FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Interaktives Dashboard für die Technische Analyse von Rohstoffmarktindikatoren

BACHELORARBEIT

zur Erlangung des akademischen Grades

Bachelor of Science

im Rahmen des Studiums

Medieninformatik und Visual Computing

eingereicht von

Dea Cizmic
Matrikelnummer 01425788

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Dr. techn. Manuela Waldner, MSc

Wien, 20. März 2018

Dea Cizmic

Manuela Waldner

Exploratory Data Visualization Dashboard for Technical Analysis of Commodity Market Indicators

BACHELOR'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science

in

Media Informatics and Visual Computing

by

Dea Cizmic
Registration Number 01425788

to the Faculty of Informatics

at the TU Wien

Advisor: Dr. techn. Manuela Waldner, MSc

Vienna, 20th March, 2018

Dea Cizmic   Manuela Waldner

Erklärung zur Verfassung der Arbeit

Dea Cizmic
Inzersdorfer Straße 76/1/9
1100 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 20. März 2018

Dea Cizmic

Danksagung

Ich möchte mich bei meiner Betreuerin Manuela Waldner für ihre Unterstützung und das regelmäßige Feedback während der Bachelorarbeit bedanken. Danke auch an Michael Pühringer und Sebastian Schrey für ihre Expertise und die Ermöglichung dieser Arbeit.

Acknowledgements

I would like to express my gratitude to my advisor Manuela Waldner for providing support and continuous feedback during the development of this thesis. I also want to thank Michael Pühringer and Sebastian Schrey for making this thesis possible and providing their domain expert knowledge.

Abstract

Companies and traders working in the commodity market encounter a variety of different data sets, including numerous economic indicators. The analysis of those indicators and their connection to certain markets can lead to important insights. The understanding of the market can be improved and predictions of the future market development can be created. However, dozens of economic indicators exist and one of the main challenges is to show a clear overview of the indicators and identify those, which show a correlation to a certain market. Software tools are often utilised in order to perform the analysis of financial markets. However, according to domain experts, they often hit the limit of human perception capabilities. This thesis focuses on the development of a prototypical web application dashboard, which enables the user to analyse the relation between a defined commodity market and different economic indicators. Besides the relation between one indicator and a given market, the possibility to interactively create one's own composite indicator, for comparison with the given market, is implemented. The process of creating a composite indicator is another challenge as it requires numerous decisions to be made. The dashboard therefore offers a platform for exploring the different composite indicator configurations. Moreover, the web-application provides also some visualization and interaction techniques, like highlighting, brushing and details-on-demand to enhance the comparison process and amplify human cognition.

Contents

CHAPTER 1

Introduction

Companies that are dependent on the development of a (certain) market have to analyse financial data in order to make the right decisions and, for example, purchase goods at a profitable rate. By performing a technical analysis of a stock market, it is possible to create a prediction for the future trend of a certain stock. Edwards et al. [EM48] describe technical analysis as

> "the science of recording, usually in graphic form, the actual history of trading (price changes, volume of transactions, etc.) in a certain stock or in 'the Averages' and then deducing from that pictured history the probable future trend."

A technical analysis includes, apart from investigating the price changes of a certain stock, also the analysis of correlations between different variables. Some stock prices can be influenced by certain economic factors, so that analysing the relationship between a stock and economic indicators can help at recognizing (price) patterns. The technical analysis requires certain visualizations and tools, which support the data analysis process. They are often crucial for the analysis and are especially needed when large data sets have to be processed. The features of such tools range from basic chart visualizations to complex software with automatic analysis and calculation functionality. However, according to experts from the financial domain, buyers from industrial companies often do not want to use complex software, which they perceive as a black box that has its own calculation algorithms and does not make use of their expertise. Thereupon, in order to address this problem, the domain experts performed an analysis of the current market of software supporting financial decision and prediction making. They revealed that there is a demand for an application, which will enable the user to explore the current commodity market situation by using his or her own domain expert knowledge (in addition) and support the technical analysis process in order to facilitate an early price pattern recognition.

The aim of this thesis is to meet this demand and create a web-based finance dashboard which provides the possibility to visually represent over 50 different indicators and analyse the relationships between them and a certain commodity stock. The features are all integrated into a prototypical application, which was designed and implemented in close collaboration with the experts, which performed the analysis of the current finance software market. Besides the functionality, that the dashboard provides, it also serves as a prototype for further developments. Users sometimes can/do not exactly state which features they need and how the features should be implemented. The developed dashboard prototype therefore also serves as an initial implementation of the user needs. The prototype can show the user the (potential) possibilities of such an application. Moreover, it can be used for user evaluations, which should reveal, in more detail, the needs of the user.

As shown in Figure 1.1 it is difficult to read a visualization of 10 indicators, represented by their time series, and extract those which resemble a reference curve the most. However, if the relationships between indicator and reference curve is expressed by calculating the correlation coefficient, like in Figure 1.2, it is possible to identify the indicators with the strongest correlation immediately.

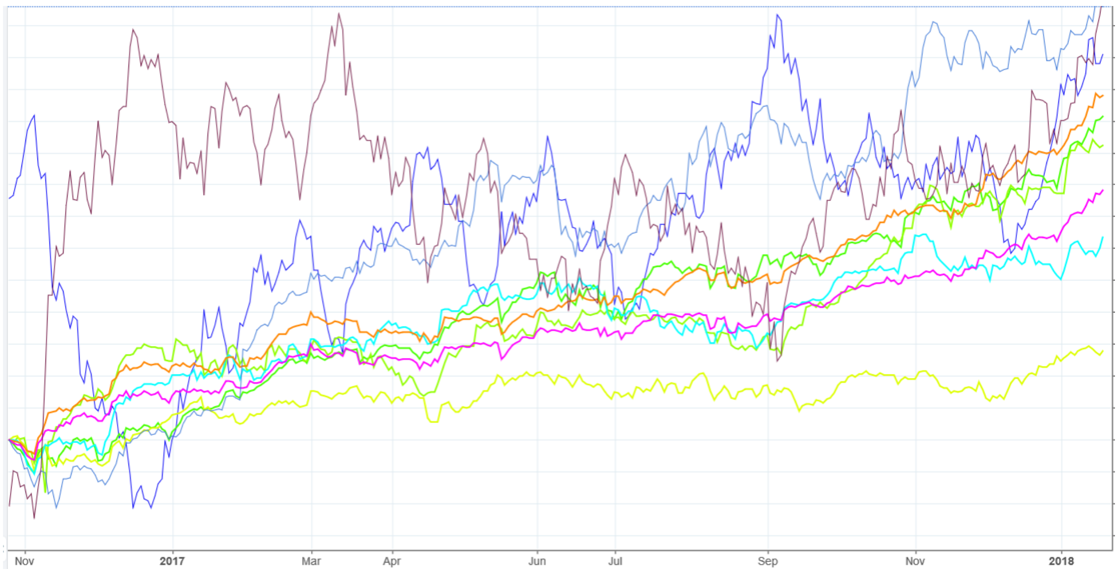


Figure 1.1: 10 different indicators visualized by lines [Tra18].

