display of the time series may lead to false impressions. And if at least two variables have a different unit a second vertical axis is needed. In most cases the arrangement of the two vertical axes make visual comparisons between different variables completely arbitrary. The results are often misleading. The visualization pioneer Jacques Bertin has studied this problem and has worked out a few suggestions for possible solutions.

Two simple solutions to reduce the limitations of visual comparison for multivariate data are juxtaposition and superimposition of time series. Juxtaposed graphs are very flexible and can be used for homogenous and for heterogeneous data series. The big drawback of this method is the reduction of the usefulness for visual comparisons. Each time series is displayed on an own graph and each graph does have a different vertical axis scale. As a consequence the gradient of the curves cannot be compared as they are not related anymore. Another consequence of juxtaposition is that the lines are often too much apart which also reduces the usefulness. Both factors add to a high vagueness of this method for visual comparison tasks.

Visual comparison of data has a lot of benefits. Visualizations in general provide a clearer picture of the development of the underlying data than numerical values in tabular form ever could. Visualizations improve the communication of data, trends and patterns with other persons.

Superimposed graphs are another useful method to display multivariate data. Encoding of homogenous data leads to problems if the values of the variables are too different. But troubles start when heterogeneous multivariate data has to be encoded. The use of multiple axes leads to confusion for the viewer. And it points to another important question. How to visually compare different units of data? In most cases there exists no natural relation between different units. The scale of each axis is more or less arbitrary, which is removing any meaning of the gradients for visual comparisons. Furthermore any intersection between time series is of no meaning at all. It is rather a product of the relation between the arbitrary selected scales of the involved axes.

The comparison by superimposition is also prone to bad tricks. [Wainer, 1997] describes how a simple visualization trick can change the meaning of the data. He explains the trick through a series of graphs about the death rates for smokers vs. non-smokers for a given age. At first a superposed line graph is used for the visualization of the data. It is clearly visible that death rates are constantly higher for every age. Through splitting of the two variables smoker and nonsmoker onto two axes with different value ranges the visualization is distorted. The difference in death rates is not so different anymore.

The French cartographer Jacques Bertin discusses in his book Semiology of Graphics [Bertin, 1983] some theoretical and many practical subjects of chart, network and map visualizations.

His impressive work on information visualization provides a lot of good advice. In the section about visual comparisons with one quantitative component he introduces a method which is called indexing. This method transforms the original values into percent changes to a reference value which is also called indexing point. This method reduces the limitations of homogenous and heterogeneous multivariate data.

The estimation of value changes between data points is essential for successful visual comparisons of multiple time series. Some tasks require the identification and comparison of absolute changes. Other tasks require the identification of relative changes. Line graphs often use vertical axes with a linear scale. This should be the best choice to estimate absolute value changes. A logarithmic scale offers improved visual comparison of relative value changes independent of the absolute values. Therefore tasks which emphasize on comparisons of relative value changes should favor axes with a logarithmic scale.

The display of relative changes is also used in a lot of other visualizations. More prominent examples are Sparklines [Tufte, 2006] and Horizon Graphs [Heer et al., 2009]. Both visualizations are promising for visual comparison tasks.

1.3 Research Objectives

A large part of this master thesis was the comparative study of different visualizations for comparisons of heterogeneous time series. The main focus of the evaluation will be the measurement of task effectiveness and time efficiency for comparison tasks.

The second part of the thesis was the implementation of a visualization prototype, based on the indexing method explained by Bertin in [Bertin, 1983]. Juxtaposed and superimposed line charts are implemented as a reference for usual comparison methods.

Another research objective will be easy to understand and extensible source code of the visualization prototype. A reuse of the prototype for further research projects should be unproblematic.

The usability test function, which was used for this master thesis, is included in the prototype software. Therefore, repeating the comparative study or extending the test should be possible without much trouble. Usability tests can be executed with the same test procedure and the same test data. Additionally changes of the test tasks, test data or answers should be very easy.

1.4 Research Questions

This section lists all research questions for the present master thesis. The questions are divided into three parts. The first part is a state of the art research on currently used visualizations for comparison of multivariate, heterogeneous, time series data.

The second part is based on the indexing method of Bertin and will be part of a comparative study which will be described in detail in a later chapter of this master thesis.

The third part will focus on the effects of linear and logarithmic scales for visual comparisons of time series.

1.4.1 State of the Art Research

Which visualization types are suited for comparing multivariate, heterogeneous data over time?
A great collection of different visualization types, for the display of multivariate data, have been created in the past years. The most interesting ones will be presented.

Which visualizations are described in the scientific literature?
The study of information visualization produces a lot of new insight and improvements for the visual representation of data.

Which visualizations are used in current applications?
This question is similar to the previous question. A collection of important applications in the field of information visualization will be presented.

1.4.2 Indexing

The following questions are closely related to Bertin's indexing method and ist effects on information visualization.

What are the benefits of the indexing method for comparison of multivariate time series data?
The indexing method is promising because it can be used to display multivariate data over time superimposed in one chart without an arbitrary selection of the unit size for different units. All time series are normalized according to the indexing point.

What are the differences in the visual display?
This goal of this question wants is to identify the differences between a simple line chart and the indexing method.

Which point of the curve should be chosen?
The indexing point is the base value of the index. It is often the starting point of the data. Other possible points can be the value of the last point of the time series or the mean value of the data series.

What are the differences in the efficiency and the accuracy?
This question will determine the efficiency and the accuracy for given user tasks.

1.4.3 Impact of the Scale

The selection of the scale does make a difference for comparing superimposed time series. Logarithmic scales do have some interesting effects on the chart.

When to use a linear scale and when to use a logarithmic scale?
In comparisons of multivariate data the scale has an influence on the visual recognition of the data. Logarithmic scales put emphasis on the differences of quantities, rather than quantities of units. Find and identify tasks which improve when using a linear scale and which task improve when using a logarithmic scale.

What are the advantages and drawbacks?
The effects, which arise when using linear and logarithmic scales, will be explored. Also general advantages and drawbacks for each scale will be determined.

1.5 Implications of Research

The present master thesis is concerned with the empirical evaluation of the indexing concept by Jacques Bertin, which is described in detail in his book Semiology of Graphics [Bertin, 1983]. The practical part of this work will be the development of a visualization prototype. The prototype incorporates the concept of the indexing method for comparing multivariate data. The prototype will be tested how well it is suited for comparisons of homogenous and heterogeneous multivariate time series data.

The evaluation of the prototype will be done by a comparative study. It will include the comparison of the indexing method and traditional line chart visualizations as reference. The setting will be multivariate time series. Additionally it will be explored how juxtaposition and superimposition of time series are affecting the comparison of line charts. The comparative study will analyze the usefulness of the indexing method for multivariate data and the consequences of a linear and a logarithmic scale for comparisons.

2 Background: Visualization Basics

"When data is communicated graphically, just like verbal communication using language, certain rules of syntax and semantics apply. If you disobey the rules, you run the risk of being misunderstood. The rules of graphical communication are rarely arbitrary, but are usually based on an understanding of visual perception—how we see and the ways in which information can be visually encoded for easy and accurate decoding by our audience." [Few, 2006b]

This chapter introduces important basics of information visualization. It is a theoretical foundation for problems in later chapters and will help the reader to gain an understanding of information visualization.

The first section of this chapter is an introduction of the three most common graphs to visually encode data. Along with a short explanation, their advantages and disadvantages will be given. The selection of the proper visualization according to the data class is the key for a successful display. Bad selections can lead to wrong insights by the viewer about the data. So the viewer might be irritated or even develop a wrong understanding of the data.

The distinction of data into data classes is an important action. With the knowledge about the data better visualization methods and visualization items can be used, which result in a clearer display.

A short explanation of multivariate data and time series data describes the specific characteristics. Both terms are important for the identification of data series for the following chapters of this master thesis.

2.1 Multivariate Data

A visualization, which displays two or more variables, is called multivariate visualization. The term is derived from mathematics and is describing data sets with more than one variable. Depending on the units of the individual data series of a data set, it is either a homogenous multivariate data set or a heterogeneous multivariate data set.

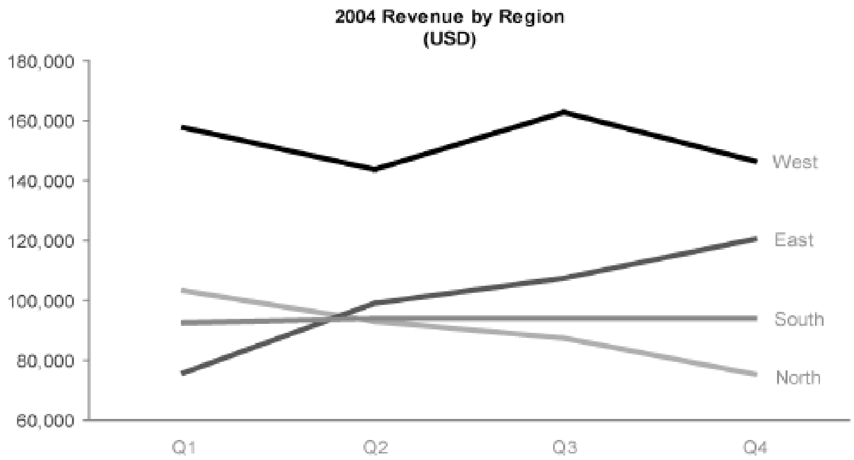


Figure 2: Revenue of four divisions in US-$ [Few, 2005]

Homogenous multivariate data sets consist of variables, which share the same unit. The following example in Figure 2 displays the last quarterly revenues in US-$ by region (north, east, south, west). The data series share the same unit US-$, so it is relatively easy to display them in one chart by superimposition of the data.

The next visualization in Figure 3 is an outstanding example for heterogeneous multivariate data visualization. This rather famous visualization presents four data sets of different units, while using only one chart. The creator of the original map was Charles Joseph Minard, a French engineer of the 19th century. He successfully integrated four different data series of Napoleon's Russian campaign from 1812 to 1813. All data series are working very well together, resulting in a stunning masterwork of information visualization.

The displayed data series are as follows:

- size of Napoleon's army [thickness of the light brown / black flow-line]
- location of Napoleon's army [coordinates spatially visualized by the light brown / black flow-line]
- direction of Napoleon's army [light brown or black colored flow-line]
- air temperature according to the location [°C]

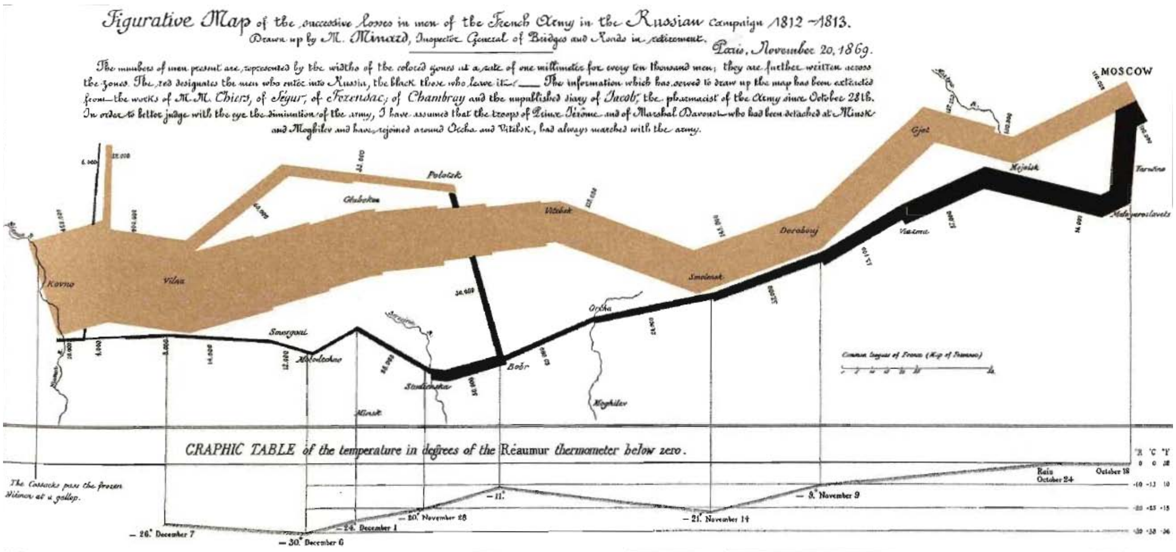


Figure 3: Map of Napoleon's Russian Campaign [Tufte, 2001]

In order to display heterogeneous multivariate data in one chart, a primary and a secondary vertical axis are used. Figure 4 shows US Consumer Price Index (CPI) and US Interest Rates for the year 2008. Both time series use different units, so two axes are used. The axis on the left is displaying values for US CPI while the axis on the right is showing values for US interest rates.

Figure 4: line chart with two vertical axes

2.2 Time Series

Time series are data series which connect data values with a specific point in time. For reasons of simplicity, the time intervals between the displayed data points are equal in most visualization. It is of importance to display the data points in chronological order or else the encoded values lose their meaning. Statistical trend analysis methods like the Moving Average can smooth data values out and show a better recognizable curve.

[Müller et al., 2003] defines time series in general:

$$d = f(t)$$

Data with discrete time stamps can be written as:

$$D = \{(t_1, d_1)(t_2, d_2), \ldots, (t_n, d_n)\}$$

where

$$d_i = f(t_i)$$

Multivariate time series are a special case:

$$D = \{(t_1, d_{1,1}, d_{1,2}, \ldots, d_{1,k})(t_2, d_{2,1}, d_{2,2}, \ldots, d_{2,k}), \ldots, (t_n, d_{n,1}, d_{n,2}, \ldots, d_{n,k})\}$$

where each data element $d_{i,j}$ is dependent on:

$$d_{i,j} = f_j(t_i)$$

Stephen Few states eight attributes in [Few, 2004] which often characterize the patterns of time series data: Change, Rise, Increase, Fluctuate, Grow, Decline, Decrease and Trend.

Time series are used in a broad spectrum of areas. Stocks are a popular example for finance applications. Other fields for time series are economic, medical, climate, logistics, data mining and business intelligence.

Time is inherently different from other units. [Schumann, 2000] describes time as additional unit which can be understood as independent variable. Value changes can easily be described by visualization of the time dimension.

[Schumann, 2000] states the problem of unequal time units of heterogeneous time series data. Heterogeneous data often has unequal measure intervals. Using time as variable can clearly describe changes in data. On the other side, a uniformly scaled time scale is implied. Especially heterogeneous data sets feature individual capturing time steps of data points, so that larger areas of the observation area are empty.

Stephen Few argues in [Few, 2004] about possible encoding methods for time series data. One axis of the graph has to handle the time dimension of the time series. This is usually the horizontal axis. The points in time have to be displayed in chronological order from left to right. This makes it very easy to compare several time series by vertically aligning the charts by date. The vertical axis is used to display the quantitative value of the data point.

Stephen Few recommends three graphical objects for encoding time series data:

- lines
- points and lines
- vertical bars

Encoding of time series data only by displaying points removes the sense of continuity between adjacent data points. Consequently floating points in space cannot represent the flow of time.

Lines connecting points of data point values form a shape. The shape of time series enables the human mind to visually process the time series. Shapes are extremely well in the ability to show trends of data values and to visualize patterns within the data.

Figure 5 shows a points and lines graphical object. Three time series are displayed by using different line colors. Each data point is represented by a black point. The display of the points enhances the comparability of the three time series.

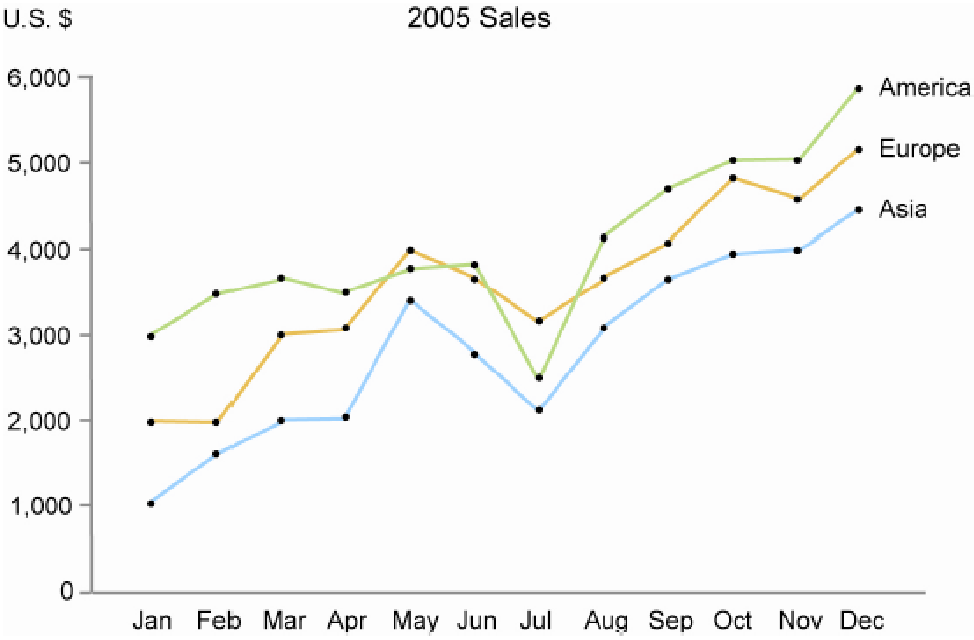


Figure 5: Points and lines [Few, 2006b]

Vertical bars can also be used to display time series. But they are more limited in their ability than the two previous graphical objects. Vertical bars emphasize on the comparison of individual values. The overall shape of the time series is not as clear.

In the application of stock market data, vertical bars are typically used for volume data.

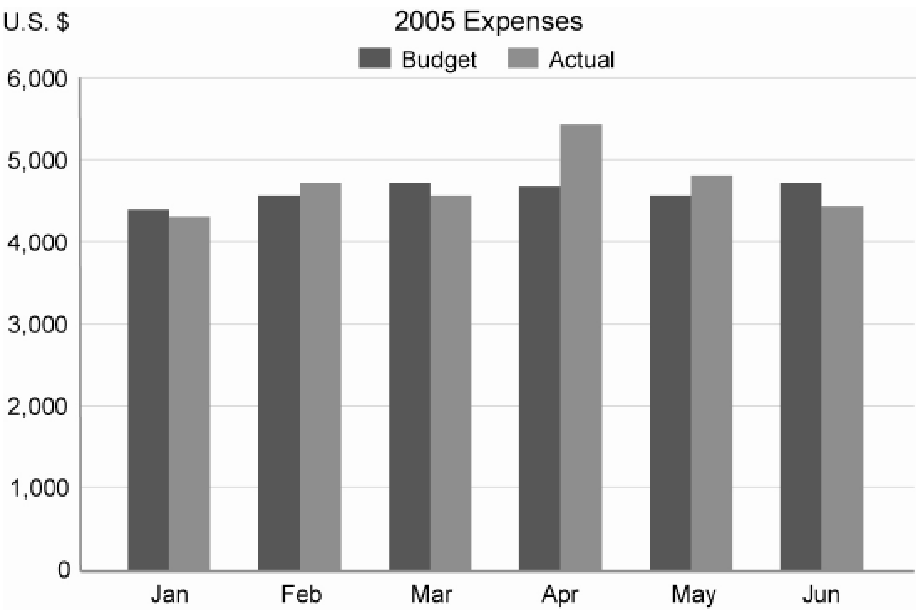


Figure 6: Vertical bars [Few, 2006b]

2.3 Data Scales

Visualized data can be usually classified into four disjoint data classes, defined by a common accepted taxonomy. The usage of the taxonomy allows a more detailed analysis and better communication of various data applications. Through the use of the taxonomy and the mapping of specific data to a data class, proven visualizations for the specific data class can be selected.

Stuart Card distinguishes in [Card, 2007] between four main classes of data: nominal, ordinal, interval and quantitative data. Table 1 summarizes the main classes of the taxonomy and gives an example for each category.

Class	Description	Example
U	Unstructured (can only distinguish presence or absence)	ErrorFlag
N	Nominal (can only distinguish whether two values are equal)	movie titles e.g. Goldfinger, Ben Hur, Star Wars
O	Ordinal (can distinguish whether one value is less or greater but not natural zero and cannot compute ratios)	sizes e.g. Small, Medium, Large
I	Interval (can do subtraction on values, but no natural zero and cannot compute ratios)	Celsius e.g. 20°C, 25°C
Q	Quantitative (can do arithmetic on values)	weight e.g. 100 kg

Table 1: Data Classes [Card, 2007]

2.4 Common Graphs

"Most graphs that are used to present quantitative business data are two-dimensional with two axes (one horizontal, called the X axis, and one vertical, called the Y axis), and use one or more of three particular objects to encode values: points, lines and bars. The choice of which one or more of these three objects to use in a graph should never be arbitrary, and need never be, because the rules are simple to understand and follow." [Few, 2006b]

This section will present three most common graphs for data visualization. All three are suitable for time series data, although not all are equally successful at the representation. The common graphs use the following visual marks to display data:

- Points
- Lines
- Bars

Figure 7 shows the common graphs or charts. The graph on the left is a scatter graph using points. The graph in the middle is a line graph. And the graph on the right is a bar graph.

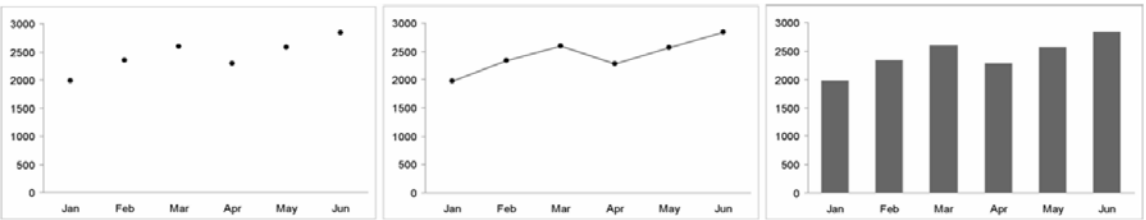


Figure 7: Scatter graph (left), line graph (middle) and bar graph (right) [Few, 2004b]

2.4.1 Scatter Graph

This term refers to a two dimensional graph, displaying data or information by using points. The points can take any shape to encode further information like subsets of the data or multivariate data. A Point Graph is a special scatter graph, which has a category scale on the horizontal axis. Three dimensional scatter graphs are also possible, but may take longer to analyze because of the increased complexity of the visualization.

Scatter Graph excel in visualizing correlations of two variables. If the data points can be clustered easily, the relationship of the data points is considered strong. Positive correlation and negative correlation are visualized by a cluster of data points forming an imaginary line. The form represents a linear correlation. Other types of correlation are exponential and U-shaped.

Additional information in more detail about scatter graphs can be found in [Harris, 1999].

2.4.2 Line Graph

The term line graph is standing for a variety of graphs that use lines to visualize information. Similar to point or scatter graphs the information is encoded by displaying the values according to the scale of the axes. Instead of displaying actual points where the data point would be, lines are displayed by connecting adjacent points.

The scale of the visualized data or information is usually quantitative. Nominal and unordered scales can also be used, but are more effective with other graphs.

The actual strength of this graph is the resulting shape of the values, which is unlike in scatter and bar graphs. By drawing lines along the points in time the data series gains a own shape, which visually summarizes all values.

“When lines are used properly in graphs, their slope is meaningful. For instance, with time-series data, the slope of the line from one value to the next represents the rate of change—the steeper the slope the greater the rate.” [Few, 2006b]

Stacked line graphs are an interesting variant of line graphs. Multivariate Data is stacked on top of one another. This is very useful to recognize percent changes of the various time series against each other.

2.4.3 Bar Graph

Bar Graphs are a variety of graphs which use horizontal rectangles to display the information. This type of graph can be used for all scales of data. Discrete data values are more effectively encoded than quantitative data values. The rectangles are usually horizontal positioned. Bar graphs with vertical rectangles are generally called column graphs.

Bars are contrary to points and lines much more dominant in a graph. The length of a bar is usually representing the value of the data point. The width of the bar should be constant for all bars.

Because bar graphs represents values by the means of a bar's length a lot on the space of the displayed bar, a quantitative scale must begin with the number zero. Scales that do not begin with zero would lead to misinterpretation of the data, because the lengths of the bars are not in relation to the encoded value.

There exist many variations of the simple bar graph. The grouped bar graph is especially well suited for the display and comparison of multidimensional or multivariate data. The data values are grouped by a common value like a shared date or another shared value of a variable. This makes it very easy to compare the data series. The disadvantage is the limitation to a few groups because of the space requirements. Also the development of a time series cannot be visualized as effectively as with line graphs, because individual values are emphasized rather than the overall shape.

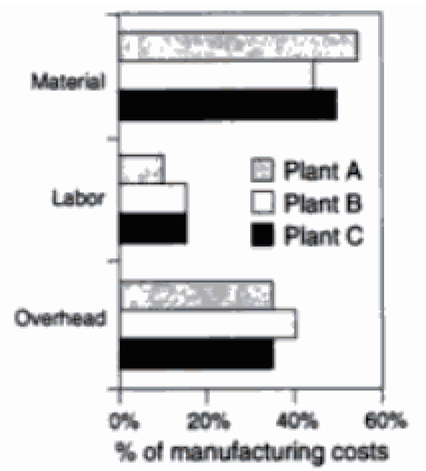


Figure 8: Grouped bar graph [Harris, 1999]

Stacked bar graph is similar to a stacked line graph. The multiple variables are stacked on top of another. Another variation is the paired bar graph, which is displaying two variables together. This graph is extensively used to display age distribution of females and males for a given population. Synonyms for this graph are pyramid graph or two way histogram.

2.5 Visual Variables

The display of multivariate data is limited by the spatial dimensions of a graph. Therefore the visualization of multiple time series requires special considerations. Normally, visual objects like points, lines or bars represent values of data series. But how can multiple data series be displayed in a graph? The different data series have to be distinguished by the viewer. The following section will introduce the concept of visual variables, which enable a distinction between different data series. The variables comply with studies and research of visual perception and human cognition studies.

The following visual variables and their ability are based on Jacques Bertin's Image Theory in [Bertin, 1983]. Bertin's work is summarized in [Green, 1998]. Both argue that visualizations are limited in their ability to encode data or information by planar and a set of retinal variables. The retinal variables include size, brightness, texture, color, orientation and shape.

Table 2 displays the visual variables and their competence for specific data classes. Bertin has divided the nominal data category into an associative and a selective category. This is necessary because the two categories differ in their ability to use visual variables.

	Associative	**Selective**	**Ordered**	**Quantitative**
Planar	Yes	Yes	Yes	Yes
Size		Yes	Yes	Yes
Brightness		Yes	Yes	
Texture	Yes	Yes	Yes	
Color	Yes	Yes		
Orientation	Yes	Yes		
Shape	Yes			

Table 2: Visual variables [Green, 1998]

Bertin and Green state that quantitative data is best encoded by planar and size variables. Both variables allow a precise distinction between data values, which is the key requirement for this data category. Good examples of the use of planar variables are scatter and line graphs. Bar graphs are a good example for the use of the size variable.

Figure 9 shows a scatter graph indicating ratios of height and weight for certain groups of the human population. The data is split into humans from Laos and from Burma. The data points are distinguished by using different colors. Data points of Laos are in standard red color. Those of Burma are in dark red. This is a good example of the selective category. The distinction between the two groups is easy for most humans.

The population is further divided into male and female persons. This is a good example of the associative category. Male and female persons of the same color belong to the same nation. The visual items are a rectangular shape for females and a circular shape for males. The distinction between the shapes is not as effective as the use of colors for representing the nation.

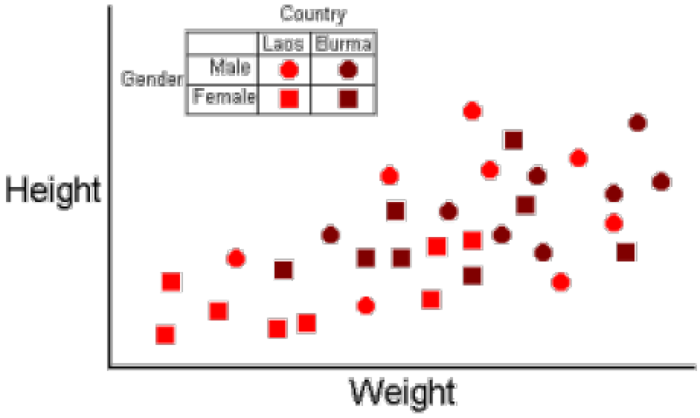


Figure 9: Demonstration of associative and selective categories [Green, 1998]

[Few, 2004] extends the classification of visual variables. Stephen Few lists four categories form, color, spatial position and motion. The category motion is a rather new technique of information visualization by using computers to animate the visual items. Table 3 lists the visual variables, which Stephen Few calls attributes, and their ability to encode quantitative data.

Category	Attribute	Quantitatively perceived?
Form	Orientation, Line length, line width, size, shape, curvature, added marks, enclosure	No Yes Yes, but limited Yes, but limited No No No No
Color	Hue, Intensity	No Yes, but limited
Spatial position	2-D position	yes
Motion	Flicker, Direction	-

Table 3: Visual attributes [Few, 2004a]

Again spatial and size variables are able to encode quantitative data. But Stephen Few also includes line length, line width and color intensity as additional techniques for encoding quantitative data.

Stephen Few differentiates between the ability of the attributes which are based on [Ware, 2004]. Only the attributes line length and 2-D position are not limited. The other three attributes line width, size and color intensity are limited. Humans are not able to estimate the underlying quantitative values precise enough.

2.6 Linear and Logarithmic scales

The scale of an axis is most of the time linear scaled. The use of a linear scale is an obvious choice, as it resembles the most natural way to visualize values onto a 2-dimensional space. The space between time units is constant along all value ranges.

The first problem of the linear scale is the display of high values. Usually axis values start with the number 0. If the data values are high enough, a lot of space of the chart is wasted. This problem can be easily solved by cutting the unneeded axis values and start with a higher number.

But a more severe problem arises when multiple data sets are involved and displayed in superimposition. When the values are too different it will be difficult to visually compare them successfully. As an example, consider the line graph in Figure 10 displaying the development of the closing price of Apple' s AAPL and Microsoft's MSFT stocks from 01.2004 to 12.2008.

It is apparent that AAPL stock is using almost the whole vertical space of the line graph with prices in the range of around 20 to 200 US-$. While MSFT stock is only using around a 10th of the whole vertical space with a price range of around 20 to 40 US-$.

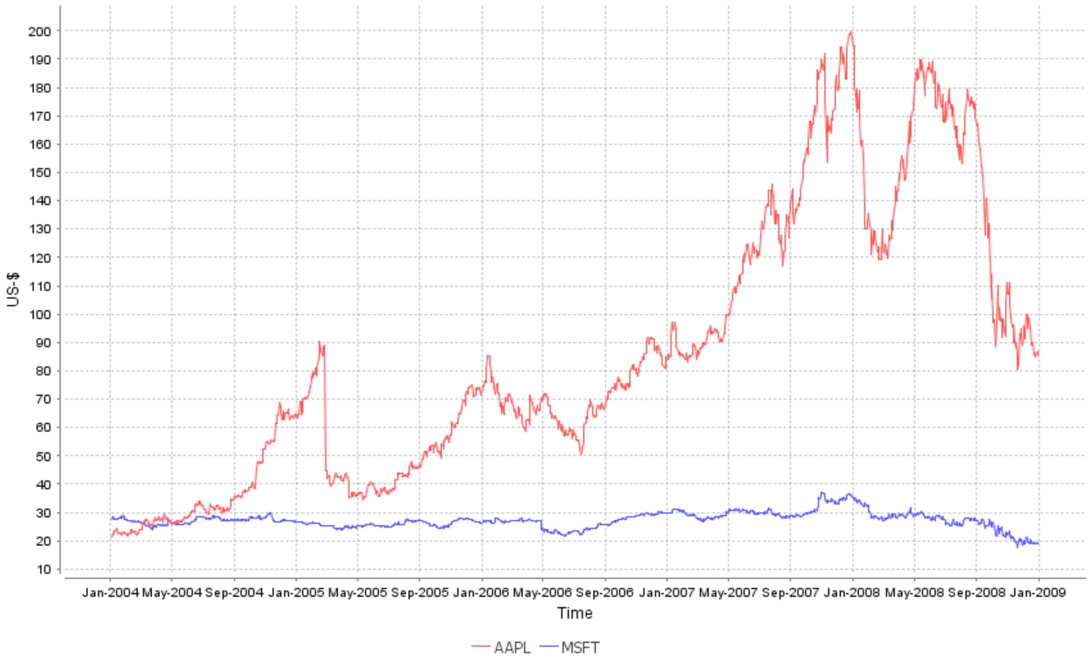


Figure 10: AAPL and MSFT - linear scale

The next graph in Figure 11 shows the same data with logarithmic scale for the y-axis. The development of the two curves is now much more visible.

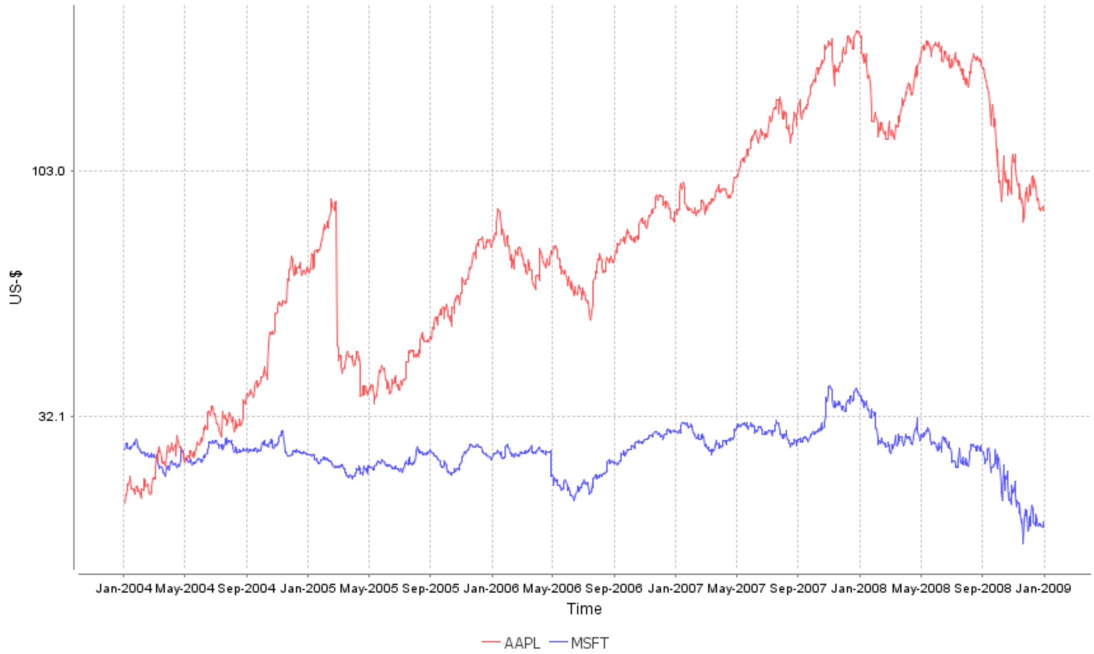


Figure 11: AAPL and MSFT - logarithmic scale

The differences of a linear and a logarithmic scale are visualized in Figure 12. A logarithmic scale is on the left side and a linear scale is located on the right side.

While the values for each tick are increasing times 10 on the log scale, the values on the linear scale are increasing constant by one unit.

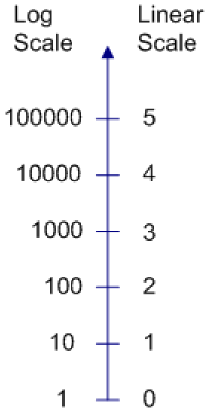


Figure 12: Linear and logarithmic scale

About the visual interpretation of (semi-) logarithmic graphs can be read in [Bertin, 1983] and in more detail in [Harris, 1999].

“The steepness of the curve on a semilog graph at any point or overall is proportional to the actual rate of change of the thing being plotted. The steeper the slope, the greater the rate of change, either positive or negative.” [Harris, 1999]

Figure 13 displays five data series. All start with the same value of 100 units. Each time series has a different percent increase. When using a simple line graph with linear scale the time series show an exponential development as expected.

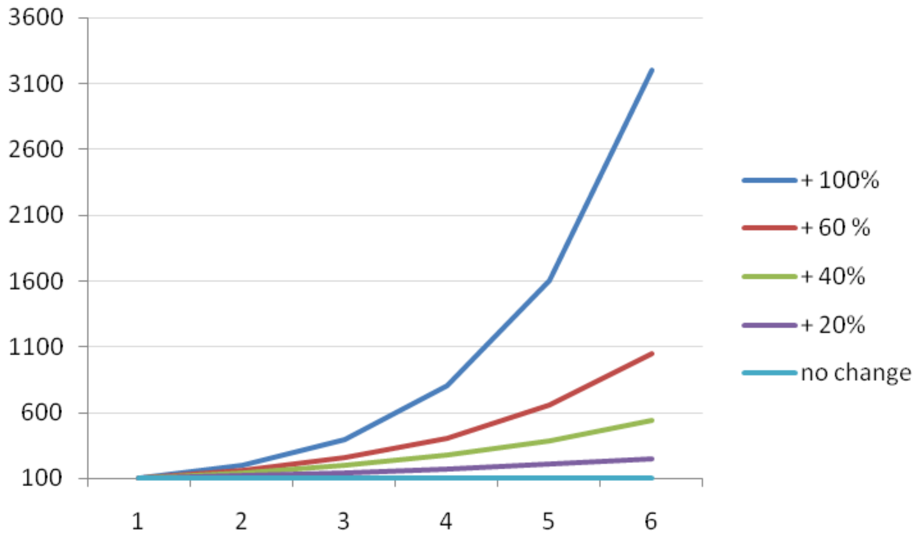


Figure 13: Linear scale

The same data series on a semi-logarithmic line graph reveals much more about the data. By using a logarithmic scale for the y-axis it is very clear to see that the percent increases are constant over time. Another effect is the ability to compare relative changes by comparing the slopes. Parallel lines have the same percent increase.

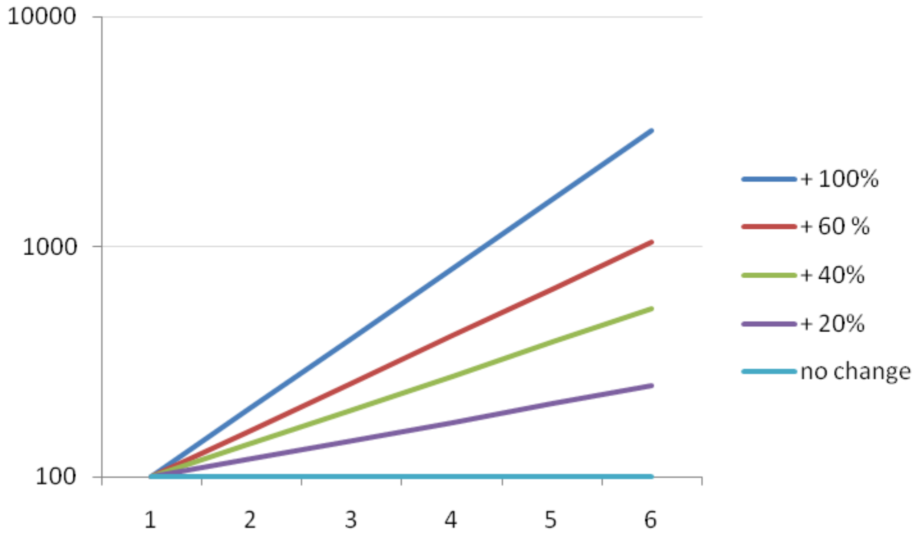


Figure 14: Logarithmic scale

3 State of the Art: Visual Comparison of Time Series

The display of only one variable is usually not enough. Often a comparison of multiple variables is needed. Consequently, the task of visual comparison is vital for visual analysis of multivariate data. Several methods are available, which can aid the viewer in the task of visual comparison.

For example, stock market applications require the comparison of different stocks and stock indices. By comparing the various data, the viewer can spot general and /or local trends. Onwards the viewer is able to extrapolate the current values, which will help with further decisions.

The first section of this chapter will explain two commonly used visual comparison methods. A few examples will be given and the main problems will be described.

Visual comparison is often used for stock market data visualizations. A small collection of example applications will be presented. This section will provide examples, which are used in practice.

At last, three advanced visual comparison methods will be presented. The first method is indexing. Even though it is just a transformation of the original values into percent values, the implications and effects of this method should be of great value for the viewer. The typical restrictions of superimposed line graphs are eliminated when using this method.

The other two advanced methods are Sparklines and Horizon Graphs. Both visualization methods are based on line graphs. But they are different from the traditional line graph in many ways. Especially Horizon Graphs are highly sophisticated.

The focus of this work lies on visual comparisons in the field of stock market data. Line graphs are often used to display stock market data.

3.1 Line graphs

In most cases line graphs are used for the display of time series data. The resulting shape of the data makes it easier to understand the historic development. Line graphs allow the viewer to quickly gain an impression of the development of the measured data.

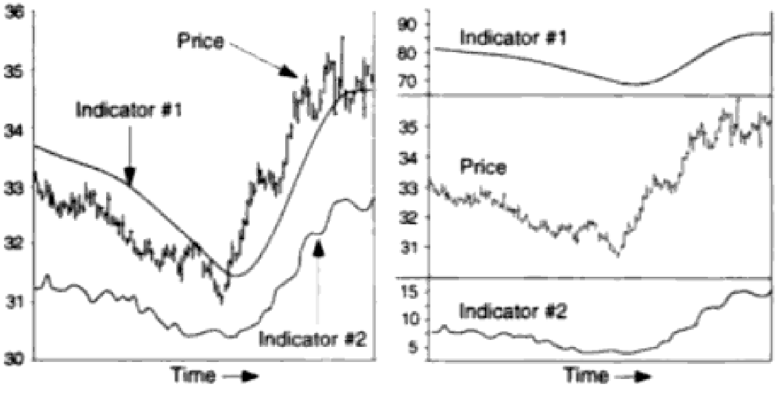


Figure 15: Stock price and indicators – juxtaposed (left) and superimposed (right) graph [Harris, 1999]

Multiple variables or multivariate time series can be visualized by juxtaposition or superimposition of the time series. Both methods have their own restrictions, which will be discussed in the following text.

3.1.1 Juxtaposition

Two or more graphs are juxtaposed, if the graphs are positioned next to each other. For time series it is best to place them above and below. Each graph should have same height and width for an optimal comparison. The time values should is typically displayed on the horizontal axis. The measured values should is typically displayed on the vertical axis.

Figure 16 shows six data series in juxtaposition. The time series are named for reasons of simplicity A, B, C, D, E and F. The advantage of this method is the independence of a common unit between all shown time series. But this also reduces the comparability, because there is no common unit for all data series.

Another disadvantage is the needed space of this method. Each data series increases the height. Thus the displayed time series are on average further away from each other, which make visual comparisons harder and inaccurate.

This is easily demonstrated by the following example in Figure 16. Compare the time series A on top against the time series F at the bottom in the figure below. This task should be much harder to accomplish than the same visual comparison of the two adjacent time series A and B.

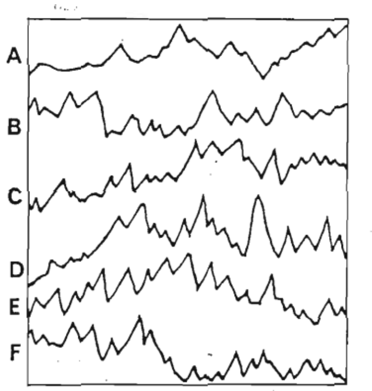


Figure 16: Juxtaposed data series [Bertin, 1983]

[Bertin, 1983] discusses the limitation of visual comparisons of time series in juxtaposition. Bertin suggests ordering the visualized time series to enhance the visual perception of the time series. He also describes the steps to create such an improved juxtaposed line graph.

First the time series have to be ranked, according to the occurrence of the highest peaks. This can be done by blackening the peaking values of each data series. After that the data series can be sorted by comparing the black filled areas. Time series, which have most peaks at the left side, are displayed at the bottom. Gradually the other time series are positioned above until the time series with the rightmost peaks, which stays on top.

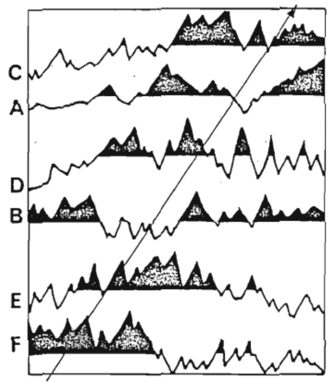


Figure 17: Juxtaposed data series, ordered by occurrence of the peak value [Bertin, 1983]

However this method is not suitable for many visual comparison tasks. The ranking of the time series is more or less arbitrary.

