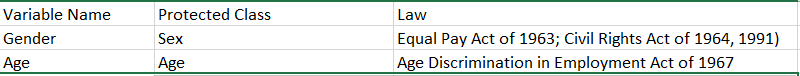
Stats 101 Project

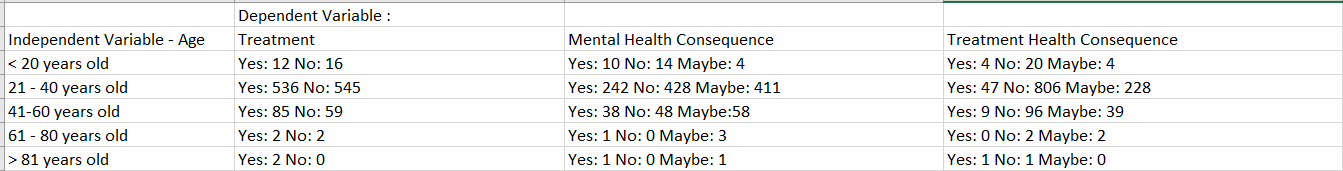
Wesley Tomjack

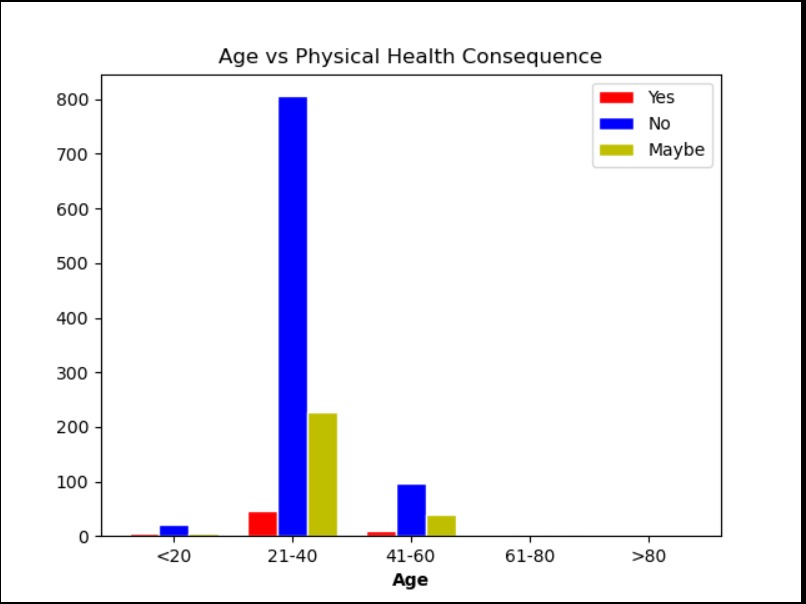
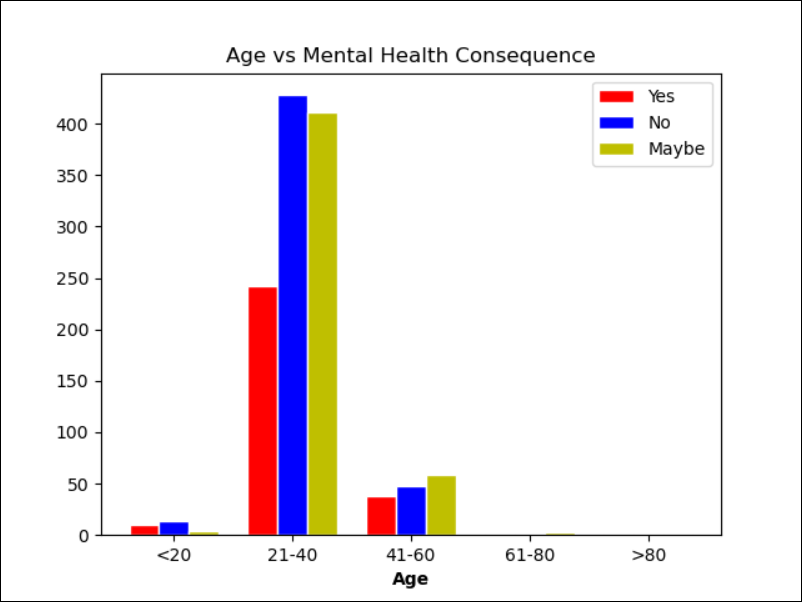
1. Which dataset did you select?
   1. **Mental Health In Tech**
2. How many observations are in the dataset?
   1. **1260 Observations**
3. How many variables in the dataset?
   1. **26 variables**
4. Does this dataset seem to belong to a regulated domain in law as discussed in the lectures? If yes, which one?
   1. **Health Information Privacy (Health Insurance Portability and Accountability Act of 1996)**
5. How many variables in the dataset are associated with a legally recognized protected class? In a table format, list those variables associated with a protected class, identify the protected class and the associated legal precedence/law as discussed in the lectures