However, it is often not sufficient to use a single indicator to create a prediction for the future development of a stock market, as it does not capture all aspects which have an impact on the underlying financial market [C+08]. Therefore, users have to form and try out different combinations of more than just one indicator, in order to create a composite indicator, which can take these different aspects into account. A composite indicator is a single index which is built of several different indicators on the basis of an economic model [C+08]. In addition to the visualization, the dashboard features

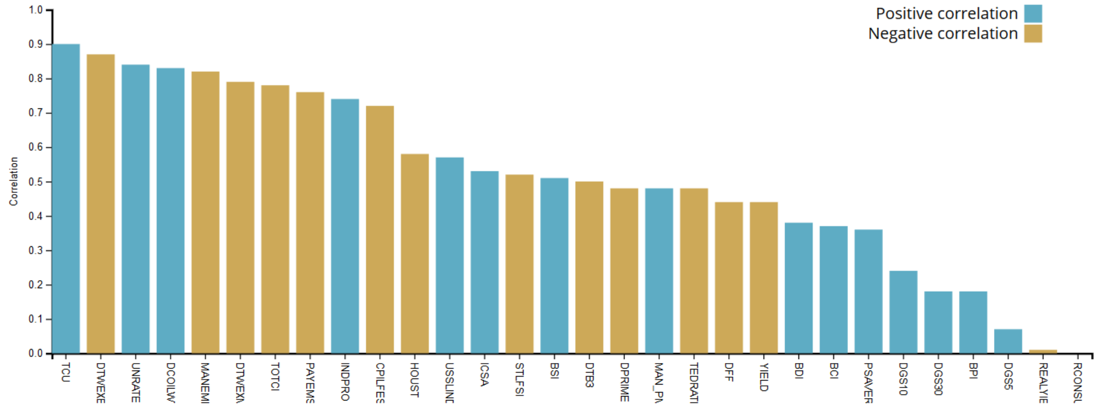
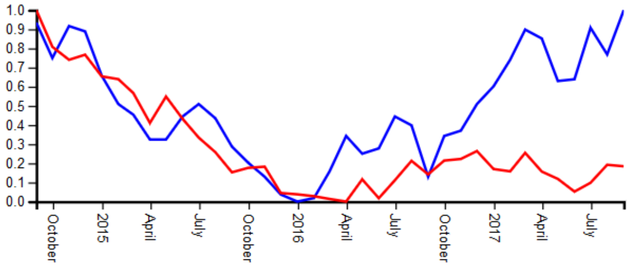


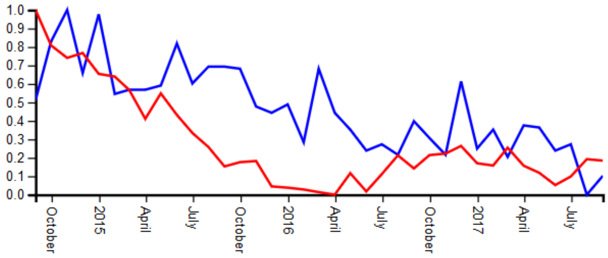
Figure 1.2: Correlation coefficient ranking of 30 indicators.

also the possibility to create own, unique composite indicators. During the creation of the composite indicator, the user can adjust the weighting of the indicators, as well as time periods, and also invert the values of the time series. Furthermore, it is possible to compare the created CI with the reference curve represented by the target stock and also calculate and display the correlation coefficient of the CI.

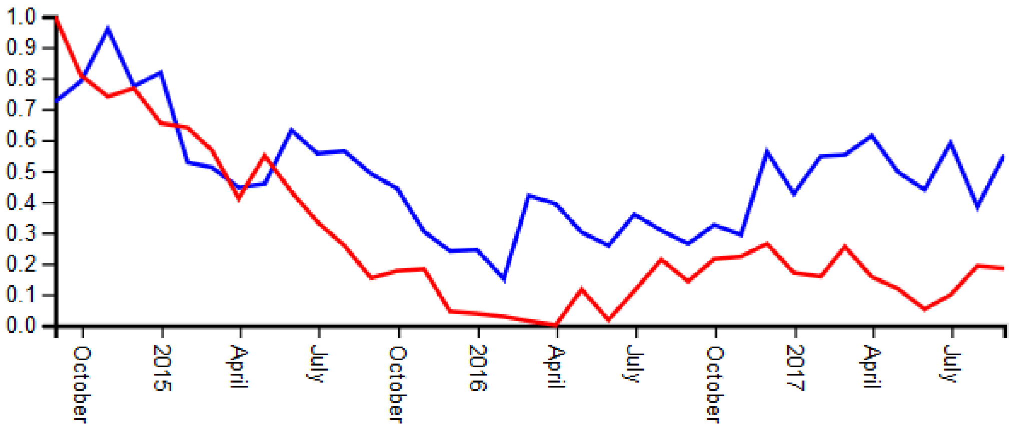
To illustrate the benefit of a composite indicator, the Figures 1.3a, 1.3b and 1.3c can be compared. In each Figure, the normalized monthly value of the indicators (blue line) and the reference curve (red line) are plotted for a given time period. Figure 1.3a and 1.3b both contain a financial indicator, like the gross domestic product (GDP), and the reference curve (red line). The indicators show some similarities to the reference curve as well as some deviations. By combining the two indicators into one composite indicator the correlation can be increased, as shown in Figure 1.3c.



(a) Indicator nr.1 with a correlation coefficient of 0.48.



(b) Indicator nr.2 with a correlation coefficient of 0.57.



(c) Composite indicator with a correlation coefficient of 0.76.

Figure 1.3: Component indicators compared to composite indicator.

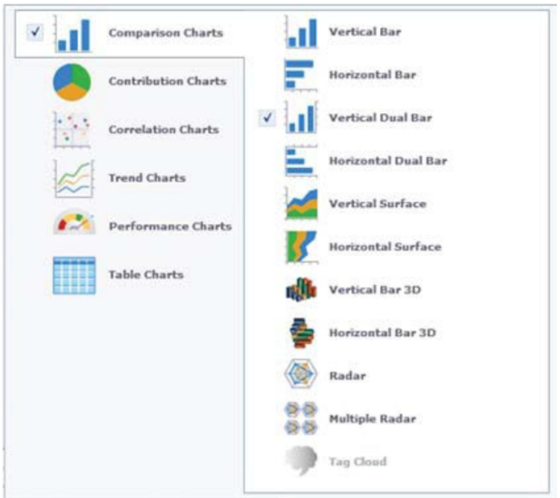
CHAPTER 2

Related Work

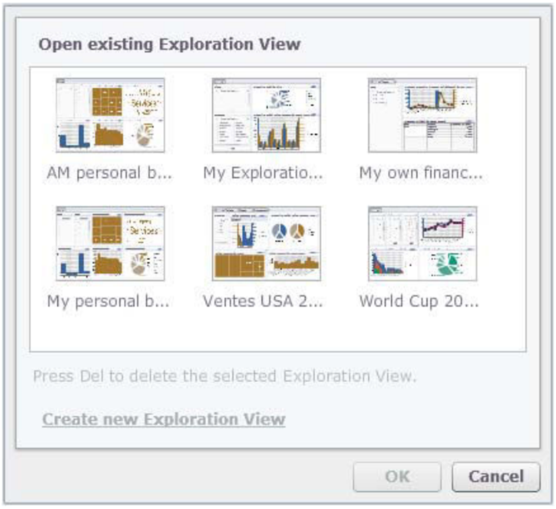
The developed application consists of a dashboard with multiple coordinated views visualizing financial data. Therefore, some examples for technical analysis tools and dashboards will be discussed.

Spritzer and Freitas [SF06] described that technical analysis requires tools with certain visualization and computation capabilities. However, they also reported that most commercial software (used for this purpose) is too complex and/or expensive for users, who are only interested in a small amount of features. Therefore, they created their own prototype, which supports the technical analysis of stock market data. The prototype includes the basic chart types and tools, which are used for a technical analysis. Furthermore, it is able to retrieve data from different sources. One of the goals was that the prototype should only provide features a non-professional user would need. Additionally, it should be easy to use and extend. Their prototype focuses on the visualization of one stock and its attributes, through different types of charts, whereas the herein described dashboard prototype allows the creation of predictions by visualizing the relationship of a certain commodity stock and different indicators relevant for this commodity market.