Juxtaposed line graphs are generally the simplest method suitable for visual comparisons and are easy to create, which is the main advantage. On the other side the visual comparisons are limited by the potential use of different units and scales.

3.1.2 Superimposition

Superimposed graphs or sometimes also called overlaid graphs are displaying multiple variables on top of each other in one graph. This method is better suited for visual comparisons of multivariate data.

Homogenous, multivariate data shares the same value unit. The display of such data should posses no problems. Otherwise, if the multivariate data does not share the same unit, the difficulty of this method is the selection of the right scale.

The two line graphs in Figure 18 below depict production and salaries data. Production is measured in thousands of tons and the variable salaries is measured in Franc. The line graph on the left and on the right use the same data. The only difference between them is the different range of the value axes. However the left graph clearly illustrates a rise in production, whereas the right graph is nearly illustrating the opposite. Salaries have greatly increased over the last time.

Heterogeneous time series require at least two vertical axes, which ranges have to be carefully chosen. The consequence of this requirement is that visual comparisons are not completely valid and the visual perception is dependent on the relation between the chosen value ranges.

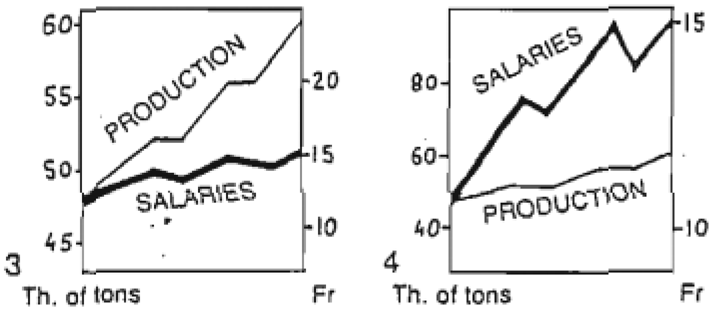


Figure 18: Superimposition [Bertin, 1983]

Figure 19 shows historic values of the US inflation (US CPI) and US Interest Rates for the year 2008. The data series are measured in different units and thus the line graph has two vertical axes. The ranges of both axes are adjusted to an equally placed display of values on the available vertical space. A visual comparison is possible, although intersections are meaningless.

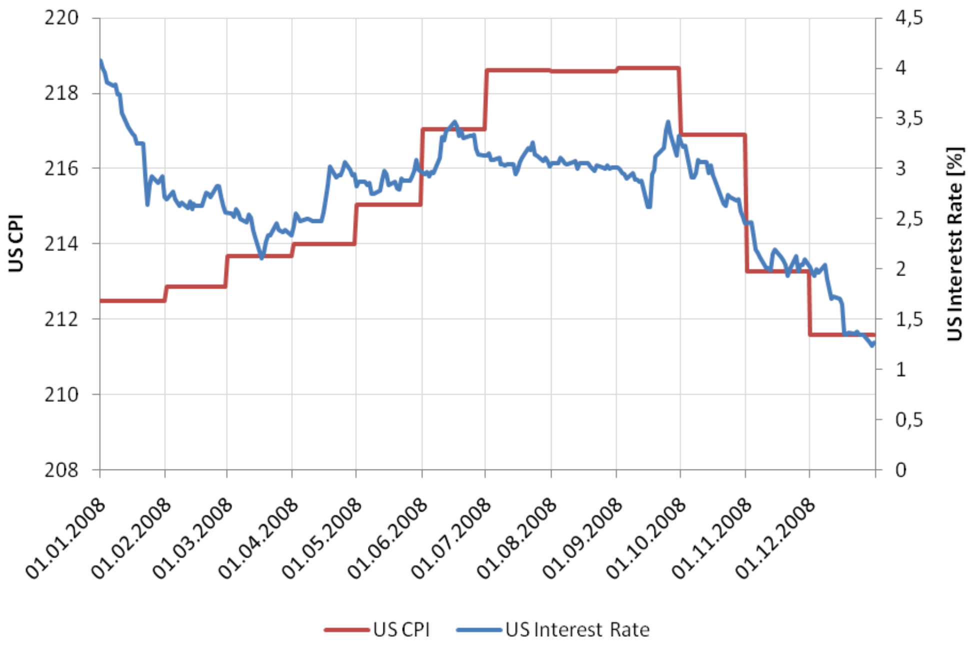


Figure 19: Inflation and interest rate at optimal axis range

The next graph in Figure 20 shows the same data. The range of the US inflation (US CPI) axis has been changed to a minimum of 0 and a maximum of 250. The graph differs greatly from the one before. Again the intersection at the start has no meaning.

The visual representation of US inflation data (US CPI) is almost flat. The viewer recognizes little to no change of the values for this time series during the whole year. The curve of US Interest Rate has not changed from the previous graph. The segmented line of US interest rate looks more volatile and unpredictable. At the same time US inflation (US CPI) looks very stable, unlike in the previous graph.

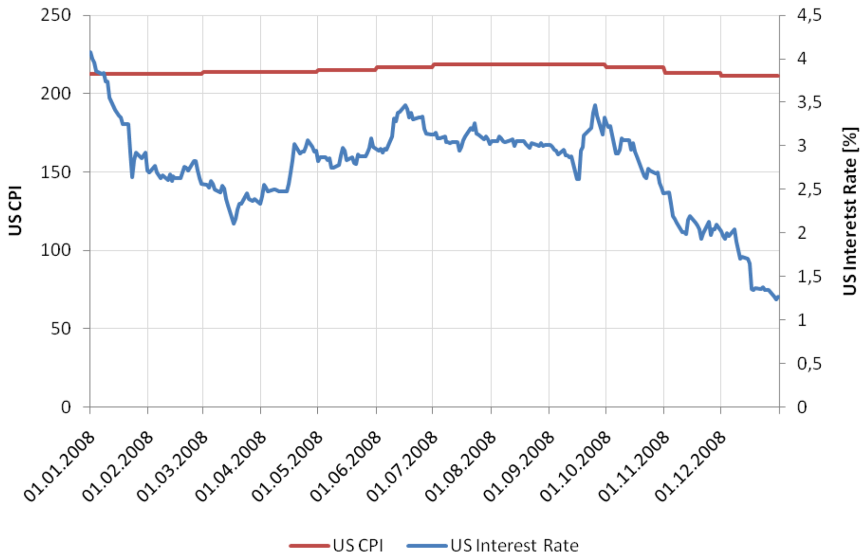


Figure 20: Visual distortion by suboptimal range of inflation axis

The graph in Figure 21 uses the same data as before. Both vertical axes have changed ranges. The vertical axis of US Interest Rate ranges from 0 to 14. As a result the segmented line of US Interest Rate seems flat and looks stable, than in the graph before. US inflation looks in the third graph more unstable.

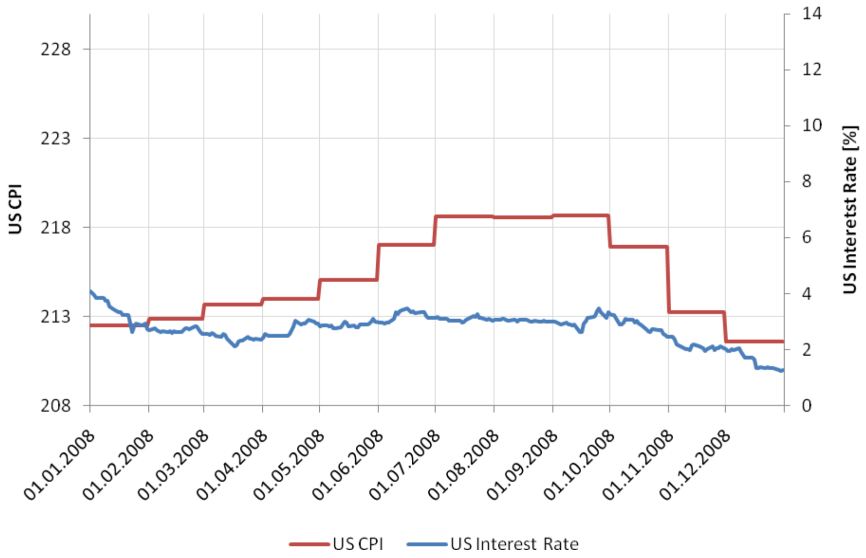


Figure 21: Visual distortion by suboptimal range of both vertical axes

Both time series are intersecting at February, which has again no meaning. The intersection may lead the viewer to false conclusions about the relationship between the time series. The intersection is only a result of arbitrary alignment and unit size of the two vertical axes.

By looking at the previous three superimposed graphs it should be clear that this method tends to have arbitrary results. The first superimposed graph is the best choice for a visual comparison. But is it valid enough to be used for serious visual comparison tasks? In any case it can be used to manipulate the visualization to display the data according to own needs. The creator and the viewer should be aware of this fact.

Superimposed graphs are in almost all situations superior to juxtaposed graphs. Only in rare occasions, if there is too much data displayed, it can be better to use juxtaposed graphs instead.

3.2 Software applications

This section presents two examples of software applications, which are suited for tasks of visual comparisons with multiple time series. These two examples are selected from a broad range of state of the art software applications. Further software applications are explained and discussed in more detail in [Ma, 2009].

The first example stands for the new form of web 2.0 services which provide a lot of interactions. The stock market application is owned by Google Inc. While other stock market visualizations are longer available, this application has a lot of interactive features to offer. There are also some innovative features, which make Google Finance special.

The second application is an interesting example from academic research. The goal of the project was to design a tool to visually spot trends and patterns in time series data.

Both examples will give the reader an overview from different views of state of the art visualizations for stock market data.

3.2.1 Google Finance

Google Inc. started this web based application in March 2006. It is available under the WWW-address http://www.finance.google.com and provides free financial information of all major US stocks and stock indices. The stock market data is updated in real-time. This should be a big advantage for a lot of small investors, who are not willing to pay extra money for up-to-date stock market data.

Although many other financial information websites are available, Google Finance provides a dynamic and highly interactive web application. As previously mentioned access is free and the application incorporates many innovative and useful functions.

By default, stock prices are displayed via line graphs. There are two other common stock market visualization types available. Both, bar charts and candlestick charts, display additional stock information: open, high, low and close prices.

The whole application is highly responsive. Changing or adding of stocks and stock indices is done by a few mouse clicks. The values in the legend are corresponding to the selected day which is set according to the current horizontal mouse position on the chart.

A quick zoom function of most common time intervals is displayed at the top left corner of the graph. Of course zooming is also possible by using the mouse wheel. It is also possible to set a individual time range by dragging the sliders at the bottom.

The legend and some other important financial or company related information are above the stock market visualization. The trading volume is displayed below the stock prices graph. Stock prices and volume are displayed at the top right corner.

Figure 22 is a typical example of the stock visualization. The blue line represents stock price of Apple Inc. from the first day to the last day of July 2009. The red line is a simple 30 day moving average of Apple's stock. The vertical price axis uses a linear scale and automatically adjusts the price range to the minimum and maximum displayed price.

Figure 22: Google Finance displaying AAPL stock and moving average [Google, 2009]

The indexing method is automatically used for visual comparison of multiple time series. All values are transformed into their representative percent changes. Unfortunately the indexing point is fixed to the first displayed date. A free selection of the indexing point could further increase the usefulness of this application.

Figure 23 shows an example of three superimposed stock market time series. Apple's AAPL stock in blue color, Microsoft's MSFT in red color and the Dow Jones stock index (DOW) in orange color are visible.

At the bottom of the figure is a list of related companies. Next to their name, stock symbol, stock price, stock price change and market capitalization are the last daily values by Sparkline visualizations drawn. This is a pretty effective visualization to compare the development of the last month.

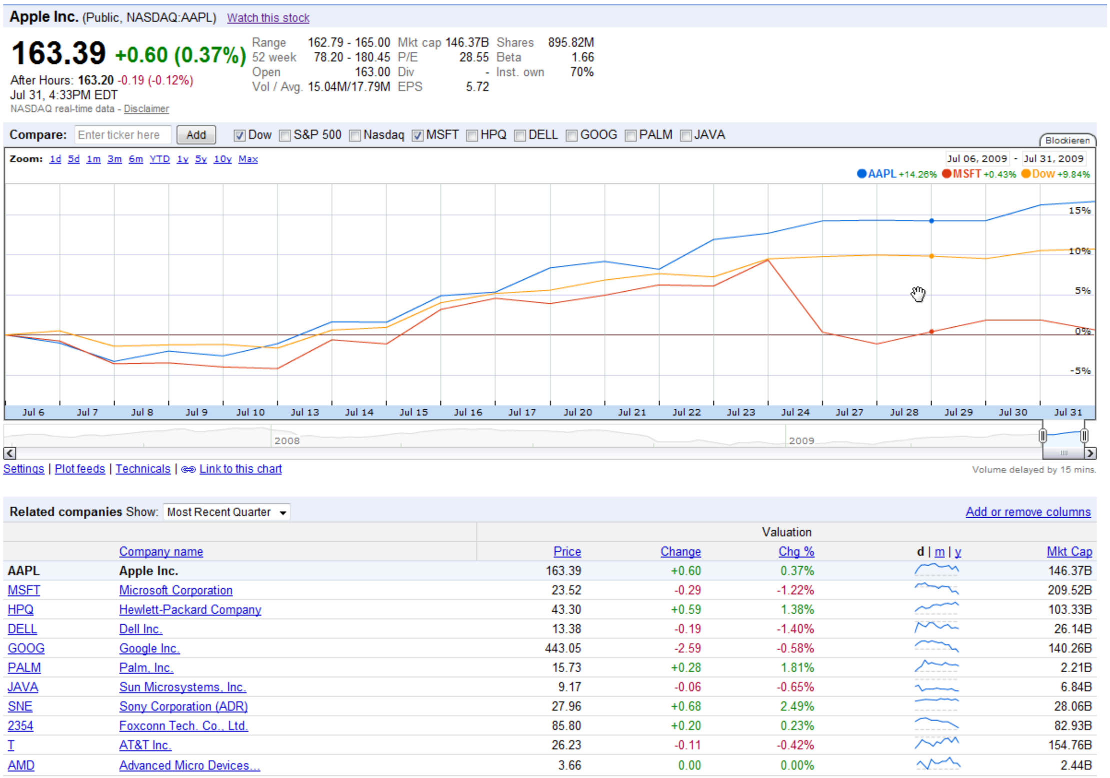


Figure 23: Google Finance displaying three time series [Google, 2009]

One clearly novel function is the direct integration of complementary blogs and news headlines into the stock market visualization. Although the news headlines may offer only fractional advice, it is a very interesting feature. Furthermore standard RSS feeds and Google spreadsheets can be added to the visualization.

The potential of this function is high. An extra application could generate RSS feeds of specific events of stock market and other economic data. These events can then be integrated in the Google Finance visualization. The comparison of time series and events can aid the process of identifying dependencies.

Figure 24 shows a one year period of the Dow Jones stock index. In addition related events are displayed. Each one of the news headlines in the top right area is mapped to a unique letter. The stock market visualization on the left shows for each news headline a flag on the specific day of the headline.

Figure 24: Google Finance mapping news events to time series [Google, 2009]

3.2.2 TimeSearcher

The aim of this project of the HCI Lab at the University of Maryland is the visual discovery of patterns and trends in time series data. Because of the success of the first prototype, a second prototype was developed. The third prototype for the TimeSearcher project is currently in progress.

The first TimeSearcher prototype was made by Harry Hochheiser and the well-known computer scientist and professor Ben Shneiderman. The research goals were to gain knowledge about interactive methods for visual discovery and exploration of multiple time series data. The prototype includes various homogenous data sets consisting of stock market data over various time periods.

Figure 25 shows a screenshot of the first prototype. The top left area shows all time series in a superimposed column graph. The turquoise colored shapes are graphical query functions, which let the user directly select certain data points instead of formulating mathematical conditions. These TimeBoxes allow an intuitive selection process, which should increase the speed of exploration and discovery tasks.

Below the superimposed graph are juxtaposed line graphs of all available time series drawn. Value range and scale of the y-axes are same for all graphs. The horizontal time axes are also linked for all graphs. Selected data points are visualized by turquoise coloring.

The table at top right shows all numerical values for the selected time series. Below is a list of all available time series. This list has the function of a time series selector for the whole application. Two sliders at the bottom right set the bounds for a selected TimeBox. The slider on top sets minimum and maximum times and the bottom slider set lowest and highest values.

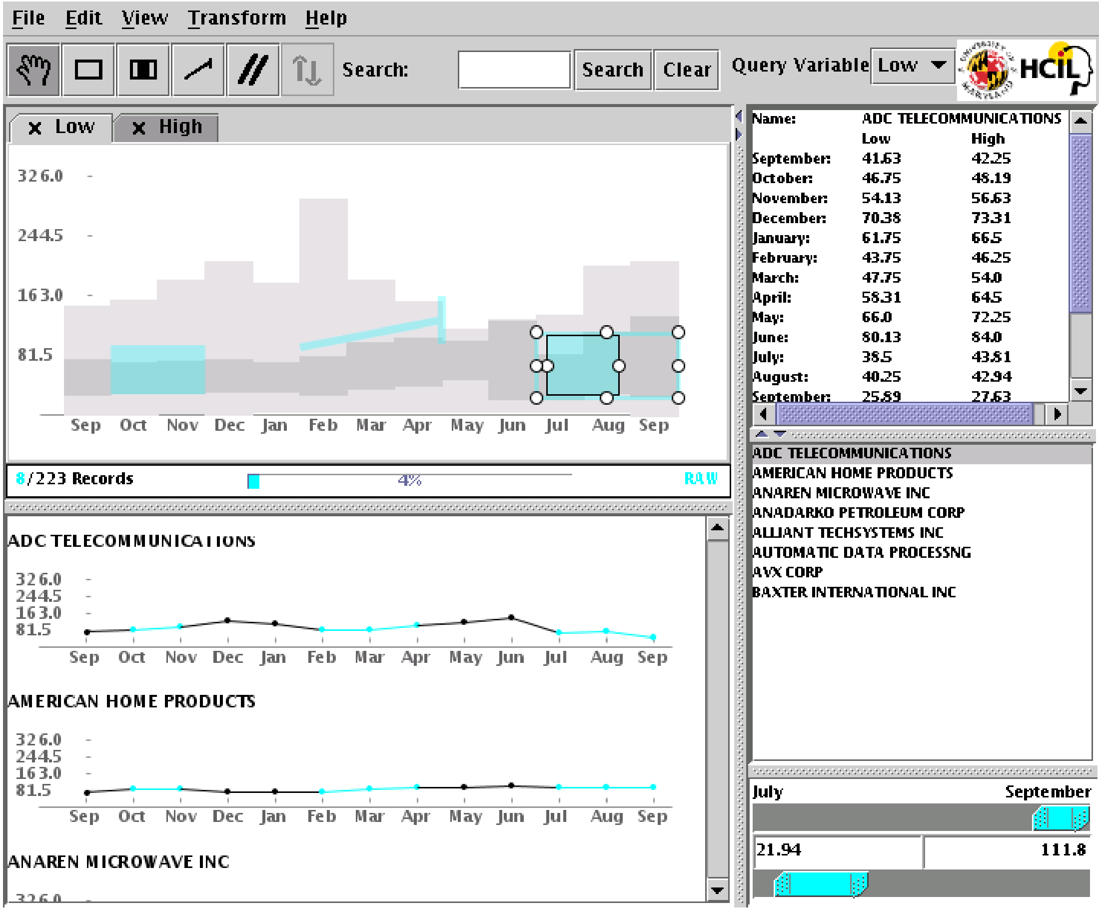


Figure 25: TimeSearcher 1 – Brushing of data through TimeBoxes in the top window [TimeSearcher, 2009]

As previously mentioned the visualization prototype has a rather unique graphical interaction technique. The user can explore the data by drawing a TimeBox over the specific area to select all relevant data. Hochheiser and Shneiderman describe the TimeBoxes as follows:

"Queries are built using timeboxes: a powerful graphical, direct-manipulation metaphor for the specification of queries over time-series datasets. These timeboxes support interactive formulation and modification of queries, thus speeding the process of exploring time-series data sets and guiding data mining." [Hochheiser et al., 2001]

The second prototype was rewritten to use the programming language C#. The main reason was to improve the performance, which was necessary to incorporate the new features.

"TimeSearcher 1's basic browsing capability was extended to include multiple heterogeneous variables and tens of thousands of time points. In addition, TimeSearcher 2's new search interface combines both filter and pattern search capability, implementing a three-step approach that can be extended to a variety of time series search interfaces." [Buono et al., 2005]

The prototype's main interface elements can be thought as overview and detail parts. The graph at the bottom shows an overview of the data. The detail views show multiple variables of one time series. The user can enable or disable the display of each variable. On the top right area of the prototype is a numerical table located. This table displays numerical values for the selected time series. At the bottom is a list of all available time series.

The time range of the detail views are altered by moving and resizing the orange rectangular box in the overview part, which is called field of view box.

The TimeBox function, a turquoise color filled rectangle, has the same purpose as in the first prototype. It acts as a dynamic filter of the available data. All not fitting time series is hidden.

Figure 26 depicts a multiple selection of time series. The variables price and velocity are displayed. Two TimeBoxes filter the displayed time series.

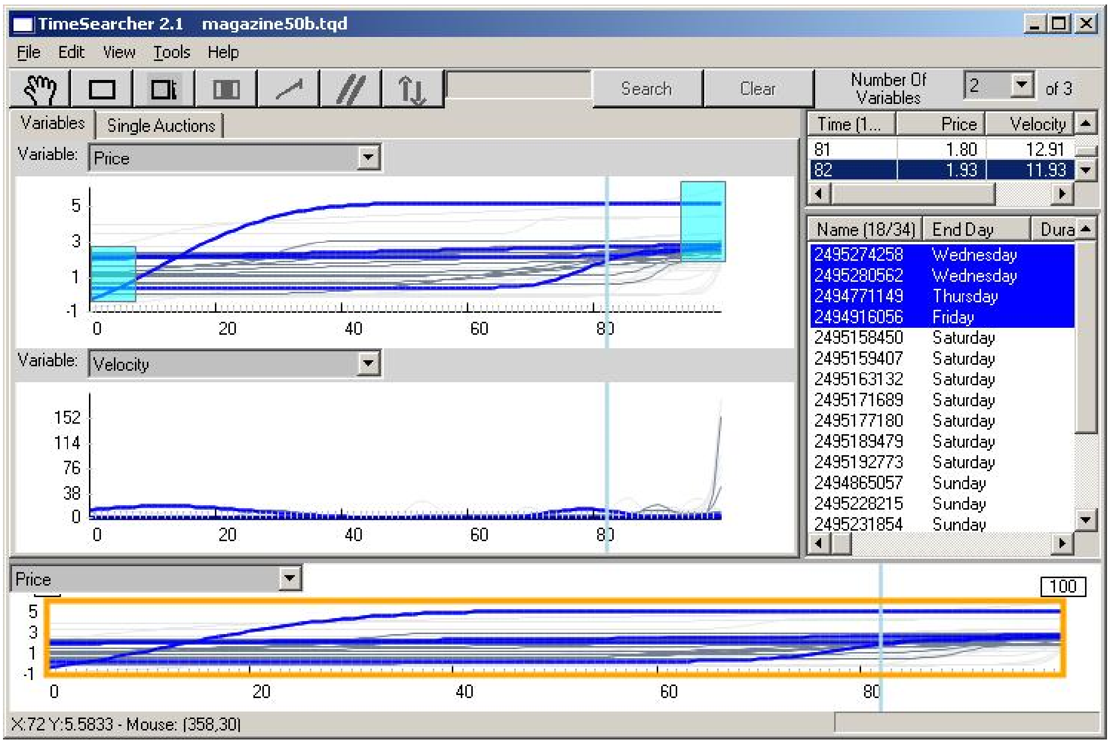


Figure 26: Multiple highlighting and TimeBox filtering [TimeSearcher, 2009]

SearchBox is a widget to find similar patterns in time series. The user has to draw a red filled rectangle over the data first. The enclosed data represents the original pattern, which is the base for the search. The search function includes four different transformations, which have different effects on the result. Results are similar data points according to the selected transformation. All matches are drawn in red color.

The screenshot in Figure 27 demonstrates how the SearchBox works. Two variables sunlight and humidity-average are displayed. A SearchBox for the variable sunlight is located at the 9/26/1996 time mark. The pattern is a horizontal mirrored V. The prototype has found a similar pattern in the data to the left. The start point of the similar data is marked with a red upward arrow. The shape of the similar pattern is also overlaid in red color.

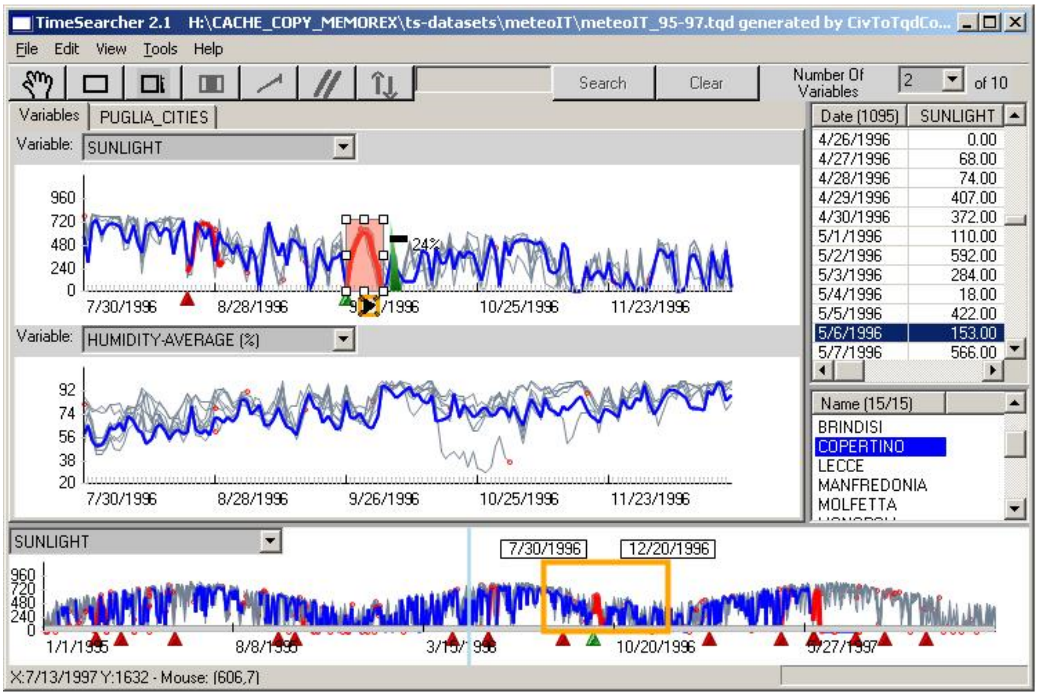


Figure 27: Overview (bottom) and detail views [TimeSearcher, 2009]

The third prototype is described in detail in [Buono et al., 2007]. The main focus lies on forecasting of time series. This function has a large number of applications such as auctions, new stock offerings or industrial processes. A data driven forecasting method and interface called Similarity-Based Forecasting (SBF) has been added.

A pattern matching search in the historical dataset produces a subset of similar curves to the partial time series. The forecast is displayed as a river plot showing statistical information about the SBF subset. A forecasting preview interface allows the users an interactive exploration with multiple simultaneous forecasts.

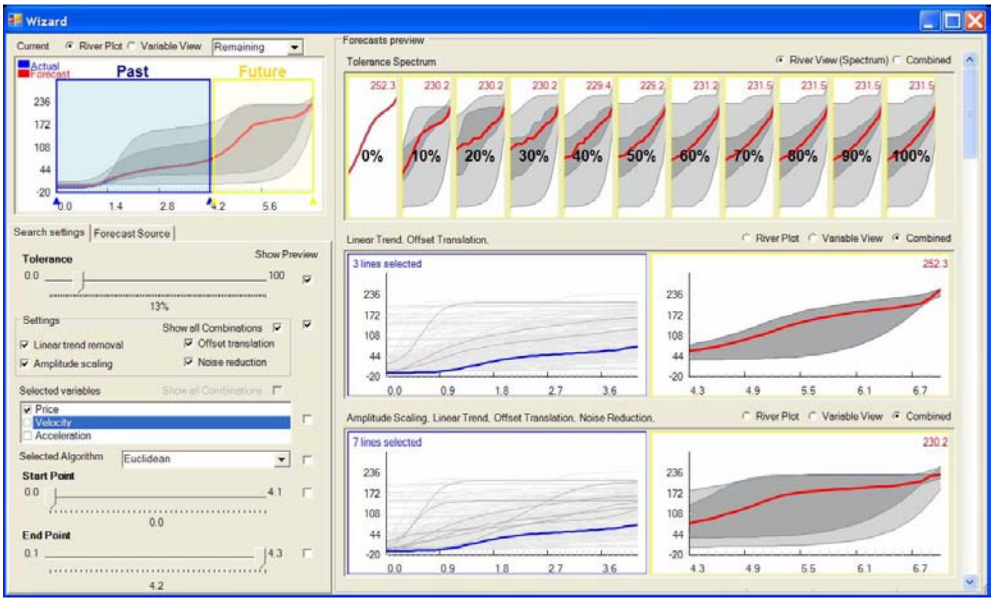


Figure 28: TimeSearcher 3 – Forecast variations (right) [TimeSearcher, 2009]

More detailed information and download links for all three prototypes and a series of related scientific articles can be found at the link: http://www.cs.umd.edu/hcil/timesearcher/

The TimeSearcher project delivers enhanced interactive methods for multivariate visualizations. The focus lies on queries and selection of time series. Discovery of similar trends and patterns can be executed quickly.

However the application does not provide enough support for visual comparison tasks. While superimposition of time series is used the benefit is limited. Unfortunately only variables of the same unit are used in the visualizations. Special comparison methods would most likely increase the results of exploration and discovery tasks.

3.3 Advanced Visual Comparison Methods

Juxtaposition and Superimposition are two simple comparison methods for multivariate time series. Unfortunately, both methods have their difficulties when using heterogeneous data.

Each of the following three advanced methods offers a better solution for visual comparisons of heterogeneous multivariate comparisons.

The indexing technique is first described. It can be used for homogenous and heterogeneous time series. This method uses a simple line graph, but transforms the values.

Sparklines is another advanced method based on line graphs. It was proposed by the visualization guru Edward Tufte.

The third advanced method is the horizon graph. This advanced comparison method seems to be very effective to compare several time series according to [Heer et al., 2009].

3.3.1 Indexing method

Jacques Bertin defines a general formula of the indexing method in [Bertin, 1983]. The formula transforms the original measured data values into indexed values. Naturally only quantitative time series data can be used.

The data points are converted into percent values, referencing relative changes between a given base or indexing point and the each other data point. The selection of the base point determines the result of the function. Therefore the base point has to be selected at first.

$$index_n = \frac{Q_n}{Q_i} * 100 \ [\%]$$

The base or index point is mathematically denoted by the index i. The indexing value $index_n$ of a point n is calculated, by dividing the original measured data value Q_n at point n through the original measured data value Q_i of the indexing point i. The result is multiplied by 100, which makes percent values more understandable to humans.

By using this method all displayed time series values use the same percent dimension. Heterogeneous time series is far easier to compare. The transformed data values can be visually compared by using simple comparison methods. For example the time series can be drawn in superimposition without any arbitrary scales and ranges of the different axes dimensions.

Each stock market includes a broad range of various stocks from companies in different industrial sectors. The price of a share is determined by the market and therefore price ranges are often diverse.

A line graph of two stocks is illustrated in Figure 29. The visualization is using the indexing method to enhance the visual comparison. Both stocks start at an index point which is the first displayed value on the chart. The value on the y-axis is 0 % and represents the change of the price at time t = 0. The y-axis has a percentage dimension. The big increase of the price of Apple's stocks form May until September 2008 is very clearly visible.

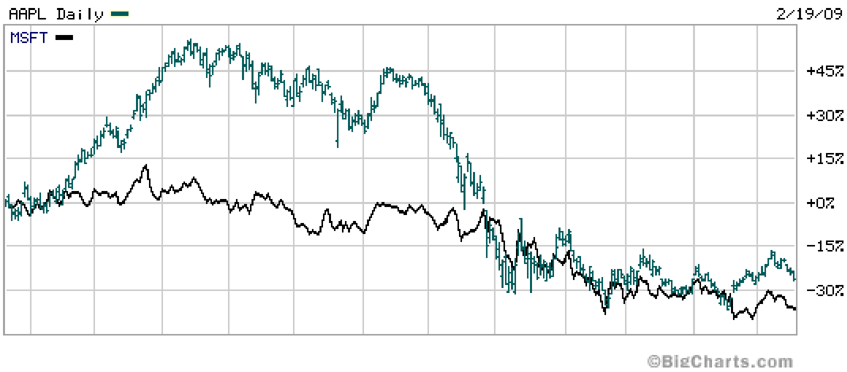


Figure 29: Comparison of Apple's and MS's stock over a 1 year period [BigCharts, 2009]

Figure 30 shows five stocks. A visual comparison of the time series is easy because the visualization displays the percent values instead of absolute values.

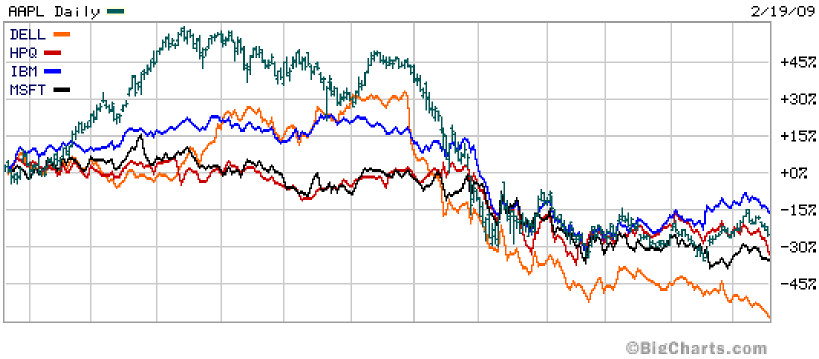


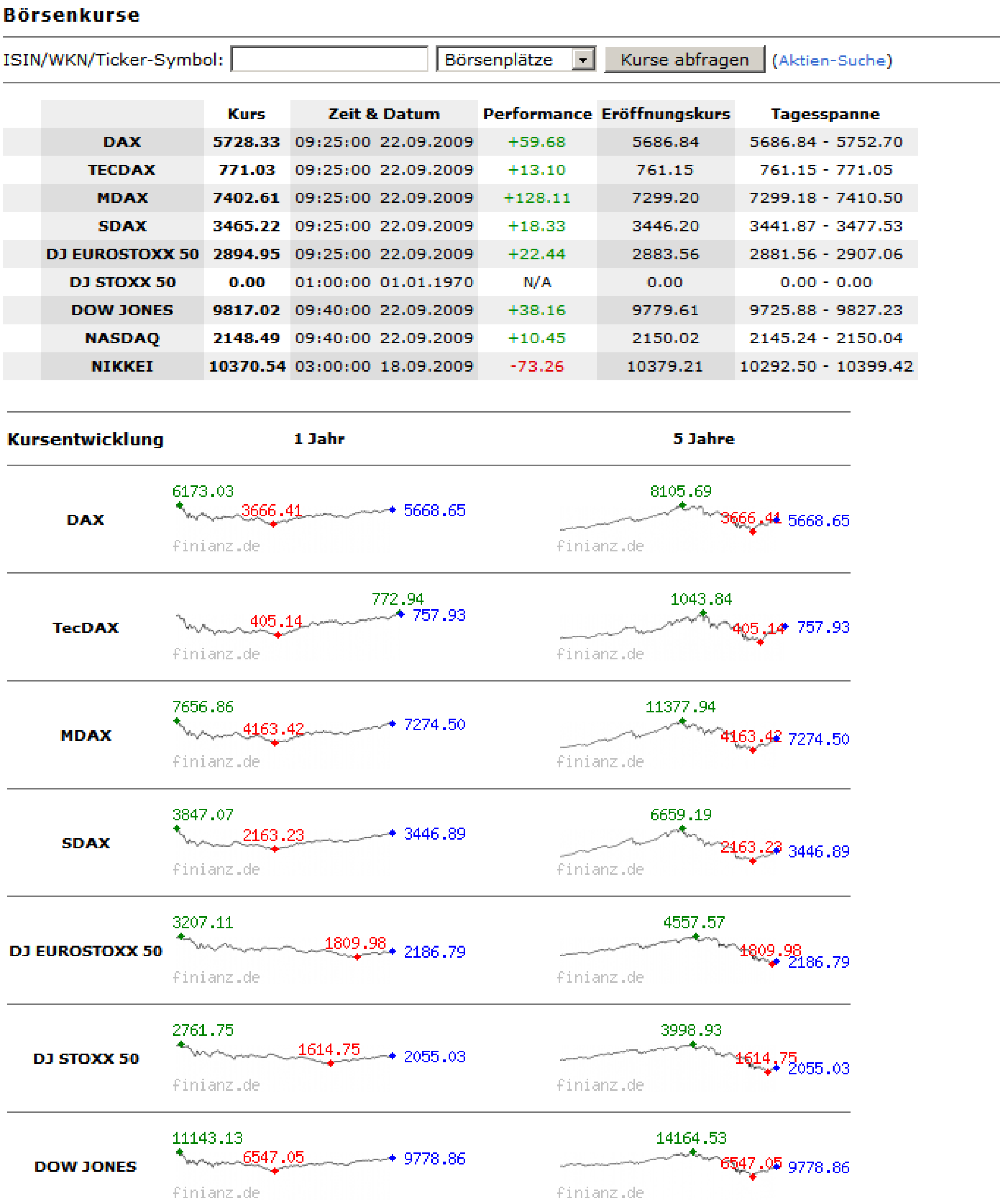
Figure 30: Comparison of 5 IT stocks [BigCharts, 2009]

3.3.2 Sparklines

Sparklines are very small line graphs for time series. This method was first introduced by Edward Tufte in his book Beautiful Evidence [Tufte, 2006]. He describes Sparklines as "*intense, simple, word-sized graphics*". Sparklines are extremely compact line graphs for encoding values of time-oriented data.

The focus of this visualization lies strongly on the development of the curve and not on specific values or dates. Comparisons of homogenous or heterogeneous time series are possible, although they are somewhat vague.

Sparklines can also be used as stock market visualizations. Figure 31 shows four important stock indices. Sparklines on the left side are representing values of a one year time period and on the right side of a five year time period. The four stock indices are above and below each other in juxtaposition. The horizontal time axes are aligned which makes visual comparisons easy. Additionally in this visualization, the highest, lowest and current values are displayed in green, red and blue color.


Börsenkurse

ISIN/WKN/Ticker-Symbol: Börsenplätze Kurse abfragen (Aktien-Suche)

	Kurs	Zeit & Datum	Performance	Eröffnungskurs	Tagesspanne
DAX	5728.33	09:25:00 22.09.2009	+59.68	5686.84	5686.84 - 5752.70
TECDAX	771.03	09:25:00 22.09.2009	+13.10	761.15	761.15 - 771.05
MDAX	7402.61	09:25:00 22.09.2009	+128.11	7299.20	7299.18 - 7410.50
SDAX	3465.22	09:25:00 22.09.2009	+18.33	3446.20	3441.87 - 3477.53
DJ EUROSTOXX 50	2894.95	09:25:00 22.09.2009	+22.44	2883.56	2881.56 - 2907.06
DJ STOXX 50	0.00	01:00:00 01.01.1970	N/A	0.00	0.00 - 0.00
DOW JONES	9817.02	09:40:00 22.09.2009	+38.16	9779.61	9725.88 - 9827.23
NASDAQ	2148.49	09:40:00 22.09.2009	+10.45	2150.02	2145.24 - 2150.04
NIKKEI	10370.54	03:00:00 18.09.2009	-73.26	10379.21	10292.50 - 10399.42

Figure 31: Sparklines of various stock indices [Ma, 2009] (Source: http://finianz.de/news/boersenkurse.php)

3.3.3 Horizon Graphs

This rather new visualization method is analyzed and discussed in detail in [Few, 2008] and [Heer et al., 2009]. The first article by Stephen Few introduces the concept of horizon graphs and explains important characteristics of this unique method.

The horizon graph increases the encoded data density by mirroring and layering bands. The term mirroring means that positive and negative values are displayed on the same side of the axis. The values can be distinguished by using two distinctive colors. In the article the color for positive values is blue and for negative values is red.

Figure 32 shows a simple evolution from a filled line graph to a horizon graph in 3 steps. The first is a filled line graph. The second is mirrored so that the negative values are also drawn on the positive side, but in order to distinguish them from positive values, the negative values have a red filling. The third graph is called horizon graph. It is derived from the second graph but divides the value areas in bands. These bands are then layered and thus reducing the needed height of the graph.

The second article takes a more pragmatic approach to this visualization method. The horizon graph is tested for speed and accuracy while performing comparison tasks. Dependent variables are three slightly different horizon graphs (line chart, mirrored line chart and mirrored layered line charts) and the chart height.

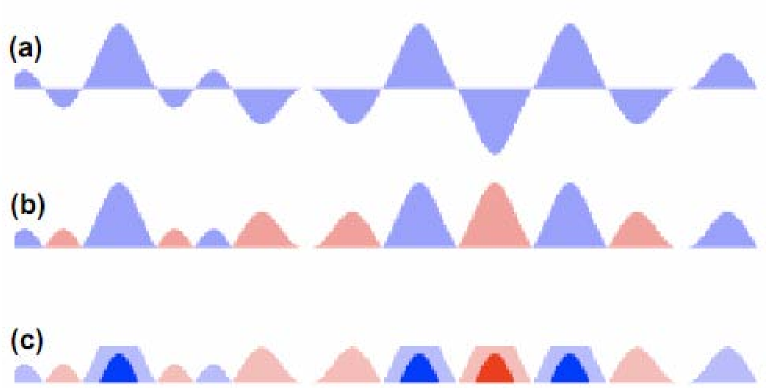


Figure 32: Evolution of the Horizon graph in 3 steps: simple line chart (a), mirroring negative values (b), layering of values (c) into bands [Heer et al., 2009]

“The first graph is a filled line chart - a line chart with the area between the data value on the line and zero filled in. The second graph “mirrors” negative values into the same region as positive values, and it relies on hue to differentiate between the two. The mirror chart doubles the data density compared to the line chart. The third chart, called a horizon graph, further reduces space use by dividing the chart into bands and layering the bands to create a nested form. With two layered bands the horizon graph doubles the data density yet again.” [Heer et al., 2009]

The test results give evidence that mirroring does not negatively affect the visual perception. Another result leads to the assumption that layering can improve the effectiveness. The chart height correlates to the task completion time. The smaller the chart height, the faster could the subjects finish their tasks. However the error count for estimations of values increased as well. The conclusion of both papers is that the horizon graph should be well suited for comparison tasks of multivariate time series.

Figure 33 shows the closing prices of thirty stocks over a one year time period. Traditional juxtaposed line charts are usually more limited by the count of displayed time series. Horizon graphs are able to display more time series because of a less needed height for each chart.

