Group Members:

- Ansam Zedan
- Daniel Wullschleger

General Idea:

We want to analyze the impact of news events on Microsoft's stock price in the past 2 years. The idea is to investigate whether the news sentiment scores (Positive / Neutral / Negative / No news) have had an impact on its stock price and if they can be used to predict market movements in the future. This is especially interesting since Microsoft openly communicated about their stake in OpenAI and there have been many events that may have had an influence. More specifically:

- Microsoft investing 10 billion dollars into OpenAI
- ChatGPT being released and created an AI-hype
- Microsoft releasing Bing Co-Pilot and Microsoft Co-Pilot
- The release of DALL-E 3 disrupting the image gen AI market.
- Sam Altman leaving OpenAI in November 2023
- Adoption of GenAI by many small and large enterprises

As for the Train/Test split, we plan to use the period of 01.01.2022 - 30.06.2023 for training, and the remaining period up until today for testing (However, this scope may be subject to change).

Data & data sources:

For the financial stock price information, we plan to make use of the yahoo finance API via the [quantmod library](#) in R, where we plan to get daily stock price movements. Using this as a basis will allow us to calculate the daily price movements as shown in class

As for the sentiment data, we plan to scrape Reddit posts of the relevant period using the python library selenium. Once we have cleaned the news texts, we can use the python library textblob. This way, we can evaluate the sentiment of each mentioned word in an article, to then aggregate them into an article-sentiment.

In the end, we will end up with 2 datasets:

- Daily MSFT stock price movements in the past 2 years
- MSFT news sentiments as granular as possible (it may be the case that there is no news on a particular day)

Methods applied:

For the time series analysis, we plan to apply a Vector Autoregression (VAR). Meaning, we aim to create a model that learns from past stock price movements while considering general news sentiment on this day.

To validate the impact of news on the stock price, we will focus on a multivariate analysis (incl. Sentiment Data) and will use a univariate analysis as a benchmark. This way, we may be able to highlight a difference in the model's accuracy.