Elias [Eli12] conducted a research on user interaction with dashboards in the Business Intelligence (BI) domain and also created her own dashboard prototype called "Exploration Views (EV)". The target audience for EV are novice and expert business users in need of a BI dashboard. Using EV, the user can create a personalised dashboard by either forming one from scratch with chart templates (Figure 2.1a) or using provided dashboard templates (Figure 2.1b). The dashboard (templates) can be customized by, for instance, adding or removing further components and filter mechanisms. Moreover, mechanisms like text search, linked visualizations and visual queries are also provided, in order to support the visual analytics process [Eli12]. The herein described dashboard includes also some customization and interaction possibilities (described in Section 4.4), but serves a more specific purpose than EV, because EV is a more general framework for creating and exploring different visualizations and dashboard configurations.



(a) Chart creation dialog [Eli12].



(b) Dashboard creation dialog [Eli12].

Figure 2.1: Exploration Views (EV) Prototype.

A versatile online platform, which does not focus on the concept of dashboards, but rather offers among others a tool for creating charts representing financial data, is TradingView [Tra18]. TradingView can be used for creating custom visualizations using various chart types, customizable time periods and intervals. It offers a broad market data coverage as well as information about fundamental company metrics. Moreover, it lets the user compare a stock to an index or other stocks. However, it can only plot the different lines in one chart, as in Figure 1.1, so that the comparison of multiple indicators becomes a difficult task. Other features of TradingView also provide the the possibility to trade directly from the charts, as well as publish and share charts.

CHAPTER 3

System Overview

The dashboard application was developed in order to provide interactive visualizations, which supports specific tasks performed by the users (buyers of commodities). In the following sections the visualized data is explained in detail. Moreover, an overview of the steps that are performed on the one hand by the user and on the other hand by the application, is introduced.

3.1 Financial Data

The key objects of observation and visualization are on the one hand an index from the commodity market, such as the stock price of copper or nickel, and on the other hand different economic indicators. An indicator is an economic measurement like the gross domestic product (GDP), consumer price index (CPI) or unemployment rate, which carries some information about the overall state of an economy [Inv18]. There are usually three types of indicators: lagging, leading and coincidental indicators. The lagging and leading indicators are shifted either forwards or backwards in time [GG12]. Leading indicators are used to predict the future development of, for instance, certain stock markets, as correlations between these indicators and other economic activities and market developments can be observed.

The user of the dashboard is able to use the economic indicators to analyse and eventually predict the development of a certain stock price. Patterns, which can lead to a more probable prediction, can be identified through the analysis of the relationship between the stock of interest and the indicators. From the various possibilities for the visualization of this relationship, the visualization and calculation of the correlation coefficient between each indicator and a stock (described in Section 4.3) was chosen for this thesis.

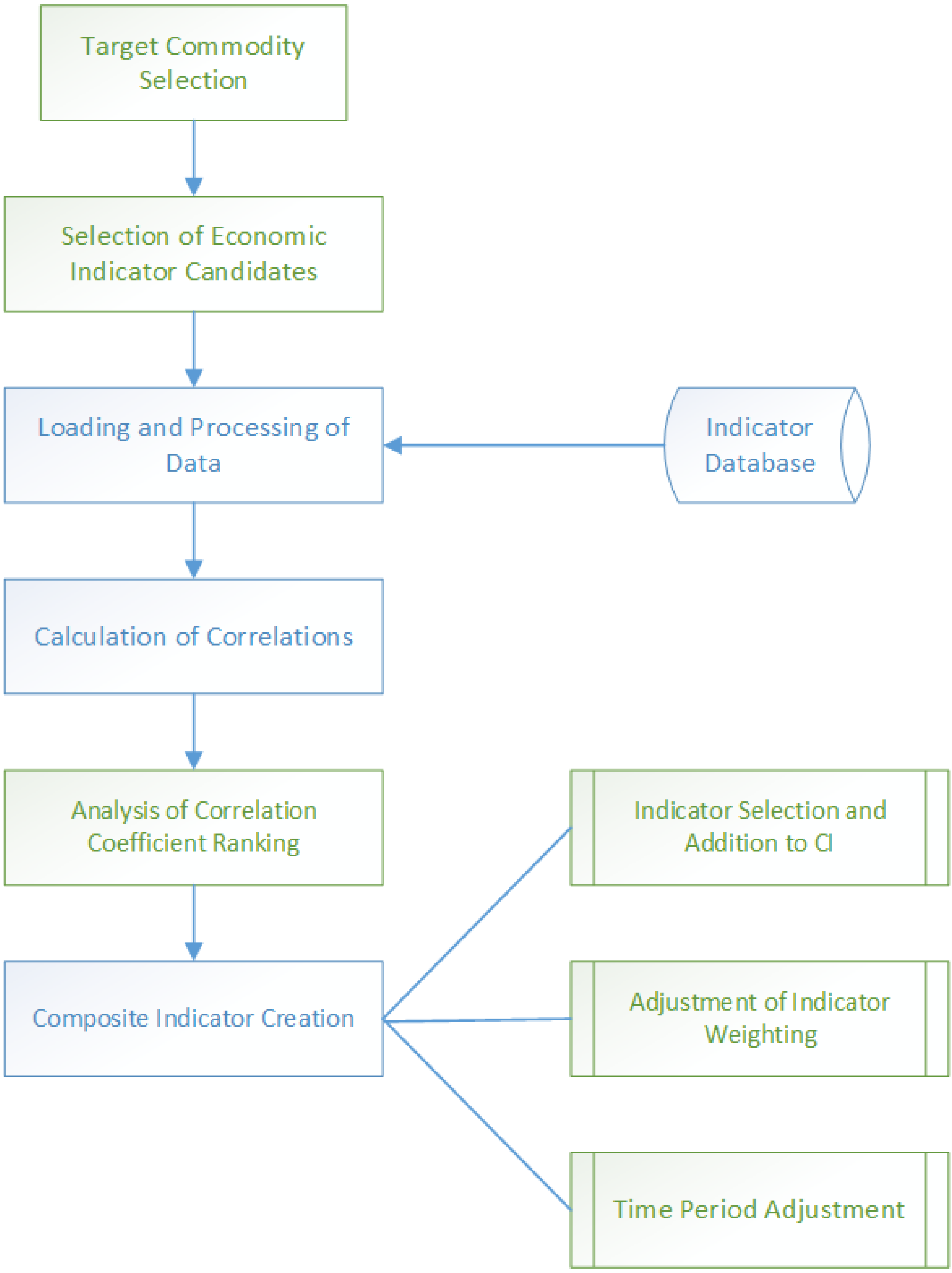


Figure 3.1: Working pipeline showing the steps performed by the user (green) and by the application (blue).

3.2 Working Pipeline

In Figure 3.1 the different steps which are performed by the user (green) and by the application (blue) during the analysis process are shown. First, a target stock (referred to as index) has to be selected for the analysis. The index is represented by the time series of the stock price for a time period chosen by the user. The default time period is three years, but can be changed through interactions described in Section 4.4. The range of the time period is only limited by the availability of the data for this period. In addition to the index, a set of indicators has to be selected. At the moment, the selection is done by domain experts, which already have performed an analysis of the different economic indicators available online. They can save a list of indicators in an external file which is processed by the application. The application can load and process over 50 indicators, but with every indicator increasing the loading time. After the data selection the correlation coefficient between the target stock and each indicator is calculated and visualized.