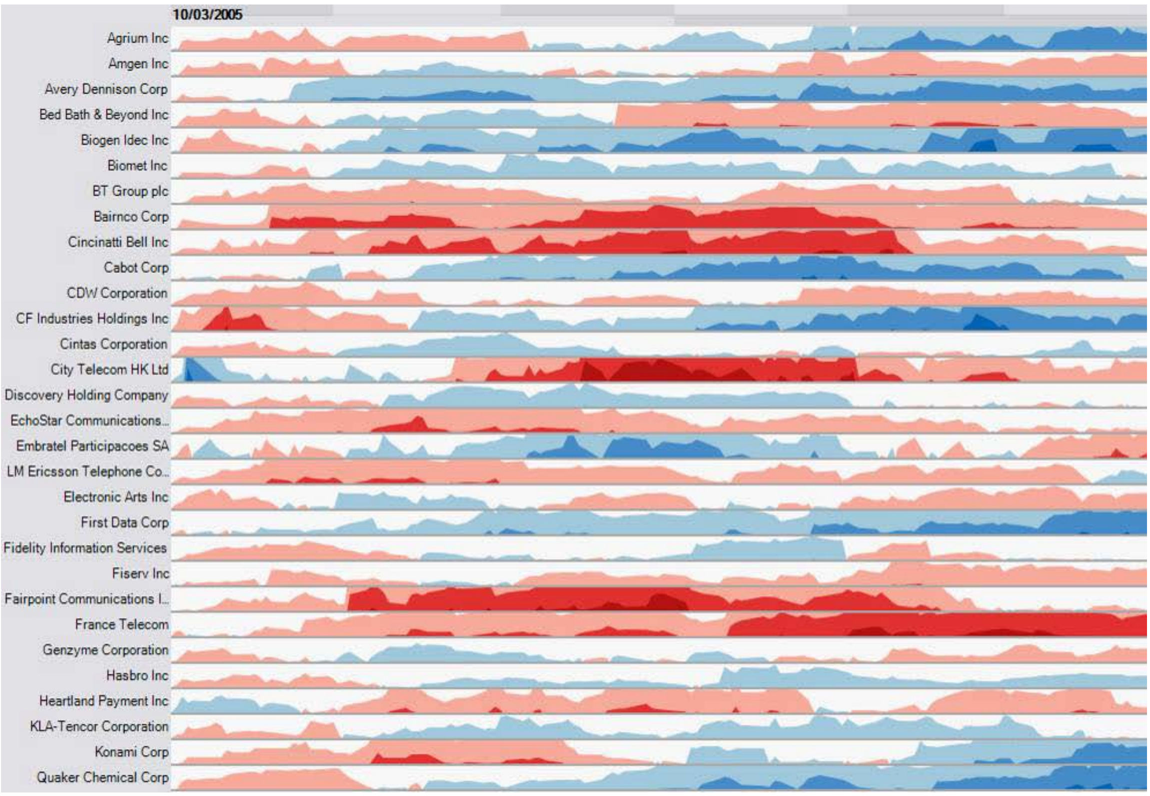


Figure 33: Horizon Graph visualizing 30 stocks [Few, 2008]

3.4 Summary

This chapter introduced two simple comparison methods, which can be used for almost any time series visualization. Juxtaposition means to align the visualizations according to the time dimension.

Superimposition can be understood as overlaying of all single graphs. This method is great for homogeneous data such as comparison of stocks. Heterogeneous data which does not share the same unit may lead to false conclusion about the data.

Better and faster comparison results may be achieved by the usage of an advanced visual comparison method. Three advanced methods were explained. Indexing method transforms the original values into percent values based on the indexing point. This makes proper comparisons by superimposition possible.

The second advanced method is sparklines. These extremely compact line graphs can visualize data trends very well. Multiple sparklines are typically compared by juxtaposition.

The last of the three presented advanced comparison methods is called horizon graph. The data is compared by small line graphs in juxtaposition. Mirroring and layering of values reduce the required height.

All three advanced comparison methods are based on the two simple comparison methods juxtaposition and superimposition. The advanced methods aim at the main weaknesses of the simple methods and solve their limitations.

Indexing enables to visualize heterogeneous data by superimposition. This avoids the main problem of arbitrary definitions for the different scales and ranges for each unit.

Comparison by juxtaposition needs a lot of space. Each additional visualized variable increases the needed space. Graphs are further away from each other, which could increase estimation errors for comparison tasks. Sparklines counter this limitation by reducing the size of each graph. Horizon graphs reduce the needed height by mirroring negative values and layering of value ranges into bands. Unlike the sparkline visualization values can be distinguished more precisely thanks to the differently colored bands.

Stock investments are a widespread investment instrument for many people. Increasing stock prices over the last few years lead many investors to the stock market. Only in the second half of 2008 the majority of stocks have lost some of their value. However stock markets have and will exert a fascination to many investors.

Stock market data is easily available and includes various data such as stocks, stock market indices and technical indicators. Stocks further include daily opening, highest, lowest and closing stock prices and volume information.

The combination of easy access to the data and relevance for many people makes this application field for comparison of heterogeneous time series promising.

The word chart will be used instead of the word graph in the succeeding chapters about stock market visualizations.

[Harris, 1999] gives a hint about the origin of the distinction. *"A chart is a vehicle for consolidating and displaying information for purposes such as analyses, planning, monitoring, communicating, etc. Previously, charts were tangible things such as single sheets of paper, display boards or flip charts."*

Nowadays both words are rather equal in their meaning regardless of their medium. In finance application areas it is common to use the word chart for stock market visualizations. This is the reason to use it in the following text.

4 Design and Architecture of the Prototype

An important part of this master thesis is the evaluation of different visualization types for visual comparisons of time series. The prototype enables to measure certain performance values for predefined tasks.

The indexing method will be evaluated against basic comparison methods such as Juxtaposition and Superimposition. The focus of the research lies on the ability to visually compare multiple time series. The design and the architecture considerations of the implemented prototype will be explained in this chapter.

The theoretical and practical foundation of software prototype is described in the previous chapter. Simple comparison methods like juxtaposition and superimposition and advanced comparison methods like indexing will be implemented. Results from the theoretical research and from software examples are considered in the implementation.

The prototype application includes most common stock market information. Daily volume and stock price information of six stocks AAPL, AMZN, CHINA_PETROLEUM, IBM, MSFT and YAHOO from the time period 2004 – 2009 are available. Four stock indices DJIA, NASDAQ, DAX and SP500 are also included. Other economic data such as consumer price index CPI, producer price index PPI, interest rate and unemployment rate are included.

4.1 Chart Types

The prototype is designed to present stock market data in various ways for visual comparison tasks. Integrated common stock visualizations are line charts, OHLC (open, high, low, close prices) and candlestick charts. These three are used in practice applications in most cases. The line chart is designed to deliver a rich set of interactions for the user. Zooming, panning, selection of the y-axis scale and dynamically changing of stock market data are a few available interactions.

There are two simple variations for the line chart. One uses juxtaposition and the other uses superimposition for the display of multivariate data. Another variation is the transformation of the values by indexing. The resulting data of the indexing method is then displayed by superimposed line charts.

These three variations of the line chart will be tested for effectiveness and efficiency.

4.1.1 Line Chart

Stock market data is in most cases visualized by line charts. Usually the daily closing prices are encoded into the graph.

Figure 34 is a screenshot of the juxtaposed line chart visualization from the prototype. The upper are of the application is showing the line chart while the bottom provides controls for display settings. Other stocks, stock indices and other economic data can be added and removed by clicking the referencing checkboxes. There is also a dropdown menu for the scale which can be set to linear or logarithmic. And there is a menu for the zoom which can be set to a other time range from 5 days to 5 years.

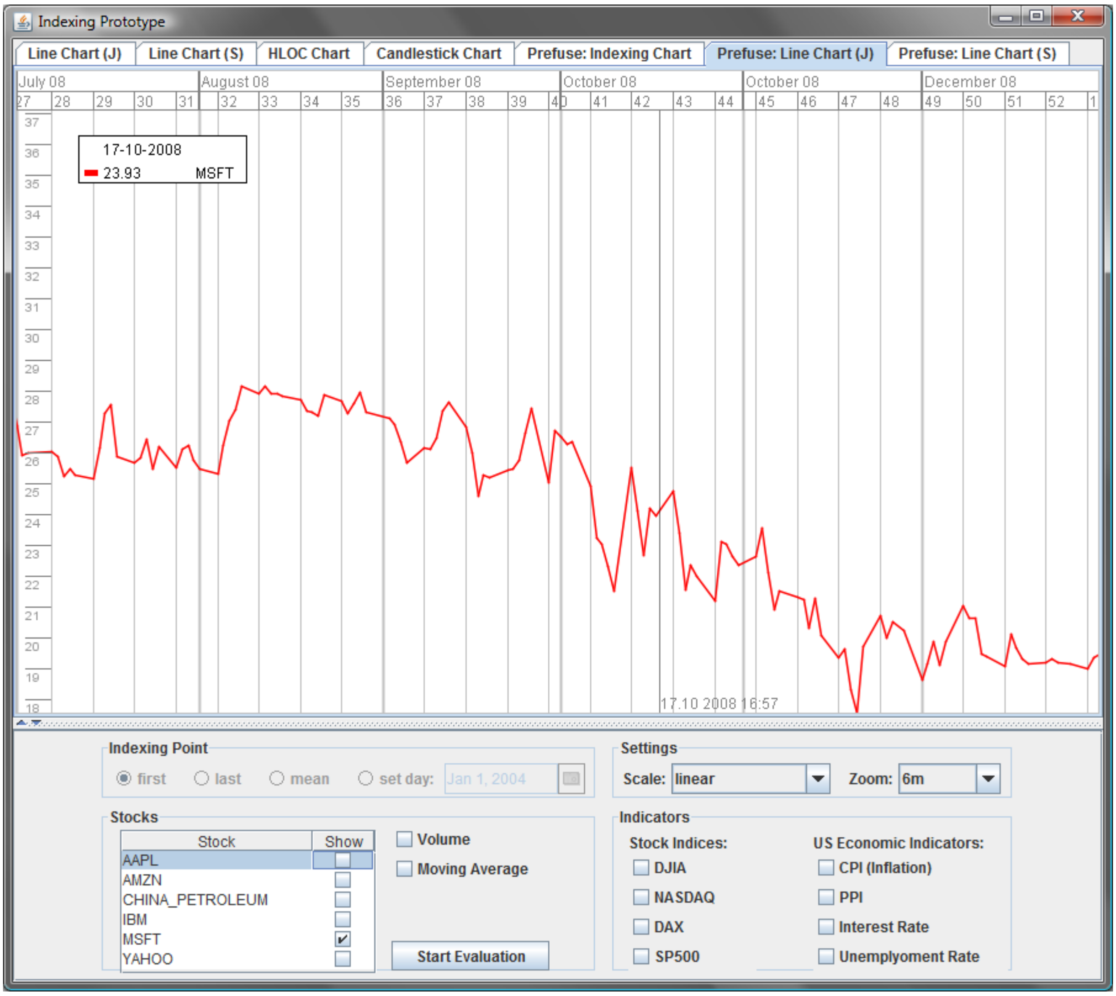


Figure 34: Line chart of MSFT stock

The segmented red line represents the daily closing prices of MSFT stock of six months from 01.07.2008 to 31.12.2008. The vertical price axis is linear scaled. The actual stock prices ranges from 17.0 US-$ to 28.5 US-$. The legend is located at the top left corner. Color and value of the displayed stock is shown.

The next screenshot in Figure 35 shows the same line chart visualization with four juxtaposed stock market series. Each time series is displayed by an own chart. The horizontal time axes are synchronized to ensure valid comparisons between all displayed time series.

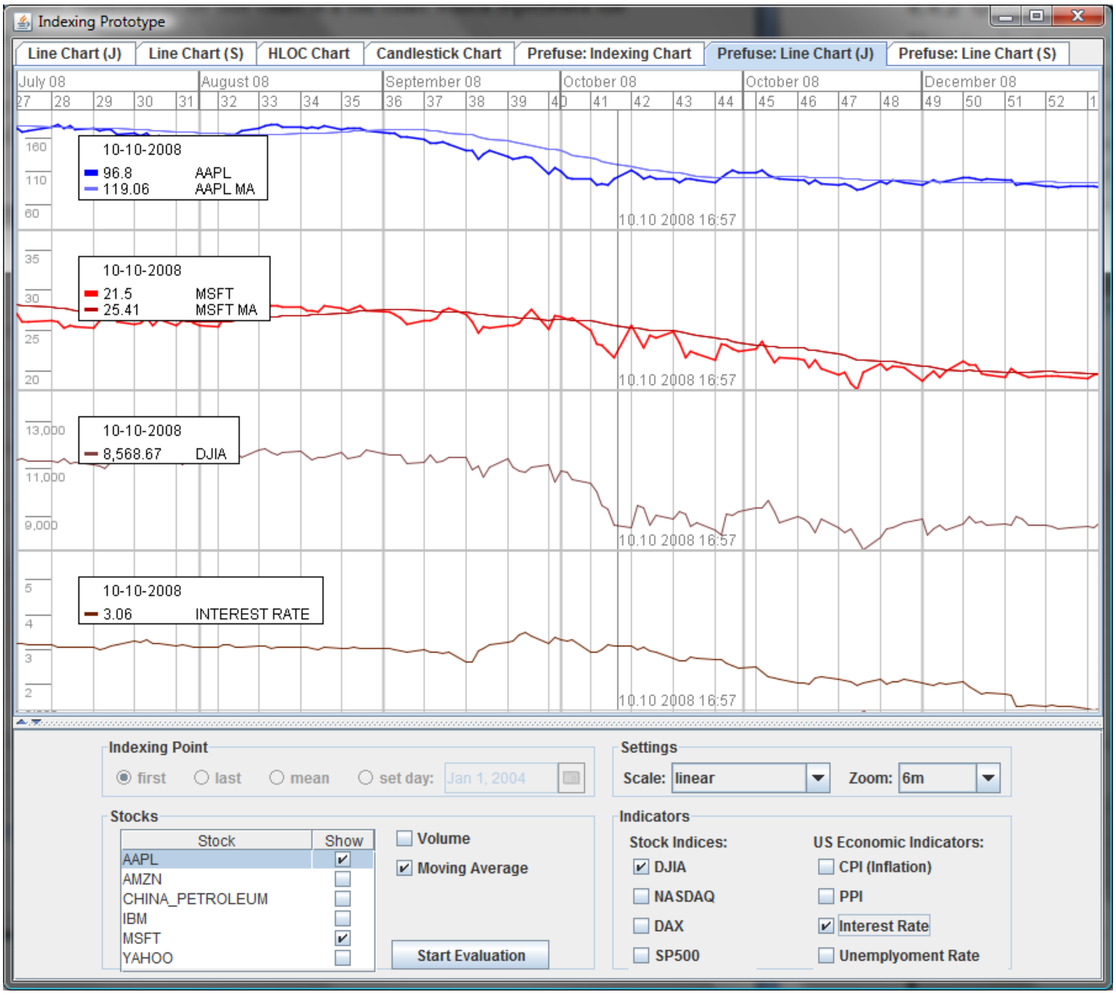


Figure 35: Juxtaposed line charts

The two upper line charts display daily closing prices of MSFT stock in red color and AAPL stock in blue color. The moving average is additionally drawn in both charts by a thinner line.

Beneath the two stocks is the stock index DJIA (Dow Jones Industrial Average). Under the DJIA is the interest rate of the US Federal Reserve drawn.

The other variant of line chart visualizations is to draw all to be displayed stocks in superimposition. This is possible because stock prices share the same unit. But other time series like stock indices and other economic data have to use own line charts. The following screenshot in Figure 36 shows the same time series as before. The difference is that the two stocks share one chart. This reduces the space because only three charts are drawn. As a consequence the three charts have more height available which results in better recognizable visualizations.

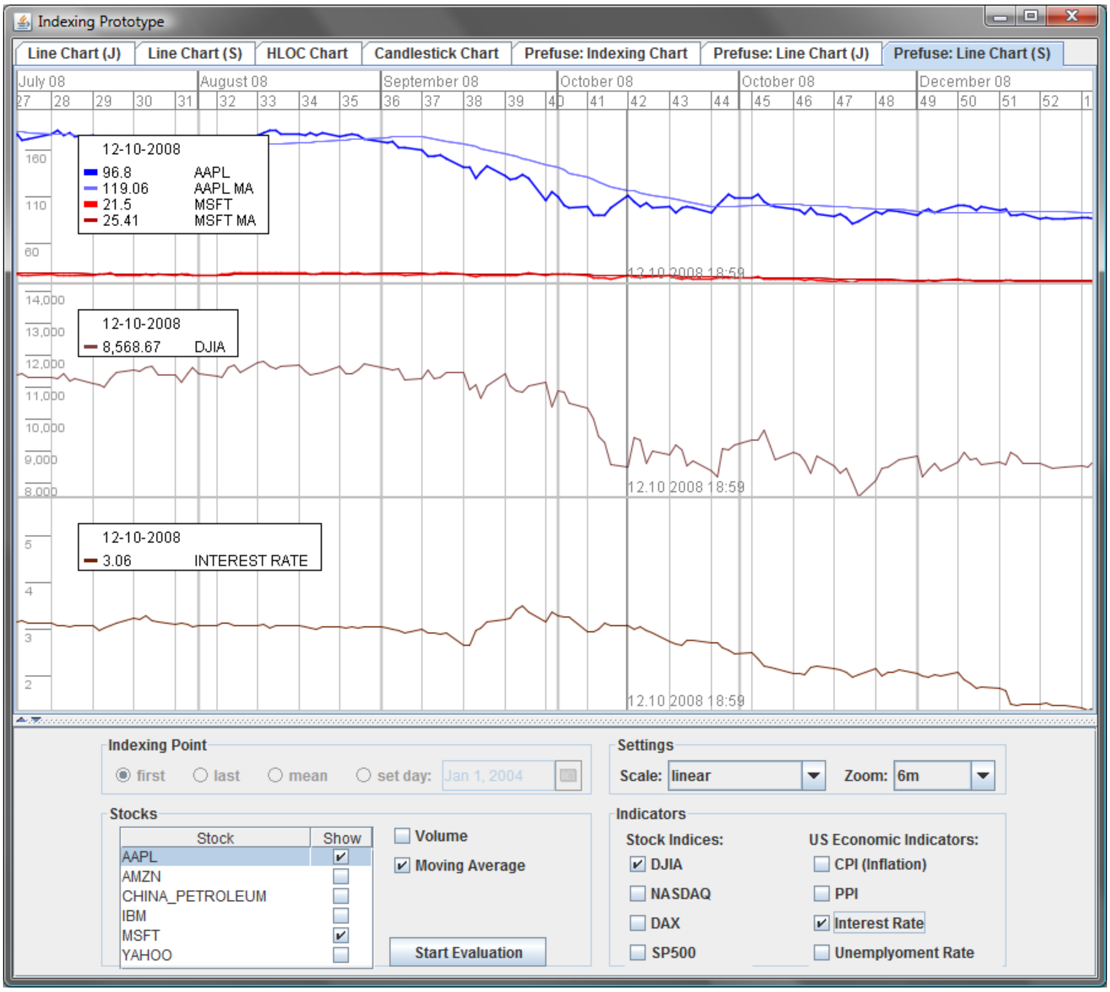


Figure 36: Superimposed line chart of stocks on top

4.1.2 Indexing Chart

The third variation of the line chart is called the indexing chart. The indexing method transforms the values of the y-axis. By using this transformation multivariate data can be displayed as superimposed line charts. Homogenous and heterogeneous data can be easily overlaid over each other.

The next screenshot in Figure 37 shows the four time series mentioned in the previous described variants of the line chart. This time all time series use the same chart, which allows maximum space allocation for one superimpose line chart.

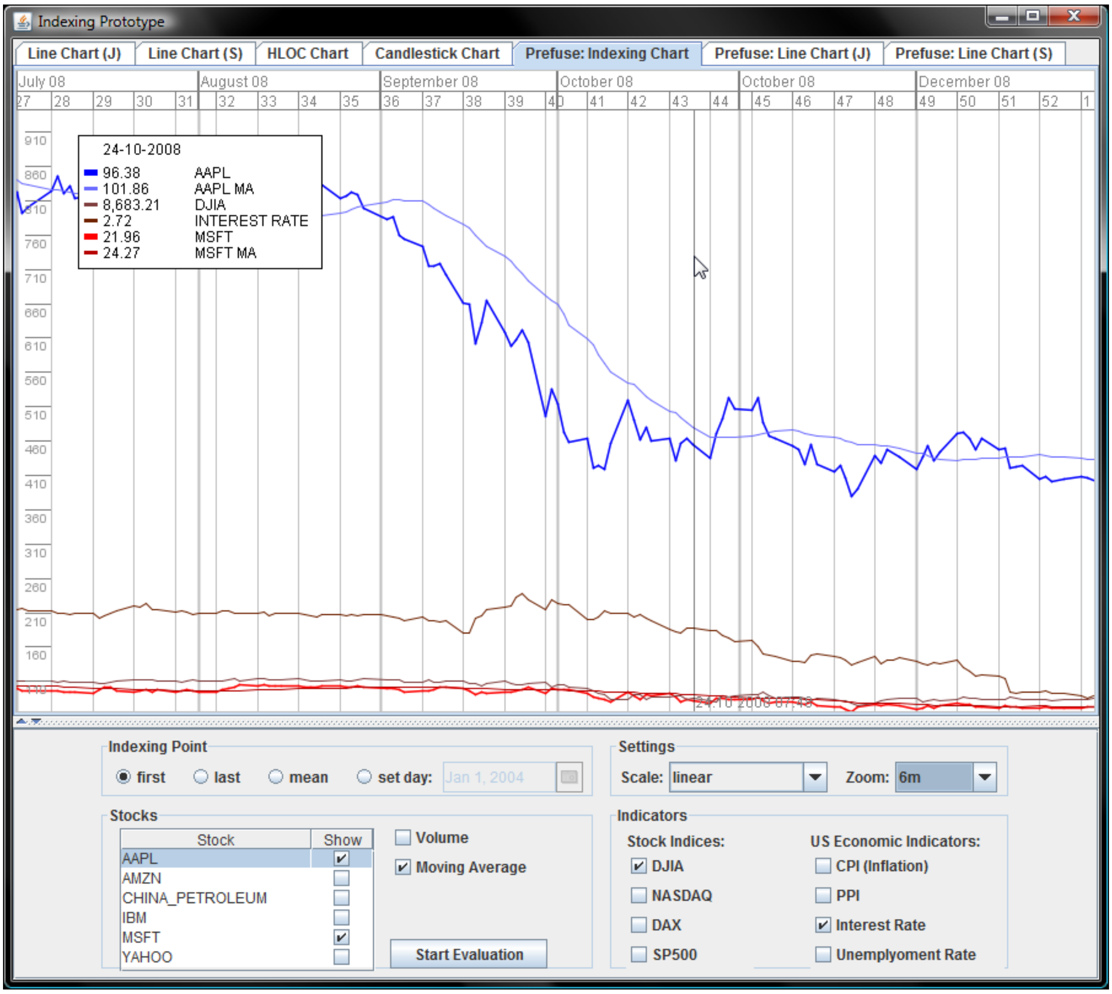


Figure 37: Indexing Chart with two stocks, a stock index and interest rate

The time series are distinguished by color and by line thickness. The legend box at top left displays values for the current mouse location on the chart. Absolute value and a relative percent change are visible. The relative change is based on the indexing point. The indexing point can be easily adjusted to the first or last date. Other possible choices are the mean value or a user-selectable date. The user can set the indexing point in the region left under the line chart.

4.1.3 OHLC Chart

This visualization provides daily opening, highest, lowest and closing prices. The information for each day is encoded by a glyph. Figure 38 illustrates the encoding of the prices.

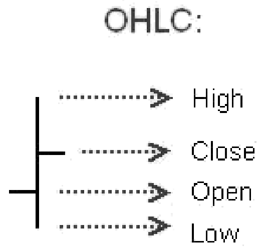


Figure 38: OHLC glyph [Ma, 2009]

The OHLC chart is not well suited for superimposing multiple stocks. Multiple stocks can be visualized by using juxtaposition only.

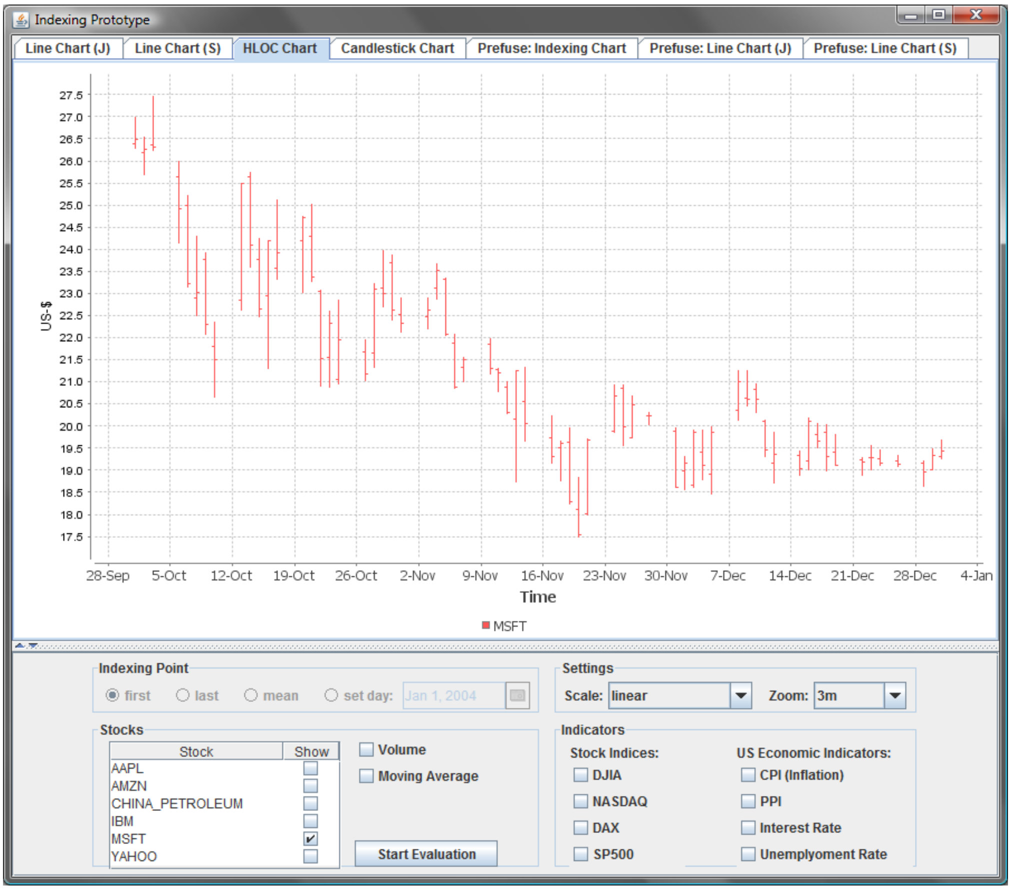


Figure 39: OHLC chart

4.1.4 Candlestick Chart

This visualization is similar to the OHLC chart. Two colors are used to distinguish positive and negative price trends. The colored glyph is a simple enhancement over the OHLC glyph but nonetheless very helpful. The user can faster estimate the development of the data through looking at the colors of the glaphs.

Figure 40 illustrates how highest, lowest, opening and closing stock price are encoded through the candlestick glyph. Red colored figures represent negative daily price trends. Green colored figures stand for positive daily price trends.

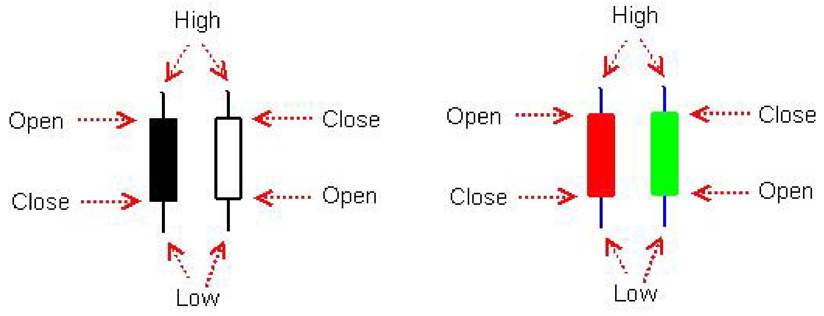


Figure 40: Candlestick glyph [Ma, 2009]

Figure 41: Candlestick chart

4.2 Interactions

The prototype offers an assortment of helpful interaction operations. Many of the implemented interactions are based on observations from other (web 2.0) applications. The successful integration of useful interactions is also of importance for the long time motivation. Otherwise users will switch to applications, where tasks are easier and more intuitive to execute. The major important features will be described in the following text.

The screenshot in Figure 42 illustrates some of the available interactions. Some interactions are missing in the figure because they cannot be illustrated. Panning relies on drag and drop and zooming requires the mouse wheel. Both interactions are not visible in the screenshot.

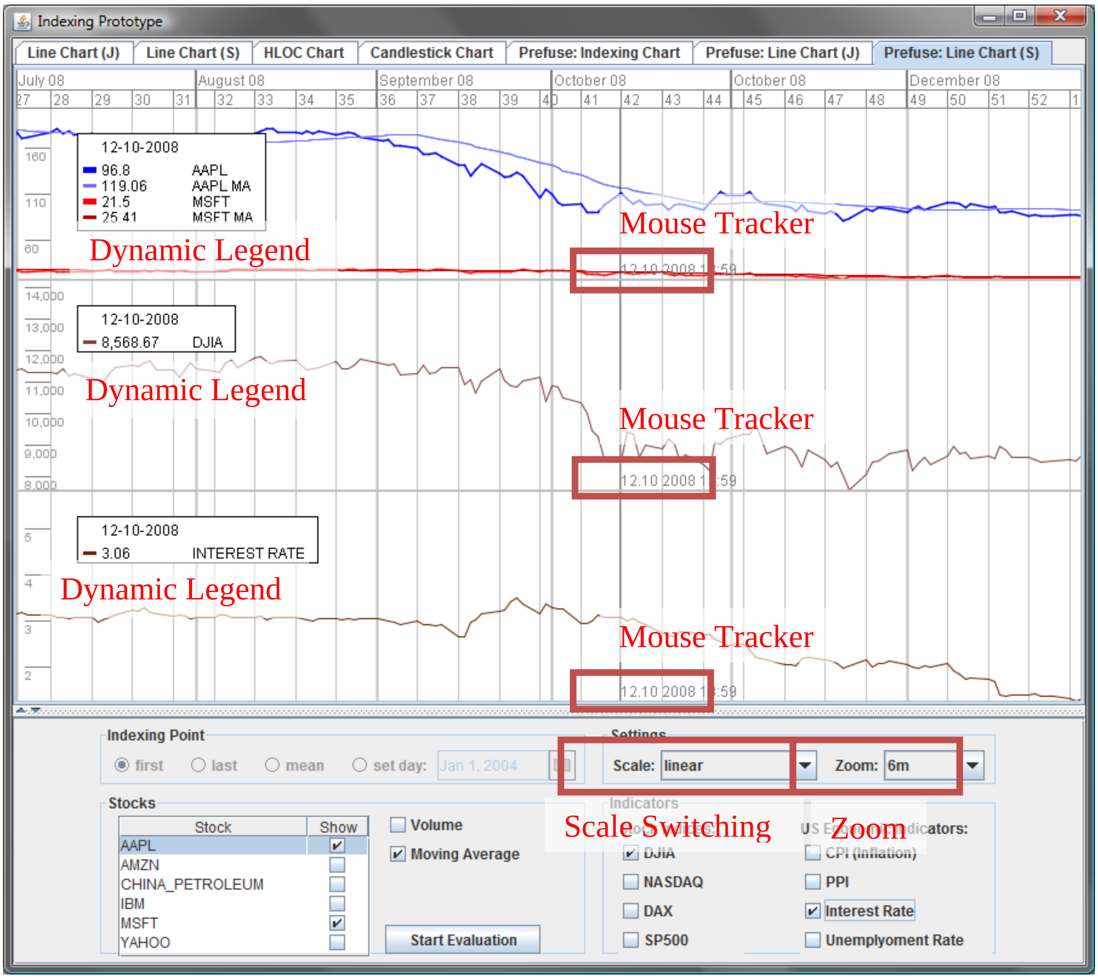


Figure 42: Position of a few available interactions

4.2.1 Zooming and Linking

Line, OHLC and candlestick charts support zooming and linking of juxtaposed charts for visual comparisons between multiple time series. Available zoom ranges are five years, two years, one year, six months, three months, one month and two weeks. A more precise zooming function is available by using the mouse wheel.

4.2.2 Panning

Panning is the ability the move the visualized data to the left or to the right similar to a picture or text within a scroll pane. This is useful for small screen respectively window sizes where the displayable size is limited.

4.2.3 Mousetracker

The Mousetracker allows the user to gather information of the displayed time series according to the horizontal mouse position. In other words, the Mousetracker follows the movement of the mouse cursor. A thin vertical line represents the current position and the referenced point in time.

The line shows the corresponding date for the horizontal position of the mouse cursor. This makes visual comparison tasks easier to the user. Any given day is easily located by using this feature. Dates of particular points in the chart can easily be identified by the user.

4.2.4 Dynamic Legend

The legend is visible in the top left corner of each line chart. The values for the legend are in accordance to the current mouse position on the horizontal time axis. This enables the user to quickly investigate the value for each displayed time series. To change the current date, the user has only to move the mouse cursor to the desired location.

This feature should improve the overall performance in visual comparison tasks. The user is able to get more precise information by looking up the actual values for a given day.

4.2.5 Scale Switching

The user can select two scales for the vertical axis. The linear scale uses a constant ratio between a dimensional unit of the axis and the required space on the chart.

The logarithmic scale can improve tasks where percent changes have to be compared. The ratio between units of the vertical axis and the space for display are not constant. Slopes of lines represent percent changes between the two data points.

4.3 Architecture and Implementation

The prototype application offers several visualizations for stock market data. Common visualizations like the line chart, OHLC chart and candlestick chart were at first implemented by using the JFreeChart library. One disadvantage of this implementation is the rather static display. Interactions are severely limited and cannot be extended. Out of the previous five described interactions only zooming and switching between scales are possible.

So the three visualization types, which will be tested in the comparative study, were implemented by using the Prefuse Toolkit. This graphical Framework was designed to realize highly interactive visualizations. Many interactions are predefined and can be customized. The framework for creating visualizations is much more flexible and adaptable. On the other hand this framework requires more time to implement. A lot of knowledge about many classes and their interconnection within the framework is needed.

4.3.1 Java

Java (http://java.sun.com/) is a popular object-oriented programming language. It was released in 1995 by Sun Microsystems. The syntax of this programming language is similar to C or C++. In contrast to C and C++ is the Java programming language interpreted at runtime by the Virtual Machine (VM). The reason behind the usage of a VM is the gained independence of the application from a specific hardware or software architecture.

The source code is only compiled to bytecode. The VM converts the bytecode at runtime into machine executable code for the given hardware architecture and operating system.

The Virtual Machine is the reason for Java’s portability. A well-known slogan of Java is: compile once, run anywhere. Windows, Mac and Linux operating systems for 32 and 64 bit are supported. Compiled source code does not need a platform-specific compilation.

Java is a very powerful language with an exhaustive library for many applications. The internet provides a lot of support via internet forums, wikis and other internet pages. This is certainly a result of the huge community behind it.

But the Java language does also have a great variety of different application fields. Supported are of course simple console applications and graphical applications. There are further client-side web applications called Applets. Servlets are server-side web applications and are also the foundation for the more advanced JSP and JSF frameworks.

Java offers several advantages to developers. Some are listed above but there are surely more. The prototype was developed in Java because of many of the above mentioned advantages. Another reason is that many graphical frameworks are available, which lighten the development of stock market visualizations. Two important frameworks for visualization were used for the implementation. Both frameworks are described in the two following sections.

4.3.2 JFreeChart

JFreeChart (http://www.jfree.org/jfreechart/) is a very popular framework for creation of various charts. It offers an extensive base of various chart types and the framework along with the source code is freely available. The framework is used in many other open source and commercial projects used. Some examples are JBoss, JIRA and NetBeans.

The advantage of this framework is the simplification of the creation of charts. A few lines of code are usually enough to create a new chart. The drawback of this is the difficulty to create individual visualizations.

Another disadvantage is the absence of advanced interactions. It is possible to zoom within the chart. But more advanced features like a mouse tracker or panning are missing.

Nevertheless it provides a lot of functionalities to customize the available charts. Axis, Gridlines, Legend, Fonts, Colors are freely customizable. It should satisfy most common needs.

Some of the more important chart types in this framework are:

- X-Y charts (line, spline and scatter charts)
- Pie charts
- Gantt charts
- Bar charts

4.3.3 Prefuse

Prefuse (http://prefuse.org/) is a toolkit for the creation of interactive and highly customizable visualizations for the JAVA programming language. The toolkit provides a polylithic design, which lets the developer more freedom to use various ways to write his / her code. The customization of the visualization is practically limitless.

In contrast to JFreeChart there are no visualizations ready for use in the prefuse toolkit. The visualization has to be built from multiple blocks.

The design model of the prefuse toolkit is strongly influenced by the two information visualization reference models by [Card et al., 1999] and [Chi, 2000].

A reference model is similar to a software architecture pattern. It is successfully describing the whole process from data acquisition to the final presentation of the data. An overview of the process is depicted in Figure 43.

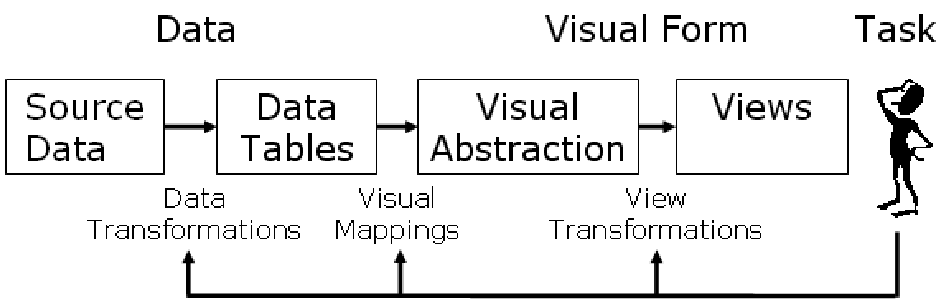


Figure 43: Information visualization reference model from [Heer, 2009] based on [Card et al., 1999] and [Chi, 2000]

The source data on the left is the input for the visualization. In the next step the data will then be transformed into data tables. Next the elements of the data tables will be mapped to specific visual abstractions like color, shape, size and so on. The last step is the view transformation of the visual abstractions to views. Interactions like zooming and panning will be defined in this step. The views are the end result of this process and will be displayed on the screen.

A more detailed explanation of the prefuse toolkit is shown in Figure 44. The most important packages are linked to the related steps of the previous described process.

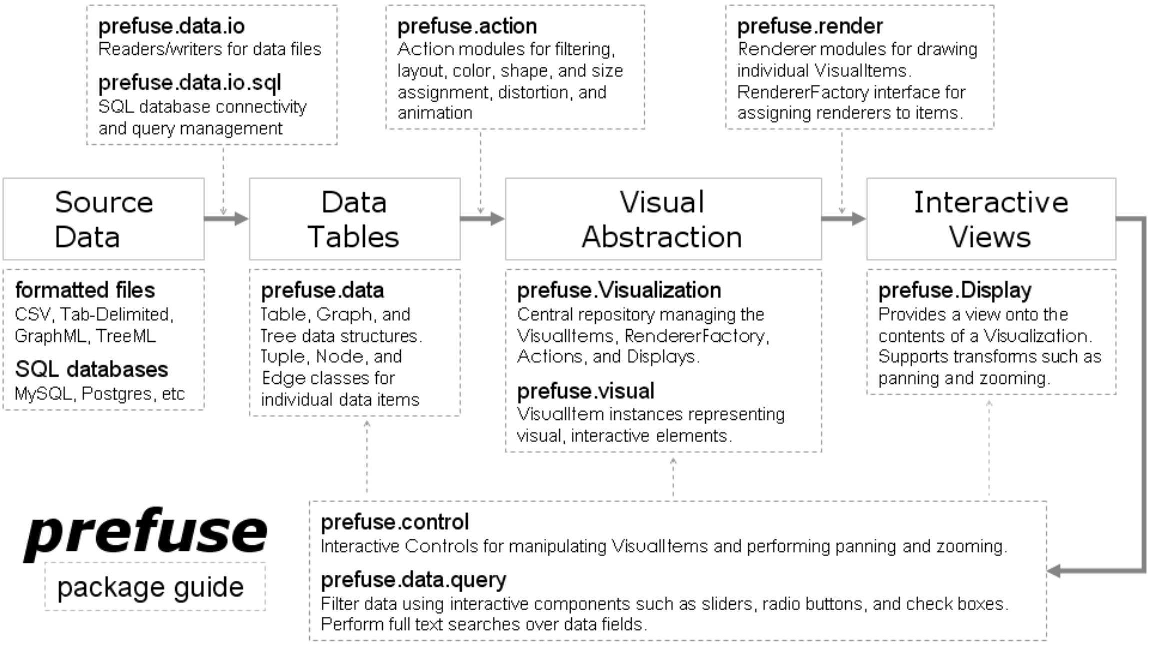


Figure 44: Relation of different prefuse packages and classes to the infovis reference model [Heer, 2009]

Source data for the visualization can be extracted from various sources. Support for CSV, GraphML and TreeML formatted files and for SQL databases is included.

4.3.4 Package Structure

The prototype derives a lot of functionality from the JFreeChart framework and the Prefuse framework. A great amount of implementation work could be saved through using these two external frameworks as a base. However a total amount of 52 source classes, divided among 17 packages, were implemented during the development phase.

The diagram in Figure 45 illustrates a simplified overview of the main package called stockvis. The class Start is the entry point of the application. The class checks before the start of the application if the amount of memory is sufficient. If the available memory is less than 200 MB, the application asks the user if the default memory can be increased.

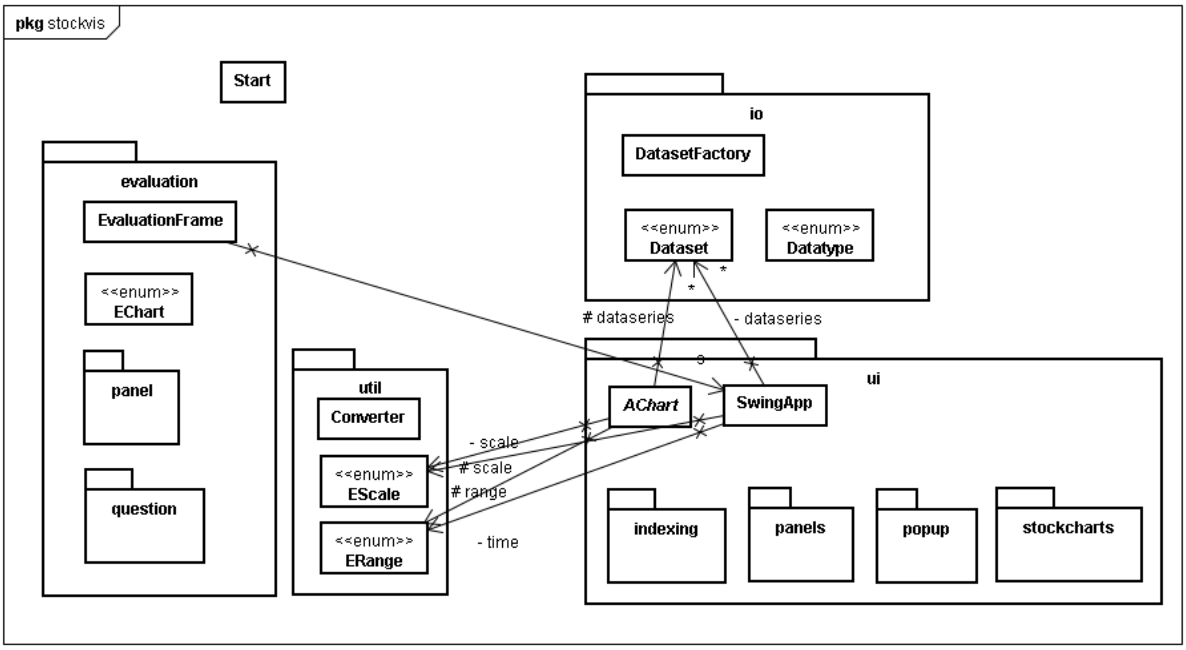


Figure 45: Simplified package diagram

The class SwingApp in the Sub-Package ui will be instantiated as next. This class initializes all required operations to present the application to the user. The package ui is grouping all classes of the user interface into one package.

The package io is responsible for gathering and processing the stock market data. Enumeration classes Dataset and Datatype distinguish between data sets. The DatasetFactory class delivers readily processed stock market data.

The package util provides a class Converter which is used for conversions between datasets. The two enumeration classes EScale and ERange are needed to transfer selection of axis scale and time range between source code classes.

The package evaluation provides classes for the evaluation mode of the prototype. The class EvaluationFrame is the main class which handles all input of the user and visualizes output like questions and takes control of the application during the evaluation process.

The diagram in Figure 46 shows the content of the sub package indexing. The purpose of the package is to provide all needed code to present line chart visualizations which use the indexing method.

