Dr. Jan Wedigo Radermacher, Ph.D.

Chair of Finance

Goethe University Frankfurt

**Financial Economics with Python**

Ein Bild, das Text enthält.

Automatisch generierte Beschreibung

Submitted on 30 January 2023 by:

Marvin Silva Fortes (..)  
Marc Eberle (7218978)

Alexander Weinbuch (7240559)

Table of Contents

[1 Team 1](#_Toc125912201)

[2 Development of Idea 1](#_Toc125912202)

[3 Product Development 2](#_Toc125912203)

[4 Application of Ideas and Features (incl. Problems) 4](#_Toc125912204)

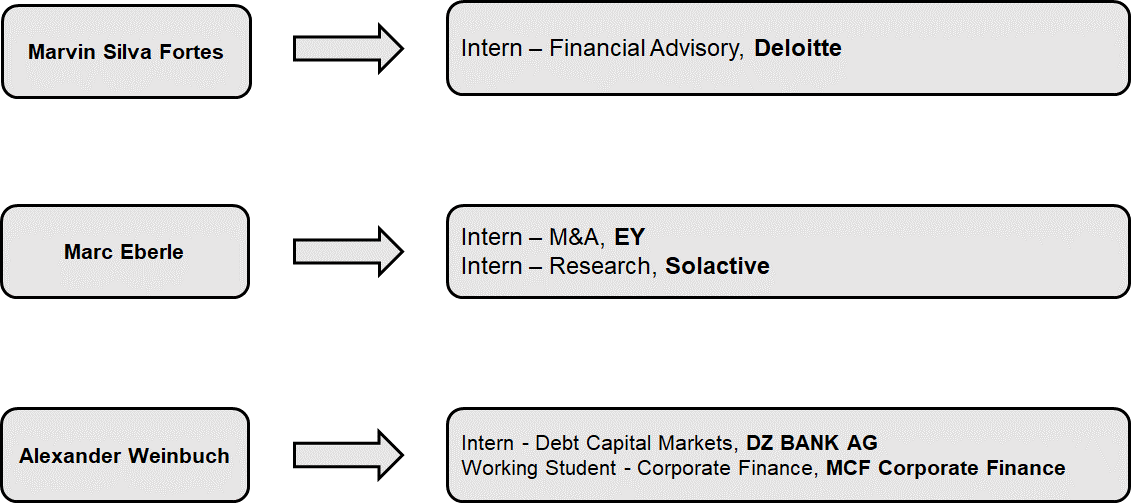
[4.1 Alex 4](#_Toc125912205)

[4.2 Marvin 9](#_Toc125912206)

[4.3 Marc 9](#_Toc125912207)

[5 Project Summary 17](#_Toc125912208)

# 1 Team



# 2 Development of Idea

As shown in the team introduction, Alex worked for DZ BANK in Debt Capital Markets. During that time, he was confronted with insider trading for the first time as he was not allowed to publish any confidential information and trade stocks only with permission. Following this experience, made him think about insider trading and wondered whether insiders are better investors than other investors. But with finishing the internship and therefore the end of mandatory reports to the Compliance team in the bank, insider investing was not a frequently appearing topic anymore.

A few days before Prof. Radermacher required the groups to select an idea for their dashboard, Alex was listening to one of his and Marc’s favorite podcasts “Doppelgänger TechTalk” by Philipp Klöckner and Philipp Glöckler and Philipp Klöckner is one of the angel investors at Gorillas (now part of Getir). The two podcasters talked about the earning published by Just Eat Takeaway and Delivery Hero and Philipp Klöckner explained to the audience that due to his investment he is not allowed to talk about competitors and developments in the industry. With the topic of insider trading coming to Alex’s mind, he asked Marvin and Marc to include the topic in our Python project.