Through the correlation coefficient ranking (described in the following section) the user can construct a composite indicator by choosing the indicators from the ranking and adding them to the composite indicator view. Subsequently, a composite indicator is constructed, as described in Section 4.3, and the user can adjust the weighting of the indicators and invert the values in case of a negative correlation. A common issue which can be observed is that indicator fluctuations sometimes have a delayed impact on stock prices. The user can shift the indicators and adjust the time periods in order to balance the lead and lag of an indicator.

After creating a composite indicator with a strong correlation coefficient the user can utilize the gained insights to conduct further analysis and eventually develop a prediction. The construction of the composite indicator is not the end of the analysis process. Through the comparison of the time series of the (composite) indicators and the index, important turning points can be detected, which can subsequently be analysed in detail. A user can, for example, retrieve further information about the economic activities which took place at the time of a turning point. Moreover, the domain expert stated that users sense that certain indicators have a higher correlation than others, but sometimes do not have an appropriate analysis which shows this correlation. Hence, the correlation coefficient ranking allows the user to check if his or her assumption is true or not.

CHAPTER 4

Dashboard Design and Components

Dashboards are, according to Stephan Few [Few06], a very popular visualization tool for Business Intelligence (BI), because they are visually appealing to the user and offer the possibility to display the data in a unique way. Few further stated that dashboards are visual displays, which aim at showing an overview of a specific set of data, displaying and communicating different information at one place for easy monitoring. The information is displayed using multiple (coordinated) views which should all fit on a single computer screen.

Roberts [Rob07] defined the term multiple coordinated views (MCV) as a visualization in which data is displayed in multiple windows, which are connected and coordinated together. Hence, when data is manipulated in one view, the corresponding data of the other windows is also updated. Multiple coordinated views aim at connecting data and viewing alternative or additional representation of the data. According to Baldonado et al. [WBWK00], they can help the user identify correlations between the displayed data, because the different aspect of the data are all displayed at once. In contrary to a single view, the user does not have to keep in mind information during the analysis process, for later comparison. Baldonado et al. furthermore stated that multiple coordinated views can be used "when there is a diversity of attributes, models, user profiles, levels of abstraction, or genres" [WBWK00]. This is because they can visualize the information to serve different needs without either overwhelming the user with its complexity or leaving out some informations. Roberts also described that MCV offer an exploratory visualization as various interaction techniques can be implemented and coordinated between multiple views. Interactions can either be performed directly, for example by manipulating the elements of the visualization or indirectly, through menus, buttons or sliders.

4.1 Correlation Coefficient Ranking

One of the requirements is that the application should allow over 50 different economic indicators as input, with every indicator having a possibly large data series. However, a visual comparison of every indicator and a target stock would be very difficult. This issue was discussed with domain experts and the discussion revealed that it would be sufficient for the analysis process if for each indicator the correlation coefficient between the target commodity market and the indicator is calculated and visualized as an initial view. Moreover, it should also facilitate the process of selecting indicators for further investigations.

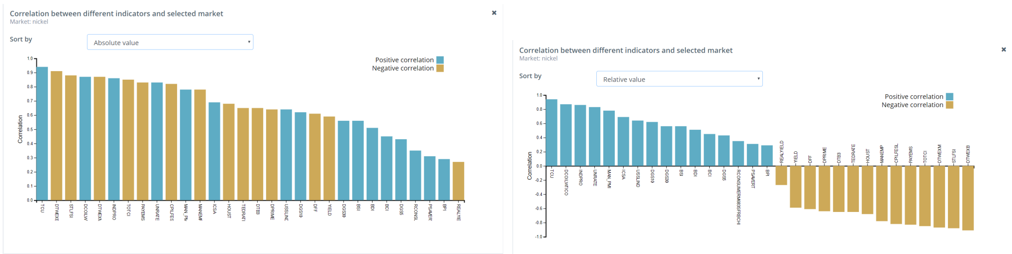
The correlation coefficient (or also called Pearson's correlation) measures the strength of the linear association between two variables. The value of the coefficient can range from -1 to 1, where both, -1 and 1, mean that the linear association is very strong. The correlation coefficient r_{xy} between two data sets x and y, where $x_i \in x$ and $y_i \in y$ with $i \in \{1, 2, ..., n\}$ (n is the size of the data sets) and $\bar{x}$, $\bar{y}$ representing the mean of the data sets, can be calculated as follows:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{4.1}$$

For the ranking, the correlation coefficient between the index and an indicator is calculated for the time period chosen by the user (default three years), where x_i and y_i are monthly values. Hence, for the default time period, the index and indicator are each represented by a time series consisting of 36 values.

Using the correlation coefficient, each indicator can be represented by one number, reducing the number of data points to a single value per indicator. The resulting values can then be visualized by a bar chart, where each correlation coefficient is depicted by a bar. A bar chart was chosen because bars can represent individual values in a clear way, also supporting the comparison of these values [Few12].

In order to facilitate the identification of the economic indicators with the highest correlation coefficient, different sorting options were implemented. The indicators can be sorted either by their relative or absolute correlation value, with the indicator with the highest correlation being the first. When the indicators are sorted by their absolute value, those indicators with an originally negative correlation get a different colour, so that negative and positive correlations are distinguishable at first sight. As an addition to the relative sorting, some indicators that are interesting because of other attributes than the correlation, such as their source or (economic) domain, can also be highlighted by applying a different colour. The colour, however, has to be defined by domain experts and manually added to the application. Figure 4.1 shows the two different sorting options.



(a) Absolute sorting. (b) Relative sorting.

Figure 4.1: Sorting options of the correlation coefficient ranking.

4.2 Indicator Details

As the indicators are only represented by their correlation coefficient, it is important to provide a way to show further information about the indicator. For this purpose, the user can obtain details-on-demand [Shn03]. For the details-on-demand interaction a separate view, with additional information about the indicator, is displayed (shown in Figure 4.2). The user can click on a bar from the bar chart visualizing the correlation coefficients in order to view the information. Currently, this information consists of the full name of the indicator, a description, the correlation coefficient as a rounded decimal number and the visualization of the data series. However, the detail view should also offer room for future information extensions, such as custom information about the indicator, added by domain experts. Furthermore, the user can add/remove the selected indicator to/from the composite indicator through the buttons provided by the detail view.

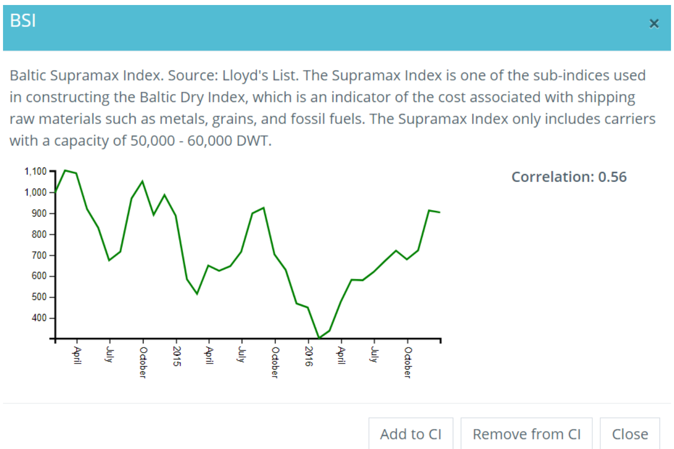


Figure 4.2: The detail view of the indicator BSI.