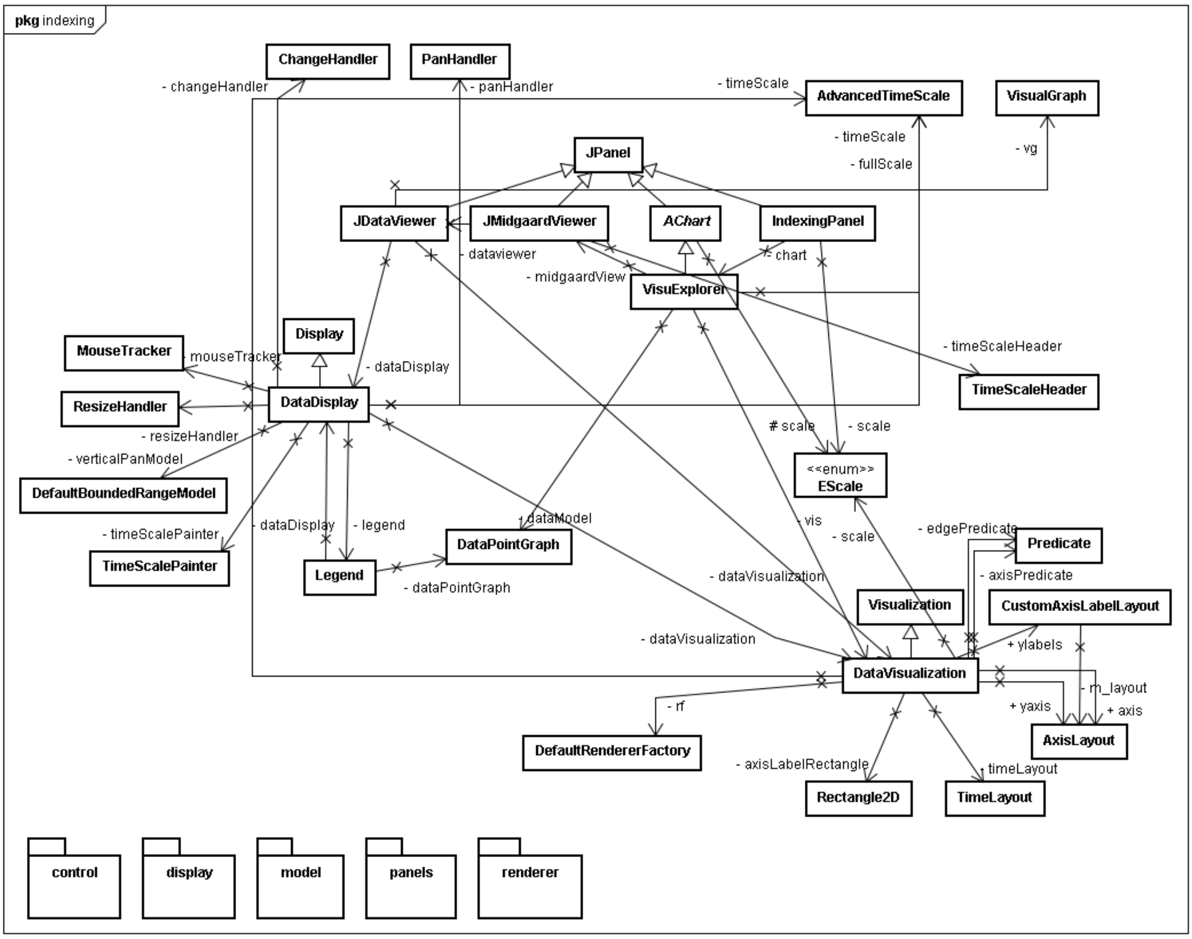


Figure 46: Package Indexing

The indexing method is an important part of the present master thesis and the prototype. Therefore the content of this package will be explained in the following paragraphs.

The class IndexingPanel encapsulates all functionality of the indexing method. It uses the class VisuExplorer. Each instance of the VisuExplorer class represents one line chart. Juxtaposition of multiple line charts is executed by using multiple VisuExplorer objects.

JMidgaardViewer and JDataViewer are responsible for the configuration of the visualization and processing of user input. Both classes are extending the JPanel class. This class is part of the Java Swing which is a toolkit for graphical user interfaces.

The class DataPointGraph is a customized data table or a data structure for storing values of the time series which are to be displayed.

The classes DataDisplay and DataVisualization are specializations of the equivalent prefuse classes. DataDisplay is responsible for handling user interactions through the display while DataVisualization is focused with the graphical output. Renderer settings and axis labels are configured by this class.

Some parts of the prototype are based on two proceeding works. The work of Peter Weishapl is a software prototype for plan visualizations with interactive features such as providing an overview (navigation by a range slider) and a detail interface (navigation by panning and zooming) and an optional fisheye view. More information about the project as well as the documentation and the executable prototype are available at http://ieg.ifs.tuwien.ac.at/projects/timeviewer-timevis/index.html.

The second work of the student Hoffman is based on Weishapl's prototype. The subject of Hoffman's project is the development of a semantic zoom prototype for time series data. The visualization is depending on the zoom level.

5 Comparative Study

This chapter covers all informational aspects of the comparative study. The goal of the comparative study is to evaluate three different line chart visualizations. Two of them are more traditional line charts, while the third is based on the indexing method.

The evaluation will measure the performance by logging the needed time and accuracy of the results. Both variables will be recorded for each task and visualization. These performance measures will be the foundation for the succeeding evaluation of the comparative study.

An important part of the study is the evaluation of the indexing method compared to the two other more traditional line chart visualizations. Another vital part is to determine the effects of the scale of the y-axis, which can be either linear or logarithmic.

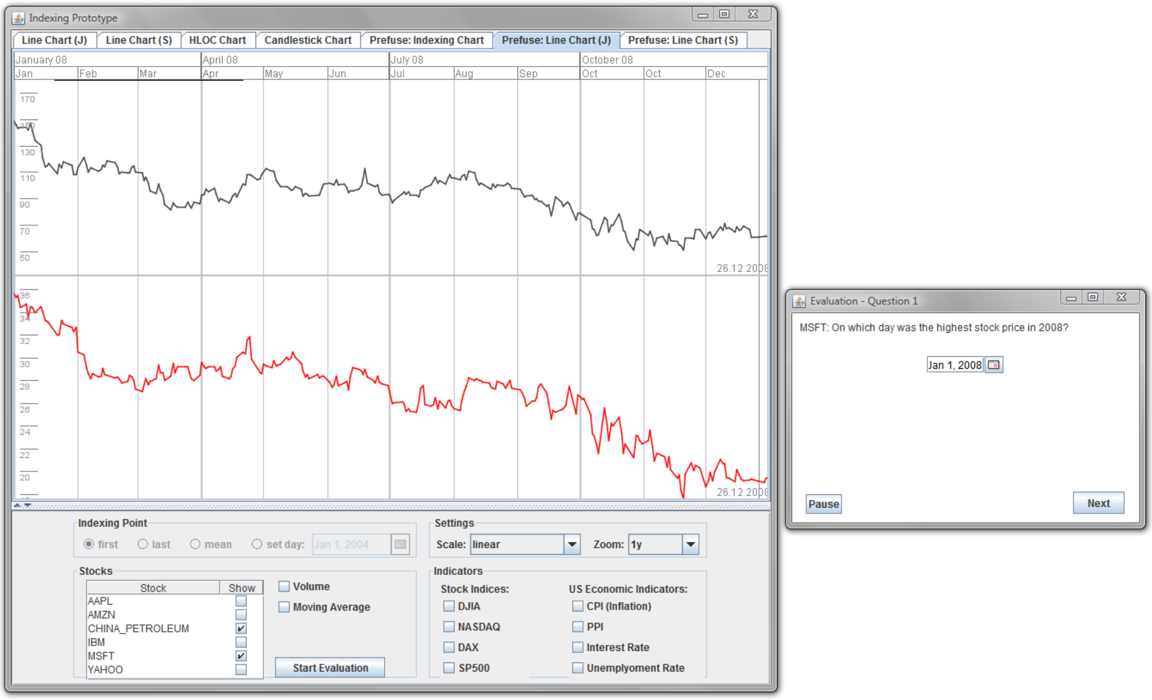


Figure 47: Screenshot of the evaluation mode

The prototype has an evaluation mode built in. This mode allows a simple execution of the usability process. The application displays the task in a separate popup window. After the user answers the question the next task will be shown. All needed information is stored automatically by the application.

5.1 Hypotheses

The goal of the evaluation will be to validate the following five hypotheses, which are derived from the research questions. The research questions and subsequently the five hypotheses focus on the ability to visually compare multivariate time series. The ability of three variations of the line charts will be evaluated.

The five hypotheses can be separated approximately into three groups. The first group (H1 and H2) focuses on questions about the visual comparison of percent changes in line charts. The second group (H3) is concerned with the visual recognition of the development of curves. The third group (H4, H5) contains question which visualization is best suited for a more generic set of tasks for most common tasks.

H1: Log Scale for percent estimation tasks

The logarithmic scale can directly visualize percent changes of the displayed data. It is predicted that estimations of percent changes are more precise and faster when using logarithmic scales compared to estimations of percent changes when using linear scales.

H2: Indexing method for percent estimation tasks

The indexing method transforms absolute values into percent changes based on the indexing point. This method should make visual comparisons of percent changes easier i.e. reduce the needed time to estimate percent changes. It is predicted that the indexing method is more effective for estimation and comparison tasks of percent changes than logarithmic scaled line charts, which display the usual absolute values.

H3: Indexing method for trend comparison

The selection of the indexing point is very useful for comparisons of time series trends. It is predicted that the test persons can make estimations and comparisons of trends for different time series more precise and faster.

H4: Superimposed, logarithmic scaled line charts are better than juxtaposed line charts for visual comparisons

Superimposed, logarithmic scaled line charts can display percent changes more directly. Comparisons by superimposition should be easier than by juxtaposition. It is predicted that comparison of absolute values, comparison of percent changes and comparisons of trends are faster and contain less errors than comparisons with juxtaposed, linear scaled line charts.

H5: Indexing method is overall better for visual comparisons

The indexing method leads to a direct display of percent changes. It is predicted that the indexing method makes comparisons of absolute values, relative values and trends faster and comparison results have higher task correctness rates.

5.2 User Tasks

The selection of the proper user tasks is critical for the relevance and also for the success of the evaluation. Each hypothesis has to be linked with adequate tasks or else the output of the hypothesis becomes useless. Therefore it is most essential to create proper user task which correspond with the hypotheses which have to be tested.

The structure of the tasks for the evaluation is strongly based on the well-thought-out task taxonomy for temporal data from the book “Exploratory Analysis of Spatial and Temporal Data - A Systematic Approach” [Andrienko, 2006].

The task taxonomy is divided into two categories: elementary tasks and synoptic tasks. The following two sections will deliver more details about both task categories.

5.2.1 Elementary Tasks

Elementary tasks set their focus on a single time series. [Andrienko, 2006] defines three elementary task types: Lookup, Comparison and Relation-Seeking.

Elementary Lookup tasks refer to seek a specific value of a single time series. An example would be to a date value for a specific point in time of a stock in a given time interval.

Elementary Comparison tasks refer to tasks which involve a comparison of time and y-axis values. An example would be to guess the difference between two parts of a given stock for a given time interval.

Elementary Relation-Seeking tasks refer to patterns within a single time series. An example would be to find the month(s), which have a higher maximum value than the value of a given date.

5.2.2 Synoptic Tasks

Synoptic tasks are centered on analyzing multiple configurations of characteristics corresponding to subsets of references. [Andrienko, 2006] defines the three following synoptic task types: Pattern Identification, Behavior (Pattern) Comparison and Relation-Seeking.

Synoptic Pattern Identification tasks refer to recognition of particular patterns in the given time series data. An example would be to distinguish if a given month of a time series has a positive or negative trend.

Synoptic Behavior (Pattern) Comparison tasks refer to identifying and comparing patterns of two time series. An example would be to decide which of two stocks has a higher volatility for a given time period.

Synoptic Relation-Seeking tasks refer to spot characteristics of a time series. An example would be to identify the year of a given time series which had the biggest percent increase from start to the end of the year.

5.3 Experiment design

The independent variable of the experiment is the visualization type. Three different visualization types will be compared against each other. The first type is the juxtaposed line chart. The second type is the superimposed line charts with a logarithmic scaled y-axis. The third type is the line chart visualization based on the indexing method. These three visualization types are noted in the following text also as visualization type Linear Scale, Juxtaposition (A), Log Scale, Superimposition (B) and Indexing (C). A full exemplary test (includes all 14 tasks for each visualization type and appropriate screenshots) is available in Appendix B – Comparative Study Tasks. The 14 tasks are listed in Table 6.

The two dependent variables of the usability test are task completion time and task accuracy. The task accuracy will be interpreted as a binary value of true or false. A more detailed analysis of the task correctness for special tasks will be made by using further information, which will be automatically recorded during the test.

The experiment takes a within-subjects approach. This increases the output of the test results, because every test person will be evaluating all three visualization types. Each test person will use the juxtaposed line chart, the superimposed line chart and the indexing chart instead of just one visualization type.

This method implies the use of a Latin square to counterbalance any learning and fatigue effects of the involved test persons. Such effects could negatively influence the test results. As a result the order of the visualization types will be assigned to each test person according to the following Latin square variation in Table 4.

Test Group	Test Persons	Visualization Variation
Group 1	TP1, TP7, TP13, TP19	A - B - C
Group 2	TP2, TP8, TP14, TP20	B - C - A
Group 3	TP3, TP9, TP15, TP21	C - A - B
Group 4	TP4, TP10, TP16, TP22	A - C - B
Group 5	TP5, TP11, TP17, TP23	B - A - C
Group 6	TP6, TP12, TP18, TP24	C - B – A

Table 4: Latin square for every test person (TP) and group

Each test group consists of four test persons. Each test person of the user group has the same sequence order of the used visualization. The different order variations reduce any negative influences.

The test person has to complete 14 tasks for every visualization type. Every task of the 14 tasks is defined for three datasets. The three datasets differ in their choice of stocks and stock index. Table 6 illustrates the datasets and their defined stock market data.

	Stock 1	Stock 2	Stock Index
Dataset 1	AAPL	IBM	NASDAQ
Dataset 2	AMZN	YAHOO	SP500
Dataset 3	MSFT	CHINA PETROLEUM	DJIA

Table 5 : Three datasets for the evaluation

The dataset for each task is randomly assigned during the test process. The evaluation application assigns a specific dataset for each task and visualization type. This is needed to remove any learning effects between visualization types, which would influence test results of later tested visualization types.

The selection of the dataset for every task is random to avoid any effects because of differences between the datasets. If each dataset is linked to a specific visualization type, the test results could also be influenced by the dataset.

The comparison tasks are defined in such a way that three different combinations of stock market data are used. Homogenous data consists of two stocks, heterogeneous data consists of one stock and one stock index and combination uses two stocks and one stock index.

	Task Type	Question
1.	Elementary Lookup Task (homogenous data)	<stock 1>: On which day was the highest stock price in <year>?
2.	Elementary Lookup Task (homogenous data)	<stock 1>: On which day was the lowest stock price in <year>?
3.	Elementary Comparison Task (homogenous data)	Compare the values of <stock 1> and <stock 2> on <date>
4.	Elementary Comparison Task (homogenous data)	Please quantify the amount of price change for the given time periods in dollars for <stock 1> and <stock 2>.
5.	Elementary Comparison Task (homogenous data)	Please quantify the amount of price change for the given time periods in percent for <stock 1> and <stock 2>.
6.	Elementary Comparison Task (heterogeneous data)	Compare the values of <stock 1> and the <stock index> index on <date>
7.	Elementary Comparison Task (combination data)	<stock index>: How much percent did the values change in <year>?
8.	Elementary Relation-seeking	<stock 1>: Which of the following months in

	Task (homogenous data)	<year> have a higher value than the value on <date>?
9.	Synoptic Pattern Identification Task (homogenous data)	<stock 1>: Which of the following months in <year> have a positive trend?
10.	Synoptic Behavior Comparison Task (homogenous data)	Which stock has a bigger percent increase from the beginning of <month> to the end of <month>?
11.	Synoptic Behavior Comparison Task (homogenous data)	Which stock has a lower percent loss in <year>?
12.	Synoptic Behavior Comparison Task (heterogeneous data)	In which months is the percent increase of <stock 1> greater than <stock index>?
13.	Synoptic Behavior Comparison Task (combination data)	Which stock or index has the highest volatility (relative variations) in September <year>?
14.	Synoptic Relation-seeking Task (homogenous data)	In which year had <stock 1> the highest percent increase from beginning to the end of the year?

Table 6: Task list

Table 7 displays a matrix which shows user tasks vertically and the five hypotheses horizontally. Thus each column represents a hypothesis and each row represents a task. Each cell defines which visualization types will be considered for the evaluation of the hypothesis. An empty cell means that the corresponding task is not needed for the particular hypothesis.

Task	H1	H2	H3	H4	H5
1				A, B	A, B, C
2				A, B	A, B, C
3				A, B	A, B, C
4				A, B	A, B, C
5	A, B	B, C	A, B, C	A, B	A, B, C
6				A, B	A, B, C
7	A, B	B, C	A, B, C	A, B	A, B, C
8			A, B, C	A, B	A, B, C
9			A, B, C	A, B	A, B, C
10	A, B	B, C	A, B, C	A, B	A, B, C
11	A, B	B, C	A, B, C	A, B	A, B, C

12	A, B	B, C	A, B, C	A, B	A, B, C
13	A, B	B, C	A, B, C	A, B	A, B, C
14	A, B	B, C	A, B, C	A, B	A, B, C

Table 7: Task - Hypothesis Matrix, Linear Scale & Juxtaposition (A), Log Scale & Superimposition (B), Indexing (C)

In order to keep track of the test results the evaluation prototype will store all needed information of a usability test in a CSV file. This file is best for storing tabular-structured values and can easily be imported by many applications. The following variables will be recorded for each task of the usability test:

- Task number: the order of the particular task
- Visualization: the used visualization for the task
- Task completion time: How long the test person used to answer the task
- Task correctness: stores whether the task was successfully answered or not
- Full task description: contains the whole task description
- Valid answers: a list of all valid answers for the task
- Given answers: a list of the answers given by the test person for the task

The test output for the three visualization types will be analyzed by using statistical tests called ANOVA (Analysis of Variance). This test determines if the performance measures are significantly different between the defined groups.

One-factorial ANOVA (Analysis of Variance) tests will be used for the evaluation of task completion times. The test results will be grouped by the displayed visualization type. Two-factorial ANOVA tests will be used for the evaluation of task correctness values. The test output is grouped by two factors, which are visualization type and task.

The user preferences for the visualization for visual comparison tasks from the post-test questionnaires will be analyzed by another statistical test. The chi-square test will be used to determine if the user preferences are significantly different from an equal distribution.

The appendix provides supplementary material to reproduce the usability test. All 42 tasks along with a screenshot of the data visualization can be found in Appendix B – Comparative Study Tasks. Because of the limitation to display all possible data and visualization combinations every visualization type is linked to a certain dataset in the appendix.

5.4 Materials

The environment for the tests is important because it could disturb the test persons and thus have a negative effect on the test results. So the tests were conducted in a quiet environment with a relaxed and friendly atmosphere. Occurrences of external influences which would disturb the test person were minimized whenever possible.

The test application was executed for all tests on the same laptop with the same computer mouse. The laptop is fast enough to run the test application without any memory or processor problems. The hardware specifications of the used laptop are a 2 GHz Dual Core processor with 2 GB RAM and Windows XP SP3 as operating system. The graphics were displayed on a 15.4 inch LCD monitor with 1280 x 800 pixels resolution. A standard symmetrical shaped Logitech optical mouse was used as input device.

Java Runtime version 1.6.0 was used to execute the evaluation application. All other programs were closed during the evaluation process. Otherwise some program might be interfering with the test application.

5.5 Procedure

Each test procedure involves a test supervisor and a test person. The duties of the test supervisor include setting up the test environment and to ensure that the test process runs accordingly to the following procedure.

Activity	**Time**
Greeting	2 min
Introduction and Orientation	5 min
Filling out the Pre-Test Questionnaire	10 min
Demonstration of the Introduction examples with the prototype	10 min
Execution of Usability Test	30 min
Filling out the Post-Test Questionnaire	5 min
Debriefing and Goodbye	3 min
total	**65 min**

Table 8: Activities of the test procedure

The procedure starts with the greeting of the test candidate. The test person will be asked if the environment is comfortable enough and if something in the surrounding is disturbing the test person.

If everything is okay, a short introduction of the subject comparisons of multivariate time series will be given by the test supervisor. Basic concepts of multivariate data in finance applications will help the test person to gain an understanding of the importance of the subject. The orientation part will describe the succeeding sequence of events of the test. The test person will be informed so that no surprises will arise, which would make the test person uncomfortable.

Afterwards the pre-test questionnaire will be given to the test person. The questions help to gather basic personal information and previous experience with data analysis, stock analysis and about most common stock visualizations. This information can be helpful later when the test results will be analyzed. The test results can then be associated to a certain group of persons.

Then the test supervisor will demonstrate the usage of the prototype application. A short introduction of the application includes three example tasks. The three tasks are similar to the real task in the usability test. The test supervisor will explain the intention of the task and how the task can be solved. The test person is encouraged to ask any questions when something is not clear.

After the introduction the usability test will be started. The test consists of 42 tasks for three visualizations. The test person can stop the test by pressing the pause button.

When the usability test is finished the post-test questionnaire will be given to the test person. The test person has to fill out which visualization would the test person prefer for visual comparison tasks.

At last the debriefing is intended to answer any further questions of the test person about the visualizations. It will also be explained how the data will be used in further steps. This is the last activity of the procedure.

It is estimated that the procedure will endure about 65 minutes. However the required time is depending on how fast the test person can give an answer to the tasks of the usability test. It is expected that the usability part will fluctuate at most between test persons.

5.6 Pilot test

Before the actual start of the usability tests a pilot test was performed. The aim of the pilot test is to find possible problems in the test design.

The test process of the pilot test did correspond in overall to the planned process. The estimation of the required time for the test process of 65 minutes was confirmed by the pilot test.

The pilot test also showed that the set of 42 tasks is demanding a lot of concentration from the user. The required effort is relatively high but should nevertheless be reachable by most test persons.

After every block of 14 questions for one of the three visualizations a short break was made to ensure that the test person could remain concentrated for the remaining tasks.

The pilot test was fully successful, so the results were added to the final test results. The test procedure was not changed from the original plan, because no problems did occur during the execution of the pilot test.

5.7 Participants

Twenty-four individuals could participate in the comparative study. The age of the test persons is within the range of 20 to 30 years. Half of the test subjects are male and the other half are female. The education of all test persons is at least a Matura which is similar to a high school graduation and allows the owner to enroll at a university. Out of 24 test subjects have thirteen persons a Bachelor's Degree and four people a Master's Degree. At the time of the test are 19 persons studying at a university.

Information about personal data, previous education and job description of the test persons are listed in Appendix A – Pre-Test Results.

Figure 48 shows a histogram of the jobs for the participants. The majority of the test persons are students (18). The remaining 6 persons are financial advisor, research assistant, software developer, electronic engineer, accountant and a sales consultant.

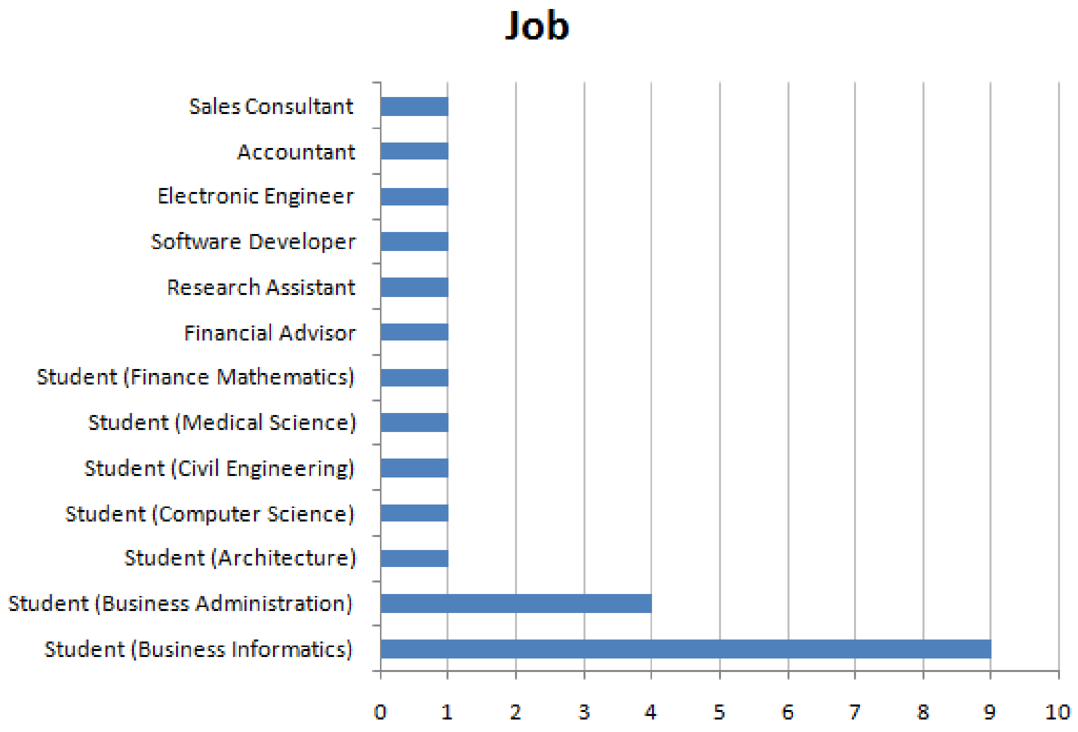


Figure 48: Job distribution of the 24 participants

Figure 49 shows the distribution of education among the test persons. 7 have a Matura, 3 have a master's degree and the majority of 13 participants have a bachelor's degree.

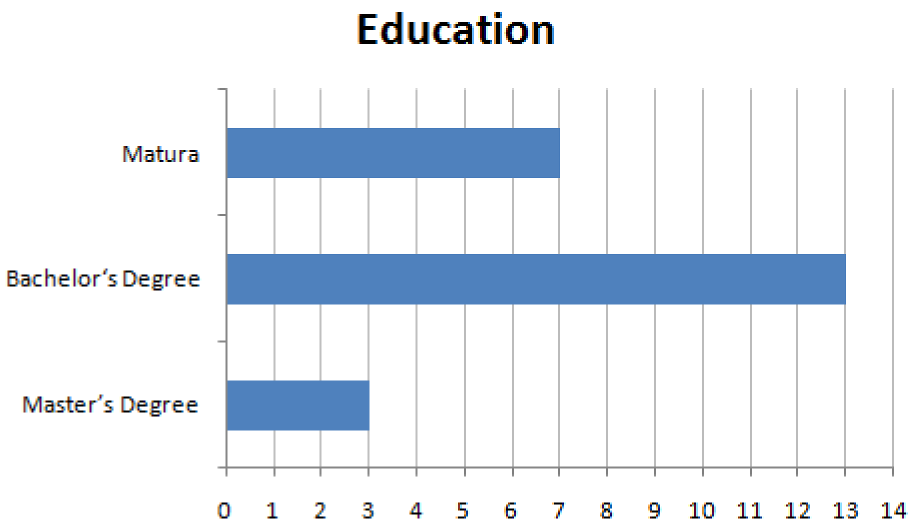


Figure 49: Last education of the 24 participants

The distribution of the participant's gender is equal. Half of the participants are female and the other participants are male.

One precondition for all participants was that they are used to work with a computer and a computer mouse. The ability to use the computer mouse as input device is essential for obtaining valid test results. All test persons were carefully selected to match this precondition.

Table 17 shows the self assigned experience levels of every test person. The experience is divided into six application fields. The application fields are experience about data analysis in general, about stocks, about line charts, about OHLC charts, about candlestick charts and about other chart types. The user had to classify their experience by selecting one of four predefined values. The user had to choose between inexperienced, average, good and very good.

The experience level of the participants for data analysis, stocks, line charts, OHLC charts and candlestick charts is listed in Appendix A – Pre-Test Results.

Figure 50 shows a histogram illustrating the distribution of the level of experience for the 24 test persons. The most selected experience level is average with around 53 %. Next follows the level inexperienced with 28,5 % and experience level good with 18 %.

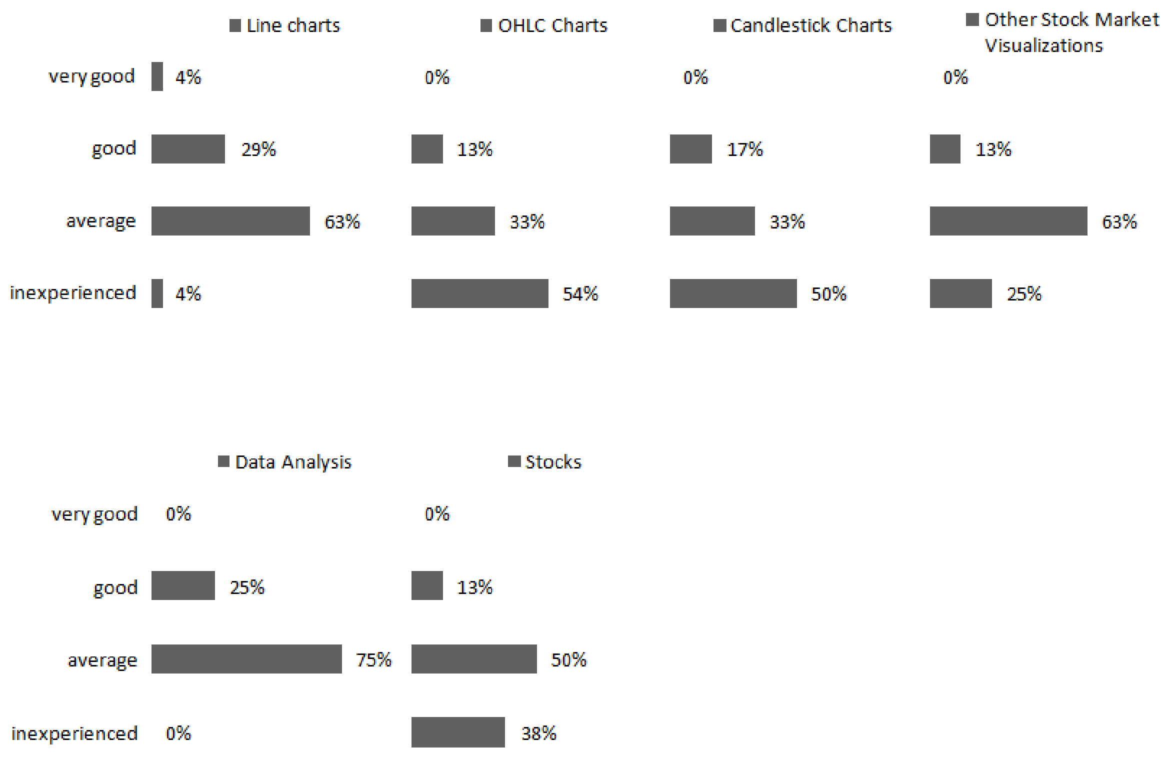


Figure 50: Distribution of the user experience

The test participants described themselves as more than average experienced with data analysis and line charts. The users had least experience with OHLC charts and candlestick charts. Many of the test subjects are also inexperienced with stocks.

6 Results

This chapter presents the test results of the usability test. 24 test persons participated in the usability test. Each test participant had to answer 42 tasks. Various input methods were used to answer the tasks. Some tasks required to select a specific date, other required to select the correct stock from a list of possible stocks. Other tasks require selecting a certain range as answer.

This master thesis wants to verify five hypotheses. The outputs of the usability test are measures of the two dependent variables task completion time and task correctness rate. A summary of the hypotheses and the results will be presented in the first subsection.

The following subsections will illustrate the results of the two dependant variables in greater detail. At last the user preference choices will be presented.

6.1 Hypotheses

The table below summarizes the statistical test results for the five hypotheses according to the values of the two dependent variables. P-values indicate the probability of the null hypothesis. The null hypothesis assumes equal distributed variables. A low p-value indicates that the dependent variable differs from the assumed distribution.

Table 9 shows the statistical results for the five hypotheses based on the task completion time. All hypotheses, except the first, do not have a significant difference in the time for the execution of the tasks. The significance level is < 0,05 for all hypotheses.

The result of the first hypothesis Log Scale - percent estimation (H1) states that the task completion time between linear scaled line charts and logarithmic scaled line charts is not equal. This means that the use of logarithmic scaled line charts results in faster task execution for tasks of visual comparisons of percentage values.

Interestingly, this result does not hold for a test consisting of a wider range of tasks between the two visualization types. As the result of hypothesis Log Scale, Superimposition (H4) shows, there is no significant difference in task completion time between the two visualization types when a set of more general tasks is used.

It seems that the task completion time for general tasks is not dependent on one of the three tested visualization types. Our test subjects seem to be using a certain amount of time for each answer for the tasks. The task completion time is rather constant, regardless whether the subjects are sure about their answer or not.

	H1 (Log Scale - percent estimation)	H2 (Indexing - percent estimation)	H3 (Indexing - trend comparison)	H4 (Log Scale, Superimposition)	H5 (Indexing)
Visualization type	A, B	B, C	A, B, C	A, B	A, B, C
Tasks	5, 7, 10 - 14	5, 7, 10 - 14	5, 7 - 14	1 - 14	1 - 14
Mean / Std. dev.	A: 42.920 / 8.057 B: 36.699 / 8.558	B: 36.699 / 8.558 C: 35.842 / 7.079	A: 40.755 / 7.632 B: 36.333 / 8.840 C: 38.026 / 7.501	A: 35.892 / 6.943 B: 33.139 / 7.980	A: 35.892 / 6.943 B: 33.139 / 7.980 C: 36.784 / 8.482
P-value	0.013	0.707	0.163	0.209	0.25
Time Distribution	Significantly different	Not significantly different	Not significantly different	Not significantly different	Not significantly different

Table 9: ANOVA test results of the five hypotheses for task completion time; Linear Scale & Juxtaposition (A), Log Scale & Superimposition (B), Indexing (C)

The test results for the five hypotheses based on the task correctness rate are quite different from the test results of the task completion time.

The first hypothesis Log Scale - percent estimation (H1) and the forth hypothesis Log Scale, Superimposition (H4) do not show any significant difference between the two visualization types linear scaled juxtaposed line chart and logarithmic scaled superimposed line charts.

The second hypothesis Indexing - percent estimation (H2), the third hypothesis Indexing - trend comparison (H3) and the fifth hypothesis Indexing (H5) give evidence for a significant difference for task correctness rates between line charts which use Indexing (C) and the other two visualization types juxtaposed, linear scaled line chart (A) and Log Scale, Superimposition (B).

The task completion time is in general equal between all visualization types. However, logarithmic scaled line charts (B) are significantly faster than linear scaled line charts (A) for comparison tasks of percent changes. The test results of the task correctness rate indicate a significant difference between the three visualization types (Linear Scale & Juxtaposition (A), Log Scale & Superimposition (B), Indexing (C)). The Indexing method produces significantly less estimation errors.

	H1 (Log Scale - percent estimation)	H2 (Indexing - percent estimation)	H3 (Indexing - trend comparison)	H4 (Log Scale, Superimposition)	H5 (Indexing)
Visualization type	A, B	B, C	A, B, C	A, B	A, B, C
Tasks	5, 7, 10 - 14	5, 7, 10 - 14	5, 7 - 14	1 - 14	1 - 14
Mean / Std. dev.	A: 0.446 / 0.212 B: 0.512 / 0.290	B: 0.512 / 0.290 C: 0,857 / 0.120	A: 0.495 / 0.243 B: 0.560 / 0.293 C: 0.829 / 0.129	A: 0.637 / 0.278 B: 0.690 / 0.296	A: 0.637 / 0.278 B: 0.690 / 0.296 C: 0.842 / 0.130
P-value	0.446	< 0.01	0.001	0.202	< 0.01
Correctness Distribution	Not significantly different	Significantly different	Significantly different	Not significantly different	Significantly different

Table 10: ANOVA test results of the five hypotheses for task correctness rate; Linear Scale & Juxtaposition (A), Log Scale & Superimposition (B), Indexing (C)

6.2 Task completion time

The chart in Figure 51 illustrates the task completion time for each of the three visualization types in distinct colors. The times are grouped by tasks. The thin black bars indicate the standard deviation of individual task completion times from the test subjects.

Task completion times of task 12 are standing out from completion times of the other tasks. The goal of the task is to visually compare the percentage increase between two time series each month of one year. The user has to identify which time series has the greater monthly percentage increase. This task is therefore consisting of twelve sub tasks. This could explain a part of the higher task completion times. Although task 8 and 9 also consist of monthly comparisons, they are less complex and involve only one time series.

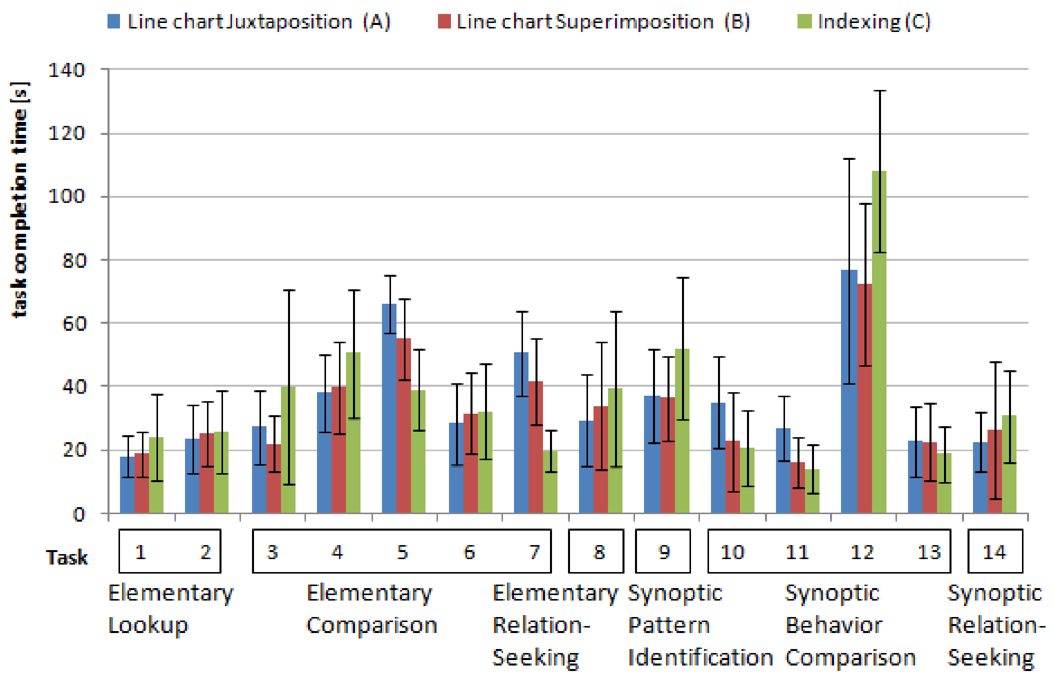


Figure 51: Task completion time for each visualization type grouped by tasks

6.3 Task correctness rate

The chart in Figure 52 depicts task correctness rates for the three visualization types in distinct colors. The results are grouped by tasks for an easier comparison.

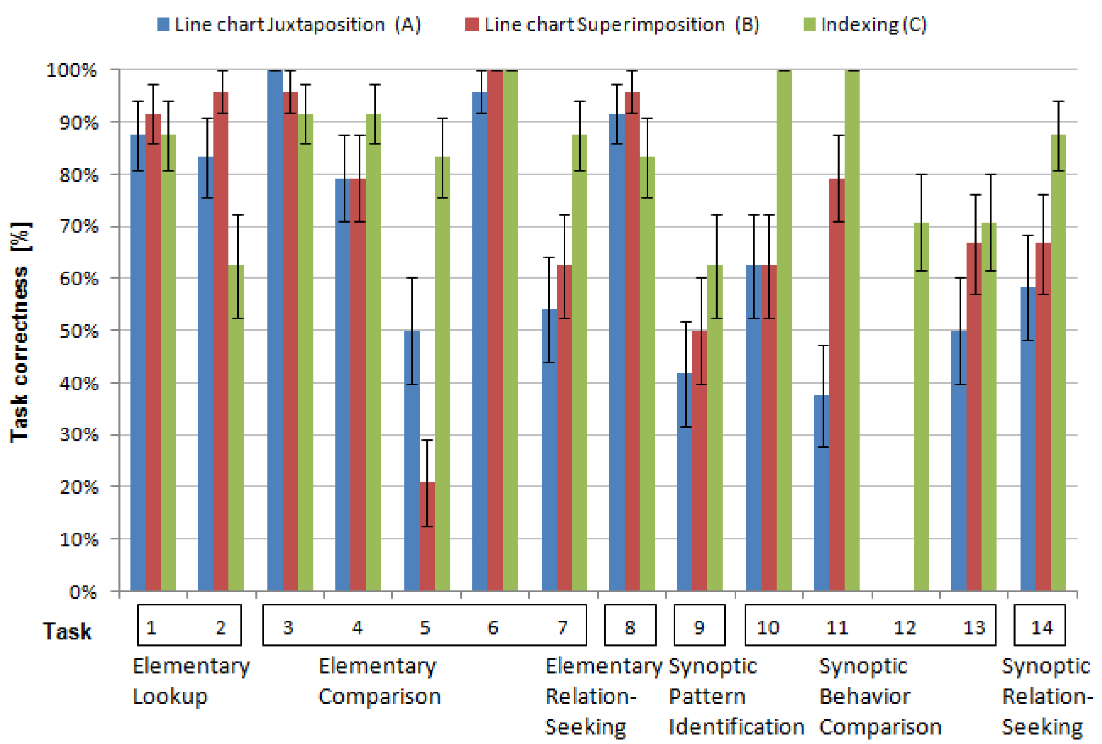


Figure 52: Task correctness rate for each visualization type grouped by tasks

Figure 53 illustrates the average task correctness between the three visualization types for every task. The correctness rate strongly differs in tasks 5 (SD: 30%), 7 (SD: 17%), 10 (SD: 21%), 11 (SD: 31%) and 12 (SD: 39%). On the other side, the correctness rates for tasks 1 (SD: 2%), 3 (SD: 4%), 6 (SD: 2%) and 8 (SD: 6%) are relative close together. Interestingly the tasks 1, 3, 6 and 8 all belong to the elementary task category.

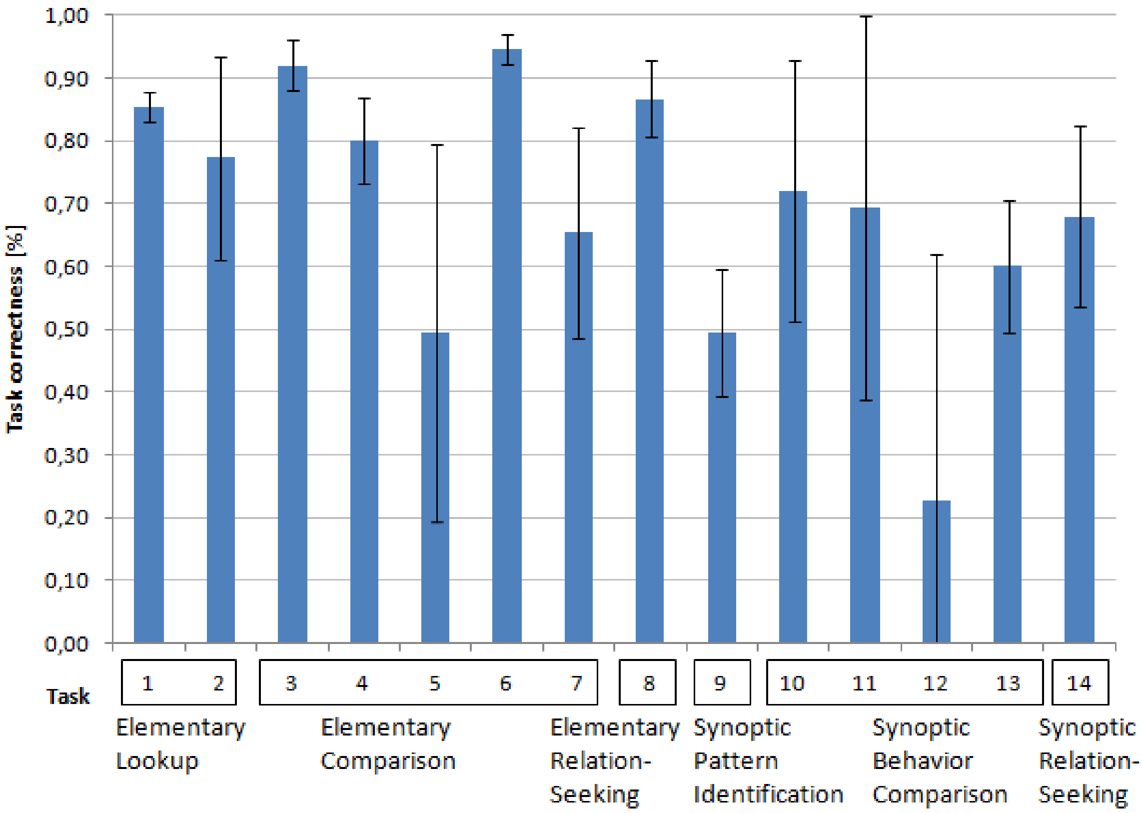


Figure 53: Task correctness rate grouped by tasks

The task correctness rates for task 5, 7, 10, 11 and 12 vary between the three visualization types at most. The following paragraphs provide more detail for the five tasks. It should be mentioned that visualization type Indexing (C) has the lowest error rate in all five tasks.

