**Research Questions**

Throughout the eight-week fellowship program I am planning to work on ”Facebook Followers dataset” provided by The Data Incubator and Thinknum. The data is obtained by subscription at <https://thedataincubator.us8.list-manage.com/subscribe/confirm?u=70e04e2160786cdebf3df2567&id=fbf1336bda&e=eb1ca5e7ba>

The date contains information about company name, number of check-ins for specific day, were- here -counts, likes and talk about. The data is used to answer the following research questions :

1. When the company is visited most and in which day?
2. From companies in the same industry which company was visited most in a specific day? Provide Realtime information for companies for their social-media share in the industry.
3. For a particular date, what portion of the check-in converts to “like” which gives information for the company how much of their visitor like them for a particular day? Again for marketing purpose to recommend visitors the more likable companies in the industry.
4. Does the proportion of likes per check\_in determine stock price for publicly available companies in a particular day? If so, companies should strategically work to get more check-in and likes.
5. Does the talk about a company has positive or negative sentiment?

**Relevance**

Answering the above questions help companies to understand their social-media share, how relevant is this share and how to increase it, what the talk about them implies, do this all affect their stock price. All this information can be used for marketing and strategic planning and allow then to play better in the market.

**Methodology**

The first part of the research question can be addressed by data exploration and analysis and last two need machine learning methods such as regression, clustering and sentiment analysis.

First group the companies into industries. Companies information could be obtained from company registry information publicly available at <https://www.uscompanieslist.com/> . Once the industry information for each company is collected, group the data by industry and see how the proportion of “check-in” a company gets from the industry. This can be streamed and can be used for marketing purpose.

If Facebook allows us to scrap its webpage or collaborate to provide us the talk about a company, the talks can be clustered into groups and further analyzed to understand the sentiment of the talk. For the first phase, cluster analysis will be performed and then sentiment analysis to classify the talk.