# 3 Product Development

Based on our understanding, the most important part of developing a dashboard for the first time is a well-defined niche which needs further applications to become accessible to the public or at least the people who are really interest in the topic. Since all group members are interested in capital markets and invest actively in different markets, we agreed on the topic of insider trading and started to look at the current situation of access to information in this niche segment. This was one of the first issues we faced in our project as most websites only display selected information on insider trading or show outdated data that is published in SEC Form 4 Filings. After several searches and general research we found some websites that display the information we are interested in. Most of the websites display the information in an extremely messy and confusing format, which makes it almost impossible for “standard” retail investors to understand who is treading how many shares, how many shares does the person hold and most importantly is this information at all relevant.

Due to this situation, we decided to focus our project on making information on insider trading accessible to retail investors, comparing the performance of insider trades to the market, and help investors to make use of this information in a broader context, e.g. understand the market sentiment.

We started looking for further websites we could use to apply Web Scraping and gather relevant information. One main obstacle we faced and still face is that those websites with most sufficient and best-formatted data are only accessible through a subscription. At the same time, these circumstances motivate us to bring (more) transparency into a very important niche segment that provides crucial information as from our perspective (potential) investors have a right to have proper information on what insiders (e.g. C-level executives, x% owners, and founders) are doing with their holdings in the companies.

What all websites, we visited during our initial research and development phase, are missing:

1. Newsletter that informs the (potential) investor on a monthly basis about what insiders of the largest US companies (investor’s companies of interest) are trading.
2. Integrated overview of a firm’s stock price development, news on the firm
3. Comparison of the performance of insider trades with the market in the relevant period (S&P500)
4. Giving information on the most relevant question: Do insiders generate Alpha? Are insider better investors?

To implement these features and give information on these questions, we started to build the tool around our initial idea.

Within the process, we even developed a company/tool name with a slogan:

* Company/Tool name: InsidersInvest
* Slogan: Transparency for all

We even implemented another graphical element as we think that this best describes the purpose of our tool. The flower we used is growing from beneath the surface, which is where information on insider trades currently is located. Our tool helps information to spread, grow, and reach sunlight to help investors make for fundamental decisions. Furthermore, investors knowledge is also growing like a flower – first there is a seed, planted beneath the surface but with water (relevant information) the seed develops into a beautiful flower.