Task 5 is an Elementary Comparison Task with homogenous data. The goal of the task is to quantify the amount of price change for each of the two displayed stocks and for a given time period in percent. The average absolute minimum error distance was calculated for each visualization type. The visualization type juxtaposed, linear scaled line chart (A) has a distance of 157.14. Visualization type superimposed, log scaled line chart (B) has an average distance of 71.49 and visualization type line chart with indexing method (C) has an average distance of 14.95. The error distance of visualization type A (Linear Scale, Juxtaposition) is more than ten times higher than visualization type C (Indexing).

Visualization type	Average distance	Compared to Indexing (C)
Linear Scale, Juxtaposition (A)	157.14 (SD: 8.47)	1051.10%
Log Scale, Superimposition (B)	71.49 (SD: 2.66)	478.19%
Indexing (C)	14.95 (SD: 1.42)	100.00%

Table 11: Average error distance for task 5

Task 7 is also an Elementary Comparison Task with combined data (two stocks and one stock index are displayed). The goal of the task is to identify the percent change for a given year of one stock. The average error distance is lower than in task 5. The distances of visualization type A (Linear Scale, Juxtaposition) and B (Log Scale, Superimposition) are also much closer together. Visualization type A (Linear Scale, Juxtaposition) has an average distance of 50.44, which is close to visualization type B (Log Scale, Superimposition) with 43.44. Visualization type C (Indexing) has an average error distance of 18.22.

Visualization type	Average distance	Compared to Indexing (C)
Linear Scale, Juxtaposition (A)	50.44 (SD: 6.72)	276.84%
Log Scale, Superimposition (B)	43.44 (SD: 3.26)	238.42%
Indexing (C)	18.22 (SD: 2.75)	100.00%

Table 12: Average error distance for task 7

Task 10 is a Synoptic Behavior Comparison Task of homogenous data. The goal of the task is to identify the stock which has a greater percent increase for a given time period. Only visualization type C (Indexing) has a correctness rate of 100%. The other two visualization types A (Linear Scale, Juxtaposition) and B (Log Scale, Superimposition) have equal correctness rates of about 60%. Visualization type C (Indexing) shows a clear advantage. This may be a result of the usage of the indexing point, which simplifies percent comparisons of multivariate data.

Task 11 belongs also to the category of Synoptic Behavior Comparison Tasks. The goal of the task is to identify the stock which has a lower percent loss for a given year. This task is similar to task 10. Visualization type A (Linear Scale, Juxtaposition) has a 40 % correctness rate and B (Log Scale, Superimposition) has an 80 % correctness rate. Visualization type C (Indexing) has again a 100 % correctness rate.

Task 12 is a Synoptic Behavior Comparison Task with heterogeneous data. The goal of this task is to identify the months where a stock has a higher percent increase than a given stock index. Visualization type A (Linear Scale, Juxtaposition) and B (Log Scale, Superimposition) have similar correctness rate of around 70%. Thus 7 of 10 months are correctly identified. Visualization type C (Indexing) has a higher rate of about 90%.

A more detailed analysis of the five tasks with the greatest standard deviation between the visualization types showed that for each task the correctness rate of visualization type C (Indexing) was higher.

Visualization type	Average correctness rate	Compared to Indexing (C)
Linear Scale, Juxtaposition (A)	68.84	76.28%
Log Scale, Superimposition (B)	67.48	74.77%
Indexing (C)	90.24	100.00%

Table 13: Average correctness rate for task 12

Line charts with Indexing method have overall a higher correctness rate compared to the other two visualization types. Especially task with identification with percent values are superior with visualization type C (Indexing).

6.4 User Preferences

After the usability test, each test subject had to select one visualization type of the three possible visualization types which was perceived as most useful for the subjects.

The visualization type C (Indexing) was chosen 19 times out of 24. Visualization type A (Linear Scale, Juxtaposition) has been chosen only once and visualization type B (Log Scale, Superimposition) has been chosen 4 times.

The preference choices of the 24 test persons were statistically tested if they are uniformly distributed. The distribution was analyzed by a chi-square test. The value for the chi-square-statistic $\boldsymbol{X^2}$ resulted to 23.25. The test statistic X^2 for 2 degrees of freedom and a significance level of 0.05 is equal to approximately 5.9915.

The preference choices of the 24 participants are therefore significantly different from a uniform distribution. The p-value is much lower than 0.01 (exact value is $8.94 * 10^{-6}$).

Test person	Preference	Test person	Preference
TP 1	C	TP 13	A
TP 2	C	TP 14	C
TP 3	C	TP 15	B
TP 4	C	TP 16	B

TP 5	C	TP 17	C
TP 6	C	TP 18	C
TP 7	C	TP 19	C
TP 8	C	TP 20	C
TP 9	C	TP 21	B
TP 10	C	TP 22	C
TP 11	C	TP 23	B
TP 12	C	TP 24	C

Table 14: Preferences of the 24 test persons (TP); Linear Scale & Juxtaposition (A), Log Scale & Superimposition (B), Indexing (C)

Table 15 summarizes the hypotheses and the test results for task completion time (TCT) and task correctness (TC). Significant differences between the visualization types are denoted as not equal (N Eq.) and hypothesis results with no significant differences are denoted as equal (Eq.). All tests are evaluated with a significance level of 0.05.

	H1 (Log Scale - percent estimation)		**H2 (Indexing - percent estimation)**		**H3 (Indexing - trend comparison)**		**H4 (Log Scale, Superimposition)**		**H5 (Indexing)**	
Visualization type	A, B		B, C		A, B, C		A, B		A, B, C	
Tasks	5, 7, 10 - 14		5, 7, 10 - 14		5, 7 - 14		1 - 14		1 - 14	
	TCT	TC	TCT	TC	TCT	TC	TCT	TC	TCT	TC
P-value	0.01	0.45	0.14	< 0.01	0.16	0.001	0.21	0.20	0.25	< 0.01
Result	N Eq.	Eq.	Eq.	N Eq.	Eq.	N Eq.	Eq.	Eq.	Eq.	N Eq.

Table 15: Overview of the five hypotheses and test results for task completion time (TCT) and task correctness (TC); Linear Scale & Juxtaposition (A), Log Scale & Superimposition (B), Indexing (C)

7 Discussion and Outlook

Three visualization types for the display of multivariate time series were examined by a series of usability tests with 24 test persons. Two dependent variables were measured to statistically compare the performance of the three visualization types A (Linear Scale, Juxtaposition), B (Log Scale, Superimposition) and C (Indexing).

One dependent variables of the test was the task completion time. The results did only show a significant difference between the tested visualization types for hypothesis H1 (Log Scale - percent estimation). The hypothesis result based on percent estimation comparison tasks showed that task completion times of juxtaposed linear scaled line charts are significant faster than of superimposed logarithmic scaled line charts.

About the rather equal results of the task completion time can only speculated. It is possible that this outcome is a psychological phenomenon. The users limit the completion time for the tasks. It is not clear whether this decision is conscious or not. Another observation is that the task completion time is not correlated to the task correctness rate.

The second dependant variable is the task correctness rate. The results show more significant differences between the three visualization types. The test results of the task correctness rate are supporting the hypotheses H2 (Indexing - percent estimation), H3 (Indexing - trend comparison) and H5 (Indexing). The indexing chart has a higher correctness rate in all hypotheses.

The superior results of the visualization type could be a consequence of the ability to superimpose multivariate data. This visualization method improves the user’s capability to perform comparison tasks. The visual comparison process of heterogeneous time series is much easier and more effective to execute, when using the indexing method. Any dimension is transformed into a percent dimension, which makes superimposition for any multivariate time series possible.

The user can select an indexing point based on a specific point in time as start for the comparison. After that all points on the chart represent relative changes in relation to the indexing point. This implies an increase in task correctness rates.

Juxtaposed linear scaled line charts and superimposed logarithmic scaled line charts did not have significant differences in their task correctness rate. So the test results give evidence that these two visualization types do not have a statistically significant effect on the correctness of the task results.

This may be because heterogeneous time series cannot be displayed in superimposition. Both visualization types need a separate axis or chart for each dimension. This implies a limited comparison because the values cannot be compared by their spatial position. This further implies a lower task correctness rate for both visualization types for heterogeneous time series.

Superimposed line charts should have an advantage for the visual comparison of homogeneous time series. Superimposition enables the user to compare time series more effective.

H1: Log Scale for percent estimation tasks

The test results of the task completion time support hypothesis Log Scale - percent estimation (H1) that percent estimations tasks in superimposed logarithmic scaled line charts are indeed significantly faster than in juxtaposed linear scaled line charts. However the task correctness rates did not show any significant differences.

In most tasks are the correctness rates close together. The values only diverge considerable in two tasks. One is task 5, which is an elementary comparison task with homogeneous data. The other is task 11, which is a synoptic behavior comparison task with homogeneous data. Visualization type A (Linear Scale, Juxtaposition) has a 30 % higher correctness in task 5, while visualization type B (Log Scale, Superimposition) reaches a higher correctness rate of additional 40 % in task 11.

H2: Indexing method for percent estimation tasks

Task completion times do not differ significantly for the two tested visualization types. But the task correctness rates are significantly different. The visualization type C (Indexing), line chart with indexing method, has a higher correctness rate than the visualization type B (Log Scale, Superimposition), superimposed logarithmic line chart.

The task correctness of visualization type C (Indexing) was in all tasks higher. This advantage should at least be partially based on the free selectable indexing point. The user could set the time according to the needs which results in more correct answers. It is interesting to note that the task completion time is not significantly increased although the user has to additionally select a specific date.

H3: Indexing method for trend comparison

The results are similar to the previous results of the hypotheses. The task completion times are not significantly different among the three tested visualization types. But the task correctness rates are significantly higher for visualization type C (Indexing).

Again, the better results for the task correctness rates are partly based on the free selectable indexing point. This hypothesis is generalizing the statement of the hypothesis H1 (Log Scale - percent estimation) and H2 (Indexing - percent estimation).

H4: Superimposed, logarithmic scaled line charts are better than juxtaposed line charts for visual comparisons

The test results for Hypothesis H4 did not show any differences for task completion times or task correctness rates between visualization types A (Linear Scale, Juxtaposition) and B (Log Scale, Superimposition).

While hypothesis H1 (Log Scale - percent estimation) is supporting that visualization type B (Log Scale, Superimposition) has faster task completion times than visualization type A (Linear Scale, Juxtaposition) for percent estimation tasks. But this result is not valid for a broader set of tasks.

H5: Indexing method is overall better for visual comparisons

Task completion time is not significantly different between the three visualization types. But the task correctness rates of visualization type C (Indexing) are significantly higher.

This result is consistent with hypotheses H2 (Indexing - percent estimation) and H3 (Indexing - trend comparison). The visualization type C (Indexing) offers a higher correctness for similar task completion times. The advantage of the indexing chart is the superimposition of homogeneous and heterogeneous time series. Comparison tasks should greatly benefit through superimposition.

Visualization type results

Logarithmic scales enable the user to execute percent estimation tasks faster than linear scales. The test results show that the scale has no significant effect on the task correctness. When performing a mixture of tasks the advantage of logarithmic scales disappears.

The usability result shows that the line chart with indexing method is superior to the other two visualization types. Performance measures and test user's subjective opinions favor this visualization method.

Juxtaposed line charts with linear scale were slower than superimposed logarithmic scaled line charts for percent estimation tasks. Other tasks did show a difference for task completion times. Task correctness rates were similar to superimpose logarithmic scaled line charts. But the task correctness rates were clearly worse than when using a line chart with indexing method.

The superimposed logarithmic scaled line chart was superior to the juxtaposed logarithmic scaled line chart in task completion times.

The line chart with indexing method was superior to the other two charts in task correctness rates. It is interesting to note that line chart with indexing method was not only superior in percent comparison tasks but also in many other tasks for task correctness rates.

The line chart with indexing method offers advantages for percent estimations and comparison tasks. Task completion times were not significant different from the other visualization types.

The 24 test users preferred the line chart with indexing method over the two other visualization types. The visualization type was selected 19 times out of 24, which shows that the users were satisfied with this visualization type.

Outlook

The following paragraphs will present a few suggestions for further studies in the area of visual comparison of multivariate time series.

One interesting part for further studies of the indexing method is to test the effects of different indexing points in more detail. How does the indexing point influence the test results? How to enhance the indexing point?

The process of selecting a specific date as indexing point is too time consuming. It is also not interactive enough. One suggestion would be to make the indexing point dependent on the xcoordinate of the mouse position. The drawback of this method is that the user can not freely move the mouse without changing the indexing point.

A better solution would be to define a vertical indexing line. The line can be dragged with the mouse along the horizontal space. The line represents the indexing point of all time series. This method should be more useable, because it lets the user use his mouse cursor without changing the indexing point all the time.

An alternative could be the usage of a small point as marker instead of the vertical line. The point is used as indexing point for all displayed time series. The user can move the point by simple drag and drop operation.

[Few, 2004a] describes the advantages of reference lines. This lines represent important values of the y – axis. This could be useful for the indexing chart. A horizontal line at the indexing value of 100 % could be a reference line. This would let the user clearly identify if a certain point is above or below the indexing point, without looking at the actual value. The user has only to look whether the point is above or below the reference line. This method would make the distinction between positive or negative development very easy.

This principle can be improved by additional reference lines for important values. For instance a reference line could be drawn for + / - 150 %, + / - 200 % and so on. Also the peak values of the displayed time series could be represented by reference lines.

Line charts which use a logarithmic scale for the y-axis could display example lines for certain percent values. 0 percent is represented by a horizontal line. 100 percent is a line with a 62 degrees angle and 200 percent is approximately a line with a 72 degrees angle. By providing important percent values the user can better estimate the percent changes on the chart. A percental change has always the same gradient in the logarithmic scale.

8 Conclusion

Comparison of time series data is a very important part of data analysis. Through the evolution of the computer more and more data can be measured, processed and stored. Therefore time series data is used in many application fields of science and economics. Stock markets are a good example where time series data is extensively used and available to the public. The aim of time series analyses can be divided into general exploration, description and prediction & forecasting.

A comparative study for three line chart visualizations was conducted to examine differences in task completion times and task correctness rates. Homogeneous and heterogeneous stock market data was used. Stocks, stock indices and other economic data are available in the evaluation.

The three observed visualization types are juxtaposed linear scaled line chart, superimposed logarithmic scaled line chart and line chart with indexing method.

Linear scaled axes are often used in charts. However, logarithmic scaled axes are better suited for the display of percent changes. A specialty of the logarithmic scale is the direct representation of percent changes through the slope. The percent change between two points is encoded by the gradient of the connection line.

Another rarely used method to compare multivariate data is indexing. This method transforms the values into percent values according to an indexing point. Through the transformation of the original data into a shared unit, the original heterogeneous data can be displayed by superimposition of the transformed time series.

A usability test was defined, which consists of 42 tasks. The tasks can be divided into elementary and synoptic task groups. Elementary tasks are concerned with identifying simple characteristics like a specific day, while synoptic tasks are more complex and involve identifying a value for a certain time period.

Two thirds of the tasks were centered on comparisons of two time series. The other tasks are of the categories lookup, pattern seeking and relation identifying. The focus of the usability test lies on comparison tasks.

The test was performed with the help of 24 test persons. The persons belong to a rather homogeneous group. Half were female and the other half were male. The age was within the range of 20 to 30 years. 19 out of 24 persons were studying at the time of the test at a university. The remaining five persons already graduated from university.

The usability tests are used to verify the five hypotheses of the master thesis. Task completion time and task correctness rate were evaluated for a performance comparison of the visualization types. Each hypothesis was evaluated by a specific ANOVA test.

The test results give evidence that the indexing chart has a higher correctness rate than the two other visualization types. Task completion times are not significantly different. One of the two main benefits is the ability to superimpose any data by transformation of values into a percent dimension. The other benefit is the arbitrary definition of an indexing point. This makes comparisons even more effective and precise.

User preferences after the test also support the indexing chart. 19 of 24 users favor the line chart with indexing method for visual comparison tasks.

An influence of the logarithmic scale could only be found for percent estimation tasks. Task completion times are lower, while task correctness rate do not show any significant differences.

Although the indexing method was proposed by Bertin in 1983, very few studies are dedicated to this subject. Future studies in this area could examine more interactive ways to set the indexing point dynamically. This could further increase the performance of the indexing chart.

Bibliography

[Andrienko, 2006] N. Andrienko, G. Andrienko. Exploratory Analysis of Spatial and Temporal Data - A Systematic Approach, Springer, 2006.

[Aigner et al., 2007] Wolfgang Aigner. Towards a Conceptual Framework for Visual Analytics of Time and Time-Oriented Data, Proceedings of the 2007 Winter Simulation Conference, pp. 721-729, 2007.

[Aigner, 2009] Wolfgang Aigner. Practical Courses and Diploma Theses. Retrieved at: Oct 22, 2009. http://ieg.ifs.tuwien.ac.at/practical.html#linegraphs.

[Bertin, 1983] Jacques Bertin. Semiology of Graphics. University of Wisconsin Press, 1983.

[Buono et al., 2005] Paolo Buono, Aleks Aris, Catherine Plaisant, Amir Khella, Ben Shneiderman. Interactive Pattern Search in Time Series. Proceedings of Conference on Visualization and Data Analysis, pp. 175-186, VDA 2005, SPIE, Washington DC, 2005.

[Buono et al., 2007] Paolo Buono, Catherine Plaisant, Adalberto Simeone, Aleks Aris, Ben Shneiderman, Galit Shmueli, Wolfgang Jank, "Similarity-Based Forecasting with Simultaneous Previews: A River Plot Interface for Time Series Forecasting," iv, pp.191-196, 11th International Conference Information Visualization (IV '07), 2007

[Card et al., 1999] Stuart K. Card, Jock Mackinlay, Ben Shneiderman. Information Visualization: Using Vision to Think. Morgan Kaufmann, 1999.

[Card, 2007] Stuart K. Card. Information .visualization. In the Human-Computer Interaction Handbook, Lawrence Erlbaum Associates, 2007.

[Chi, 2000] Ed H. Chi, "A taxonomy of visualization techniques using the data state reference model," in Proceedings of the IEEE Information Visualization Symposium, pp. 69-76, 2000.

[Fekete et al., 2008] J. Fekete, J. J. Wijk, J. T. Stasko, C. North. The Value of Information Visualization. In Information Visualization: Human-Centered Issues and Perspectives, Springer Verlag Berlin Heidelberg, 2008.

[Few, 2004a] Stephen Few. Show me the numbers. Analytics Press Oakland, 2004.

[Few, 2004b] Stephen Few. Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message. Retrieved at: Oct 22, 2009. http://www.perceptualedge.com/articles/b-eye/the_right_graph.pdf.

[Few, 2005] Stephen Few. Effectively Communicating Numbers: Selecting the Best Means and Manner of Display. Retrieved at: Oct 22, 2009. http://www.perceptualedge.com/articles/b-eye/Communicating_Numbers.pdf.

[Few, 2006a] Stephen Few. An Introduction to Visual Multivariate Analysis. Retrieved at: Oct 22, 2009. http://www.perceptualedge.com/articles/b-eye/visual_multivariate_analysis.pdf.

[Few, 2006b] Stephen Few. Data Visualization: Rules for Encoding Values in Graph. Retrieved at: Oct 22, 2009. http://www.perceptualedge.com/articles/b-eye/encoding_values_in_graph.pdf.

[Few, 2008] Stephen Few. Time on the Horizon. Retrieved at: Oct 22, 2009. http://www.perceptualedge.com/articles/b-eye/time_on_the_horizon.pdf.

[Google, 2009] Google Finance. Retrieved at: Oct 22, 2009. http://www.google.com/finance

[Harris, 1999] Robert L. Harris. Information Graphics: A Comprehensive Illustrated Reference. Oxford University Press. 1999.

[Heer, 2009] Jeffrey Heer. prefuse | documentation > manual > introduction > toolkit structure. Retrieved at: Oct 22, 2009. http://prefuse.org/doc/manual/introduction/structure/

[Heer et al., 2009] Jeffrey Heer, Nicholas Kong, and Maneesh Agrawala. Sizing the Horizon: The Effects of Chart Size and Layering on the Graphical Perception of Time Series Visualizations. ACM Human Factors in Computing Systems (CHI), 2009.

[Hochheiser et al., 2001] H. Hochheiser, B. Shneiderman. Visual Specification of Queries for Finding Patterns in Time-Series Data. Proceedings of Discovery Science 2001.

[Ma, 2009] Rui Ma. Designing Interactive Visualiza-tion Methods for Comparing Mul-tivariate Stock Market Data over Time .Master Thesis. Technical University of Vienna, 2009.

[Müller et al., 2003] Wolfgang Müller, Heidrun Schumann. Visualization methods for time-dependent data – an overview. Simulation Conference, 2003. Proceedings of the 2003 Winter, pp. 737-745, 2003.

[Shneiderman, 1996] Ben Shneiderman. The Eyes have it: A Task by Data Type Taxonomy for Information Visualizations. IEEE Symposium on Visual Languages, 1996.

[Schumann, 2000] Heidrun Schumann, Wolfgang Müller. Visualisierung: Grundlagen und allgemeine Methoden. Springer Verlag Berlin Heidelberg, 2000.

[Silva et al., 2000] S. Silva, T. Catarci. Visualization of linear time-oriented data: a survey. Web Information Systems Engineering, 2000 Proceedings of the First International Conference, pp. 310, 2000.

[TimeSearcher, 2009] Visual Exploration of Time-Series Data. Retrieved at: Oct 22, 2009. http://www.cs.umd.edu/hcil/timesearcher/

[Tufte, 2001] Edward Tufte. Visual Display of Quantitative Information, Second Edition. Graphics Press Connecticut, 2001.

[Tufte, 2006] Edward Tufte. Beautiful Evidence. Graphics Press Connecticut, 2006.

[Ware, 2004] Colin Ware. Information Visualization. Perception for Design. Morgan Kaufmann Publishers, 2004.

[Wattenberg, 1999] Martin Wattenberg. Visualizing the Stock Market. ACM, 1999.

[Wainer, 1997] Howard Wainer. Visual revelations: graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot. Lawrence Erlbaum Associates, 1997.

List of figures

List of tables

Appendix A – Pre-Test Results

Table 16 gives a tabular view of the personal information, last education and job position for the 24 test persons.

Test Person	Age	Gender	Education	Job
TP 1	20 - 30	Male	Bachelor‘s Degree	Student (Business Informatics)
TP 2	20 - 30	Female	Bachelor‘s Degree	Student (Business Informatics)
TP 3	20 - 30	Male	Bachelor‘s Degree	Student (Business Informatics)
TP 4	20 - 30	Male	Bachelor‘s Degree	Student (Business Informatics)
TP 5	20 - 30	Female	Master’s Degree	Financial Adviser
TP 6	20 - 30	Male	Master’s Degree	Research Assistant
TP 7	20 - 30	Female	Matura	Student (Business Administration)
TP 8	20 - 30	Female	Bachelor‘s Degree	Student (Business Informatics)
TP 9	20 - 30	Male	Bachelor‘s Degree	Student (Business Administration)
TP 10	20 - 30	Male	Bachelor‘s Degree	Student (Business Informatics)
TP 11	20 - 30	Female	Bachelor‘s Degree	Student (Architecture)
TP 12	20 - 30	Female	Matura	Student (Computer Science)
TP 13	20 - 30	Male	Matura	Software Developer
TP 14	20 - 30	Female	Matura	Student (Business Informatics)
TP 15	20 - 30	Male	Master‘s Degree	Electronic Engineer
TP 16	20 - 30	Male	Bachelor‘s Degree	Student (Civil Engineering)
TP 17	20 - 30	Female	Matura	Student (Medical Science)
TP 18	20 - 30	Male	Bachelor‘s Degree	Student (Business Informatics)
TP 19	20 - 30	Female	Bachelor‘s Degree	Student (Business Administration)
TP 20	20 - 30	Female	Bachelor‘s Degree	Student (Business Informatics)
TP 21	20 - 30	Male	Matura	Student (Business Administration)
TP 22	20 - 30	Female	Matura	Student (Finance Mathematics)
TP 23	20 - 30	Female	Master’s Degree	Accountant
TP 24	20 - 30	Male	Bachelor‘s Degree	Sales Consultant

Table 16: Personal information of the 24 participants (TP)

Table 17 shows the self-assigned experience levels about data analysis, stocks and common stock market visualizations for the 24 test persons.

Experience with …						
	Data Analysis	**Stocks**	**Line charts**	**OHLC Charts**	**Candlestick Charts**	**Other**
TP 1	good	good	good	good	good	average
TP 2	good	good	good	good	good	average
TP 3	good	Inexperienced	inexperienced	Inexperienced	Inexperienced	Inexperienced
TP 4	good	average	good	average	average	average
TP 5	average	average	good	average	average	good
TP 6	average	average	good	average	average	good
TP 7	average	inexperienced	average	inexperienced	inexperienced	inexperienced
TP 8	good	average	average	average	average	inexperienced
TP 9	average	inexperienced	average	inexperienced	inexperienced	average
TP 10	average	average	average	average	good	average
TP 11	average	average	average	inexperienced	average	average
TP 12	average	inexperienced	average	inexperienced	inexperienced	inexperienced
TP 13	average	inexperienced	very good	inexperienced	inexperienced	average
TP 14	average	inexperienced	average	inexperienced	inexperienced	average
TP 15	average	average	good	average	average	good
TP 16	good	good	good	good	good	average
TP 17	average	inexperienced	average	inexperienced	inexperienced	average
TP 18	average	average	average	inexperienced	inexperienced	average
TP 19	average	average	average	average	average	average
TP 20	average	average	average	inexperienced	inexperienced	average

TP 21	average	inexpe-rienced	average	inexperienced	inexperienced	inexpe-rienced
TP 22	average	average	average	inexperienced	inexperienced	average
TP 23	average	average	average	average	average	average
TP 24	average	inexpe-rienced	average	inexperienced	inexperienced	inexpe-rienced

Table 17: Experience of participants (TP)

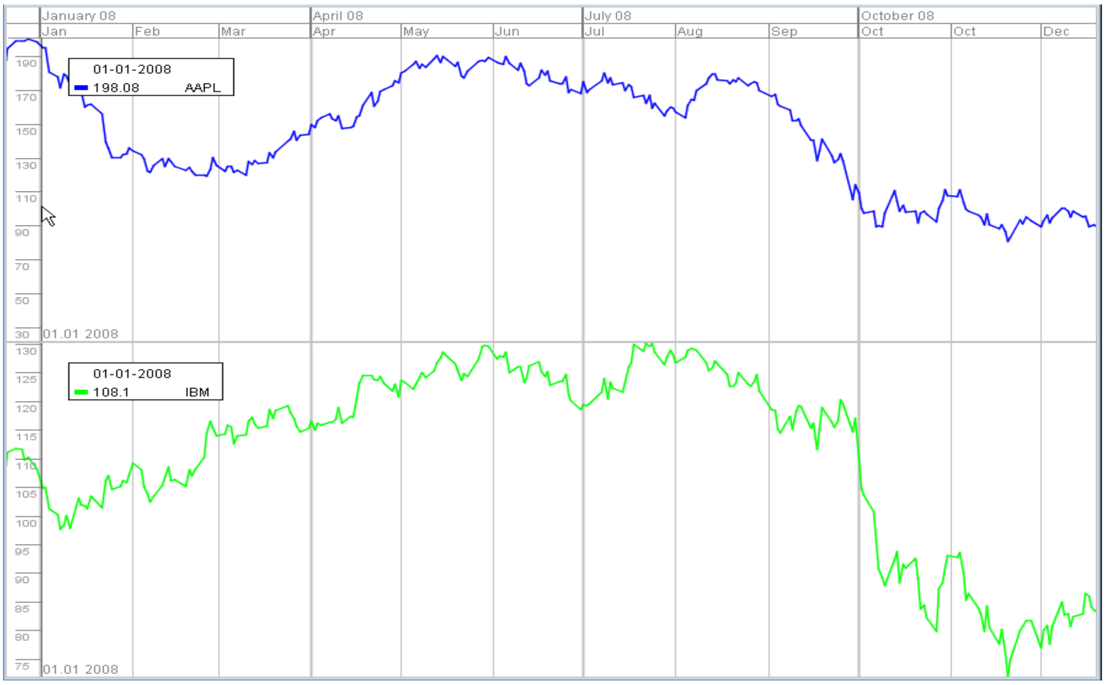
Appendix B – Comparative Study Tasks

Bold typed answers represent the valid task answers.

A) Line charts with juxtaposition

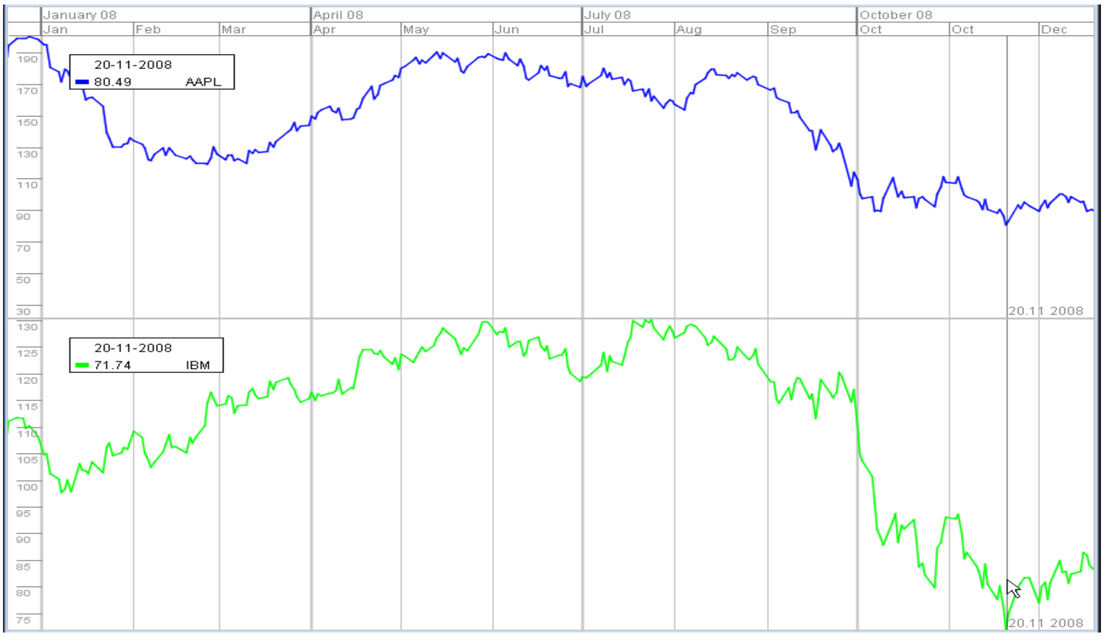
AAPL vs. IBM [NASDAQ]

Lookup:



1. AAPL: On which day was the highest stock price in 2008?

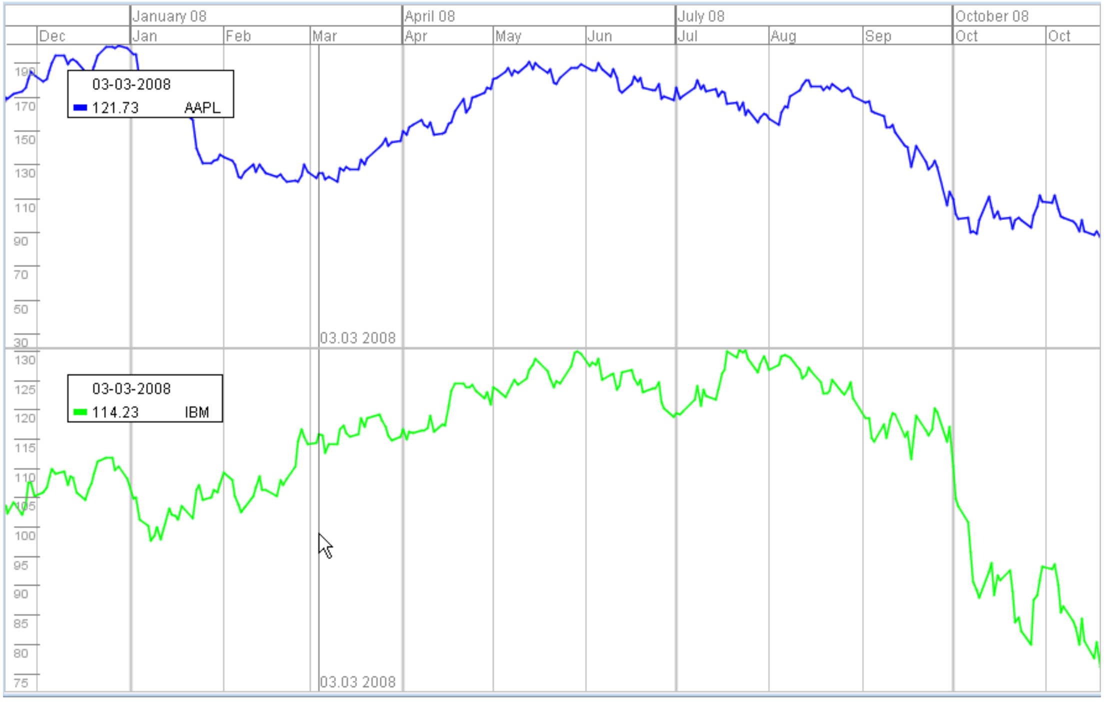
   (01.01.2008)



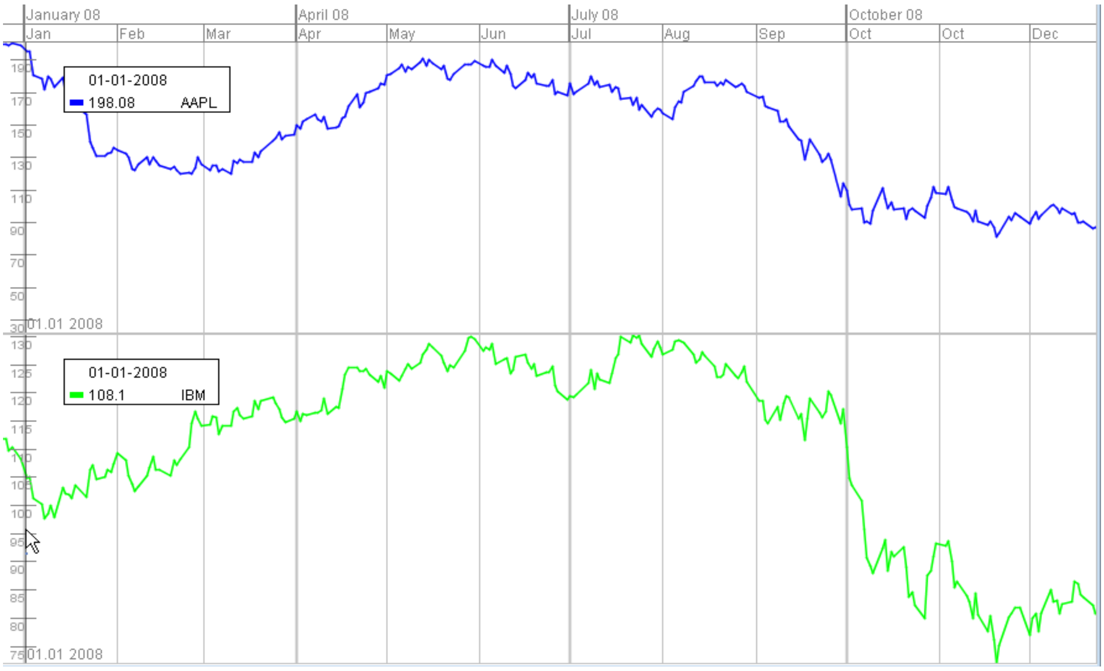
2. IBM: On which day was the lowest stock price in 2008?

   (**20.11.2008)**

Comparison homogenous



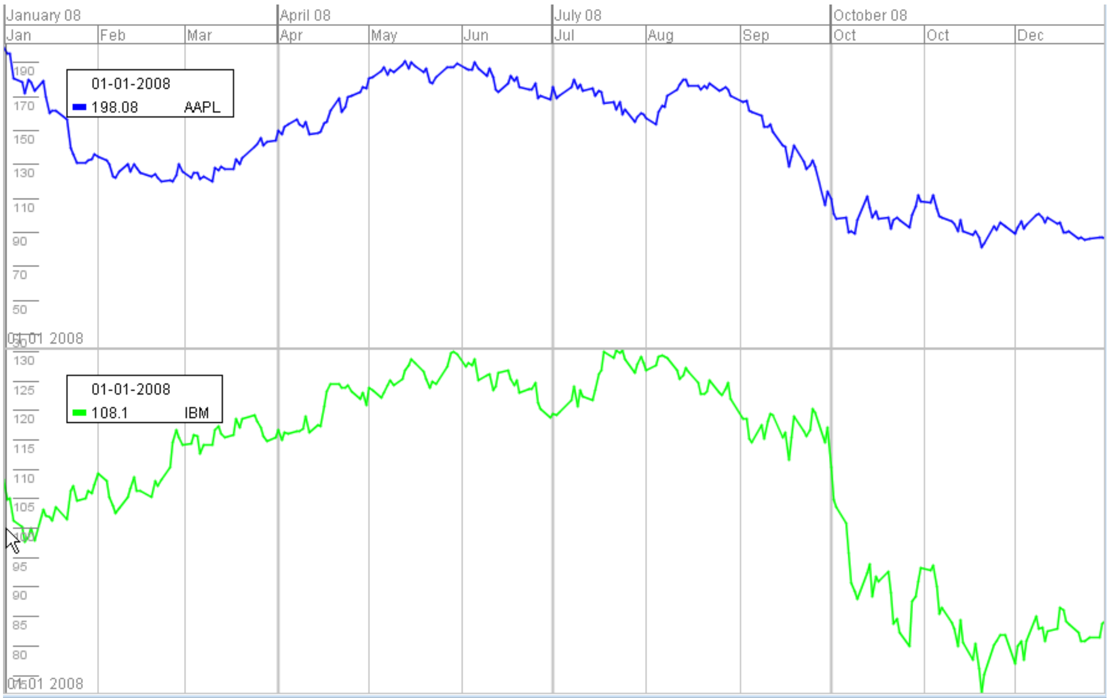
3. Compare the values of AAPL and IBM on the given dates - Which of the following statements are valid?
   1) **On 3.3.2008, AAPL was higher than IBM**
   2) On 16.10.20008, AAPL was lower than IBM



4. Please quantify the amount of price change for the given time periods in dollars:

   AAPL, May 2008: **(8.75 US-$)**

   IBM, May 2008: **(5.82 US-$)**

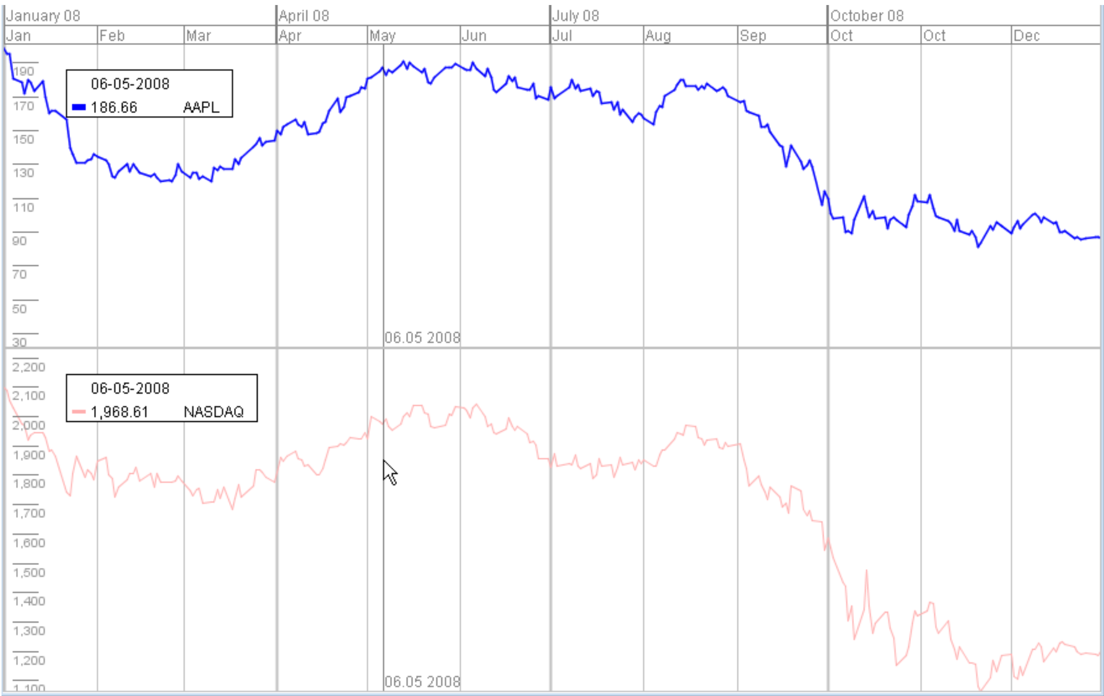


5. Please quantify the amount of price change for the given time periods in percent:

    AAPL, January - May 2008: **(-4.71 %)**

    IBM, January - May 2008: **(19.73 %)**

Comparison heterogeneous

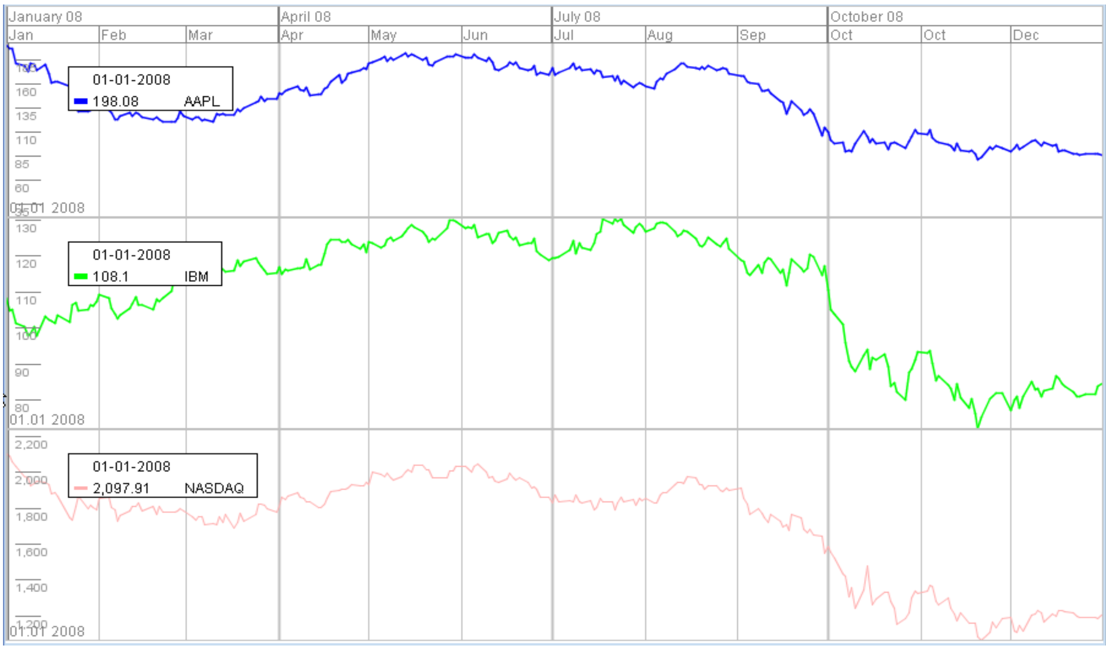


6. Compare the values of AAPL and the NASDAQ index on the given dates - Which of the following statements are valid?

    1) On 05.06.2008 was the value of the NASDAQ over 2000

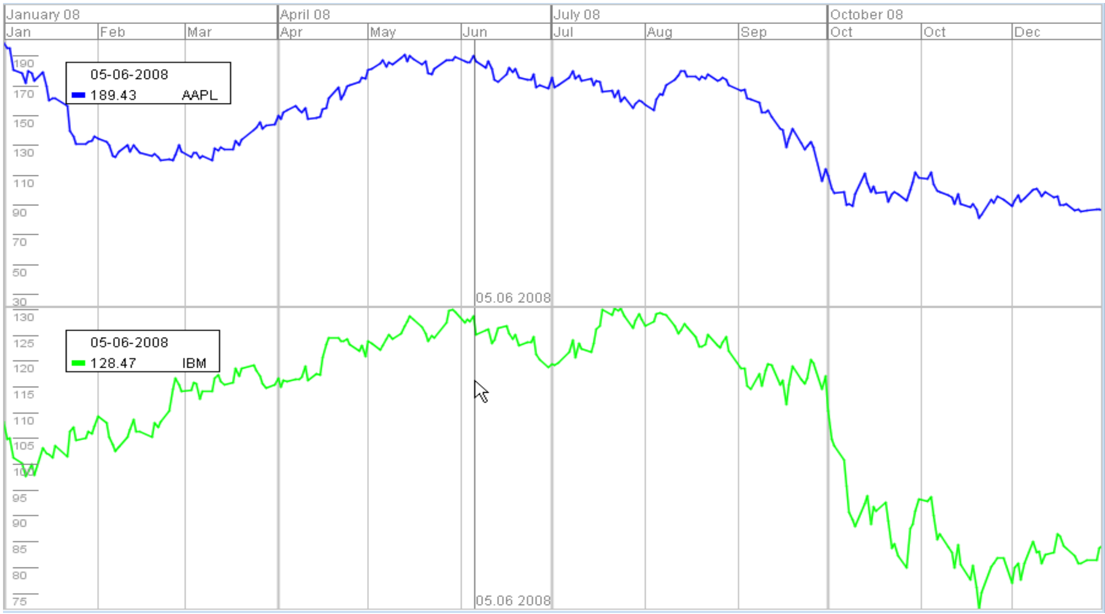
    2) On 20.05.2008 was the value of the AAPL stock under 175