4.3 Composite Indicator

As described in Section 1, it is also important to combine multiple indicators into a composite indicator and use it for the analysis process. The dashboard therefore provides the possibility to interactively construct a composite (leading) indicator and also compare it to the chosen index. The OECD defines a composite indicator as follows:

> "A composite indicator is formed when individual indicators are compiled into a single index, on the basis of an underlying model." [C^+08]

The construction of a composite indicator is a complex process, because of the various decisions, which have to be made carefully. The OECD published a Handbook [C^+08] in which some recommendations for the construction are described. A part of the Handbook is a ten step guideline, which consists of the following, briefly described, steps:

1. In order to define the context for which a composite indicator should be created, it is important to develop a theoretical framework that answers some questions like which phenomenon should be measured and how the selection of data should be performed. The herein described dashboard application provides a basic framework for the manual selection and combination of different indicators, which should also answer some of these questions.

2. The second step consists of choosing a set of indicators, which will be used for the combination process. Domain experts should decide, using their expert knowledge, which indicators are appropriate for the future combination and analysis process. The decision should be based on properties like relevance, analytical soundness, timeliness, or accessibility of an indicator. In the case of the dashboard prototype, a set of indicators relevant for the commodity market, which are selected by experts from the financial domain, are used.

3. A decision should be made on how to replace missing data parts. This decision is not yet made and has to be considered for the future development of the application. Currently, the application is implemented as if all data sets are complete.

4. The structure of the indicator data should be analysed with the aim to evaluate the most suitable combination techniques. For example, a principal components analysis (PCA) or a factor analysis (FA) can be performed for this. However, this step is currently not supported by the dashboard and has to be performed by domain experts separately.

5. A normalization appropriate for the data domain should be chosen and applied on the data, in order to enable the combination of different data sets. For the dashboard prototype, a min/max normalization (Equation 4.2) was implemented which results in data ranging from zero to one.

$$x'_i = \frac{x_i - min(x)}{max(x) - min(x)} \tag{4.2}$$

As described in the Handbook [C+08], it is possible to use a different normalization option like the standardization using the average and standard deviation or a normalization using the distance to a reference point. However, in this case, it is important that the normalization results in a uniform scale, so that the indicators and index can be compared and viewed in one chart with one scale.

6. A formula for weighting and aggregation of the indicators should be chosen. For this step, a linear combination as in Equation 4.3, with the possibility to adjust the individual weights of the indicators, is implemented for the dashboard prototype. The Equation 4.3 shows the linear combination of two indicators x and y and with their time series being represented as vectors and their corresponding weights as w_x and w_y.

$$\begin{pmatrix} ci_1 \\ ci_2 \\ ci_3 \\ ci_4 \\ ... \\ ci_n \end{pmatrix} = w_x \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ ... \\ x_n \end{pmatrix} + w_y \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ ... \\ y_n \end{pmatrix} \tag{4.3}$$

7. The developed composite indicator should be analysed and its robustness regarding the previously made choices should be assessed.

8. It has to be made traceable, which indicators are contained in the composite indicator. The herein described application therefore provides a visualization of the composite indicator, which assures the transparency by viewing the individual indicators which are part of the CI.

9. A possibility to analyse the correlation between the created composite indicators and other measures or indicators should be provided. Using the dashboard, the user can compare the CI to the target commodity market.

10. A suitable visualization, which facilitates the interpretation process, should be implemented. The dashboard prototype provides this visualization.

Beside this ten step guideline, the OECD published also a guideline [GG12] on how to construct a leading CI, which includes some additional steps. There are also additional criteria for selecting indicators, based on their suitability as leading indicators. One of the additional steps is, for example, the lag-shifting and inversion. In this step, the indicator components are categorized either as leading, lagging or coincidental (see Section 3.1 for definition). Furthermore, it is observed, whether the time series of the indicator shows an inverse behaviour, compared to the reference curve, or not. In order to balance some disparities, the lagging indicators are shifted so that a composite leading indicator with better turning points is created.

4.4 User Interface and Interaction

The user interface (UI) is a dashboard, which includes different components and visualizations. Figure 4.3 shows an overview of the default dashboard configuration, which can be changed by the user by utilising the Admin View (which is described later). Number one is the correlation coefficient ranking chart. This chart is visually and programmatically connected with the composite indicator component (nr. 2), as the selected and highlighted indicators from this chart are displayed in the composite indicator component. By clicking on a bar, the details component with the additional indicator information is displayed in a dialogue window. Through this window, the selected indicator can be added to the composite indicator. Each indicator contained in the CI is also visualized as a line chart (nr. 4).

For the general overview of the target commodity market - the main subject of analysis - a line chart (nr. 3) visualizing the time series of the commodity stock is added to the dashboard. This line chart, and also the line charts of the indicators, are all accompanied with navigation charts. The navigation chart (nr. 5) is a smaller line chart, which offers the possibility to edit the time periods of the data by brushing. Brushing is the act of selecting certain entities by choosing them directly from a data display, for example, by point and click interactions. One of the benefits of brushing is that it can be used to link multiple views, because data that is selected in one view is also highlighted in the other connected views [AMST11]. In the case of the dashboard, the user can select a different time period by brushing over one of the provided navigation charts. The change of the time period is then also applied to the other line charts and their navigation charts, so that the visualized time period of the different line charts is synchronized. This interaction will be explained in more detail in the next section.

4.4.1 Lead and Lag Balancing

The lead and lag of some indicators can be approximately balanced through the navigation charts of the indicators. For this task, the selected time period has to be changed by reducing the selected area of the navigation chart, in order to create room for the shifting. This is necessary because the user can not brush over data which is not visible in the navigation chart, in order to keep the data consistent and avoid input lag caused by loading of additional data. Subsequently, in each navigation chart, a different section has to be selected, while the size of the selection has to remain the same, as shown in Figure 4.4c. Hence, a composite indicator is constructed with the shifted time series. Figure 4.4a and 4.4b shows on the example of one indicator, how the shifting can result in a better correlation coefficient. These figures also show the difference view between indicator and index, which can be made visible through a checkbox.