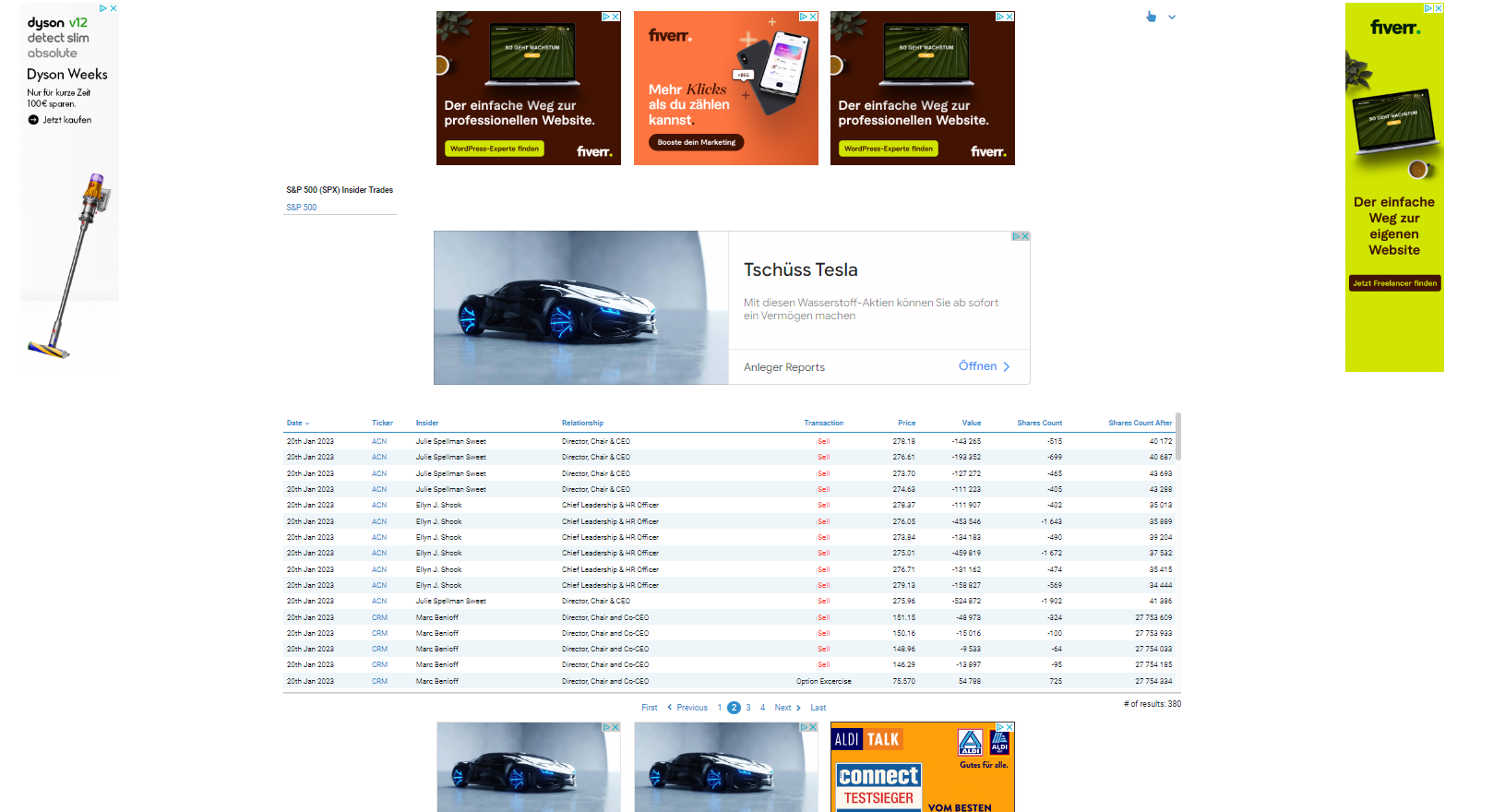
In the following sections, we will explain our different tasks, challenges, and results.

# 4 Application of Ideas and Features (incl. Problems)

In the following sections Alex, Marvin and Marc will describe their individual tasks, problems, learnings and present their results.

## 4.1 Alex

After developing and formulating the idea, we started to look for websites that provide us with reliable and recent data. As described above, one of the main problems was that most websites that provide promising content are paid services. On the other hand, most free websites are extremely messy and only provide outdated or parts of relevant data. We found a website that seemed to fit our needs and help us to gather relevant data, but we faced several obstacles as there are irregular pop-ups and advertisements on the website which prevented us from scraping data.



<https://www.finscreener.org/insider/insider-trades/sp500?o=1002&pg=2>

Furthermore, we wanted to use <https://finviz.com/insidertrading.ashx> but on the website only the 200 most recent trades are displayed, which is not enough data for our analysis. Therefore, we had to start again and spent time on finding the best possible website for our data. After checking dozens of websites, we found <http://www.openinsider.com/> , which is the website we use to scrape the data on insider trading.

Following this, we started to think about special features that our tool should provide to differentiate from other tools/websites that only show data without “guiding” the (possible) investor through the information. In this process, we had the idea to implement a newsletter and “develop” the InsidersInvest Alpha as an indicator for insider’s trading performance. The measure is constructed based on the following idea: As our data is focused on companies that are part of S&P500 index, we use the performance of the index as our benchmark.

Comparing these two values leads us to InsidersInvest Alpha:

This performance measurement is calculated and displayed for all of the trades we provide data for when filtered through the filter option. Problems, (e.g. data format) arising during the implementation of the InsidersInvest Alpha are described by Marvin.

Ein Bild, das Text enthält.

