Introduction to Social Data Science: Supervision Sheet

Department of Economics, University of Copenhagen

Summer 2023

This document asks you a series of questions about your exam project that you should try to answer. At this point, we do not expect you to have complete idea about what you will eventually be working on. However, we are hoping that — by filling in this document — you will become motivated to start brainstorming and look for ideas for your exam project that both (i) are interesting to you, (ii) allows you to demonstrate that you master the learning outcomes that were stated in the **course description**, and (iii) are feasible. Furthermore, this document will also hopefully help you make most out of the time with the TA at your supervision meeting.

You should have prepared and printed this for your first supervision meeting.

1. What is your 'research question'? That is, which 'phenomenon' will you try to understand with your project?

We want to research the correlation (if there is one) between finance news and the stock market. Initially we want to look at NASDAQ companies to shorten our research. Therefore, the phenomenon we try to understand is how people (hence the market) react to news about stocks.

2. What data are you going to use? And how will you get your hands on these data?

We are planning to use an API to collect stock prices (either google finance or yahoo finance). This is quite simple. The challenge lies in scraping news articles and rating them (negative or positive). To do this we consider using openAI's api to get chatgpt to rate the articles

3. What theoretical and practical concepts are you going to work with?

Natural language processing and text sentiment analysis(with openAI api). We also want to use a simple machine learning model to see if it is possible to predict future movements in the market

Lasso & Ridge regressions

4. Who will be responsible for what (answer if possible)?

We don't know yet, we are still trying to see if the project is doable