Abstract

Recent studies and news articles have come out and shown how high workplace stress is inversely correlated with productivity. This decrease in productivity can come from high turnover to a lower quality of work. Gaining an understanding of these stressors can help a company create programs that would relieve some workplace stresses for employees in a company. Which in turn could increase a company's profits or help a company save money from tasks such as hiring and training new employees due to high turnovers. A linear regression model will be built to gain an understanding of general stressors and programs will be made to address these problems.

Design

To gain an understanding of the what are the key stressors of people's lives we will take responses from the publicly available Depression, Anxiety, Stress Scale test, conduct EDA and create a baseline linear regression model. This will give us a baseline understanding of general stressors in people's lives. From the results of the first linear regression model, we will create a new questionnaire that includes these general stressors with the addition of questions that ask about aspects of the workplace. Questions would include: "How long does it take to commute to work?", "How welcoming is the team you are currently working with?", etc… We then will build a new linear regression model to gain an understanding of these possible other stressors and finally create programs that will aid in relieving these stresses.

Data

Data was gathered from an open-source website that has public responses from the Depression, Anxiety, Stress Scale test. It has a total of 39,776 responses from a variety of participants around the world.

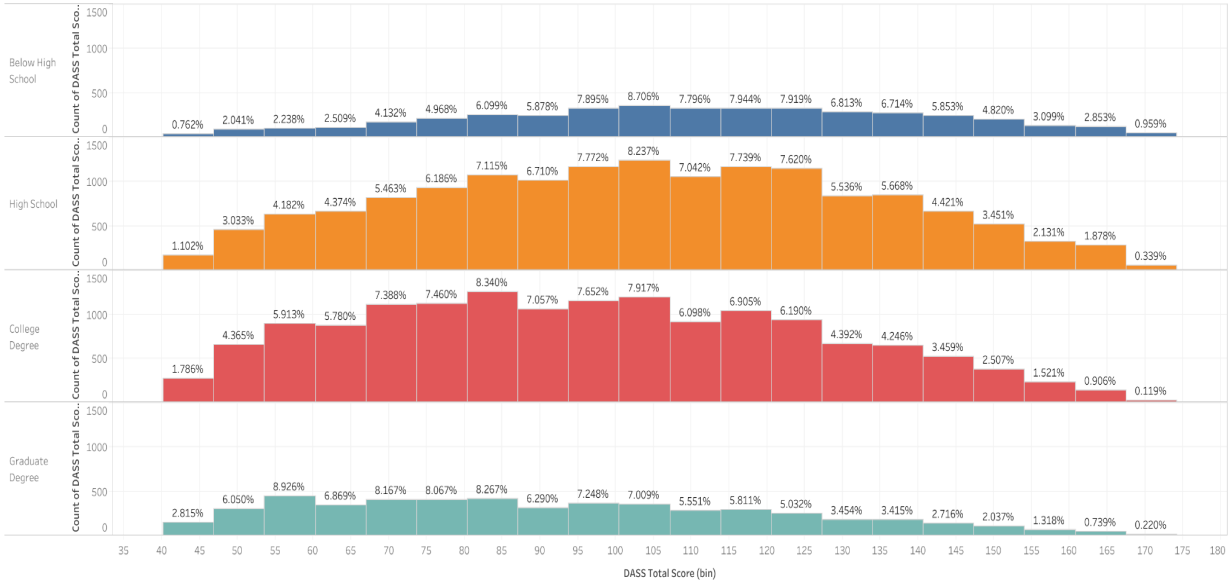
Algorithms

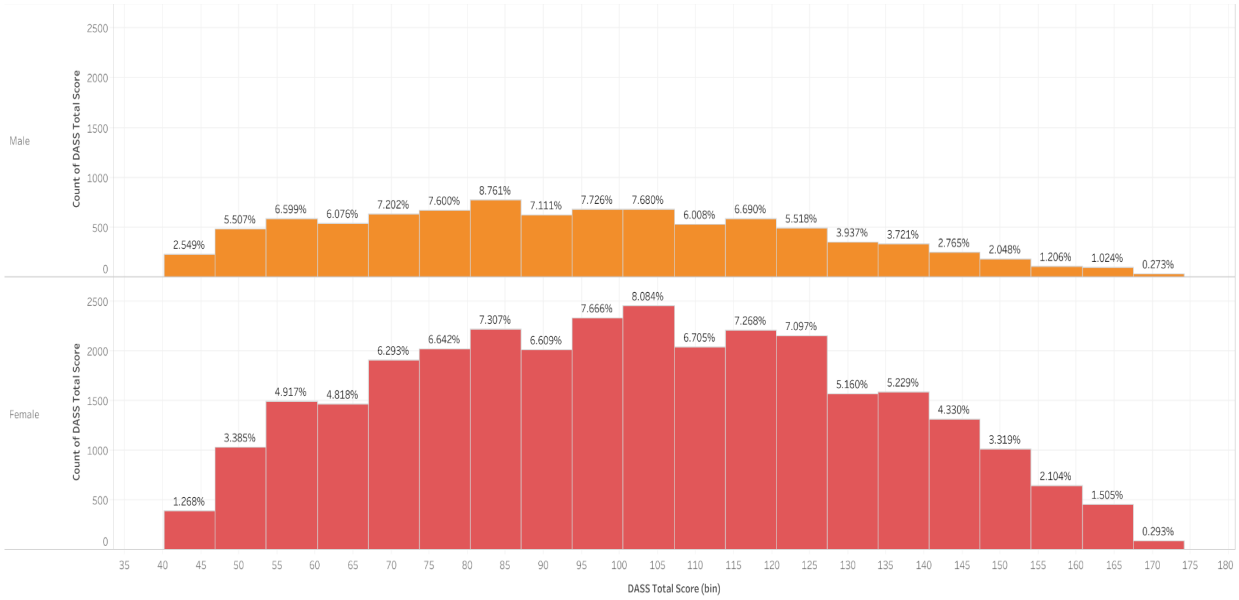
- Exploratory data analysis was done on the raw data to view basic relationships between total score and different aspects of the participant's lives such as race, gender, or education level.
- A basic Linear Regression Model will be created to show proof of concept

Tools

- Google Sheets was used for performing basic EDA on the data
- Tableau Public was used to create data visualizations to see the distribution of data
- Pandas, NumPy, Scikit-Learn, and Statsmodels.api were used to create a simple linear regression model on the data.

Communication




Male
Count of DASS Total Score
2500
2000
1500
1000
500
0
2.549%
5.507%
6.599%
6.076%
7.202%
7.600%
8.761%
7.111%
7.726%
7.680%
6.008%
6.690%
5.518%
3.937%
3.721%
2.765%
2.048%
1.206%
1.024%
0.273%
Female
Count of DASS Total Score
2500
2000
1500
1000
500
0
1.268%
3.385%
4.917%
4.818%
6.293%
6.642%
7.307%
6.609%
7.666%
8.084%
6.705%
7.268%
7.097%
5.160%
5.229%
4.330%
3.319%
2.104%
1.505%
0.293%
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110
115
120
125
130
135
140
145
150
155
160
165
170
175
180
DASS Total Score (bin)