Erin tansley – davita – 720.224.1442 – ettansley@outlook.com  
Predicting returning to the office post-COVID-19

05.04.2021

# Description of project

## What question or problem are you addressing?

|  |  |
| --- | --- |
|  | I would like to answer the question – can you predict if office workers will want to return to the office after the COVID pandemic has subsided based on external factors? |

## What type of data science task is it?

|  |  |
| --- | --- |
|  | Prediction based on supervised learning. |

## Brief description of data:

|  |  |
| --- | --- |
|  | I am considering creating my own questionnaire and gathering data via that route so I will have to figure out how much training data and test data would be needed overall. For the training data I would have to interview X number of people based on an answer of “yes I want to return to the office” and subsequent answers based on a variety of different environmental, social, economical, and political factors. From there, I am considering creating randomized test data with a version of every scenario for each question asked. This has the potential to be pretty big if I have a lot of questions, I want answered. Gathering the data for the training set could potentially be a huge hang-up, however, I have some potential sources (mturk, individuals at my company, Reddit sample sizes) that may work. |

## How will you analyze the data? What machine learning methods do you plan to use, and/or what business intelligence aspect do you plan on incorporating?

|  |  |
| --- | --- |
|  | I will be analyzing the data in R and will be using classification methods, most likely a mix of ISLR, Tibble, Caret and Naïve Bayes but the method could change based on the data I receive and if a different method appears to be more appropriate for my data. I am currently planning on using Plotly or ggplot2 to visualize the data. I do not currently know if I will present my findings to my business or not but any further clarification on what aspects are good for incorporating would be welcome. |

## Describe any anticipated difficulties and problems. Discuss how you may overcome the problems.

|  |  |
| --- | --- |
|  | My biggest difficulty or problem is going to be getting the data I need since I am asking extremely specific questions both in terms of getting enough responses that indicate that they want to go back to the office and in a timely manner. To overcome the problem, I might have to pay people out of my pocket, or I might have to aggressively approach individuals in settings that might not be comfortable to me such as reaching out to people at work that I know or via LinkedIn. I may also have to investigate seeing if there are any additional existing data sets that match close to my desired questions if I cannot get enough responses. I also anticipate problems with the process as a collective event since I have never completed a project from scratch start to finish. The visualization packages in R have not felt that intuitive to me and have felt difficult to navigate in terms of text size or data point size. Lastly, I do anticipate problems with understanding and being able to clearly communicate the results in a meaningful way. I hope to overcome these problems by reaching out for help via the instructor or potential online resources that may be able to walk me through interpretation and I hope to overcome these problems by starting as early as possible on the analysis portion. |

## Timeline:

|  |  |
| --- | --- |
|  | **Week 1 – 2:** Survey creation, Data Capture, Data Cleaning.  **Week 2 - 3:** Package and Model finalization, preparing any final scientific question outlines, testing data for imbalance and usability, data exploration.  **Week 4 – 5:** Modeling the data, working through the experiment, coming to conclusions, creating the base outline of the final project and base presentation.  **Week 6 – 8:** Finalizing the presentation, Finalizing the visualizations, Creating the explanations of the concepts and methods, Ensuring code is documented in GitHub and accessible, Creating a basic website page to house all of this on/landing page, engage with other students to provide feedback, Decide on if I want to present this to my team or company – likely not formally but may at least show my supervisor what I did and if she feels it could be useful to pass on I would give permission at that time. |