Comparison combination



7. NASDAQ: How much percent did the values change in 2008?
   1) NASDAQ: **(-42.7%)**

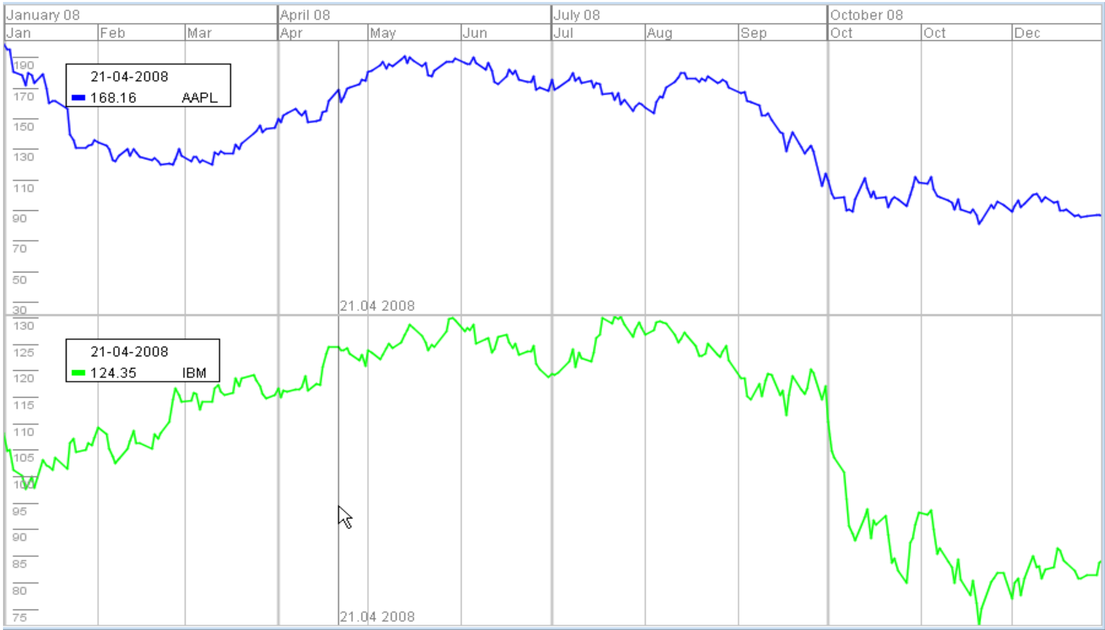
Relation-seeking



8. AAPL: Which of the following months in 2008 have a higher value than the value on 05.06.2008?

   1) January
   2) February
   3) March
   4) April
   5) May
   6) June
   7) July
   8) August
   9) September
   10) October
   11) November
   12) December

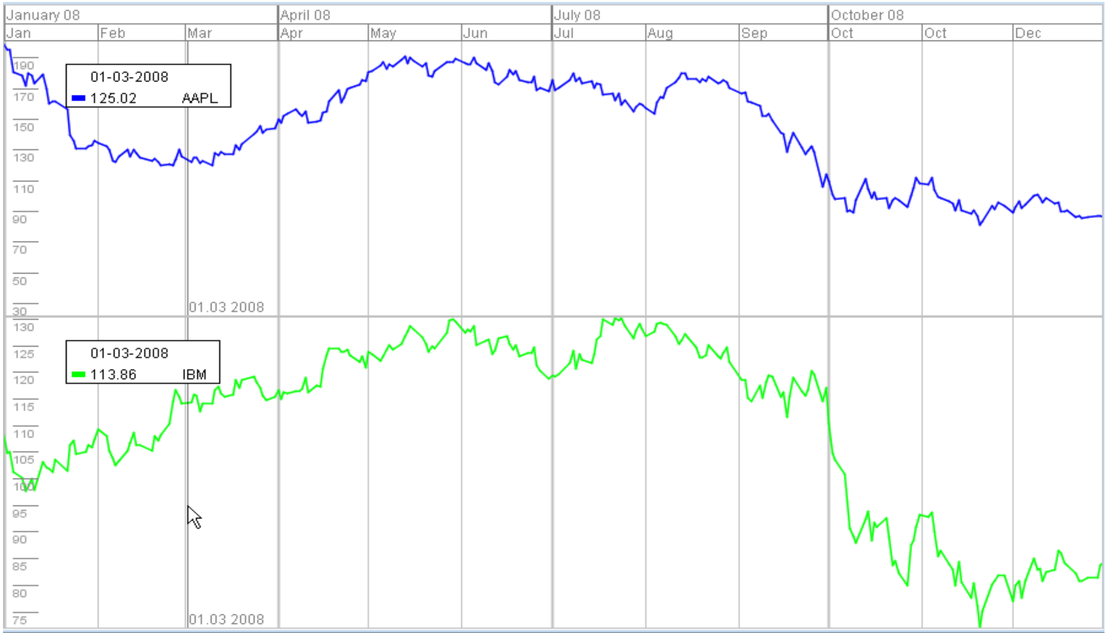
Pattern identification



9. AAPL: Which of the following months in 2008 have a positive trend?

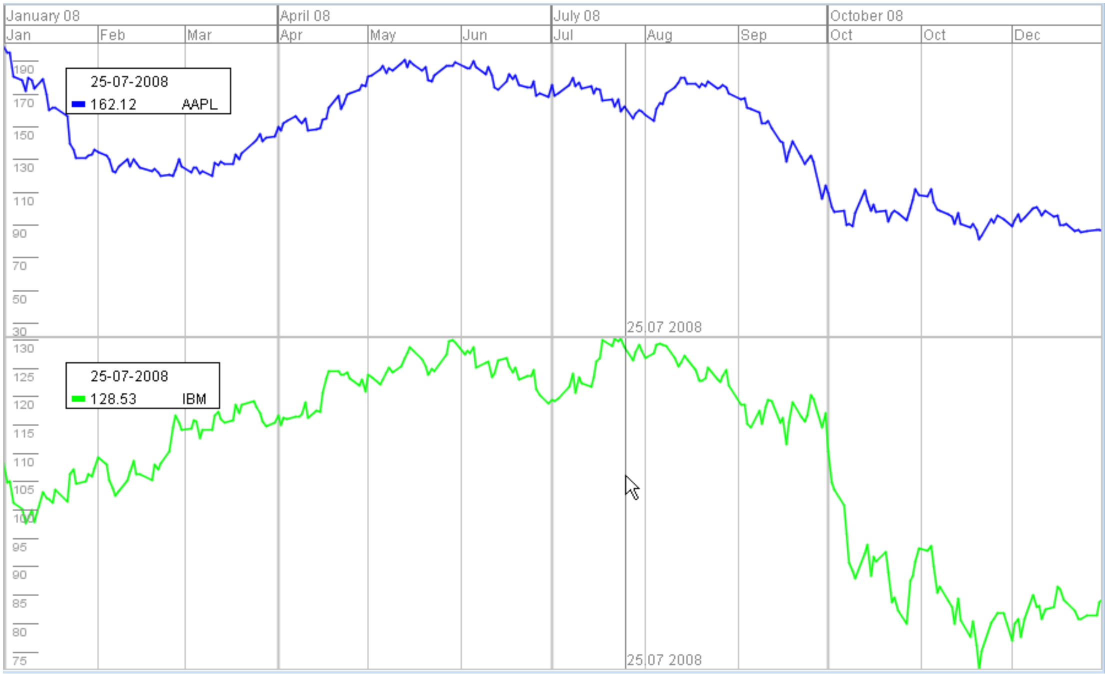
1) January	7) July
2) February	**8) August**
3) March	9) September
4) April	10) October
5) May	11) November
6) June	12) December

Behavior comparison homogenous



10. Which stock has a bigger percent increase from the beginning of 03.2008 to the end of 04.2008?

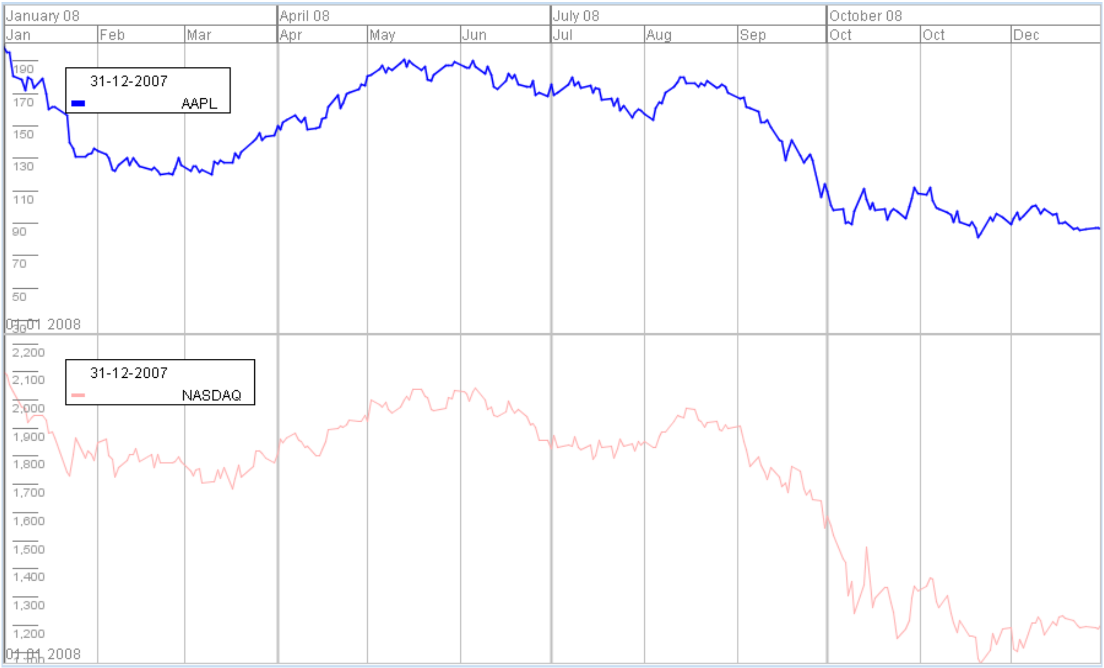
1) AAPL	2) IBM



11. Which stock has a lower percent loss in 2008?

1) AAPL
2) IBM

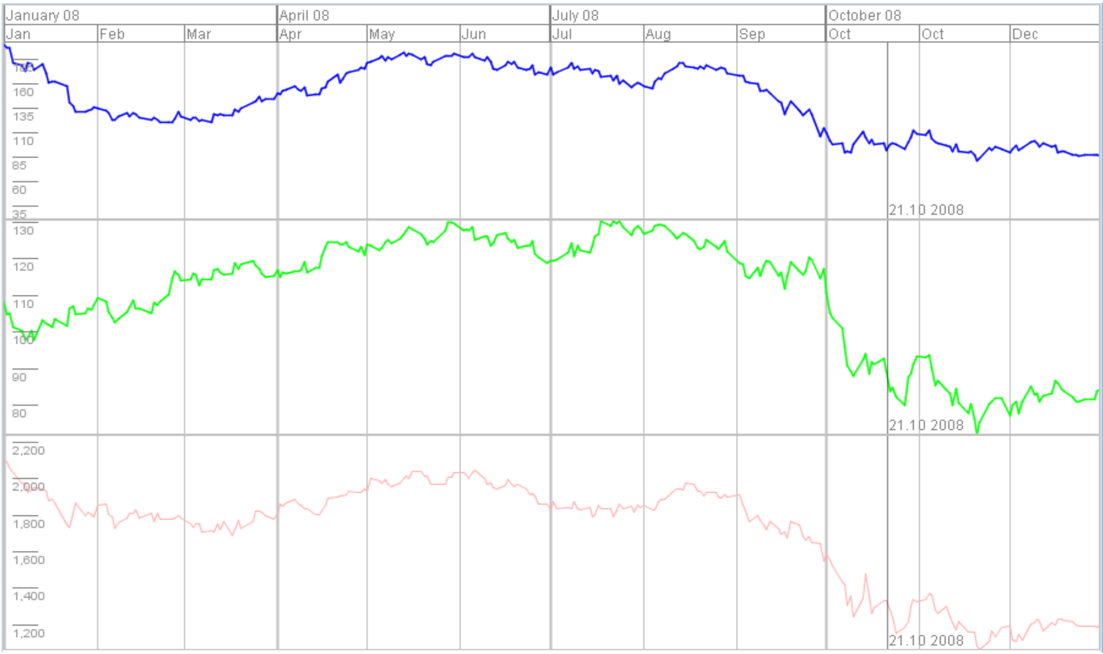
Behavior comparison heterogeneous



12. In which months is the percent increase of AAPL greater than NASDAQ?

1) January
2) February
3) March
4) April
5) May
6) June
7) July
8) August
9) September
10) October
11) November
12) December

Behavior comparison combination



13. Which stock or index has the highest volatility (relative variations) in September 2008?

1) **AAPL**
2) IBM
3) NASDAQ

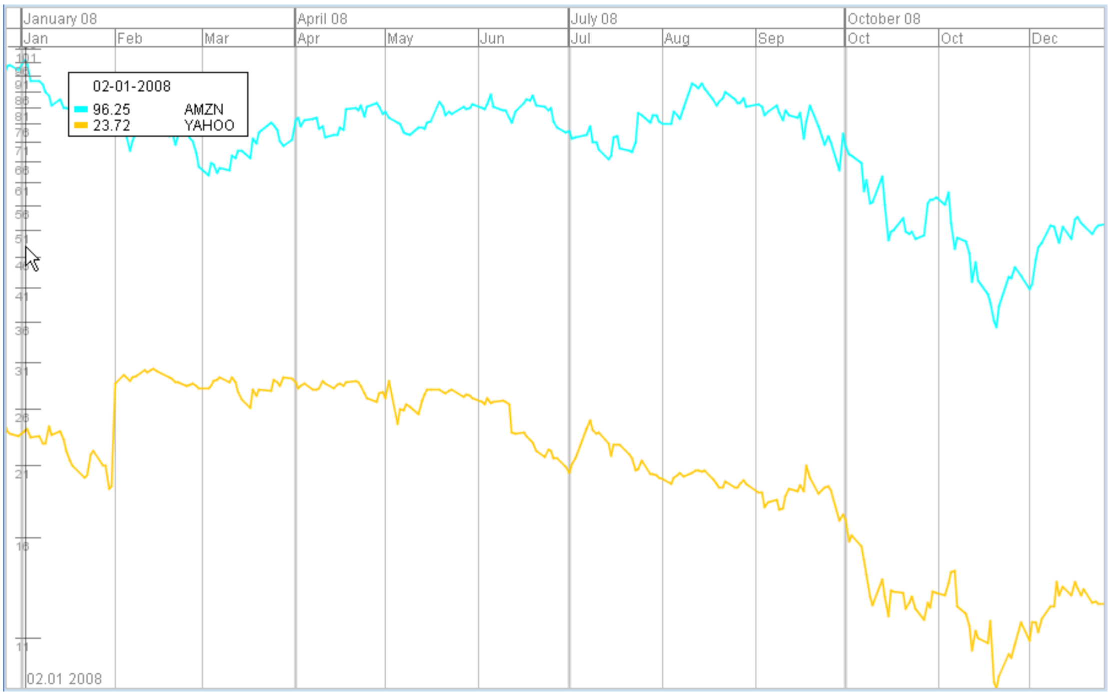
Relation-seeking juxtaposition

14. In which year had AAPL the highest percent increase from beginning to the end of the year?

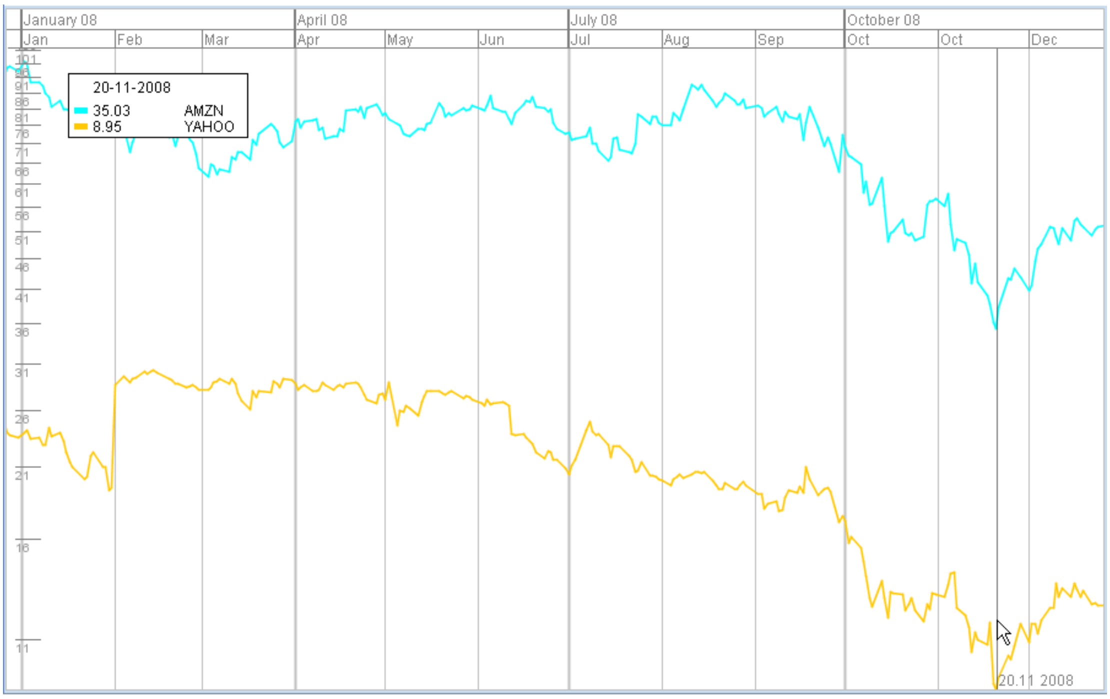
1) 2004
2) 2005
3) 2006
4) 2007

B) Line charts with superimposition und log scale
AMZN vs. YAHOO [SP500]

Lookup homogenous

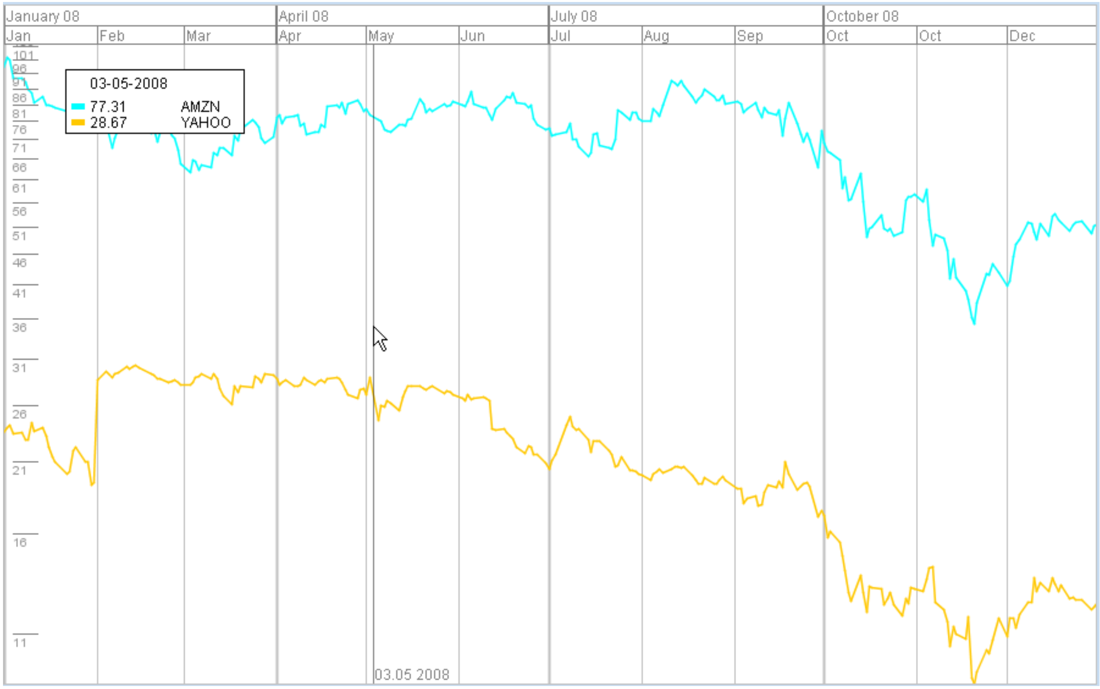


15. AMZN: On which day was the highest stock price in 2008?
(02.01.2008)

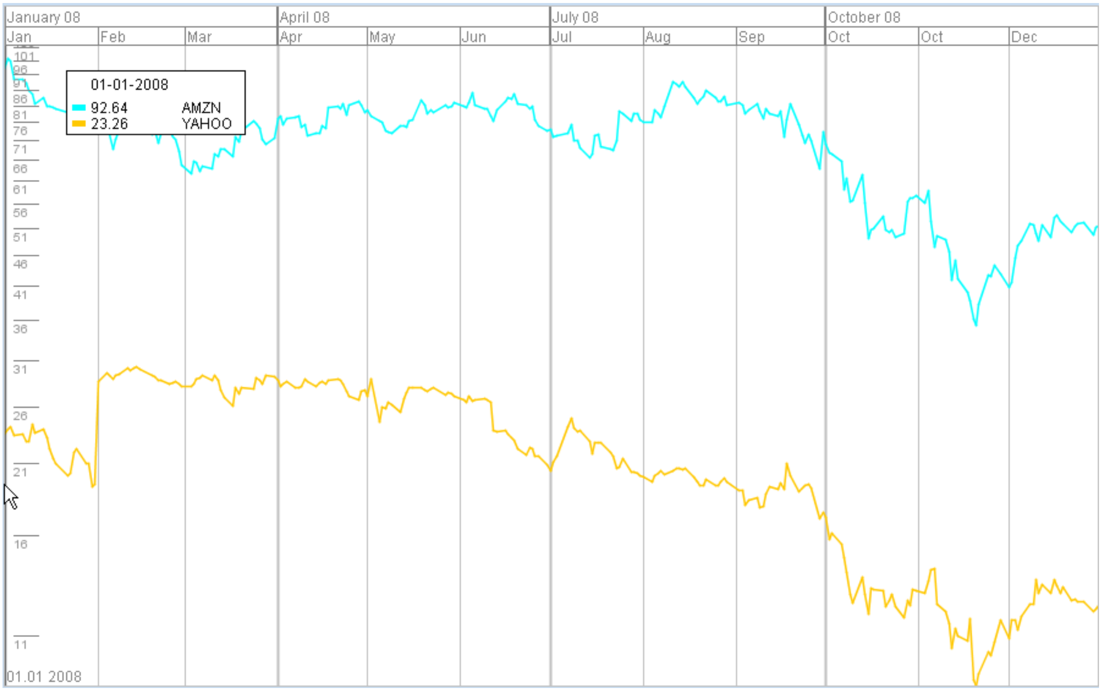


16. YAHOO: On which day was the lowest stock price in 2008?
(20.11.2008)

Comparison homogenous



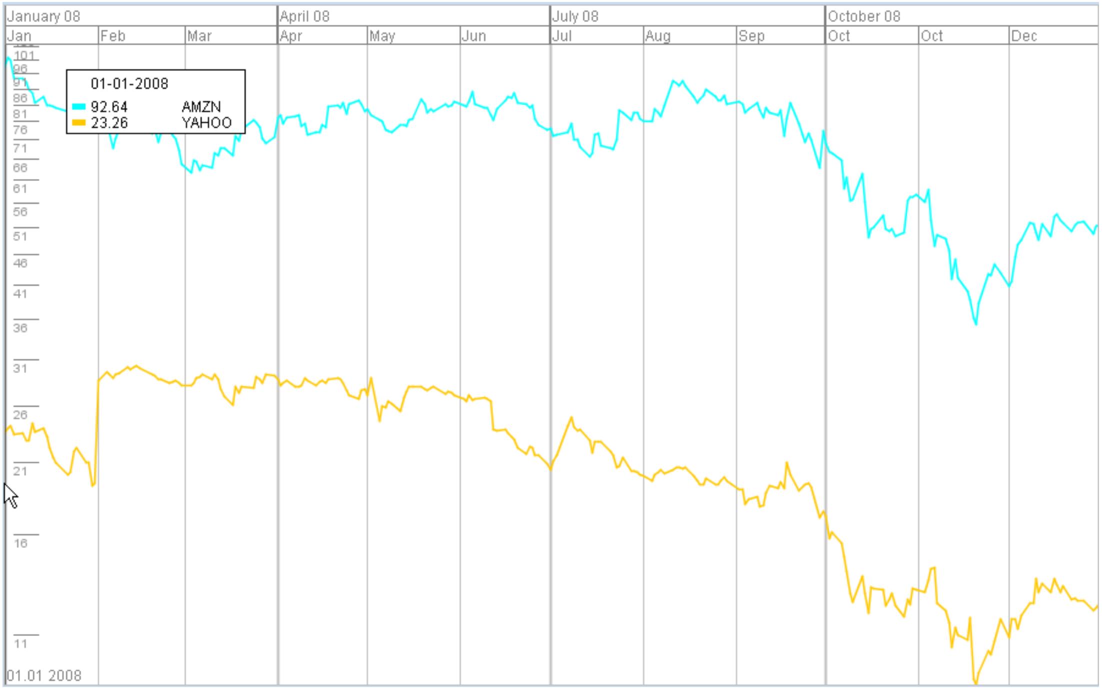
17. Compare the values of AMZN and YAHOO on the given dates - Which of the following statements are valid?
    1) **On 03.05.2008, AMZN was higher than YAHOO**
    2) On 16.10.2008, AMZN was lower than YAHOO



18. Please quantify the amount of price change for the given time periods in dollars:

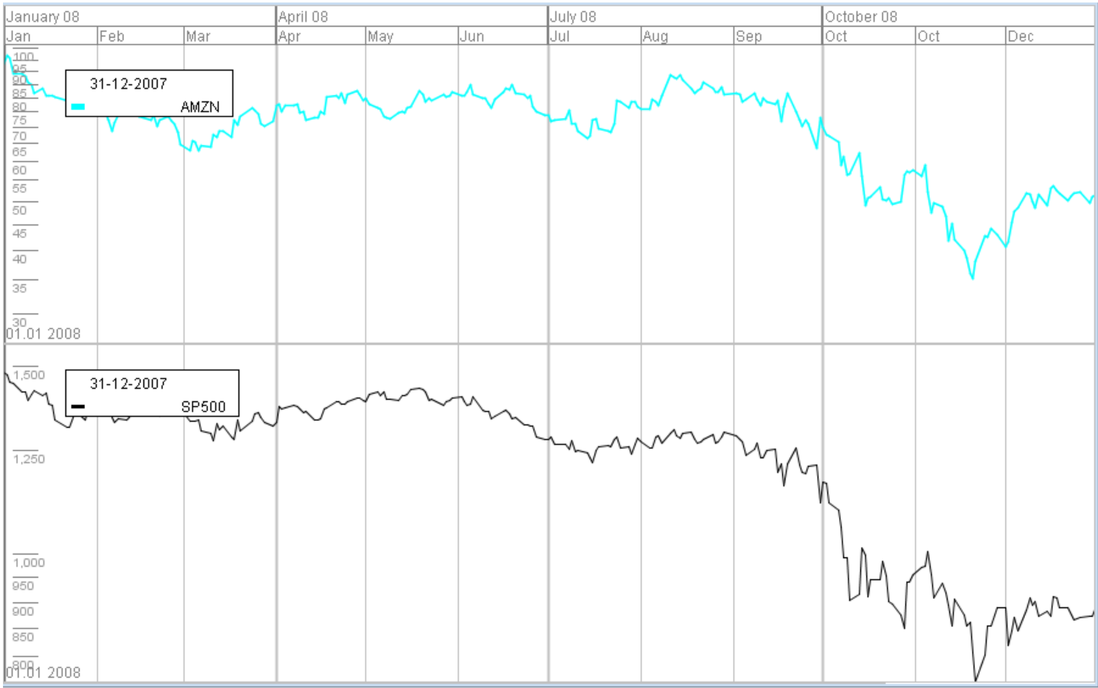
AMZN, January 2008: **(-14.94 US-$)**

YAHOO, January 2008: **(-4.08 US-$)**



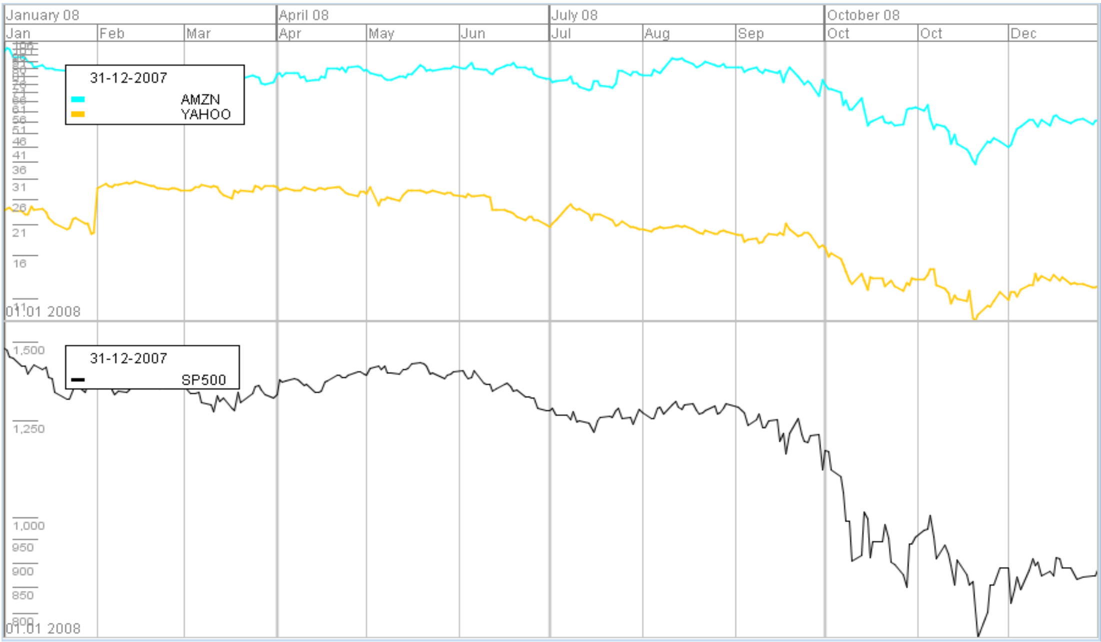
19. Please quantify the amount of price change for the given time periods in percent:
AMZN, January - June 2008: **(-20.84 %)**
YAHOO, January - June 2008: **(-11.18 %)**

Comparison heterogeneous



20. Compare the values of AMZN and the SP500 index on the given dates - Which of the following statements are valid?
    1) On 02.01.2008 was the value of the SP500 over 1500
    2) On 20.11.2008 was the value of the AMZN under 100

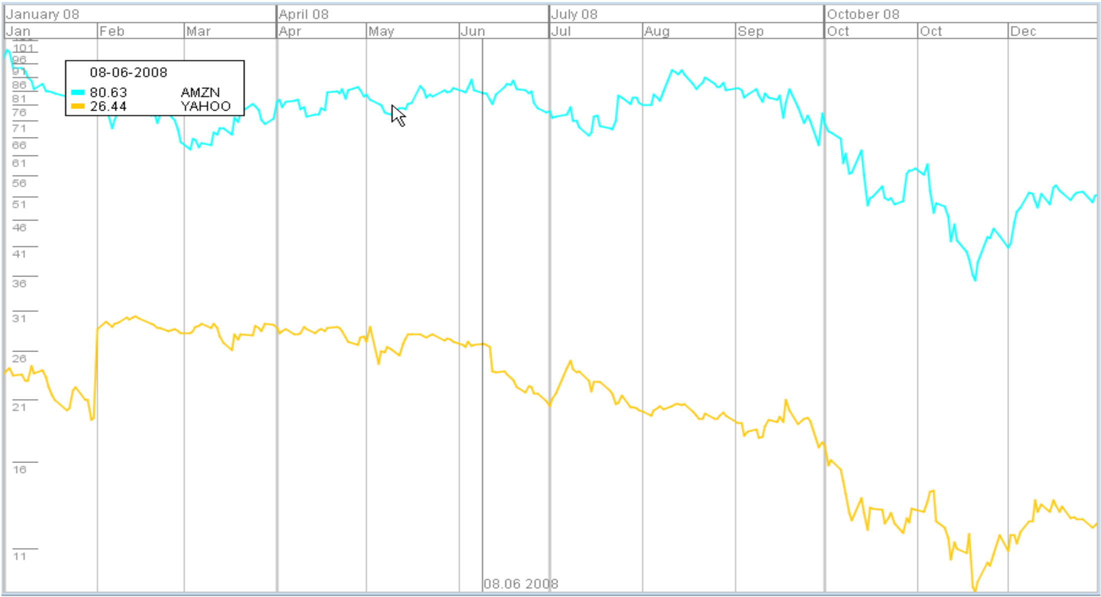
Comparison combination



21. SP500: How much percent did the values change in 2008?

SP500: **(-39.63 %)**

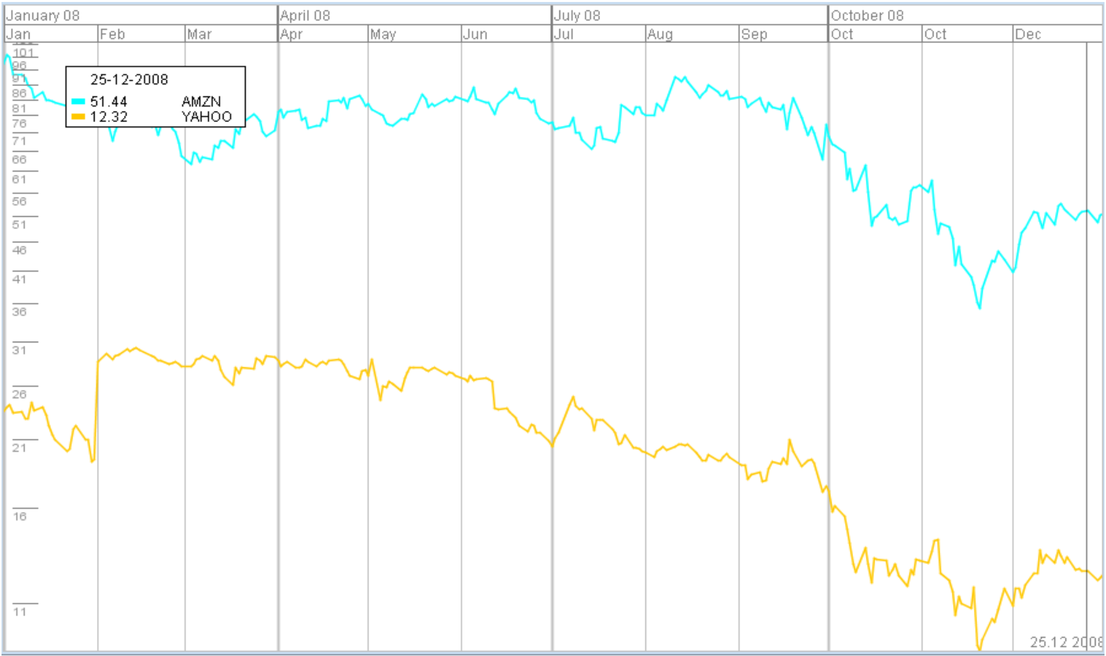
Relation-seeking



22. AMZN: Which of the following months have a higher value than the value on 11.08.2008?

1) **January**
2) February
3) March
4) April
5) May
6) June
7) July
8) August
9) September
10) October
11) November
12) December

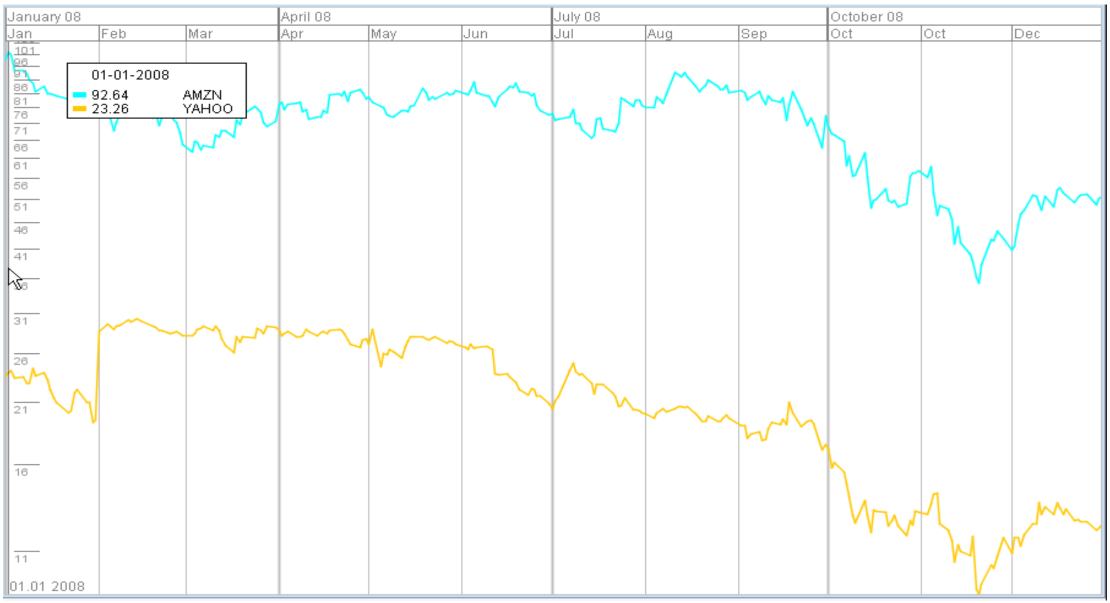
Pattern identification



23. AMZN: Which months in 2008have a positive trend?

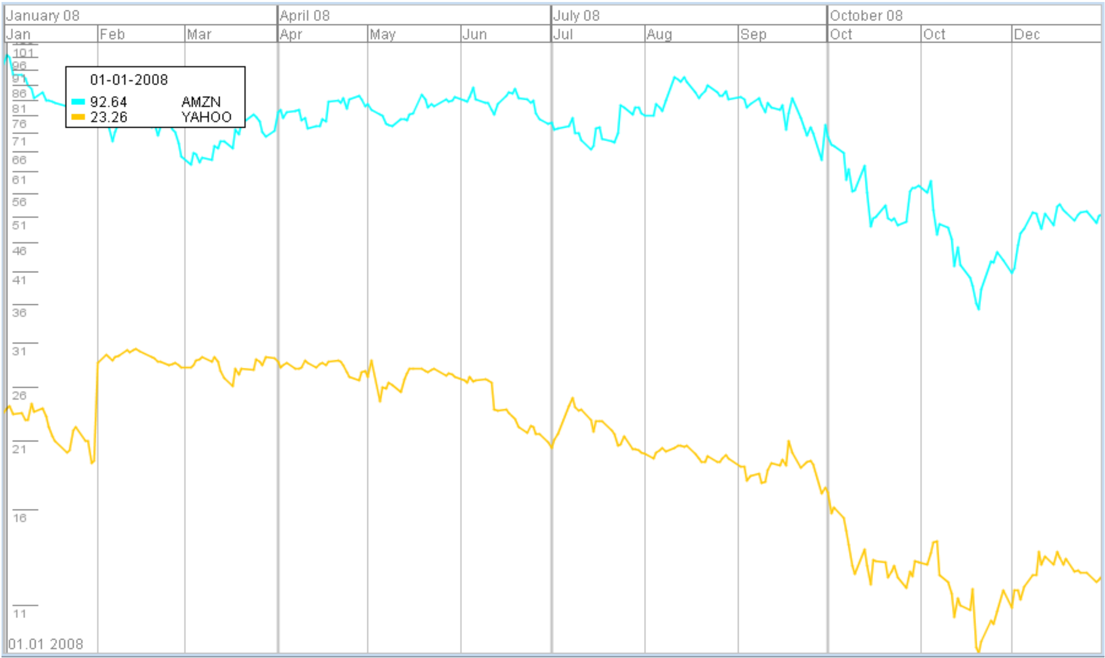
1) January
2) February
3) March
4) April
5) May
6) June
7) July
8) August
9) September
10) October
11) November
12) December

Behavior comparison homogenous



24. Which stock has a bigger percent increase in 01.2008?

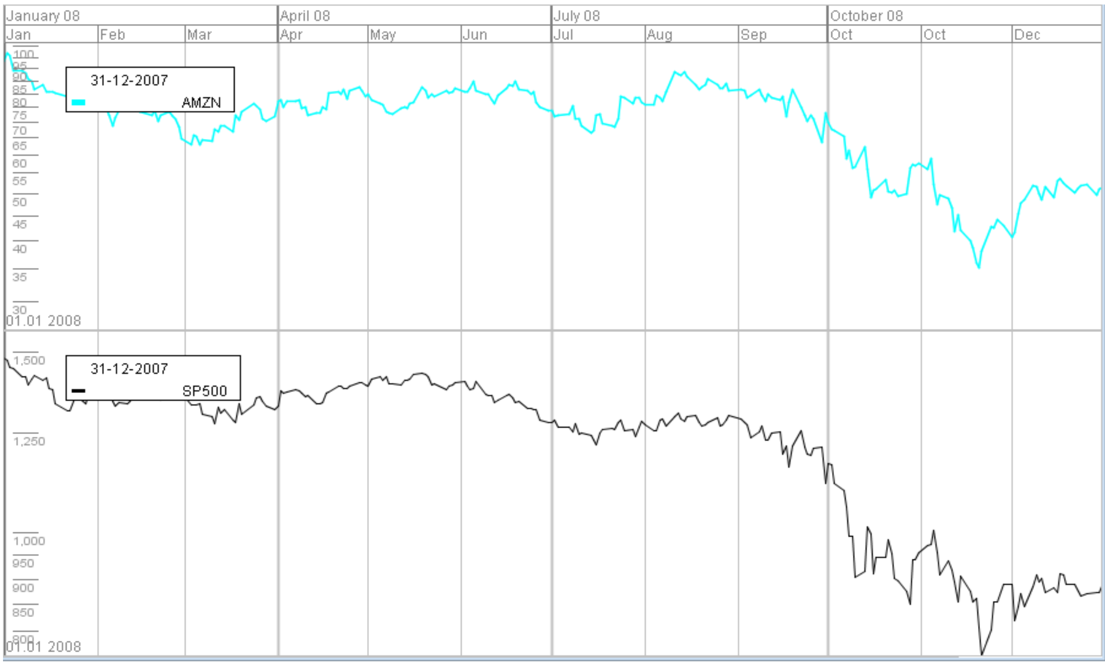
1) **AMZN**
2) YAHOO



25. Which stock has a lower percent loss in 2008?

1) **AMZN** 2) YAHOO

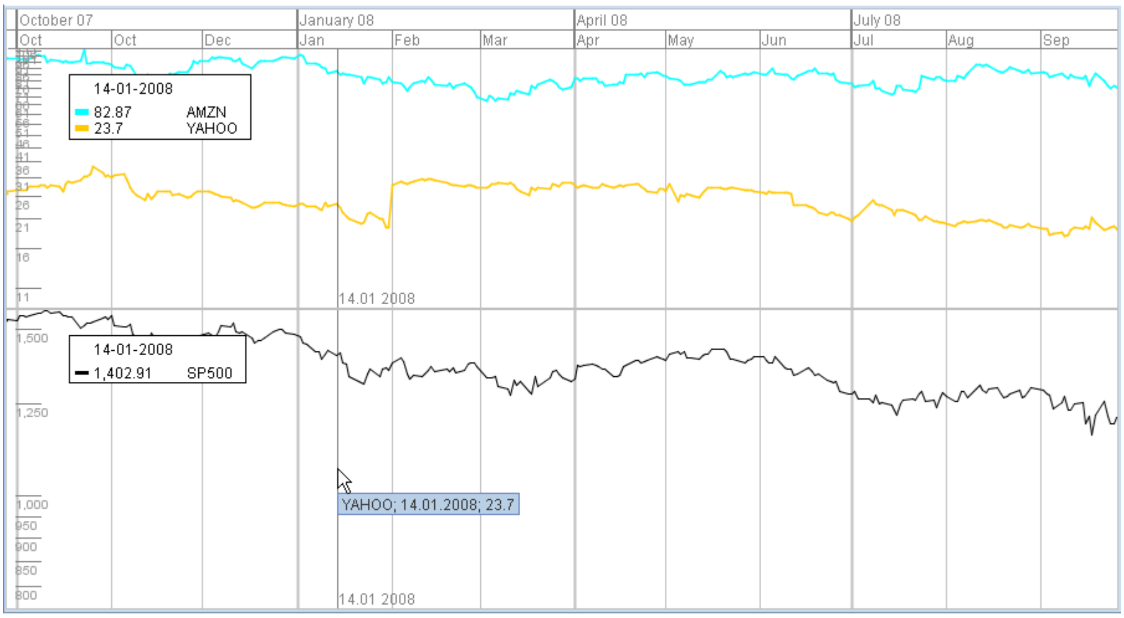
Behavior comparison heterogeneous



26. In which months of 2008 is the percent increase of AMZN greater than SP500?

1) January
2) February
3) March
4) April
5) May
6) June
7) July
8) August
9) September
10) October
11) November
12) December