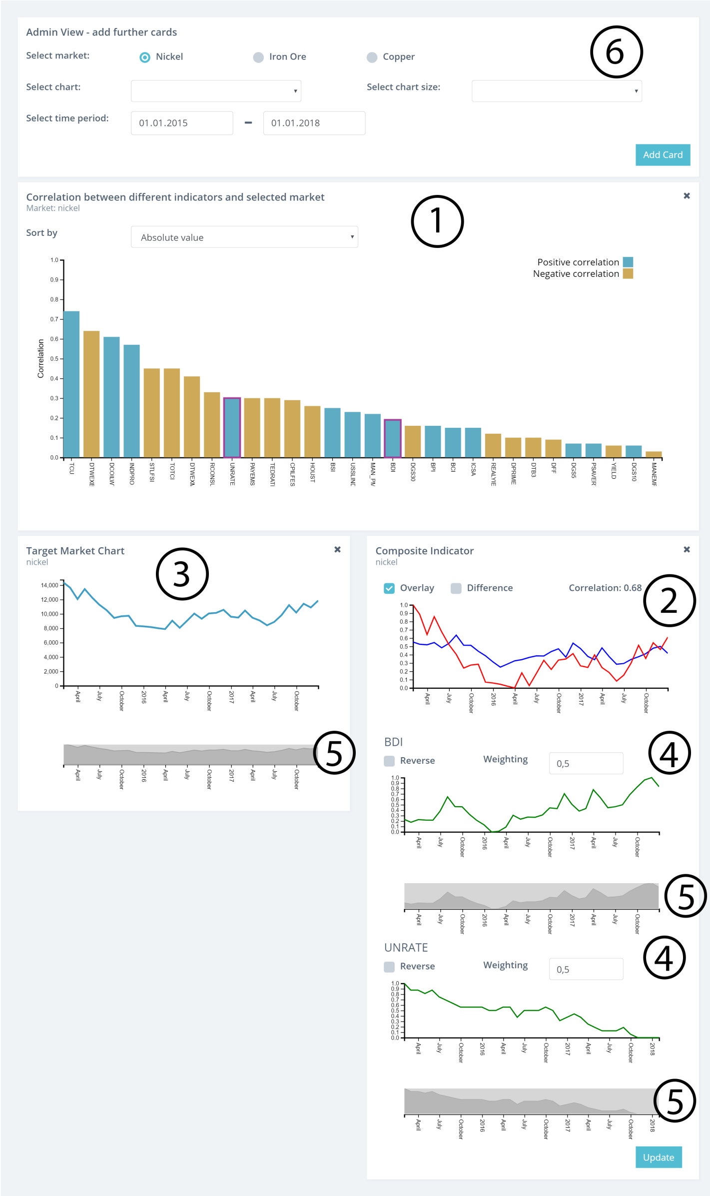
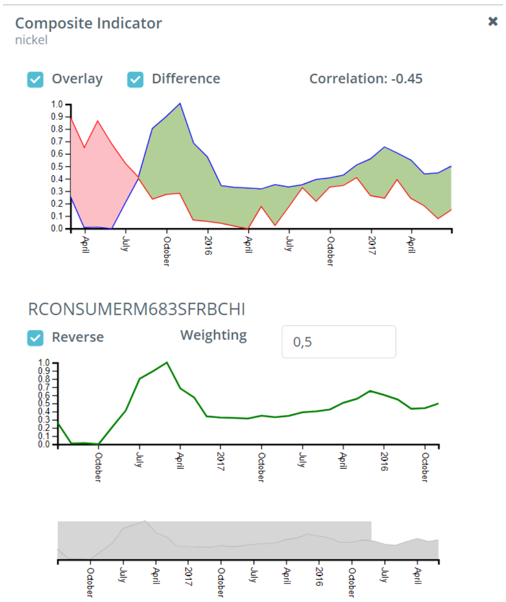


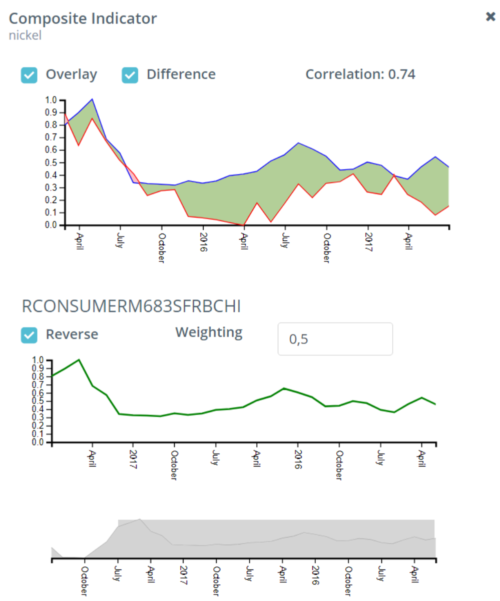
Figure 4.3: Default dashboard configuration including (1) CC ranking, (2) CI, (3) target commodity market chart, (4) indicator charts, (5) navigation charts and (6) Admin View.

4.4.2 UI customization

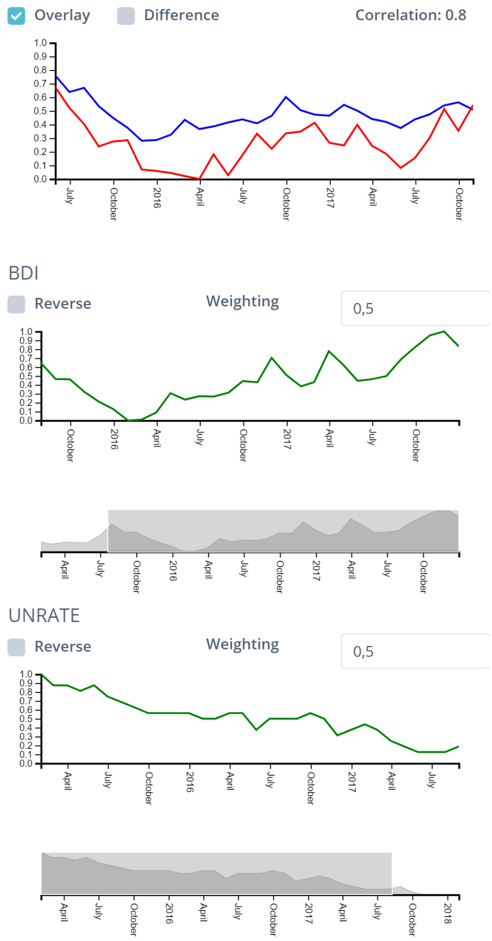
The web-application includes also some features (nr.6) for configuring the user interface and creating different views for different users and commodity markets. Through the Admin View card, one of the components (cc ranking, target commodity market chart or CI) can be added to the dashboard either as a small or large card. A large card takes up the whole page width, whereas the small card only is half as wide as the large one. Moreover, it can be decided for each component which commodity should be the target. Currently, the user can choose between three commodities. Additionally, the initial time period can be set through the provided input fields. The different card components can also be removed from the view, in order to be added subsequently in a different size or with a different time period.



(a) Indicator before shifting.



(b) Indicator after shifting.



(c) Two indicators with different time period selections combined into a CI.

Figure 4.4: Visualization of indicator shifting.

CHAPTER 5

Implementation

The web application consists of a client and server part, but the main program logic is located on the client side. The server part is a NodeJS server which is mainly used for user authentication purposes and loading of files, which include the list of indicators that have to be retrieved from an external database. The client side of the web application was created using Angular 4 and TypeScript. In addition, D3 [BOH11] was used for the implementation of the visualizations.

5.1 Frameworks and Data Source

In the following section the main frameworks and technologies used for the web application will be briefly explained. The data source Quandl will be also introduced. In addition to Quandl, it is possible to add further data sources, but some small adaptations of the code have to be performed because of the differing structure of the data. Every data set retrieved from a data source has to be adapted to fit a certain data structure (which is described in Section 5.1.3).

5.1.1 Angular

Angular is a framework for the development of client applications with HTML and JavaScript or a language like TypeScript (which compiles to JavaScript) [Goo18]. It provides many features which are beneficial for building a structured web application. One of these features is the Angular Command Line Interface (CLI), which can be used to easily set up Angular projects and add further components to the project. Angular furthermore offers code generation and code splitting through HTML templates and Angular components. Moreover, data binding and dependency injection are provided.

5.1.2 D3

D3 is a JavaScript library which provides functionality for creating visualizations for web applications by using HTML, CSS and SVG web standards [BOH11] [Bos17]. It offers various ways of manipulating the Document Object Model (DOM) as well as binding data to the DOM in order to perform changes of DOM elements based on this data. For example, it is possible to attach a data set to a DOM element, like a SVG element, and add for each value of the data set a rectangle (rect) element to the graphics element. This is how a bar chart can be constructed. Figure 5.1 shows a code snippet for the previously described example, which shows that it is possible to implement the example without the usual constructs, like a for loop.

```
let bars = barHolder.selectAll( selector: 'rect.bar')
    .data(correlations)
    .enter().append('rect')
    .classed( names: 'bar',  value: true)
    .attr( name: 'x',  value: function (entry, i) {
        return indicatorScale(entry.indicator)
    })
    .attr( name: 'y',  value: function (entry) {
        return entry.cor < 0 ? corScale(0) : corScale(entry.cor);
    })
    .attr( name: 'width', bandwidth)
    .attr( name: 'height',  value: function (entry) {
        return Math.abs( x: corScale(entry.cor) - corScale(0));
    });
```

Figure 5.1: A D3 code snippet showing one part of a bar chart implementation.