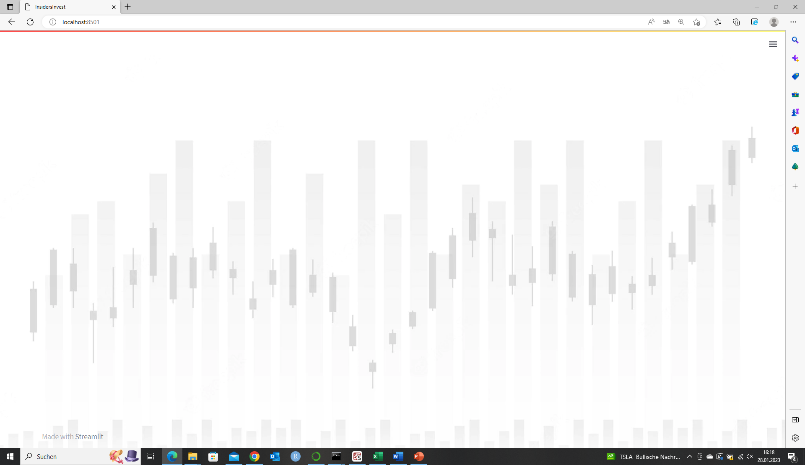
Automatisch generierte Beschreibung

As our tool needs a structured and clearly laid out landing page/homepage, we used streamlit to create one.

To make our website look professional, we decided to implement a background which was a problem at first. Based on different suggestions out of stackoverflow/github, we had to use a CSS file to implement a background. After working with the CSS file and making first steps with HTML as a language, we found a work-around:

Ein Bild, das Text enthält.

Automatisch generierte Beschreibung

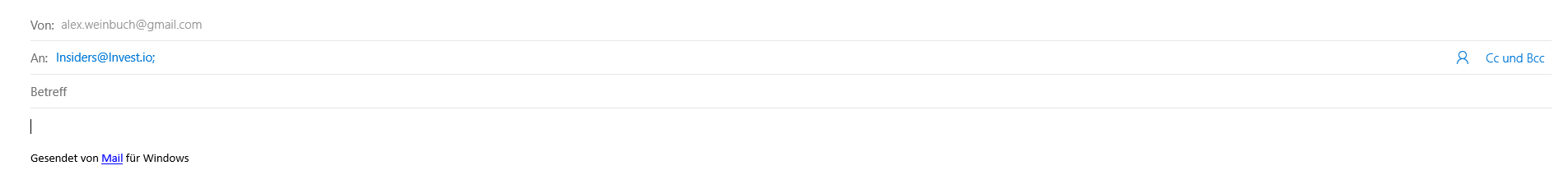


After setting a background, we started to work on the structure of our website and thought about first details on our design. Due to that, we agreed on a name for our tool and a slogan. As described above, we created a logo that fits our purpose.

In addition, we created a button to sign-up for our newsletter. After clicking this button our business e-mail pops up. When a user clicks on the e-mail address, the default e-mail application with our mail address already set as a default value opens. Writing an e-mail to us registers users for our monthly newsletter.

Ein Bild, das Text enthält.

Automatisch generierte Beschreibung



When implementing the logo and the newsletter button, we struggled with centering the image and placing the button next to it. We could use a CSS file to solve the problem. But we worked with columns instead:  
Ein Bild, das Text enthält.