Behavior comparison combined



27. Which stock or index has the highest volatility (relative variations) in January 2008?

1) AMZN

2) YAHOO

3) SP500

Relation-seeking

28. In which year had AMZN the highest percent increase from beginning to the end of the year?

1) 2004

2) 2005

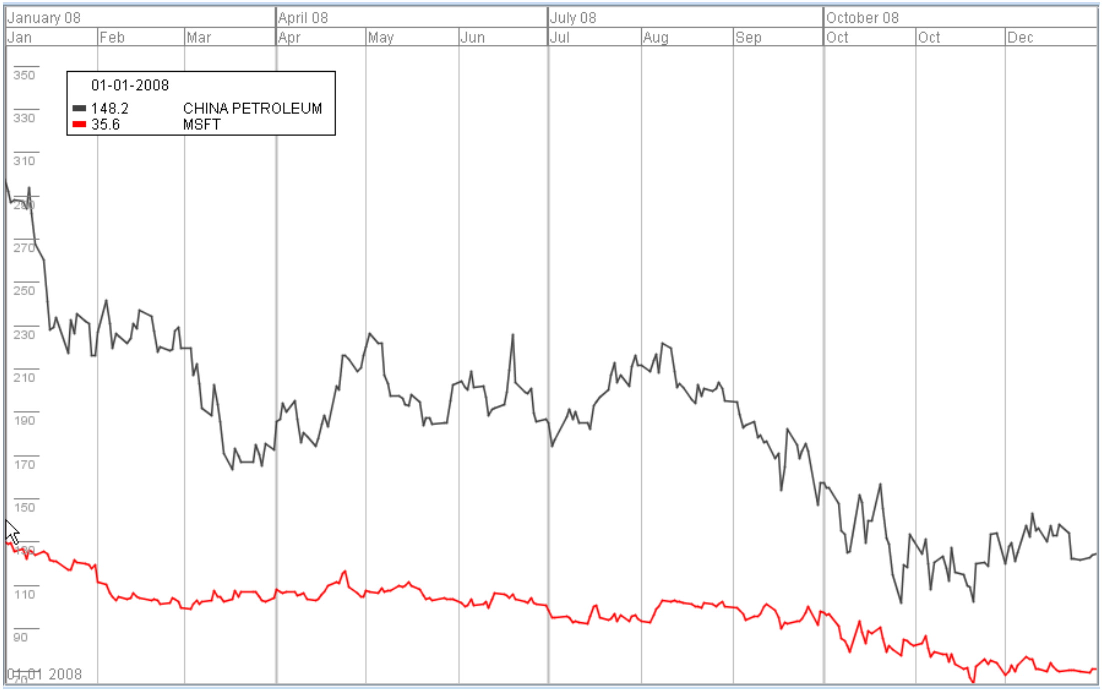
3) 2006

4) 2007

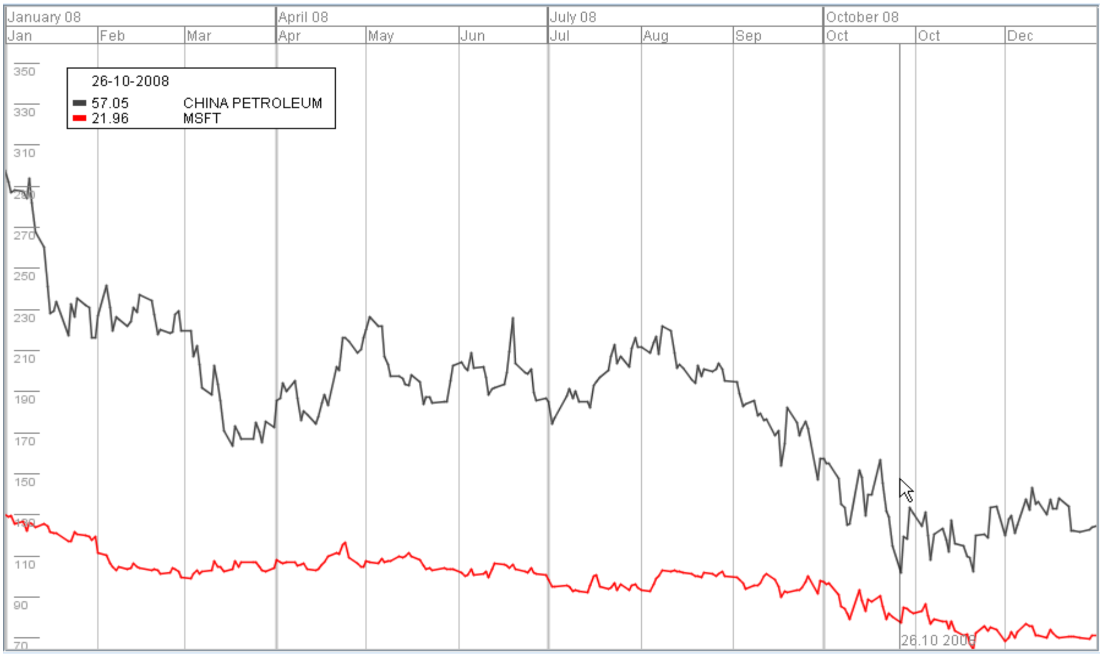
C) Indexing

MSFT vs. CHINA PETROLEUM [DJIA]

Lookup homogenous

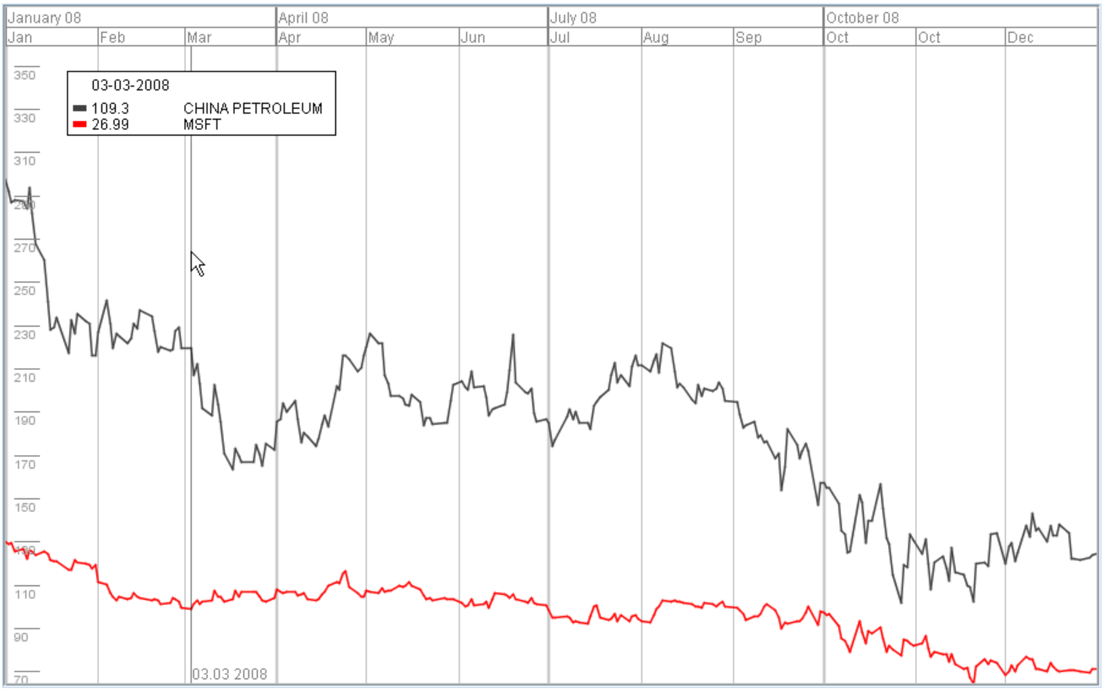


29. MSFT: On which day was the highest stock price in 2008?
(01.01.2008)

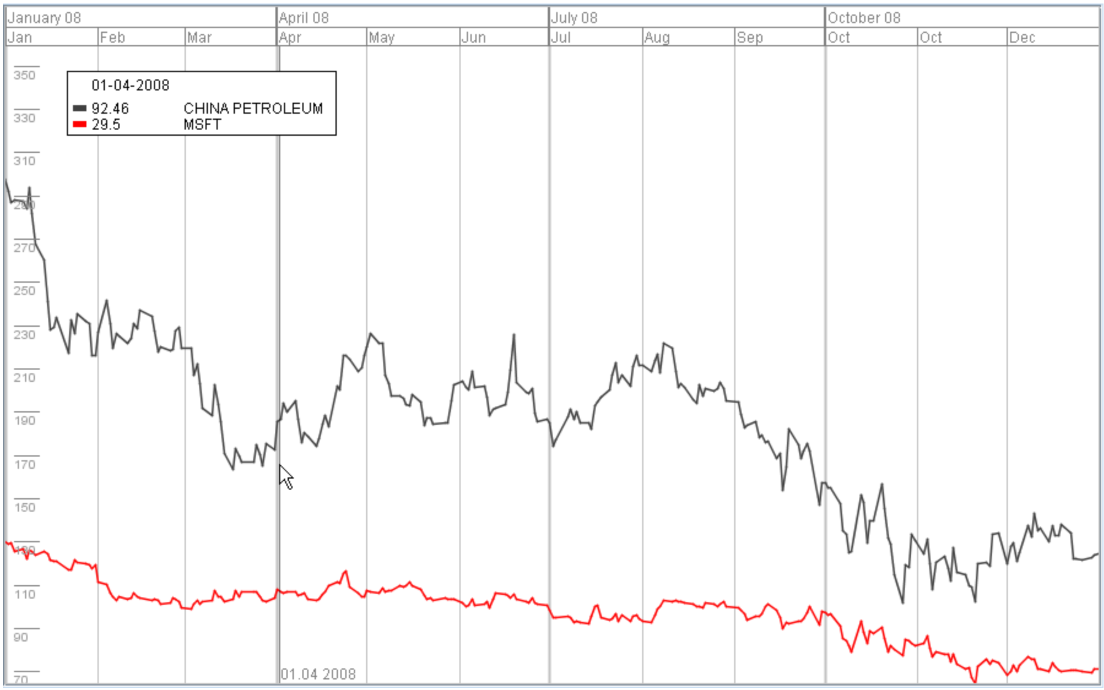


30. CHINA PETROLEUM: On which day was the lowest stock price in 2008?
(27.10.2008)

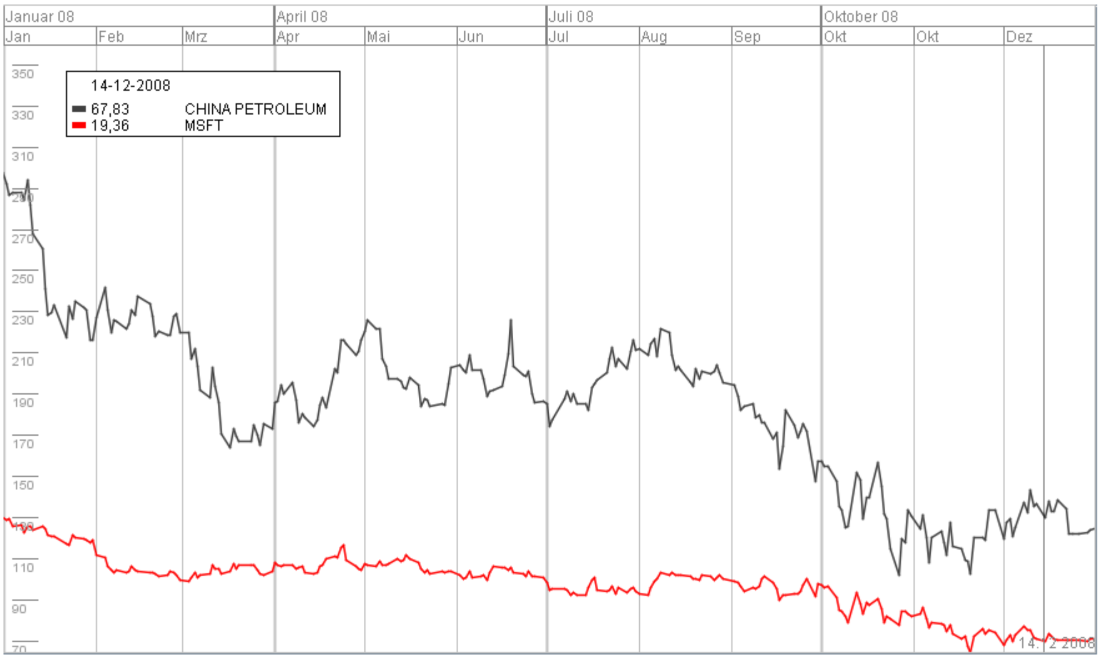
Comparison homogenous



31. Compare the values of CHINA PETROLEUM and MSFT on the given dates - Which of the following statements are valid?
   1) **On 03.03.2008, CHINA** PETROLEUM **was higher than MSFT**
   2) On 16.10.2008, CHINA PETROLEUM was lower than MSFT

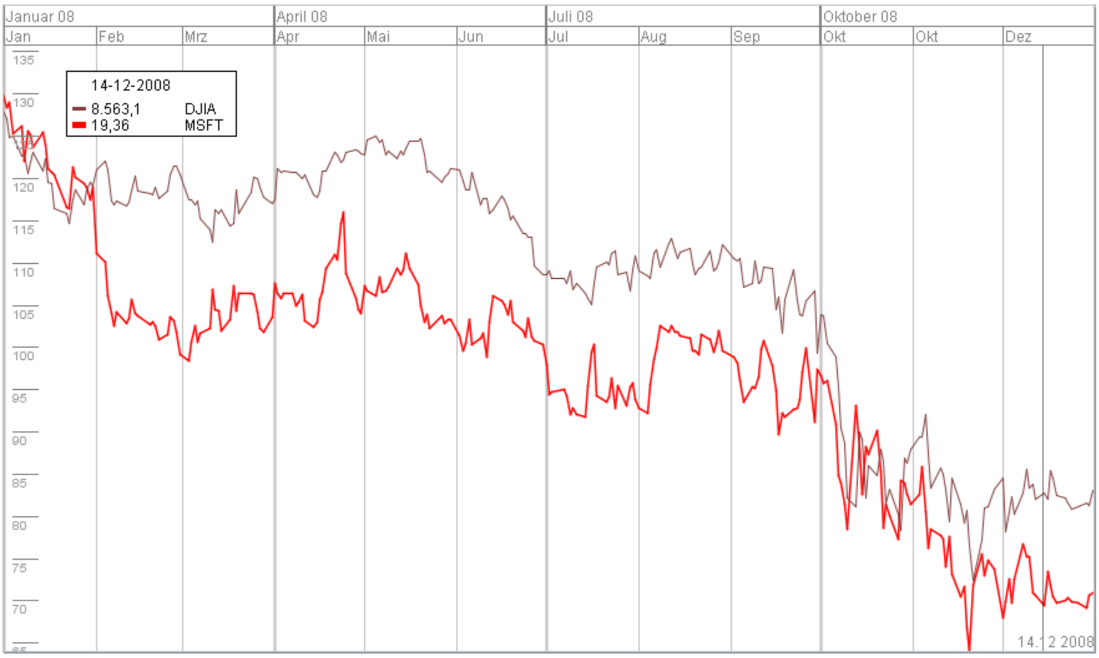


32. Please quantify the amount of price change for the given time periods in dollars:
MSFT, April 2008: **(-0.98 US-$)**
CHINA PETROLEUM, April 2008: **(15.05 US-$)**



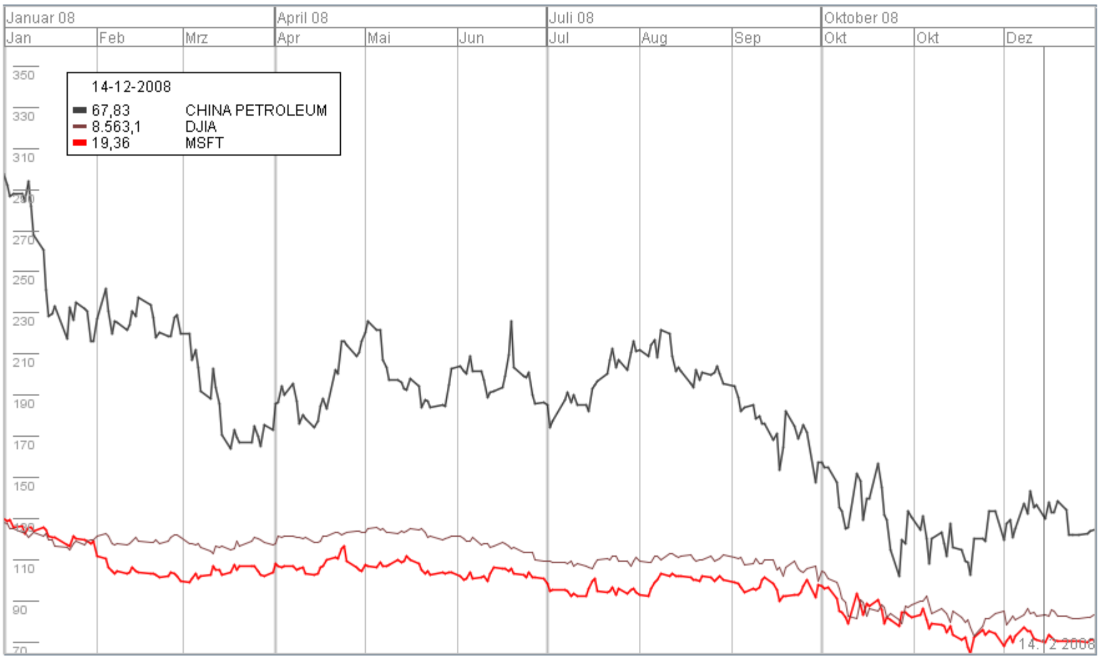
33. Please quantify the amount of price change for the given time periods in percent:
MSFT, January - June 2008: **(-22.72 %)**
CHINA PETROLEUM, January - June 2008: **(-37.32 %)**

Comparison heterogeneous



34. Compare the values of MSFT and the DJIA index on the given dates - Which of the following statements are valid?
    1) On 31.12.2008 was the value of DJIA under 10000
    2) **On 06.07.2008 was the value of MSFT over 45**

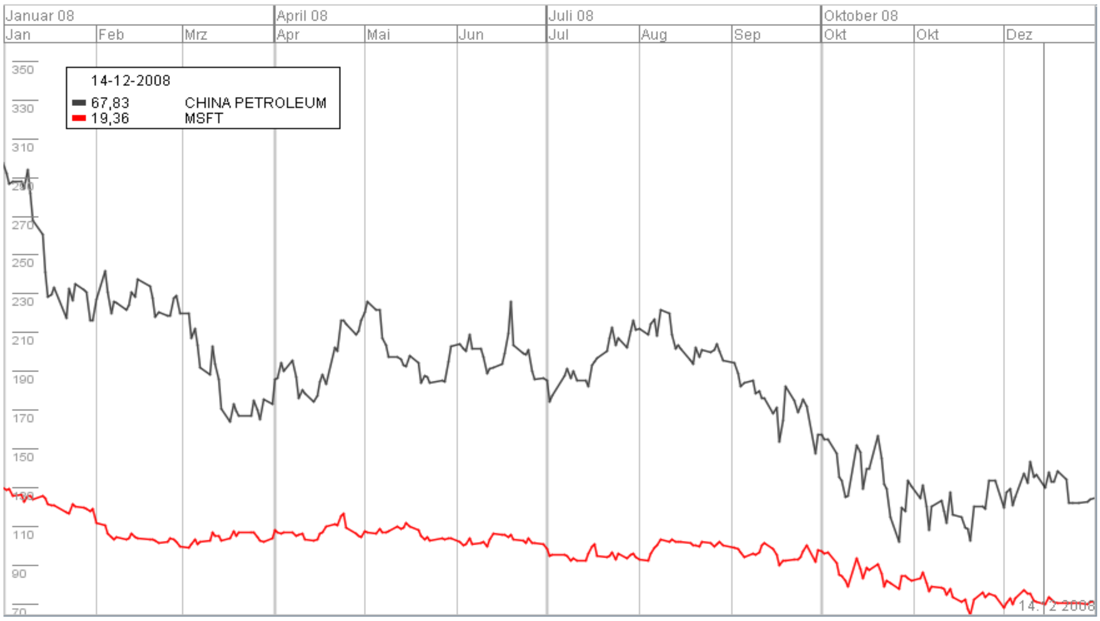
Comparison combination



35. How much percent did the DJIA change in 2008?

(-35.15 %)

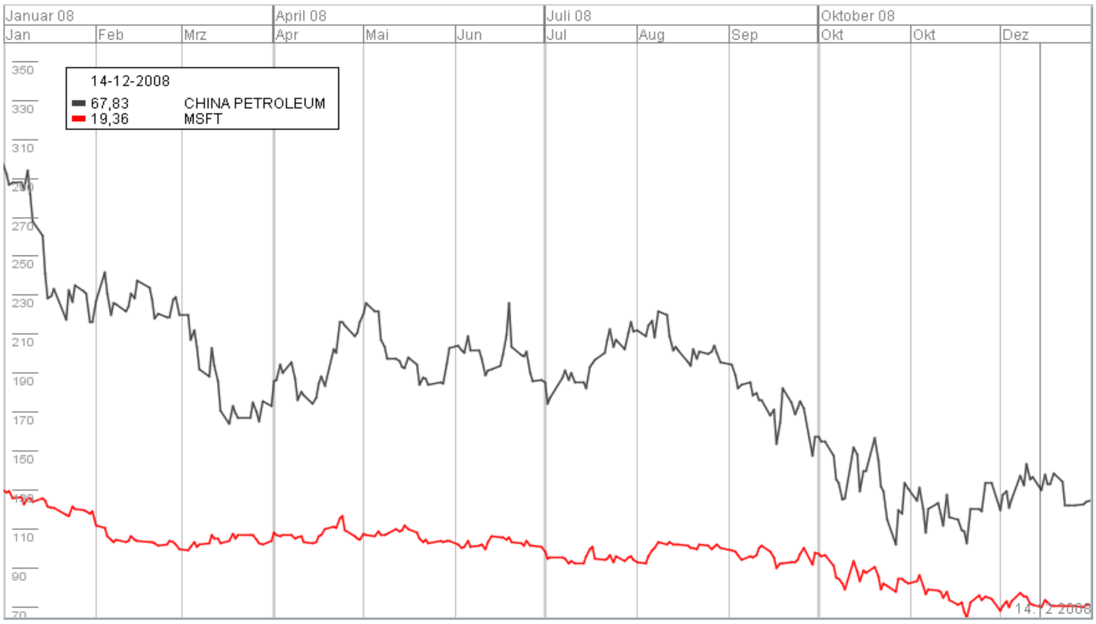
Relation-seeking



36. MSFT: Which of the following months have a higher value than the value on 24.04.2008?

1) **January**
2) February
3) March
4) April
5) May
6) June
7) July
8) August
9) September
10) October
11) November
12) December

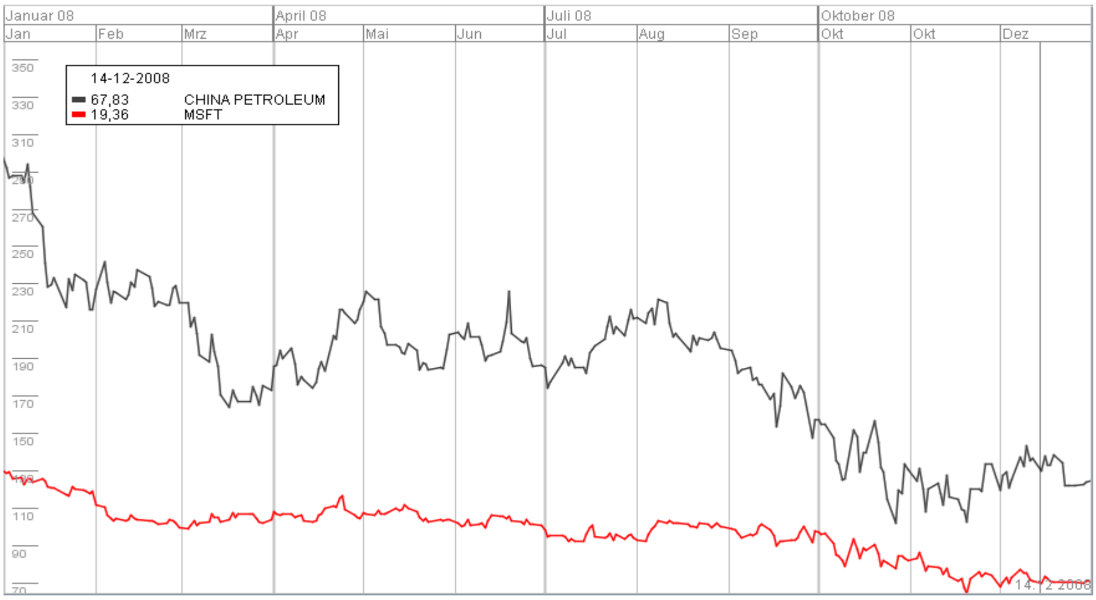
Pattern identification



37. MSFT: Which months in 2008 have a positive trend?

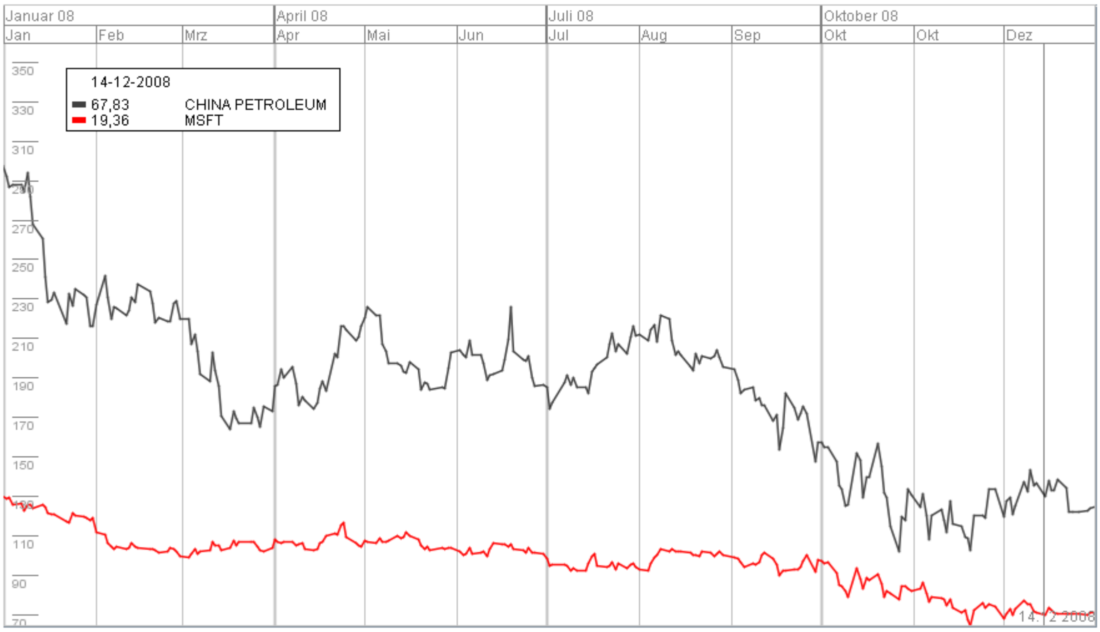
1) January
2) February
3) **March**
4) April
5) May
6) June
7) July
8) **August**
9) September
10) October
11) November
12) **December**

Behavior comparison homogenous



38. Which stock has a bigger percent increase from the beginning of 01.2008 to the end of 03.2008?

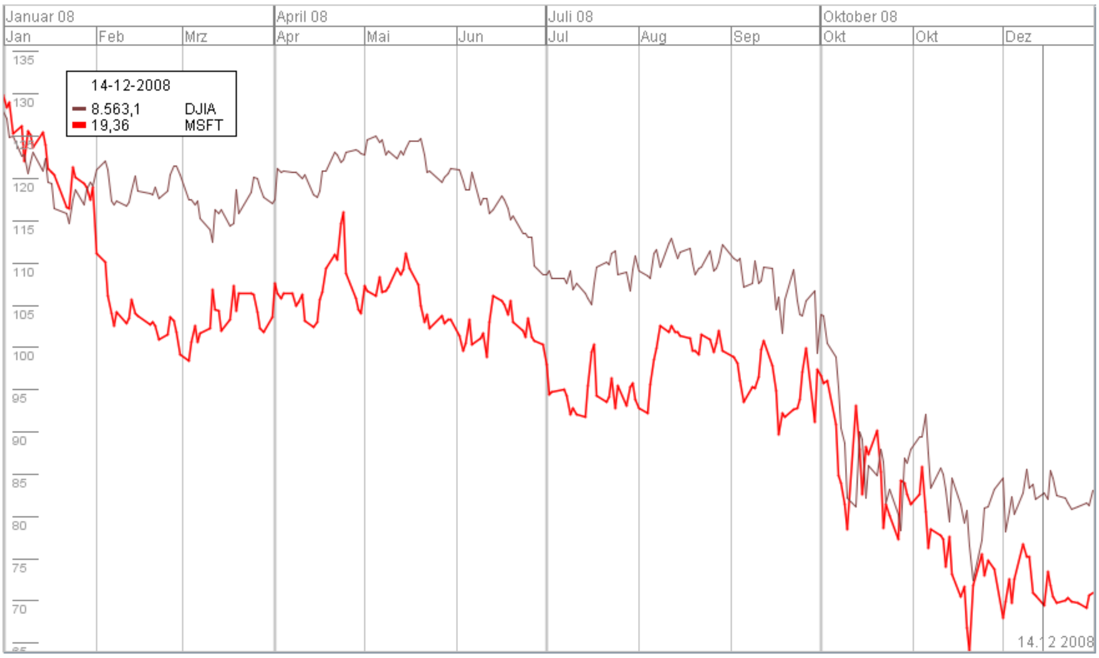
1) CHINA PETROLEUM
2) **MSFT**



39. Which stock has a lower percent loss in 2008?

1) CHINA PETROLEUM
2) **MSFT**

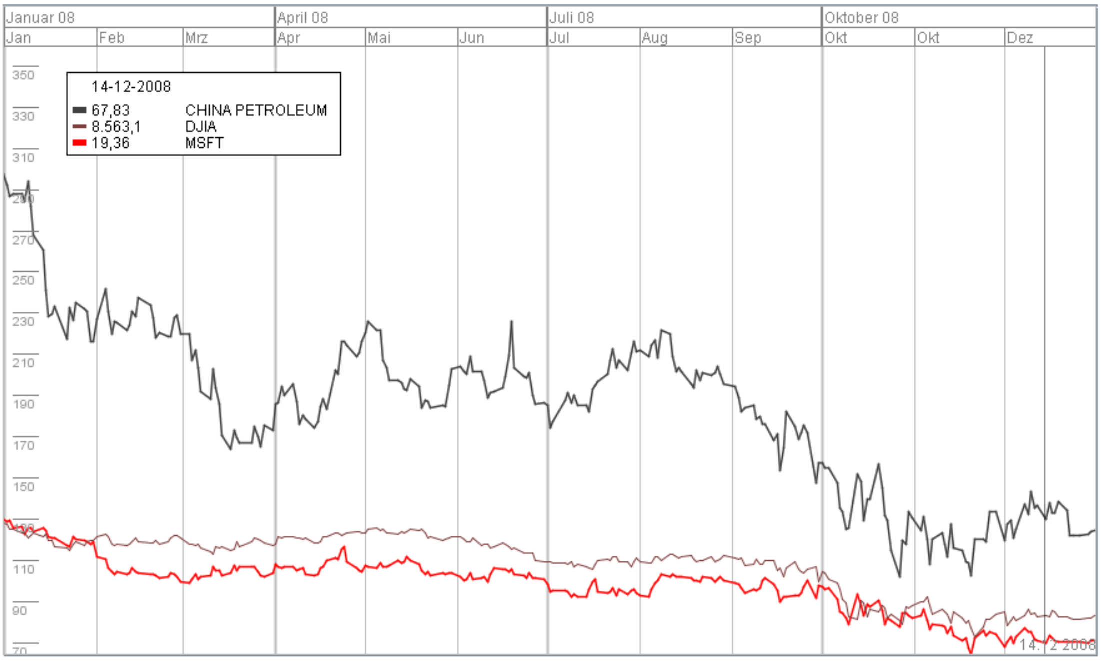
Behavior comparison heterogeneous



40. In which months of 2008 is the percent increase of MSFT greater than DJIA?

1) January
2) February
3) **March**
4) April
5) May
6) **June**
7) July
8) **August**
9) **September**
10) October
11) November
12) **December**

Behavior comparison combined



41. Which stock or index has the highest volatility (relative variations) in January 2008?

1) **CHINA PETROLEUM**
2) MSFT
3) DJIA

Relation-seeking

42. In which year had MSFT the highest percent increase from beginning to the end of the year?

1) 2004
2) 2005
3) 2006
4) **2007**

Appendix C – Raw Data

Test Person 1

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	A	5812	Q1	true	01.01.2008	01.01.2008
2	A	25890	Q30	true	27.10.2008	27.10.2008
3	A	27078	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
4	A	48343	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
5	A	60766	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
6	A	27312	Q34	false	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000 On 06.07.2008 was the value of MSFT over 45
7	A	35672	Q35	false	-35.15 %	-20 < ... <= 30 %
8	A	18984	Q36	true	January	January
9	A	22297	Q23	false	March April May July August December	March April July August
10	A	13906	Q24	false	AMZN	YAHOO
11	A	20766	Q25	true	AMZN	AMZN
12	A	57000	Q26	false	March May July August September October December	March April July October December
13	A	19610	Q41	false	CHINA_PETROLEUM	DJIA
14	A	13735	Q28	true	2007	2007
15	B	23984	Q15	true	02.01.2008	02.01.2008
16	B	16813	Q2	true	20.11.2008	20.11.2008
17	B	25875	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	B	23000	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
19	B	40171	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
20	B	24234	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	B	13265	Q7	true	-42.7 %	-40 < ... <= 50 %
22	B	17985	Q22	true	January	January
23	B	23203	Q37	false	March August December	March April
24	B	19250	Q10	true	AAPL	AAPL
25	B	7203	Q11	true	IBM	IBM
26	B	26891	Q12	false	March April August October	April August November
27	B	11375	Q13	false	AAPL	NASDAQ
28	B	57625	Q42	true	2007	2007
29	C	21250	Q29	true	01.01.2008	01.01.2008
30	C	18844	Q16	true	20.11.2008	20.11.2008
31	C	26500	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
32	C	54766	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
33	C	47734	Q19	true	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -10 < ... <= 20 %
34	C	40672	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
35	C	26110	Q21	true	-39.63 %	-30 < ... <= 40 %
36	C	33203	Q8	true	January May	January May
37	C	24391	Q9	false	March April May August	March April May August October

38	C	26156	Q38	true	MSFT	MSFT
39	C	12282	Q39	true	MSFT	MSFT
40	C	118969	Q40	false	March June August September December	March June August September October December
41	C	27125	Q27	false	AMZN	YAHOO
42	C	24422	Q14	false	2004	2007

Test Person 2

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	B	16485	Q15	true	02.01.2008	02.01.2008
2	B	24484	Q2	true	20.11.2008	20.11.2008
3	B	11204	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	B	34703	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	B	60844	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -10 <= ... <= 20 %
6	B	28109	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	B	55672	Q7	true	-42.7 %	-40 < ... <= 50 %
8	B	16375	Q22	true	January	January
9	B	32594	Q23	true	March April May July August December	March April May July August December
10	B	20000	Q10	true	AAPL	AAPL
11	B	28641	Q39	true	MSFT	MSFT
12	B	94188	Q26	false	March May July August September October December	March May June July August October December
13	B	28687	Q13	false	AAPL	NASDAQ
14	B	36937	Q42	true	2007	2007
15	C	19281	Q1	true	01.01.2008	01.01.2008
16	C	28547	Q30	true	27.10.2008	27.10.2008
17	C	59359	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
18	C	58375	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
19	C	43953	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
20	C	41344	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	C	27454	Q35	true	-35.15 %	-30 < ... <= 40 %
22	C	20203	Q36	true	January	January
23	C	113032	Q37	true	March August December	March August December
24	C	21735	Q38	true	MSFT	MSFT
25	C	14875	Q11	true	IBM	IBM
26	C	110813	Q12	true	March April August October	March April August October
27	C	24765	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
28	C	58688	Q28	true	2007	2007
29	A	21266	Q29	true	01.01.2008	01.01.2008
30	A	27828	Q16	true	20.11.2008	20.11.2008
31	A	28000	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
32	A	30188	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
33	A	76359	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -0 <= ... <= 10 %

34	A	28296	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	A	54125	Q21	false	-39.63 %	-40 < ... <= 50 %
36	A	22579	Q8	false	January May	January
37	A	27281	Q9	false	March April May August	March April May August October
38	A	27093	Q24	false	AMZN	YAHOO
39	A	36110	Q25	false	AMZN	YAHOO
40	A	61281	Q40	false	March June August September December	March June August September October December
41	A	16391	Q27	false	AMZN	YAHOO
42	A	30921	Q14	false	2004	2007

Test Person 3

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	C	43234	Q1	false	01.01.2008	13.01.2008
2	C	37907	Q16	false	20.11.2008	01.07.2008
3	C	51421	Q31	false	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 16.10.2008 CHINA PETROLEUM was lower than MSF
4	C	62766	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	C	58578	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -0 <= ... <= 10 %
6	C	47828	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	C	30922	Q7	false	-42.7 %	-20 < ... <= 30 %
8	C	74625	Q8	true	January May	January May
9	C	48219	Q23	false	March April May July August December	March May July August
10	C	52500	Q10	true	AAPL	AAPL
11	C	19984	Q39	true	MSFT	MSFT
12	C	177047	Q26	false	March May July August September October December	August September
13	C	24297	Q27	false	AMZN	YAHOO
14	C	42016	Q14	false	2004	2007
15	A	29312	Q29	false	01.01.2008	24.04.2008
16	A	19375	Q30	true	27.10.2008	27.10.2008
17	A	30109	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	A	40906	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
19	A	46422	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
20	A	29406	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	A	64719	Q21	false	-39.63 %	-40 < ... <= 50 %
22	A	37781	Q36	false	January	January February
23	A	44141	Q9	false	March April May August	March April May August October
24	A	25859	Q24	false	AMZN	YAHOO
25	A	38094	Q11	true	IBM	IBM
26	A	171062	Q12	false	March April August October	March April May August
27	A	42547	Q13	true	AAPL	AAPL
28	A	23204	Q28	true	2007	2007
29	B	13437	Q15	false	02.01.2008	01.01.2008
30	B	25125	Q2	true	20.11.2008	20.11.2008
31	B	25282	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO

32	B	54359	Q18	false	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -20 < ... <= 30 US-$ YAHOO January 2008: +20 < ... <= 30 US-$
33	B	45047	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -10 <= ... <= 20 %
34	B	74812	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	B	57110	Q35	true	-35.15 %	-30 < ... <= 40 %
36	B	36515	Q22	true	January	January
37	B	59625	Q37	false	March August December	March April August December
38	B	34063	Q38	true	MSFT	MSFT
39	B	36469	Q25	true	AMZN	AMZN
40	B	78391	Q40	false	March June August September December	March August September October
41	B	61265	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
42	B	37938	Q42	true	2007	2007

Test Person 4

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	A	10500	Q1	true	01.01.2008	01.01.2008
2	A	11592	Q30	true	27.10.2008	27.10.2008
3	A	15098	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	A	48431	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
5	A	74955	Q19	true	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -10 < ... <= 20 %
6	A	31683	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
7	A	58731	Q7	true	-42.7 %	-40 < ... <= 50 %
8	A	18975	Q22	true	January	January
9	A	38936	Q37	true	March August December	March August December
10	A	24890	Q10	true	AAPL	AAPL
11	A	16910	Q25	true	AMZN	AMZN
12	A	72027	Q40	false	March June August September December	June July
13	A	18519	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
14	A	40739	Q14	true	2004	2004
15	C	14695	Q29	true	01.01.2008	01.01.2008
16	C	18925	Q2	true	20.11.2008	20.11.2008
17	C	33571	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	C	30430	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	C	20944	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
20	C	26112	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	C	22719	Q21	true	-39.63 %	-30 < ... <= 40 %
22	C	21059	Q36	true	January	January
23	C	30566	Q9	true	March April May August	March April May August
24	C	12430	Q24	true	AMZN	AMZN
25	C	6187	Q39	true	MSFT	MSFT
26	C	121534	Q26	true	March May July August September October December	March May July August September October December
27	C	15404	Q13	true	AAPL	AAPL

28	C	39107	Q42	true	2007	2007
29	B	13557	Q15	true	02.01.2008	02.01.2008
30	B	14926	Q16	true	20.11.2008	20.11.2008
31	B	8434	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	B	30366	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
33	B	81115	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -40 < ... <= 50 % MSFT January - June 2008: -30 < ... <= 40 %
34	B	20180	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
35	B	51147	Q35	true	-35.15 %	-30 < ... <= 40 %
36	B	23722	Q8	true	January May	January May
37	B	31527	Q23	true	March April May July August December	March April May July August December
38	B	75211	Q38	true	MSFT	MSFT
39	B	15432	Q11	true	IBM	IBM
40	B	52654	Q12	false	March April August October	March April June October November
41	B	14599	Q27	false	AMZN	SP500
42	B	7721	Q28	true	2007	2007

Test Person 5

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	C	13296	Q29	true	01.01.2008	01.01.2008
2	C	16516	Q16	true	20.11.2008	20.11.2008
3	C	34672	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	C	35985	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	C	28923	Q19	true	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -10 < ... <= 20 %
6	C	22172	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	C	17376	Q21	true	-39.63 %	-30 < ... <= 40 %
8	C	35406	Q8	true	January May	January May
9	C	58079	Q37	true	March August December	March August December
10	C	25172	Q10	true	AAPL	AAPL
11	C	12374	Q11	true	IBM	IBM
12	C	88406	Q12	true	March April August October	March April August October
13	C	12172	Q13	true	AAPL	AAPL
14	C	31640	Q42	true	2007	2007
15	B	15297	Q15	true	02.01.2008	02.01.2008
16	B	14093	Q2	true	20.11.2008	20.11.2008
17	B	21047	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	B	26453	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	B	40704	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -40 < ... <= 50 % MSFT January - June 2008: -20 < ... <= 30 %
20	B	22625	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	B	64484	Q7	false	-42.7 %	-50 < ... <= 60 %
22	B	40844	Q22	true	January	January
23	B	26515	Q9	true	March April May August	March April May August
24	B	13953	Q38	true	MSFT	MSFT
25	B	9406	Q39	true	MSFT	MSFT