The figure shows how a bar is created for each entry of the "correlations" array, containing the indicator (entry.indicator) and the corresponding correlation coefficient value (entry.cor). In order to determine the size and position of the bar, previously created scales for the x (indicatorScale) and y (corScale) axis are utilized.

5.1.3 Quandl

Quandl is an online platform offering access to financial, economic, and alternative datasets [Qua18]. It provides different access possibilities including an API and extensions for R, Python or Excel. Among different stock market data, it also offers access to some economic indicators. The data for the herein described application is retrieved from the Quandl database using the Quandl API. The data is accessed through HTTPS request and the response contains the data as a JSON object. The structure of the data can be seen in Figure 5.2. Quandl offers a variety of free data sets, but has a concurrency limit of one, meaning that only one dataset at a time can be requested. Hence, the data sets have to be loaded sequentially, slowing down the data loading process. In order to make it visible that the data is being loaded, the bar chart for the correlation coefficient ranking is first rendered only with one indicator and then updated every time a new indicator data set is loaded and added to the ranking. This results in a progressive rendering of the bar chart.

(a) The whole dataset. (b) Structure of the time series.

Figure 5.2: Dataset for the indicator BSI retrieved from Quandl in JSON format.

5.2 Web-Application Architecture

Angular is based on the model view viewmodel (MVVM) pattern [Mar16] and so is also the architecture of the web-application. The model contains information about the content of the entities but does not include any logic. A model is usually described with Typescript. The view is represented by a template, which is a piece of code written in HTML and complemented with Angular expressions. Each template is connected to a component (written in TypeScript) via data binding. A component manages input and output of the templates and also includes further program logic.

Hence, every dashboard card was implemented as an Angular component in combination with an HTML template. In addition to these card components, a root component was created to manage the creation and deletion of the dashboard cards and pass the data from the Admin View (described in Section 4.4) to the underlying components of these cards. Each card is responsible for the rendering of its visualization and managing the interactions with the card content. Apart from the components, there are two models in use. One of them is a chart model which contains an ID, the name of the target market for this chart, the type and width of the chart as well as the initial time period defined by the user through the Admin View. The other model is the indicator model, which contains information about an indicator, that is needed for all of the processing steps. The model includes information like name and description of the indicator as well as the URL for retrieving the data and also the retrieved data set. In addition to the components and models, two services were implemented, which provide further functions to all of the components. One service serves only as a provider of different mathematical calculation functions, like the calculation of the correlation coefficient or the normalization. The other service is responsible for the communication of events and data between the components, allowing the coordination of multiple views. Every data change or event relevant for an other component is passed from one component, where the change or event is located, to the other component using Angular Observer and Observables.

CHAPTER 6

Results

Together with two experts from the financial domain a review of the web application was performed. The main goal was to examine the dashboard functionality and evaluate if the dashboard meets the needs of the future users (buyers from companies in need of certain commodities). The review took approximately one hour and was conducted in an informal way, similar to a cognitive walkthrough. The domain experts were already familiar with an earlier build of the application, so that not every interaction had to be explained in detail. They performed the tasks together on one computer, which had the application running in Chrome Web Browser. The review session was also recorded for later revision.

The review started with the default dashboard layout configuration (Figure 4.3), which was changed and adapted during the review. At the beginning, the cards visualized the stock for nickel as the target market and 30 indicators for the time period from 01.01.2015 to 01.01.2018. The overall task of the review was to examine the correlation coefficient ranking, identify significant indicators and use the interaction possibilities to construct a composite indicator (CI). This CI should possibly feature a correlation coefficient (CC) significantly higher than the ones of its constituents. After performing a task, like constructing the CI, the domain experts evaluated if the part of the dashboard functionality needed for this task was implemented in a way that is useful for the future users. The domain experts also asked questions from time to time during a task, which led to further discussions. Questions like "which combination method for the CI is used" or "how many indicators are supported" were asked and discussed.

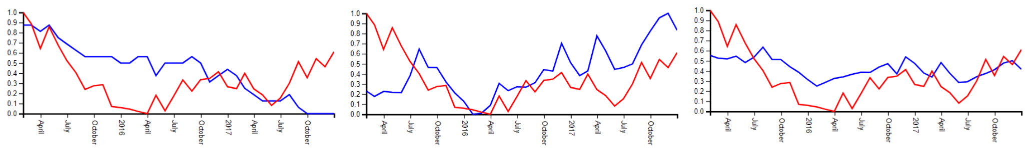
The main task was to find two indicators which combined into a composite indicator offered a higher correlation coefficient than the individual ones of the indicators. This task was separated into further subtasks in order to show different aspects of the combination process. This subtasks along with the information and results of the created composite indicators can be seen in Table 6.1. The respective line charts and the positioning in the ranking of the indicators can be seen in Figures 6.4, 6.1, 6.2 and 6.3. In order to try out

Subtask	CC of 1st indicator	CC of 2nd indicator	CC of CI
Combine two indicators with a low to average positive CC	0.3	0.19	0.68
Combine two indicators with a low to average CC, with one being positive and one reversed negative	0.25	original CC: -0.3 reversed CC: 0.23	0.53
Combine two indicators so that the resulting CC is greater than the best CC (0.69) from the ranking	0.69	0.61	0.74

Table 6.1: Subtask description and corresponding component and composite indicators.

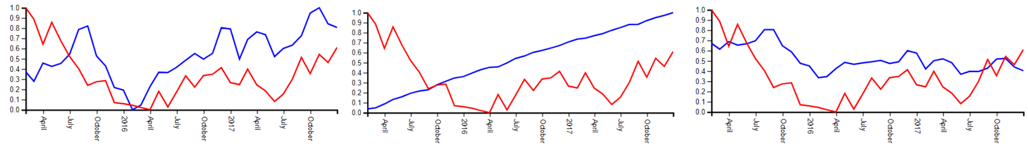
multiple combinations, the buttons of the indicator detail view, which added/removed indicators to the composite indicator, were used. As the first indicator, a random indicator suitable for the current subtask was added to the CI card. The second indicator was selected by iterating over all indicators (matching the subtask requirements) and adding/removing them to the CI until an increase of the correlation coefficient was achieved. Besides this approach, another one for finding the second indicator was performed as follows: view the chart of an indicator through the detail view and select the indicator which could balance some of the strongest disparities between the already selected indicator and the index. Again, proceed until a matching indicator is found.

As shown in Table 6.1, for each subtask, an increase of the CC has been achieved (compared to the highest CC of the 1. and 2. indicator). For instance, for the first subtask, an improvement of over 100% was reached. For the second subtask, a similarly high improvement was achieved by combining one indicator with a positive CC together with an indicator whose time series was reversed due to its original negative CC. It can be observed that by combining indicators with a low CC, a significant increase of the CC of the CI can be achieved. When using indicators with a CC higher than 0.5, the difference between the CC of the CI and of the component indicators is not as big. Nevertheless, as shown through the third subtask, even a CC higher than the best CC from the ranking can be created.