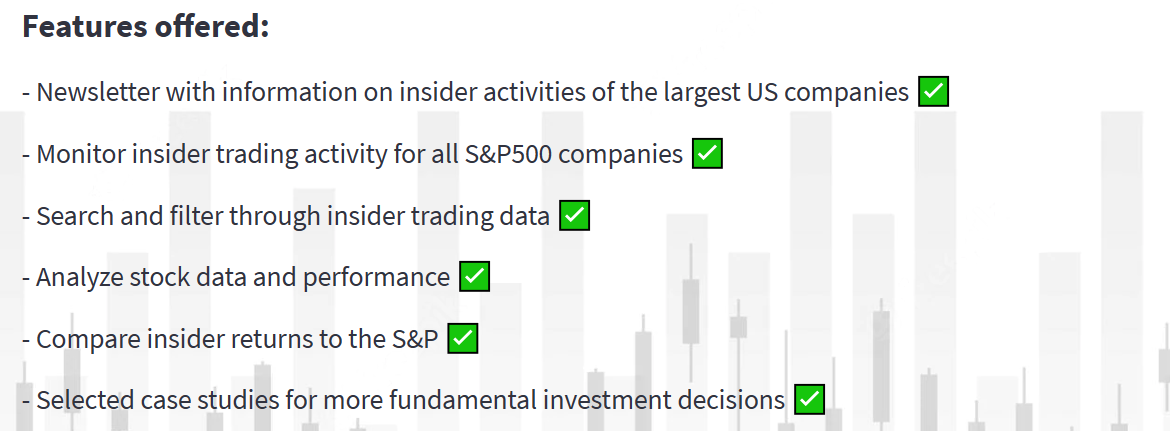
Automatisch generierte Beschreibung

Our homepage gives a quick overview on what our tool is designed for and why an investor should use it.

Ein Bild, das Text enthält.

Automatisch generierte Beschreibung

After “introducing” our product, our homepage gives an overview on our features and what to expect from using our services. Furthermore, we give an insight on the importance of insider trading by implementing a button that opens an article by New York Post about Nancy Pelosi outperforming the S&P500 since 2019. We chose this article as it shows that insiders’ portfolio (not only corporate insiders) deserve to be looked at. In addition, we show a graph that compares the performance of a portfolio based on Congress buys and S&P500 in 2021/22. This portfolio is created by “Quiver Quantitative” whose website can be reached by clicking on the button next to the graph.



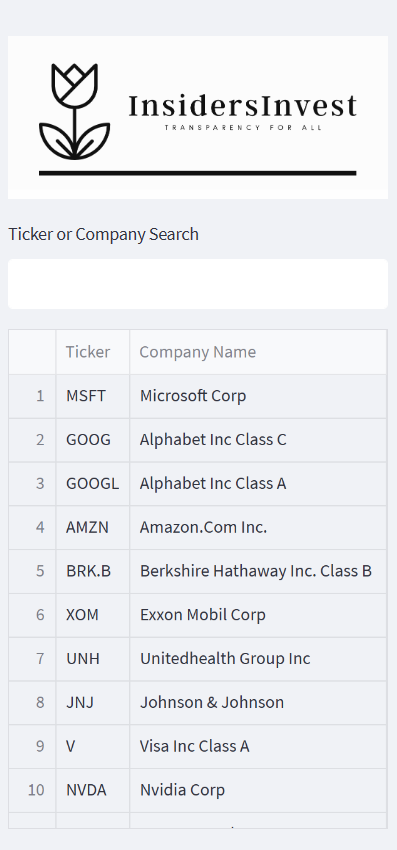


A user might think “Why are they showing data/providing links to information that refers to politicians as insiders?” – we do this because there is substantially more information on insider trading by politicians but not on corporate insiders. Our purpose is to bring more transparency into the world of insider trading by managers, owners and other investors connected to the largest companies of the United States. That is why we included the following two sentences:  
Ein Bild, das Text enthält.

Automatisch generierte Beschreibung

The ”Try it now”-button is linked to our tap “Most recent trades”.

Furthermore, we included a sidebar that allows the user to choose from all of the S&P500 companies. The selections made become effective when using the other tabs.

Ein Bild, das Text, Bildschirm, Monitor, Screenshot enthält.

