Term Project Data

Jordan Montgomery

QMB6358

Due 11/8/2020

**Are Emily and Greg More Employable Than Lakisha and Jamal?**

Building a regression model to identify resume attributes which correlate to receiving a call from an employer.

**Project Scope – UPDATE:**

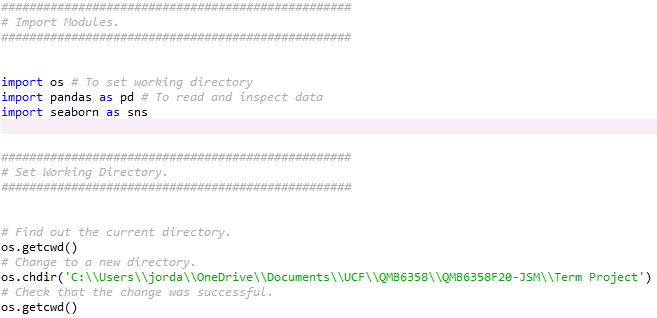
Upon review of the data using Python, I have decided to only use 17 of the 26 exploratory variables for the analysis with ‘call’ – whether or not the resume garnered a call back – being the dependent variable:

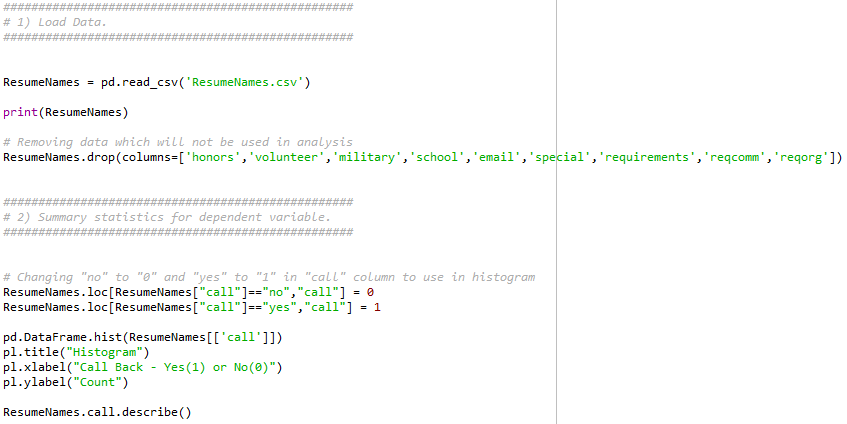
|  |  |
| --- | --- |
| 1 | **name** |
| 2 | **gender** |
| 3 | **ethnicity** |
| 4 | **quality** |
| 5 | **city** |
| 6 | **jobs** |
| 7 | **experience** |
| 8 | **holes** |
| 9 | **computer** |
| 10 | **college** |
| 11 | **minimum** |
| 12 | **equal** |
| 13 | **wanted** |
| 14 | **reqexp** |
| 15 | **reqeduc** |
| 16 | **reqcomp** |
| 17 | **industry** |

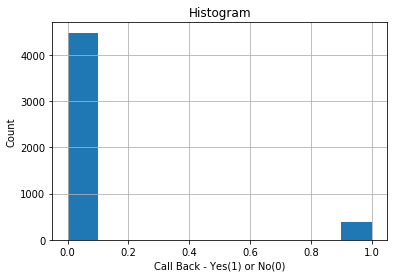
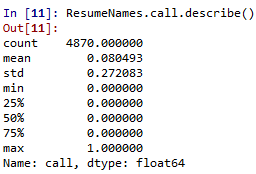
|  |  |
| --- | --- |
| 18 | **honors** |
| 19 | **volunteer** |
| 20 | **military** |
| 21 | **school** |
| 22 | **email** |
| 23 | **special** |
| 24 | **requirements** |
| 25 | **reqcomm** |
| 26 | **reqorg** |

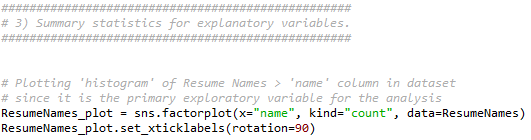
In the **ResumeNames.xlsx** file, the data is defined and sorted on the **Variable Details** tab.

The **Project\_Data\_-\_Analysis.py** file in the repository contains code to display some summary statistics and analysis of the data in Python. Screenshots of the code and charts are below as well:



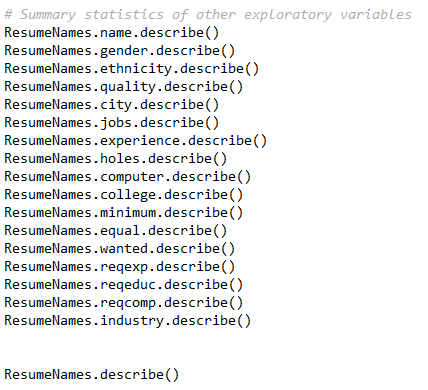




Chart, bar chart, histogram

Description automatically generated



# References

Arel-Bundock, V. (2007). *Vincent Arel-Bundock's Github projects.* Retrieved from R Datasets: https://vincentarelbundock.github.io/Rdatasets/datasets.html