(a) Indicator UNRATE with CC of 0.3 (b) Indicator BDI with CC of 0.19 (c) CI of UNRATE and BDI with CC of 0.68.

Figure 6.1: First subtask.

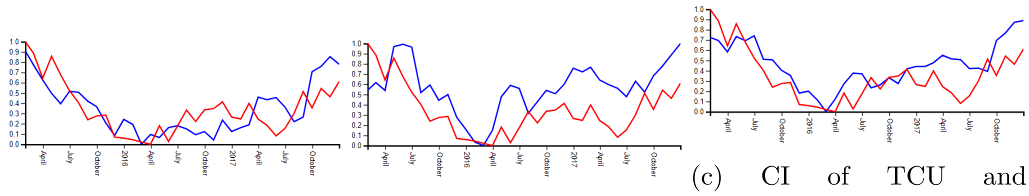


(a) Indicator BSI with CC of 0.25

(b) Indicator PAYEMS with negative CC of -0.3.

(c) CI of BSI and PAYEMS with CC of 0.53.

Figure 6.2: Second subtask.



(a) Indicator TCU with CC of 0.69

(b) Indicator DCOILWTICO with CC of 0.61.

(c) CI of TCU and DCOILWTICO with CC of 0.74.

Figure 6.3: Third subtask.

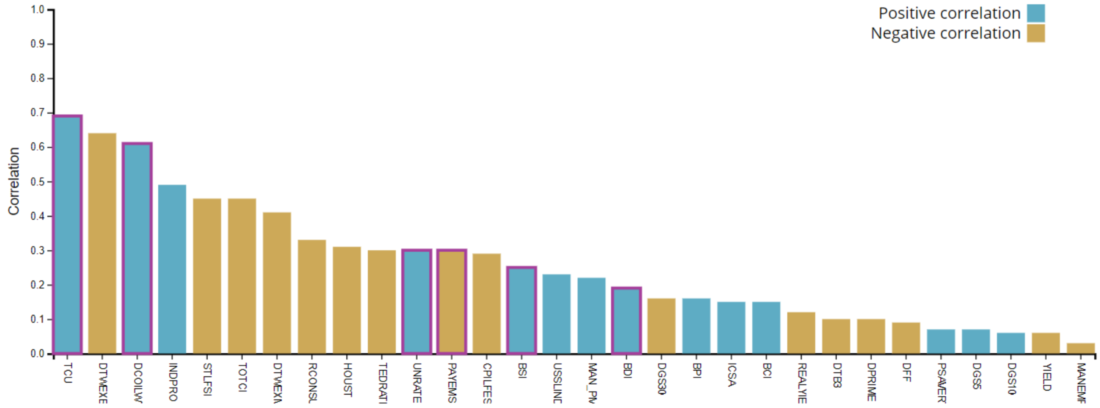


Figure 6.4: All component indicators highlighted in the ranking.

Some of the other task performed during the review were the assessment of the expressiveness of the correlation coefficient by viewing the indicators with the highest, lowest and average correlation coefficient. Furthermore, the possibility to customize the dashboard by changing the cards and creating a different dashboard configuration for the target market "iron ore" was evaluated. The interaction possibilities of the dashboard were also discussed in general. The overall feedback of the domain experts for these tasks was that they match the requirements and that they are (at the moment) sufficient for the dashboard prototype.

Besides the review performed with the domain experts, the overall performance of the dashboard application was tested. One requirement was that the dashboard application can process over 50 indicators. Therefore, the dashboard was tested with 60 indicators, which took approximately 20 seconds from the user login till the correlation coefficient ranking was fully rendered. Compared to this, the loading of the default set of 30 indicators takes about 10-12 seconds. The loading time was measured with the Chrome Web Browser but the dashboard application was also tested using Mozilla Firefox and Microsoft Edge. The loading times using these other web browsers were also similar. During the review sessions, only one to three indicators were combined into a CI, but the dashboard application was also tested with a composite indicator consisting of up to ten single indicators.

CHAPTER 7

Conclusion and Future Work

The developed dashboard application prototype is a unique tool for the exploration of the relationship between multiple different economic indicators and a selected stock market. The user can compare the time series of the indicators and the stock, view the calculated correlation coefficient and interactively compose further composite indicators through combining the existing ones. The resulting insights can then be used to additionally perform a more detailed and specific analysis of the commodity market and verify established assumptions.

During the evaluation of the dashboard application, it was revealed that not only the indicators can be efficiently compared to a certain index but also that composite indicators can be created with strong correlation coefficients. Moreover, the pipeline for the construction of the composite indicators as well as the visualizations were created and designed on the basis of the OECD Handbook [C+08] [GG12] for the creation of composite (leading) indicators. However, not all creation steps are performed yet, as the dashboard is a prototype. The missing steps will be addressed during future developments, after (further) user needs are identified using the dashboard prototype. For the future development of the application, it can be also considered to include a computational approach for detecting lead and lag of the indicators and automatically computing composite indicators with a strong correlation coefficient. Overall, a more advanced composite indicator creation mechanism can be implemented in future.

List of Figures

List of Tables

Bibliography

[AMST11] Wolfgang Aigner, Silvia Miksch, Heidrun Schumann, and Christian Tominski. *Visualization of time-oriented data.* Springer Science & Business Media, 2011.

[BOH11] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. D3: Data-driven documents. *IEEE Trans. Visualization & Comp. Graphics (Proc. InfoVis)*, 2011.

[Bos17] Mike Bostock. D3 Data Driven Documents. `https://d3js.org/`, 2017. [Online; accessed Mar-2018].

[C+08] Joint Research Centre-European Commission et al. *Handbook on constructing composite indicators: Methodology and user guide.* OECD publishing, 2008.

[Eli12] Micheline Elias. *Enhancing User Interaction with Business Intelligence Dashboards.* PhD thesis, Ecole Centrale Paris, 2012.

[EM48] R.D. Edwards and J. Magee. *Technical Analysis of Stock Trends.* Stock Trend Service, 1948.

[Few06] Stephen Few. *Information Dashboard Design: The Effective Visual Communication of Data.* O'Reilly Media, Inc., 2006.

[Few12] Stephen Few. *Show Me the Numbers: Designing Tables and Graphs to Enlighten.* Analytics Press, USA, 2nd edition, 2012.

[GG12] Gyorgy Gyomai and Emmanuelle Guidetti. *OECD System of Composite Leading Indicators.* OECD Publishing, 2012.

[Goo18] Google. Angular docs. `https://angular.io/docs`, 2018. [Online; accessed Mar-2018].

[Inv18] Investopedia. Economic Indicator. `https://www.investopedia.com/terms/e/economic_indicator.asp`, 2018. [Online; accessed Jan-2018].

[Mar16] Lukas Marx. Is Angular 2+ MVVM? https://malcoded.com/posts/angular-2-components-and-mvvm, 2016. [Online; accessed Mar-2018].

[Qua18] Quandl. https://www.quandl.com/, 2018. [Online; accessed Jan-2018].

[Rob07] J. C. Roberts. State of the art: Coordinated multiple views in exploratory visualization. In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*, pages 61–71, July 2007.

[SF06] Andre Suslik Spritzer and Carla M. D. S. Freitas. A visual tool to support technical analysis of stock market data. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '06, pages 512–515, New York, NY, USA, 2006. ACM.

[Shn03] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *The Craft of Information Visualization*, pages 364–371. Elsevier, 2003.

[Tra18] TradingView. https://www.tradingview.com/, 2018. [Online; accessed Jan-2018].

[WBWK00] Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '00, pages 110–119, New York, NY, USA, 2000. ACM.