26	B	76781	Q40	false	March June August September December	January March April June July August September
27	B	12047	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
28	B	32703	Q14	false	2004	2007
29	A	15516	Q1	true	01.01.2008	01.01.2008
30	A	23937	Q30	true	27.10.2008	27.10.2008
31	A	22563	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	A	34625	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
33	A	60469	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
34	A	19344	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	A	54141	Q35	true	-35.15 %	-30 < ... <= 40 %
36	A	18375	Q36	true	January	January
37	A	39392	Q23	true	March April May July August December	March April May July August December
38	A	21641	Q24	false	AMZN	YAHOO
39	A	21500	Q25	false	AMZN	YAHOO
40	A	73578	Q26	false	March May July August September October December	March April May July August December
41	A	14219	Q27	true	AMZN	AMZN
42	A	18922	Q28	true	2007	2007

Test Person 6

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	B	14193	Q29	true	01.01.2008	01.01.2008
2	B	25697	Q30	true	27.10.2008	27.10.2008
3	B	47700	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
4	B	31479	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
5	B	33738	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +20 < ... <= 30 %
6	B	30076	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	B	67705	Q7	true	-42.7 %	-40 < ... <= 50 %
8	B	40014	Q36	true	January	January
9	B	28900	Q23	true	March April May July August December	March April May July August December
10	B	15431	Q10	true	AAPL	AAPL
11	B	13707	Q25	false	AMZN	YAHOO
12	B	85926	Q40	false	March June August September December	June July August September October
13	B	16804	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
14	B	23903	Q42	true	2007	2007
15	A	26451	Q15	false	02.01.2008	01.01.2008
16	A	18550	Q16	true	20.11.2008	20.11.2008
17	A	41412	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	A	34628	Q32	false	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: +0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
19	A	55076	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -50 < ... <= 60 % MSFT January - June 2008: -10 < ... <= 20 %

20	A	22597	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	A	52914	Q35	true	-35.15 %	-30 < ... <= 40 %
22	A	20361	Q22	true	January	January
23	A	31441	Q9	false	March April May August	March April May August October
24	A	59575	Q38	true	MSFT	MSFT
25	A	14064	Q39	true	MSFT	MSFT
26	A	149858	Q26	false	March May July August September October December	January February March April May July August December
27	A	14552	Q13	false	AAPL	IBM
28	A	12494	Q14	false	2004	2007
29	C	16078	Q1	true	01.01.2008	01.01.2008
30	C	13223	Q2	false	20.11.2008	19.11.2008
31	C	35531	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
32	C	32064	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
33	C	28882	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -10 < ... <= 20 % YAHOO January - June 2008: -0 <= ... <= 10 %
34	C	37998	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	C	21562	Q21	true	-39.63 %	-30 < ... <= 40 %
36	C	28275	Q8	true	January May	January May
37	C	27974	Q37	false	March August December	March April August December
38	C	10438	Q24	true	YAHOO	YAHOO
39	C	17276	Q11	true	IBM	IBM
40	C	109240	Q12	false	March April August October	March April May June August October
41	C	10766	Q27	true	AMZN	AMZN
42	C	8192	Q28	true	2007	2007

Test Person 7

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	A	13562	Q15	true	02.01.2008	02.01.2008
2	A	15375	Q16	true	20.11.2008	20.11.2008
3	A	21110	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
4	A	30719	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	A	65296	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
6	A	18250	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	A	90109	Q35	false	-35.15 %	-40 < ... <= 50 %
8	A	25344	Q36	true	January	January
9	A	43984	Q23	true	March April May July August December	March April May July August December
10	A	34906	Q10	true	AAPL	AAPL
11	A	14563	Q25	false	AMZN	YAHOO
12	A	104407	Q12	false	March April August October	March April June October November
13	A	18187	Q27	false	AMZN	SP500
14	A	15187	Q28	true	2007	2007
15	B	15516	Q1	true	01.01.2008	01.01.2008
16	B	19359	Q2	true	20.11.2008	20.11.2008
17	B	20281	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	B	38906	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$

19	B	44110	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -10 <= ... <=20 % YAHOO January - June 2008: -10 < ... <= 20 %
20	B	17781	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	B	39484	Q7	true	-42.7 %	-40 < ... <= 50 %
22	B	28875	Q8	true	January May	January May
23	B	34093	Q37	false	March August December	March April August December
24	B	20203	Q24	false	AMZN	YAHOO
25	B	15906	Q39	true	MSFT	MSFT
26	B	53344	Q26	false	March May July August September October December	March May June October December
27	B	21954	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
28	B	14781	Q14	false	2004	2007
29	C	13657	Q29	true	01.01.2008	01.01.2008
30	C	23047	Q30	true	27.10.2008	27.10.2008
31	C	16687	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	C	41343	Q18	false	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -10 < ... <= 20 US-$
33	C	38265	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
34	C	19860	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	C	13078	Q21	true	-39.63 %	-30 < ... <= 40 %
36	C	12797	Q22	true	January	January
37	C	45922	Q9	true	March April May August	March April May August
38	C	12297	Q38	true	MSFT	MSFT
39	C	9844	Q11	true	IBM	IBM
40	C	129953	Q40	true	March June August September December	March June August September December
41	C	12546	Q13	true	AAPL	AAPL
42	C	14406	Q42	true	2007	2007

Test Person 8

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	B	16500	Q29	true	01.01.2008	01.01.2008
2	B	29750	Q16	true	20.11.2008	20.11.2008
3	B	20313	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
4	B	32750	Q32	false	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +0 <= ... <= 10 US-$
5	B	36719	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 < ... <= 10 % IBM January - May 2008: +20 <= ... <= 30 %
6	B	35234	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	B	29390	Q7	true	-42.7 %	-40 < ... <= 50 %
8	B	14484	Q22	true	January	January
9	B	29188	Q9	true	March April May August	March April May August
10	B	11422	Q38	true	MSFT	MSFT
11	B	9718	Q39	true	MSFT	MSFT
12	B	33781	Q12	false	March April August October	February March April June October November
13	B	17735	Q27	true	AMZN	AMZN
14	B	12875	Q28	true	2007	2007
15	C	27781	Q1	true	01.01.2008	01.01.2008
16	C	22907	Q30	true	27.10.2008	27.10.2008

17	C	61140	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	C	31187	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	C	13313	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
20	C	22765	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	C	10281	Q21	true	-39.63 %	-30 < ... <= 40 %
22	C	24515	Q8	true	January May	January May
23	C	51453	Q37	true	March August December	March August December
24	C	48500	Q24	true	AMZN	AMZN
25	C	7328	Q11	true	IBM	IBM
26	C	64437	Q40	true	March June August September December	March June August September December
27	C	10531	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
28	C	36344	Q42	true	2007	2007
29	A	13703	Q15	true	02.01.2008	02.01.2008
30	A	26906	Q2	false	20.11.2008	19.11.2008
31	A	21406	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	A	28141	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
33	A	47828	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: +0 <= ... <= 10 %
34	A	18032	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	A	64125	Q35	true	-35.15 %	-30 < ... <= 40 %
36	A	26078	Q36	true	January	January
37	A	25547	Q23	true	March April May July August December	March April May July August December
38	A	19766	Q10	true	AAPL	AAPL
39	A	36062	Q25	false	AMZN	YAHOO
40	A	36828	Q26	false	March May July August September October December	March May July August December
41	A	7078	Q13	false	AAPL	IBM
42	A	12204	Q14	false	2004	2007

Test Person 9

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	C	2641	Q29	true	01.01.2008	01.01.2008
2	C	17470	Q16	true	20.11.2008	20.11.2008
3	C	16671	Q3	false	On 3.3.2008 AAPL was higher than IBM	On 16.10.20008 AAPL was lower than IBM
4	C	66016	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	C	43079	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -10 < ... <= 20 % YAHOO January - June 2008: -0 <= ... <= 10 %
6	C	31437	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	C	13469	Q7	true	-42.7 %	-40 < ... <= 50 %
8	C	3484	Q8	true	January May	January May
9	C	46453	Q37	true	March August December	March August December
10	C	12655	Q38	true	MSFT	MSFT
11	C	17516	Q39	true	MSFT	MSFT
12	C	138344	Q12	true	March April August October	March April August October
13	C	18671	Q27	true	AMZN	AMZN

14	C	9095	Q28	true	2007	2007
15	A	16266	Q15	true	02.01.2008	02.01.2008
16	A	22984	Q30	true	27.10.2008	27.10.2008
17	A	31501	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	A	24547	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	A	65328	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
20	A	23204	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
21	A	58811	Q21	false	-39.63 %	-40 < ... <= 50 %
22	A	23234	Q22	true	January	January
23	A	18735	Q9	true	March April May August	March April May August
24	A	62531	Q24	true	AMZN	AMZN
25	A	19906	Q11	true	IBM	IBM
26	A	59532	Q26	false	March May July August September October December	March April May June July August November December
27	A	35796	Q41	false	CHINA_PETROLEUM	MSFT
28	A	15125	Q14	false	2004	2007
29	B	9532	Q1	true	01.01.2008	01.01.2008
30	B	12360	Q2	true	20.11.2008	20.11.2008
31	B	22531	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	B	29782	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: 10 < ... <= 20 US-$
33	B	52375	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
34	B	23859	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
35	B	31500	Q35	false	-35.15 %	-40 < ... <= 50 %
36	B	19515	Q36	true	January	January
37	B	26671	Q23	true	March April May July August December	March April May July August December
38	B	28344	Q10	true	AAPL	AAPL
39	B	26203	Q25	false	AMZN	YAHOO
40	B	56235	Q40	false	March June August September December	March June August September October
41	B	27375	Q13	true	AAPL	AAPL
42	B	9610	Q42	true	2007	2007

Test Person 10

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	A	17016	Q1	true	01.01.2008	01.01.2008
2	A	25235	Q16	true	20.11.2008	20.11.2008
3	A	20266	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
4	A	39892	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	A	67437	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
6	A	20703	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	A	43531	Q35	false	-35.15 %	-20 < ... <= 30 %
8	A	26390	Q8	true	January May	January May
9	A	23641	Q37	false	March August December	March April August December

10	A	34109	Q10	true	AAPL	AAPL
11	A	27625	Q25	true	AMZN	AMZN
12	A	26188	Q26	false	March May July August September October December	March August December
13	A	15406	Q27	true	AMZN	AMZN
14	A	14515	Q28	true	2007	2007
15	C	18984	Q15	false	02.01.2008	01.01.2008
16	C	23423	Q30	true	27.10.2008	27.10.2008
17	C	30265	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	C	62532	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
19	C	46188	Q19	true	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -10 < ... <= 20 %
20	C	20547	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	C	13797	Q21	true	-39.63 %	-30 < ... <= 40 %
22	C	21374	Q36	true	January	January
23	C	69577	Q23	true	March April May July August December	March April May July August December
24	C	16266	Q38	true	MSFT	MSFT
25	C	8047	Q39	true	MSFT	MSFT
26	C	81219	Q40	true	March June August September December	March June August September December
27	C	12688	Q13	true	AAPL	AAPL
28	C	51484	Q14	false	2004	2007
29	B	22360	Q29	true	01.01.2008	01.01.2008
30	B	14391	Q2	true	20.11.2008	20.11.2008
31	B	19454	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
32	B	32047	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
33	B	55874	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +20 < ... <= 30 %
34	B	32719	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	B	21735	Q7	true	-42.7 %	-40 < ... <= 50 %
36	B	11954	Q22	true	January	January
37	B	21093	Q9	false	March April May August	March April May August October
38	B	23219	Q24	false	AMZN	YAHOO
39	B	9828	Q11	true	IBM	IBM
40	B	86922	Q12	false	March April August October	March October
41	B	21078	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
42	B	22375	Q42	true	2007	2007

Test Person 11

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	C	34969	Q15	true	02.01.2008	02.01.2008
2	C	69844	Q30	false	27.10.2008	26.10.2008
3	C	17329	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
4	C	33186	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	C	23485	Q19	true	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -10 < ... <= 20 %

6	C	24516	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	C	28890	Q7	true	-42.7 %	-40 < ... <= 50 %
8	C	72781	Q36	false	January	January February
9	C	92313	Q23	true	March April May July August December	March April May July August December
10	C	11188	Q38	true	MSFT	MSFT
11	C	10719	Q11	true	IBM	IBM
12	C	90328	Q40	true	March June August September December	March June August September December
13	C	25452	Q13	true	AAPL	AAPL
14	C	34187	Q28	true	2007	2007
15	B	8469	Q29	true	01.01.2008	01.01.2008
16	B	34952	Q2	true	20.11.2008	20.11.2008
17	B	19312	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	B	29906	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	B	51438	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -10 < ... <= 20 %
20	B	31188	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
21	B	45547	Q35	false	-35.15 %	-20 < ... <= 30 %
22	B	13937	Q22	true	January	January
23	B	30828	Q9	false	March April May August	March April May August October
24	B	12094	Q10	true	AAPL	AAPL
25	B	8703	Q39	true	MSFT	MSFT
26	B	99250	Q12	false	March April August October	February March April June October November
27	B	25782	Q27	true	AMZN	AMZN
28	B	15140	Q14	false	2004	2007
29	A	12187	Q1	true	01.01.2008	01.01.2008
30	A	14890	Q16	true	20.11.2008	20.11.2008
31	A	17469	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	A	33548	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
33	A	66547	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +0 <= ... <= 10 %
34	A	17031	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
35	A	43969	Q21	false	-39.63 %	-40 < ... <= 50 %
36	A	31422	Q8	true	January May	January May
37	A	65843	Q37	false	March August December	March April June August December
38	A	22266	Q24	false	AMZN	YAHOO
39	A	26140	Q25	false	AMZN	YAHOO
40	A	89266	Q26	false	March May July August September October December	March April May June July August October December
41	A	11172	Q41	false	CHINA_PETROLEUM	MSFT
42	A	13782	Q42	true	2007	2007

Test Person 12

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	B	29172	Q15	false	02.01.2008	01.01.2008
2	B	39359	Q16	true	20.11.2008	20.11.2008
3	B	32328	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM

4	B	48625	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	B	69360	Q19	true	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -10 < ... <= 20 %
6	B	25203	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	B	37562	Q35	true	-35.15 %	-30 < ... <= 40 %
8	B	54625	Q22	true	January	January
9	B	52968	Q23	true	March April May July August December	March April May July August December
10	B	18953	Q10	true	AAPL	AAPL
11	B	13546	Q25	false	AMZN	YAHOO
12	B	134797	Q40	false	March June August September December	March June August September October
13	B	17750	Q27	false	AMZN	SP500
14	B	34032	Q14	false	2004	2007
15	A	22203	Q1	true	01.01.2008	01.01.2008
16	A	54766	Q30	true	27.10.2008	27.10.2008
17	A	23875	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	A	42500	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	A	64453	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
20	A	25359	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	A	42766	Q7	true	-42.7 %	-40 < ... <= 50 %
22	A	23734	Q36	true	January	January
23	A	39688	Q37	false	March August December	March August
24	A	45938	Q38	false	MSFT	CHINA PETROLEUM
25	A	16031	Q39	false	MSFT	CHINA PETROLEUM
26	A	107969	Q12	false	March April August October	March April October November
27	A	15985	Q13	false	AAPL	IBM
28	A	35281	Q42	true	2007	2007
29	C	17360	Q29	true	01.01.2008	01.01.2008
30	C	17843	Q2	true	20.11.2008	20.11.2008
31	C	40969	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	C	32688	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
33	C	43187	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
34	C	26000	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	C	13297	Q21	true	-39.63 %	-30 < ... <= 40 %
36	C	70610	Q8	true	January May	January May
37	C	34406	Q9	false	March April May August	March April May August October
38	C	28688	Q24	true	AMZN	AMZN
39	C	32375	Q11	true	IBM	IBM
40	C	124203	Q26	true	March May July August September October December	March May July August September October December
41	C	19219	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
42	C	43110	Q28	true	2007	2007

Test Person 13

task	chart	time [ms]	question	task correctness	correct answers	given answers

1	A	28391	Q15	false	02.01.2008	01.01.2008
2	A	32609	Q16	false	20.11.2008	11.11.2008
3	A	73907	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
4	A	73453	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
5	A	74000	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -10 < ... <= 20 % YAHOO January - June 2008: +20 < ... <= 30 %
6	A	72937	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	A	40296	Q21	true	-39.63 %	-30 < ... <= 40 %
8	A	61532	Q22	true	January	January
9	A	36171	Q23	false	March April May July August December	March April July August December
10	A	53984	Q10	true	AAPL	AAPL
11	A	56187	Q39	false	MSFT	CHINA PETROLEUM
12	A	106422	Q26	false	March May July August September October December	March July August December
13	A	42593	Q13	true	AAPL	AAPL
14	A	33047	Q28	true	2007	2007
15	B	37906	Q1	true	01.01.2008	01.01.2008
16	B	46343	Q30	false	27.10.2008	26.01.2008
17	B	27407	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	B	78126	Q18	false	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -20 < ... <= 30 US-$ YAHOO January 2008: +0 <= ... <= 10 US-$
19	B	53718	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -10 < ... <= 20 %
20	B	53031	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	B	51500	Q7	true	-42.7 %	-40 < ... <= 50 %
22	B	59891	Q36	false	January	January February March May June August
23	B	49406	Q9	false	March April May August	March April May June October
24	B	62266	Q24	false	AMZN	YAHOO
25	B	25047	Q25	true	AMZN	AMZN
26	B	97906	Q12	false	March April August October	March April May August October
27	B	14827	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
28	B	113719	Q14	true	2004	2004
29	C	19312	Q29	true	01.01.2008	01.01.2008
30	C	45939	Q2	false	20.11.2008	31.01.2008
31	C	169047	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	C	124687	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
33	C	39766	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: +0 <= ... <= 10 % IBM January - May 2008: +0 <= ... <= 10 %
34	C	85625	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	C	21734	Q35	false	-35.15 %	-20 < ... <= 30 %
36	C	104812	Q8	false	January May	May July August
37	C	18375	Q37	false	March April August	
38	C	41547	Q38	true	MSFT	MSFT
39	C	16609	Q11	true	IBM	IBM
40	C	111531	Q40	false	March June August September December	June July August September
41	C	16593	Q27	false	AMZN	YAHOO

42	C	57423	Q42	true	2007	2007

Test Person 14

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	B	21813	Q29	true	01.01.2008	01.01.2008
2	B	16750	Q2	true	20.11.2008	20.11.2008
3	B	26546	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	B	56110	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	B	65219	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -40 < ... <= 50 % MSFT January - June 2008: -10 <= ... <= 20 %
6	B	31250	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
7	B	48687	Q21	false	-39.63 %	-40 < ... <= 50 %
8	B	62391	Q8	true	January May	January May
9	B	62656	Q37	false	March August December	March August September December
10	B	24344	Q24	false	AMZN	YAHOO
11	B	26234	Q11	true	IBM	IBM
12	B	72954	Q12	false	March April August October	March April May June October November
13	B	21031	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
14	B	17484	Q14	false	2004	2007
15	C	44266	Q15	true	02.01.2008	02.01.2008
16	C	28297	Q16	false	20.11.2008	20.10.2008
17	C	31766	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
18	C	51235	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
19	C	28609	Q19	true	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -10 < ... <= 20 %
20	C	21828	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	C	21391	Q7	true	-42.7 %	-40 < ... <= 50 %
22	C	34781	Q22	true	January	January
23	C	55813	Q23	true	March April May July August December	March April May July August December
24	C	18657	Q10	true	AAPL	AAPL
25	C	36703	Q25	true	AMZN	AMZN
26	C	90718	Q26	true	March May July August September October December	March May July August September October December
27	C	18030	Q13	true	AAPL	AAPL
28	C	25109	Q28	true	2007	2007
29	A	14078	Q1	true	01.01.2008	01.01.2008
30	A	12559	Q30	true	27.10.2008	27.10.2008
31	A	30468	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
32	A	31015	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
33	A	68844	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
34	A	28812	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
35	A	50281	Q35	true	-35.15 %	-30 < ... <= 40 %
36	A	42484	Q36	true	January	January
37	A	27297	Q9	false	March April May August	March April May August October

38	A	33765	Q38	false	MSFT	CHINA PETROLEUM
39	A	30422	Q39	false	MSFT	CHINA PETROLEUM
40	A	45281	Q40	false	March June August September December	March June July August September October
41	A	17391	Q27	true	AMZN	AMZN
42	A	36281	Q42	true	2007	2007

Test Person 15

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	C	66967	Q15	true	02.01.2008	02.01.2008
2	C	44297	Q16	true	20.11.2008	20.11.2008
3	C	43968	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	C	57250	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
5	C	45765	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
6	C	27297	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	C	17000	Q21	false	-39.63 %	-40 < ... <= 50 %
8	C	26656	Q36	true	January	January
9	C	65390	Q9	false	March April May August	March April May August October
10	C	21578	Q38	true	MSFT	MSFT
11	C	13750	Q11	true	IBM	IBM
12	C	135594	Q12	true	March April August October	March April August October
13	C	15390	Q27	false	AMZN	YAHOO
14	C	30593	Q14	true	2004	2004
15	A	19907	Q1	true	01.01.2008	01.01.2008
16	A	21532	Q2	true	20.11.2008	20.11.2008
17	A	24265	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	A	41844	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	A	58422	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
20	A	22625	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	A	36500	Q7	true	-42.7 %	-40 < ... <= 50 %
22	A	16015	Q22	true	January	January
23	A	45328	Q23	true	March April May July August December	March April May July August December
24	A	23156	Q24	false	AMZN	YAHOO
25	A	24250	Q25	true	AMZN	AMZN
26	A	54359	Q40	false	March June August September December	March June August September October
27	A	15360	Q41	false	CHINA_PETROLEUM	MSFT
28	A	12156	Q28	true	2007	2007
29	B	26484	Q29	true	01.01.2008	01.01.2008
30	B	44157	Q30	true	27.10.2008	27.10.2008
31	B	13875	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	B	29937	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
33	B	41125	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -10 < ... <= 20 % YAHOO January - June 2008: -10 < ... <= 20 %

34	B	29625	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	B	33047	Q35	false	-35.15 %	-40 < ... <= 50 %
36	B	48390	Q8	true	January May	January May
37	B	61984	Q37	true	March August December	March August December
38	B	11281	Q10	true	AAPL	AAPL
39	B	9016	Q39	true	MSFT	MSFT
40	B	43938	Q26	false	March May July August September October December	March April May June July August October December
41	B	12891	Q13	true	AAPL	AAPL
42	B	31609	Q42	true	2007	2007

Test Person 16

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	A	14766	Q29	true	01.01.2008	01.01.2008
2	A	29781	Q30	true	27.10.2008	27.10.2008
3	A	29750	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
4	A	28375	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	A	49562	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -10 < ... <= 20 % YAHOO January - June 2008: -0 <= ... <= 10 %
6	A	20890	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	A	72907	Q35	true	-35.15 %	-30 < ... <= 40 %
8	A	11156	Q36	true	January	January
9	A	31703	Q23	true	March April May July August Decem-ber	March April May July August Decem-ber
10	A	65610	Q38	true	MSFT	MSFT
11	A	41359	Q11	true	IBM	IBM
12	A	114219	Q12	false	March April August October	March April June October November
13	A	26469	Q27	true	AMZN	AMZN
14	A	18375	Q14	false	2004	2007
15	C	13078	Q1	true	01.01.2008	01.01.2008
16	C	16875	Q16	true	20.11.2008	20.11.2008
17	C	36266	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	C	52297	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
19	C	19735	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
20	C	23063	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	C	13922	Q21	true	-39.63 %	-30 < ... <= 40 %
22	C	25812	Q8	true	January May	January May
23	C	41718	Q9	false	March April May August	March April May August October
24	C	10843	Q24	true	AMZN	AMZN
25	C	7203	Q25	true	AMZN	AMZN
26	C	83500	Q40	false	March June August September De-cember	March June July August September December
27	C	11281	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
28	C	40688	Q28	true	2007	2007
29	B	17625	Q15	true	02.01.2008	02.01.2008
30	B	19437	Q2	true	20.11.2008	20.11.2008
31	B	12110	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO

32	B	20328	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
33	B	40828	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
34	B	22719	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	B	26672	Q7	false	-42.7 %	-20 < ... <= 30 %
36	B	19000	Q22	true	January	January
37	B	29703	Q37	true	March August December	March August December
38	B	9516	Q10	true	AAPL	AAPL
39	B	7078	Q39	true	MSFT	MSFT
40	B	56609	Q26	false	March May July August September October December	March July August December
41	B	33609	Q13	false	AAPL	NASDAQ
42	B	19422	Q42	true	2007	2007

Test Person 17

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	C	23296	Q29	true	01.01.2008	01.01.2008
2	C	18250	Q16	true	20.11.2008	20.11.2008
3	C	43562	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
4	C	72140	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
5	C	49031	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
6	C	40343	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
7	C	16735	Q35	true	-35.15 %	-30 < ... <= 40 %
8	C	25766	Q22	true	January	January
9	C	49719	Q23	true	March April May July August Decem-ber	March April May July August Decem-ber
10	C	17641	Q24	true	AMZN	AMZN
11	C	12844	Q11	true	IBM	IBM
12	C	91609	Q12	true	March April August October	March April August October
13	C	15203	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
14	C	22172	Q28	true	2007	2007
15	B	16656	Q1	true	01.01.2008	01.01.2008
16	B	26406	Q30	true	27.10.2008	27.10.2008
17	B	19719	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	B	26359	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	B	65281	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -40 < ... <= 50 % MSFT January - June 2008: -10 <= ... <= 20 %
20	B	23109	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	B	36125	Q21	false	-39.63 %	-40 < ... <= 50 %
22	B	29391	Q8	true	January May	January May
23	B	28250	Q9	false	March April May August	March April May August October
24	B	14110	Q10	true	AAPL	AAPL
25	B	10500	Q39	true	MSFT	MSFT
26	B	86422	Q40	false	March June August September De-cember	June August October December
27	B	12031	Q13	true	AAPL	AAPL

28	B	13781	Q14	false	2004	2007
29	A	18422	Q15	true	02.01.2008	02.01.2008
30	A	19266	Q2	true	20.11.2008	20.11.2008
31	A	19203	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
32	A	26344	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
33	A	71109	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -0 <= ... <= 10 %
34	A	22844	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
35	A	37845	Q7	true	-42.7 %	-40 < ... <= 50 %
36	A	18219	Q36	true	January	January
37	A	29656	Q37	true	March August December	March August December
38	A	33375	Q38	true	MSFT	MSFT
39	A	18938	Q25	false	AMZN	YAHOO
40	A	89875	Q26	false	March May July August September October December	March May June July August October December
41	A	24266	Q27	false	AMZN	YAHOO
42	A	30453	Q42	true	2007	2007

Test Person 18

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	B	22515	Q15	true	02.01.2008	02.01.2008
2	B	16297	Q2	true	20.11.2008	20.11.2008
3	B	9109	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	B	49890	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	B	52391	Q19	true	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -10 < ... <= 20 %
6	B	25641	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	B	27515	Q35	true	-35.15 %	-30 < ... <= 40 %
8	B	10109	Q22	true	January	January
9	B	20344	Q23	true	March April May July August December	March April May July August December
10	B	13485	Q38	true	MSFT	MSFT
11	B	13797	Q25	false	AMZN	YAHOO
12	B	61985	Q40	false	March June August September December	January February August November
13	B	32156	Q27	false	AMZN	SP500
14	B	14078	Q14	false	2004	2007
15	A	14375	Q29	true	01.01.2008	01.01.2008
16	A	10828	Q16	true	20.11.2008	20.11.2008
17	A	22172	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	A	26063	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	A	68094	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
20	A	25827	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	A	34719	Q7	true	-42.7 %	-40 < ... <= 50 %
22	A	26562	Q36	true	January	January
23	A	30327	Q37	false	March August December	March May August December
24	A	22203	Q10	true	AAPL	AAPL

25	A	26219	Q39	false	MSFT	CHINA PETROLEUM
26	A	48078	Q26	false	March May July August September October December	January March May June July August December
27	A	15876	Q41	false	CHINA_PETROLEUM	MSFT
28	A	23812	Q42	true	2007	2007
29	C	16797	Q1	true	01.01.2008	01.01.2008
30	C	14641	Q30	false	27.10.2008	20.11.2008
31	C	22218	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	C	45563	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
33	C	39141	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
34	C	22765	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	C	11125	Q21	true	-39.63 %	-30 < ... <= 40 %
36	C	37374	Q8	true	January May	January May
37	C	62985	Q9	true	March April May August	March April May August
38	C	9531	Q24	true	AMZN	AMZN
39	C	8469	Q11	true	IBM	IBM
40	C	80984	Q12	true	March April August October	March April August October
41	C	19078	Q13	true	AAPL	AAPL
42	C	20344	Q28	true	2007	2007

Test Person 19

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	A	33015	Q15	true	02.01.2008	02.01.2008
2	A	28016	Q2	true	20.11.2008	20.11.2008
3	A	19501	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
4	A	54079	Q32	false	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: +0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
5	A	71015	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -50 < ... <= 60 % IBM January - May 2008: +20 < ... <= 30 %
6	A	38938	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	A	47421	Q21	false	-39.63 %	-40 < ... <= 50 %
8	A	26765	Q36	true	January	January
9	A	31671	Q9	false	March April May August	March April May
10	A	21485	Q10	true	AAPL	AAPL
11	A	32672	Q25	false	AMZN	YAHOO
12	A	92062	Q12	false	March April August October	March April May October November
13	A	21750	Q13	true	AAPL	AAPL
14	A	13281	Q14	false	2004	2007
15	B	30594	Q1	true	01.01.2008	01.01.2008
16	B	23000	Q30	true	27.10.2008	27.10.2008
17	B	22219	Q31	false	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT On 16.10.2008 CHINA PETROLEUM was lower than MSF
18	B	61375	Q18	false	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -20 < ... <= 30 US-$ YAHOO January 2008: +0 <= ... <= 10 US-$
19	B	65077	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -0 <= ... <= 10 %

20	B	46781	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
21	B	44938	Q7	true	-42.7 %	-40 < ... <= 50 %
22	B	74891	Q22	true	January	January
23	B	34750	Q37	false	March August December	March August
24	B	20734	Q24	false	AMZN	YAHOO
25	B	21765	Q11	true	IBM	IBM
26	B	115515	Q26	false	March May July August September October December	January February March August Sep-tember
27	B	10391	Q41	false	CHINA_PETROLEUM	MSFT
28	B	27140	Q28	true	2007	2007
29	C	23094	Q29	true	01.01.2008	01.01.2008
30	C	26422	Q16	false	20.11.2008	19.11.2008
31	C	39531	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
32	C	58922	Q4	false	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: -0 <= ... <= 10 US-$
33	C	56343	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
34	C	21516	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
35	C	16172	Q35	true	-35.15 %	-30 < ... <= 40 %
36	C	63813	Q8	false	January May	January
37	C	32234	Q23	true	March April May July August Decem-ber	March April May July August Decem-ber
38	C	16234	Q38	true	MSFT	MSFT
39	C	10577	Q39	true	MSFT	MSFT
40	C	139437	Q40	true	March June August September De-cember	March June August September De-cember
41	C	13923	Q27	false	AMZN	YAHOO
42	C	45766	Q42	true	2007	2007

Test Person 20

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	B	20265	Q1	true	01.01.2008	01.01.2008
2	B	34906	Q16	true	20.11.2008	20.11.2008
3	B	33844	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	B	38813	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	B	61234	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -0 <= ... <= 10 % YAHOO January - June 2008: -10 < ... <= 20 %
6	B	41516	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	B	46234	Q35	false	-35.15 %	-20 < ... <= 30 %
8	B	44453	Q8	true	January May	January May
9	B	45328	Q37	false	March August December	March April August
10	B	13844	Q24	false	AMZN	YAHOO
11	B	13391	Q39	true	MSFT	MSFT
12	B	82234	Q26	false	March May July August September October December	March April June July August October December
13	B	27687	Q27	true	AMZN	AMZN
14	B	19625	Q28	true	2007	2007
15	C	13390	Q15	false	02.01.2008	01.01.2008
16	C	16422	Q2	true	20.11.2008	20.11.2008
17	C	41218	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM

18	C	45219	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	C	38078	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
20	C	25187	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
21	C	13266	Q21	true	-39.63 %	-30 < ... <= 40 %
22	C	25672	Q36	true	January	January
23	C	26015	Q9	false	March April May August	March April May August October
24	C	15125	Q38	true	MSFT	MSFT
25	C	15921	Q25	true	AMZN	AMZN
26	C	118235	Q12	true	March April August October	March April August October
27	C	10578	Q13	true	AAPL	AAPL
28	C	16360	Q42	true	2007	2007
29	A	11860	Q29	true	01.01.2008	01.01.2008
30	A	22312	Q30	false	27.10.2008	26.10.2008
31	A	20063	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	A	55203	Q32	false	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -10 < ... <= 20 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
33	A	69765	Q33	false	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -50 < ... <= 60 % MSFT January - June 2008: -0 <= ... <= 10 %
34	A	23422	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	A	42750	Q7	false	-42.7 %	-0 <= ... <= 10 %
36	A	28344	Q22	true	January	January
37	A	32687	Q23	true	March April May July August Decem-ber	March April May July August Decem-ber
38	A	38828	Q10	true	AAPL	AAPL
39	A	16875	Q11	false	IBM	AAPL
40	A	66797	Q40	false	March June August September De-cember	January June August September Octo-ber November
41	A	15562	Q41	false	CHINA_PETROLEUM	MSFT
42	A	17938	Q14	false	2004	2007

Test Person 21

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	C	26358	Q29	true	01.01.2008	01.01.2008
2	C	13125	Q16	true	20.11.2008	20.11.2008
3	C	20641	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	C	56126	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	C	36718	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
6	C	25032	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	C	26015	Q7	true	-42.7 %	-40 < ... <= 50 %
8	C	20126	Q8	false	January May	January
9	C	64483	Q37	true	March August December	March August December
10	C	12376	Q38	true	MSFT	MSFT
11	C	10437	Q39	true	MSFT	MSFT
12	C	113280	Q26	false	March May July August September October December	March May July August September December
13	C	18327	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
14	C	29640	Q14	true	2004	2004

15	A	19640	Q1	true	01.01.2008	01.01.2008
16	A	17156	Q16	true	20.11.2008	20.11.2008
17	A	24156	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	A	45438	Q32	false	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: -0 <= ... <= 10 US-$
19	A	73204	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -0 <= ... <= 10 %
20	A	24203	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	A	44921	Q21	false	-39.63 %	-40 < ... <= 50 %
22	A	39625	Q22	true	January	January
23	A	20093	Q23	false	March April May July August December	March April July August December
24	A	32546	Q10	false	AAPL	IBM
25	A	20203	Q11	false	IBM	AAPL
26	A	42031	Q12	false	March April August October	January March April July August
27	A	28406	Q13	true	AAPL	AAPL
28	A	28985	Q28	true	2007	2007
29	B	20751	Q29	true	01.01.2008	01.01.2008
30	B	22375	Q2	true	20.11.2008	20.11.2008
31	B	21875	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
32	B	38797	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
33	B	64875	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: -0 <= ... <= 10 %
34	B	28828	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	B	35578	Q7	true	-42.7 %	-40 < ... <= 50 %
36	B	21437	Q36	true	January	January
37	B	29610	Q9	true	March April May August	March April May August
38	B	26515	Q24	false	AMZN	YAHOO
39	B	12484	Q25	false	AMZN	YAHOO
40	B	58844	Q40	false	March June August September December	March May August October
41	B	19578	Q27	true	AMZN	AMZN
42	B	18422	Q42	true	2007	2007

Test Person 22

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	A	14891	Q29	true	01.01.2008	01.01.2008
2	A	16344	Q16	true	20.11.2008	20.11.2008
3	A	31124	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
4	A	29173	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	A	77312	Q5	true	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +10 < ... <= 20 %
6	A	32671	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	A	40547	Q21	false	-39.63 %	-40 < ... <= 50 %
8	A	21578	Q22	true	January	January
9	A	42562	Q9	false	March April May August	March April May August October
10	A	32906	Q10	true	AAPL	AAPL
11	A	26548	Q25	false	AMZN	YAHOO

12	A	57390	Q26	false	March May July August September October December	March May July August December
13	A	23719	Q13	true	AAPL	AAPL
14	A	13953	Q14	false	2004	2007
15	C	31203	Q15	true	02.01.2008	02.01.2008
16	C	29501	Q30	false	27.10.2008	20.11.2008
17	C	15844	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	C	37157	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
19	C	42140	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
20	C	59282	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
21	C	17657	Q7	true	-42.7 %	-40 < ... <= 50 %
22	C	39015	Q8	true	January May	January May
23	C	52281	Q23	true	March April May July August December	March April May July August December
24	C	9609	Q38	true	MSFT	MSFT
25	C	10187	Q39	true	MSFT	MSFT
26	C	83969	Q12	true	March April August October	March April August October
27	C	13907	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
28	C	26891	Q28	true	2007	2007
29	B	10625	Q1	true	01.01.2008	01.01.2008
30	B	12532	Q2	true	20.11.2008	20.11.2008
31	B	19953	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	B	51266	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$
33	B	65015	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -10 < ... <= 20 % YAHOO January - June 2008: -0 <= ... <= 10 %
34	B	16094	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
35	B	36859	Q35	true	-35.15 %	-30 < ... <= 40 %
36	B	15829	Q36	true	January	January
37	B	22422	Q37	false	March August December	March August
38	B	12204	Q24	false	AMZN	YAHOO
39	B	10187	Q11	true	IBM	IBM
40	B	77172	Q40	false	March June August September December	March June August September
41	B	12797	Q27	true	AMZN	AMZN
42	B	17000	Q42	true	2007	2007

Test Person 23

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	C	20047	Q29	true	01.01.2008	01.01.2008
2	C	25281	Q30	true	27.10.2008	27.10.2008
3	C	20297	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
4	C	34437	Q4	true	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +0 <= ... <= 10 US-$ IBM May 2008: +0 <= ... <= 10 US-$
5	C	67094	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 < ... <= 30 %
6	C	25781	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
7	C	18453	Q35	true	-35.15 %	-30 < ... <= 40 %

8	C	48766	Q22	true	January	January
9	C	71718	Q9	false	March April May August	March April May August October
10	C	17218	Q38	true	MSFT	MSFT
11	C	6485	Q39	true	MSFT	MSFT
12	C	104438	Q26	true	March May July August September October December	March May July August September October December
13	C	51109	Q13	false	AAPL	IBM
14	C	12312	Q28	true	2007	2007
15	B	10718	Q1	true	01.01.2008	01.01.2008
16	B	26578	Q16	true	20.11.2008	20.11.2008
17	B	14250	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
18	B	26828	Q32	false	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: +10 < ... <= 20 US-$ CHINA PETROLEUM April 2008: -0 <= ... <= 10 US-$
19	B	74140	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -20 < ... <= 30 % YAHOO January - June 2008: -0 <= ... <= 10 %
20	B	29656	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
21	B	59281	Q7	false	-42.7 %	-30 < ... <= 40 %
22	B	29750	Q36	true	January	January
23	B	53907	Q37	true	March August December	March August December
24	B	16203	Q24	false	AMZN	YAHOO
25	B	14609	Q11	true	IBM	IBM
26	B	54735	Q12	false	March April August October	March April August October November
27	B	17782	Q27	false	AMZN	SP500
28	B	18953	Q14	false	2004	2007
29	A	13875	Q15	true	02.01.2008	02.01.2008
30	A	16843	Q2	true	20.11.2008	20.11.2008
31	A	31266	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	A	21312	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
33	A	76703	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: -0 <= ... <= 10 % IBM January - May 2008: +0 <= ... <= 10 %
34	A	32140	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
35	A	59281	Q21	true	-39.63 %	-30 < ... <= 40 %
36	A	43985	Q8	true	January May	January May
37	A	60687	Q23	false	March April May July August December	March April May June August December
38	A	51281	Q10	true	AAPL	AAPL
39	A	35110	Q25	true	AMZN	AMZN
40	A	68766	Q40	false	March June August September December	March June August September October
41	A	28860	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
42	A	37859	Q42	false	2007	2006

Test Person 24

task	chart	time [ms]	question	task correctness	correct answers	given answers
1	B	13594	Q29	true	01.01.2008	01.01.2008
2	B	41093	Q16	true	20.11.2008	20.11.2008
3	B	29765	Q31	true	On 03.03.2008 CHINA PETROLEUM was higher than MSFT	On 03.03.2008 CHINA PETROLEUM was higher than MSFT
4	B	62375	Q32	true	MSFT April 2008: -0.98 US-$ CHINA PETROLEUM April 2008: 15.05 US-$	MSFT April 2008: -0 <= ... <= 10 US-$ CHINA PETROLEUM April 2008: +10 < ... <= 20 US-$

5	B	64281	Q5	false	AAPL January - May 2008: -4.71 % IBM January - May 2008: 19.73 %	AAPL January - May 2008: +0 <= ... <= 10 % IBM January - May 2008: +0 <= ... <= 10 %
6	B	46281	Q20:	true	On 20.11.2008 was the value of the AMZN under 100	On 20.11.2008 was the value of the AMZN under 100
7	B	38234	Q7	true	-42.7 %	-40 < ... <= 50 %
8	B	77156	Q36	true	January	January
9	B	35391	Q9	false	March April May August	March April August
10	B	28141	Q38	false	MSFT	CHINA PETROLEUM
11	B	25515	Q39	true	MSFT	MSFT
12	B	52563	Q12	false	March April August October	March April May August
13	B	48515	Q13	true	AAPL	AAPL
14	B	15703	Q28	true	2007	2007
15	A	25688	Q15	true	02.01.2008	02.01.2008
16	A	52234	Q30	false	27.10.2008	26.10.2008
17	A	31328	Q3	true	On 3.3.2008 AAPL was higher than IBM	On 3.3.2008 AAPL was higher than IBM
18	A	46203	Q4	false	AAPL May 2008: 8.75 US-$ IBM May 2008: 5.82 US-$	AAPL May 2008: +10 < ... <= 20 US-$ IBM May 2008: +0 <= ... <= 10 US-$
19	A	76437	Q19	false	AMZN January - June 2008: -20.84 % YAHOO January - June 2008: -11.18 %	AMZN January - June 2008: -10 < ... <= 20 % YAHOO January - June 2008: -0 <= ... <= 10 %
20	A	58422	Q34	true	On 31.12.2008 was the value of DJIA under 10000	On 31.12.2008 was the value of DJIA under 10000
21	A	42687	Q35	true	-35.15 %	-30 < ... <= 40 %
22	A	74609	Q8	true	January May	January May
23	A	79313	Q23	true	March April May July August December	March April May July August December
24	A	37844	Q10	true	AAPL	AAPL
25	A	30797	Q25	false	AMZN	YAHOO
26	A	45547	Q26	false	March May July August September October December	March April August December
27	A	54453	Q41	true	CHINA_PETROLEUM	CHINA_PETROLEUM
28	A	28484	Q14	false	2004	2007
29	C	39407	Q1	true	01.01.2008	01.01.2008
30	C	28109	Q2	false	20.11.2008	19.11.2008
31	C	54063	Q17	true	On 03.05.2008 AMZN was higher than YAHOO	On 03.05.2008 AMZN was higher than YAHOO
32	C	40266	Q18	true	AMZN January 2008: -14.94 US-$ YAHOO January 2008: -4.08 US-$	AMZN January 2008: -10 < ... <= 20 US-$ YAHOO January 2008: -0 <= ... <= 10 US-$
33	C	34641	Q33	true	CHINA PET. January - June 2008: -37.32 % MSFT January - June 2008: -22.72 %	CHINA PET. January - June 2008: -30 < ... <= 40 % MSFT January - June 2008: -20 <= ... <= 30 %
34	C	31687	Q6	true	On 05.06.2008 was the value of the NASDAQ over 2000	On 05.06.2008 was the value of the NASDAQ over 2000
35	C	31765	Q21	true	-39.63 %	-30 < ... <= 40 %
36	C	69281	Q22	true	January	January
37	C	68969	Q37	true	March August December	March August December
38	C	26093	Q24	true	AMZN	AMZN
39	C	20407	Q11	true	IBM	IBM
40	C	86078	Q40	false	March June August September December	August
41	C	32031	Q27	false	AMZN	YAHOO
42	C	17907	Q42	true	2007	2007