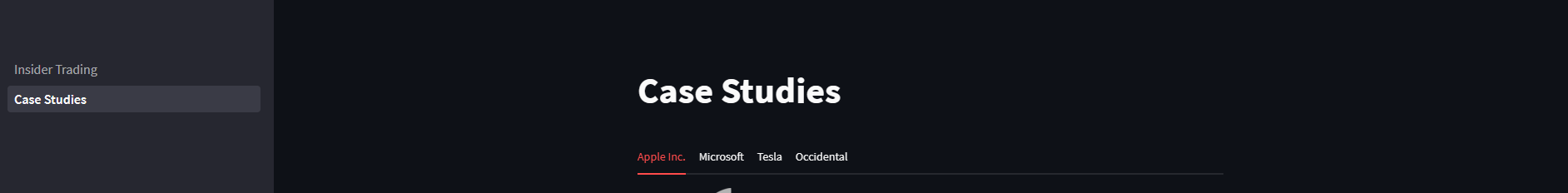
Automatisch generierte Beschreibung

## 4.2 Marvin

XXXX

## 4.3 Marc

One of our main goals is to give the user a deep insight into the insider trading activities of certain companies. This is the reason why we decided to implement a case studies section in our dashboard. The section is subdivided into different pages, so that the user can choose from interactively. On each page, we give a short introduction to the company and show different graphs, providing different important information on the activity of insiders of a specific company. Finally, we added a conclusive part, where we provide our insight into the specific case to give the user a certain guidance.



As mentioned before, we implemented an additional section for the case studies for the case studies. Using the tools which the streamlit package provides, we create 4 tabs for each company.

To start off with each case study, we give the user an introduction to the company. Each case study starts with a company logo and a short description of the business to give the user an understanding of the business. Additionally, a map with the location of the headquarters is included. The map is interactive and extremely user friendly and by that, increases user attention and interaction with our dashboard.

Text

Description automatically generated

Another crucial element of each case study section is the share price development over the last 12 months. Providing the user with the share price gives him the chance to compare the share price development with the insider activity, which is also part of our final analysis.

Chart

Description automatically generated

To access the necessary data, we used the yahoo finance API. The API provides us with daily closing prices of the stock over the last year. The data is stored in a pandas DataFrame. An initial problem was the date format, and its took a lot of effort to convert the dates to the user-friendly format that is now implementation. For the stock prices, we use the line graph provided by the streamlit package. The advantage of the graph is that it is interactive, and users can zoom to specific dates to get a closer look at the share price.

Text

Description automatically generated

In the next step, a scrape is necessary to implement and we decided to use the selenium package. The scraping is the process we are most proud of and which was very difficult to implement. In the screenshot provided above, you can see a part of the code responsible for our scraping process. We use the Nasdaq website to scrape our data from, since they provide the latest data on firm specific insider information. After accessing the website, we immediately minimize the window to ensure user-friendliness of our dashboard. Another preparational step is accepting the cookies for the website and our code automatically accepts the cookie settings. To ensure a smooth and reliable scrape, we implement waiting times. We take the hit of longer waiting times to ensure the stability of our dashboard. In the process of creating the code, we struggled several times with errors resulting from missing waiting times.

In the next step, the table that contains the relevant information needs to be identified.

The most difficult step was to understand how the table which contains the information is structured and how we can access the data. This process required a lot of sweat and tears.

We finally came up with a solution, that finds all the elements with the tag “tr” (table row) in the table element and returns all the matching elements. A loop is started afterwards that iterates through each row element in the table. All elements with the tag “td” (table data) are then found in the current row element. A new list called “row\_data” is created and filled with table data which was identified in the previous step. The last line of the loop adds the “row\_data” list to the “data” list. This will append the data of each row to the “data” list, so that the final “data” list will contain a list of all the rows in the table, and each row will contain a list of the data in the cells of that specific row.

