Trevor McGlaflin

April 1, 2022

STAT 287

Project Check-in #3

**Working Title:** What Factors Drive Employee Turnover?

**Summary**

The goal of our project is to figure out which factors about employment lead to a higher employee turnover. Our data set contains fields such as salary, time since last promotion, time with same manager, stock option amount, and other benefits. Then, the same people were asked later if they stayed at that company. This field will be used as our dependent variable in our machine learning models.

**Major Activities Planned for This Week**

1. Finalize 5 or 6 good-looking plots that show major trends

2.   Explore and query Twitter API

**Major Activities Accomplished This Week**

1. Made 5 or 6 good-looking plots that show major trends

2.   Explored and queried Twitter API

**Open Challenges and Questions**

1. Is our data set large enough with only 1400 records?
2. There is risk for company bias because data is collected from only IBM
3. Could we potentially add a dimension to our project that took data from twitter to pinpoint popular work vacation spots by finding tweets with vacation key words and using the location those tweets were posted from.

**Major Changes to Research Plan**

We have decided to incoordinate some exploratory data analysis from the twitter API into our project. Not sure if this data will be used in the ML algorithm yet.

**Revised Week by Week Timeline**

**Week of Mar 28**

1.     Finalize 5 or 6 good-looking plots that show major trends

2.   Explore and query Twitter API

**Week of Apr 4**

1.     Start experimenting with machine learning algorithms

**Week of Apr 11**

1.     Finalize one or two machine learning algorithms to include in our final analysis

2.     Include a comparison of the accuracy of these algorithms

**Week of Apr 18**

1.     Prepare for presentation