The data is then transformed into a pandas DataFrame with the column names “Insider Trades”, “3 Months” and “12 Months” since it is way easier to work with the DataFrame. After, the first row of the DataFrame is dropped since it does not contain relevant information. The whole process provides a DataFrame that looks like the following for Apple:

Text

Description automatically generated

Not only do we scrape the number of insider trades, but also the number of insider shares traded. The overall scraping process remains similar, but the location of the table needs to be redefined. Unfortunately, the class name for both tables is identical, that’s why we need to pre-specify the section in which the code is supposed to locate the table. An example of insider shares traded is the following, this time from Tesla Motors:

Text

Description automatically generated

In the following step, we plot the data that we gathered. Again, we encountered another major issue: Python could not recognize the data as integers or floats, but instead as strings. To solve this problem, we defined a new function to convert the strings to integers. It is way more difficult than initially expected, since the parentheses and commas in the numbers result in errors. That’s why we first need to get rid of the parentheses and commas, before returning an integer.

Text

Description automatically generated

We then apply the function to the DataFrames to ensure the correct data type.

Chart

Description automatically generated

We then use our scraped data to nicely plot the data in our dashboard. We decided to use matplotlib graphs, since they provide the necessary versatility to present our data and bring our intended message across to the user. The advantage of an interactive streamlit graph does not justify the huge decrease in information density. It is important to plot two bars for each category to make users aware that insider trades are usually not equally distributed over time. Another detail of the graphs is the data labels. Data labels are used instead of a y-axis to make the graph more compact and better looking.

Chart, waterfall chart

Description automatically generated

The next graph shows information on how many shares were traded by insiders over the last 3 and 12 months. The graph follows a similar structure to the first one, since a bar graph is the best way to present this kind of data. This way, it is easiest to understand for the user. The graph includes four categories, the number of shares bought, number of shares sold, total shares trade and the net activity. In our opinion, it is important to show the difference between the number of trades and number of shares traded to make the user aware of the different sizes of trades. Further, the number of shares traded gives the user an indication of how much volume was actually traded. The net activity is an important indicator to set into relation with the share price over the specific time frame.

Chart, pie chart, bubble chart

Description automatically generated

Finally, to additionally illustrate the distribution of shares sold to shares bought, we include a pie chart plotting this information. With this graph, we want to make sure that the user can easily get a feeling for the sentiment of the insiders.

A screenshot of a computer

Description automatically generated with medium confidence

For the interested users, we include detailed information of the last 15 insider trades. An additional scrape was necessary to generate the table. Again, the Nasdaq website is our source. We generate a pandas DataFrame with the underlying data and then generate an interactive table.

Text

Description automatically generated

As a last part, we include a short opinion on the company and the insider’s activity, as it can be seen above for Tesla. With this part, we want to give users an additional insight into the specific case and help users to put the information into context.

# 5 Project Summary

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Marvin** | **Marc** | **Alex** |
| **Main Field of Study** |  | Finance and Accounting | Finance and Accounting |
| **Interests** |  | Corporate Finance, Capital Markets, Tennis and Cycling | Football, Hiking, and Capital Markets |
| **Professional Experience** |  | M&A, Research | Debt Capital Markets, Corporate Finance |
| **Prior Experiences with Programming** *On Scale from 1 to 10 (with 1 as lowest)* |  | 1 | 1 |
| **Own project started before  (not only in the context of Programming)** |  | Co-founder of a student company | No |
| **Programming Languages (before)** |  | None | None |
| **Programming Languages (after)** |  | Python | Python, HTML |
| **Biggest Challenge(s)** |  | Implementing the web scrape, create user friendly graphs, time management | Understanding Web Scraping; Using CSS flies; Start thinking like an project co-manager; Time management (Bachelor Thesis) |
| **Biggest Struggle(s)** |  | Number formats | Relying on blog entries about Yahoo Finance API which was discontiued |
| **Biggest Learning(s)** |  | Resilience pays off | There is always someone who had the exact same problem before - just google it! |
| **What do you like about the product?** |  | Interactive and user-friendly interface that provides unique information | Easy to access and understand, filter options, InsidersInvest Alpha (never seen before) |
| **What will be your next improvement?** |  | Add additional case studies, improve the scraping process to make it faster and more efficient | More Case Studies, Weekly Newsletter, UI optimization |