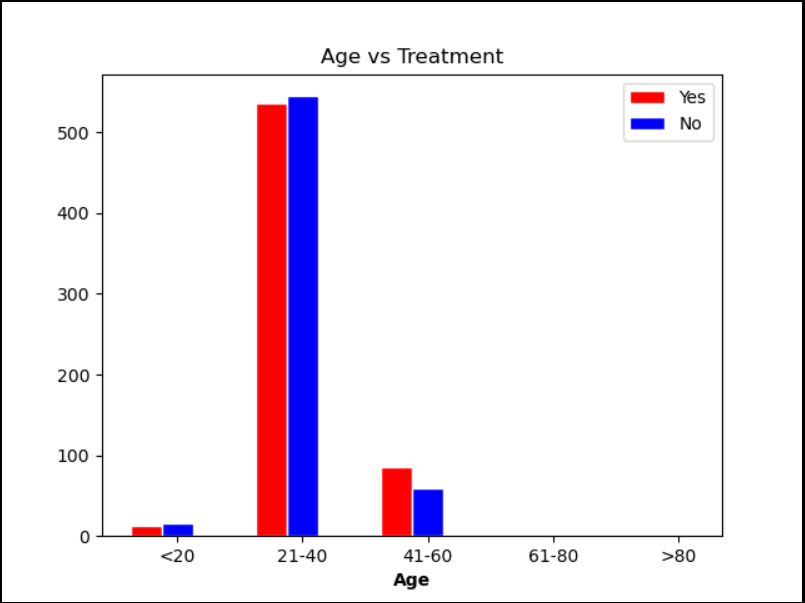


**Step 3**

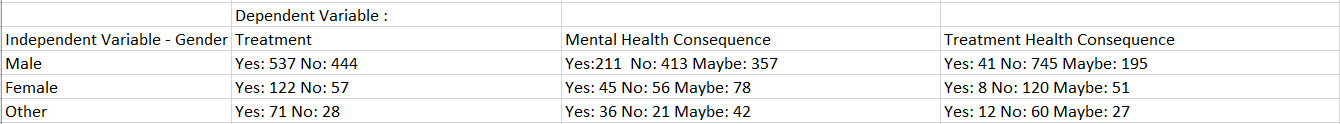
Independent Variable: Age

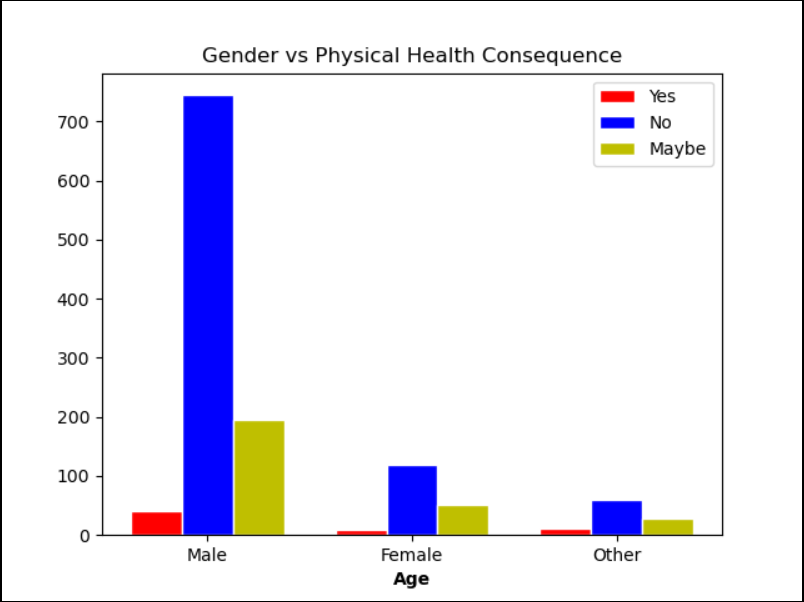
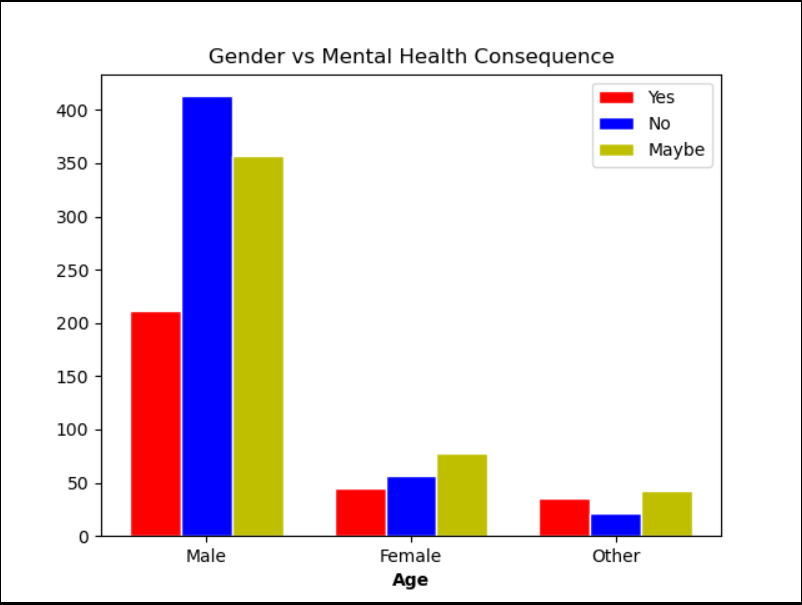


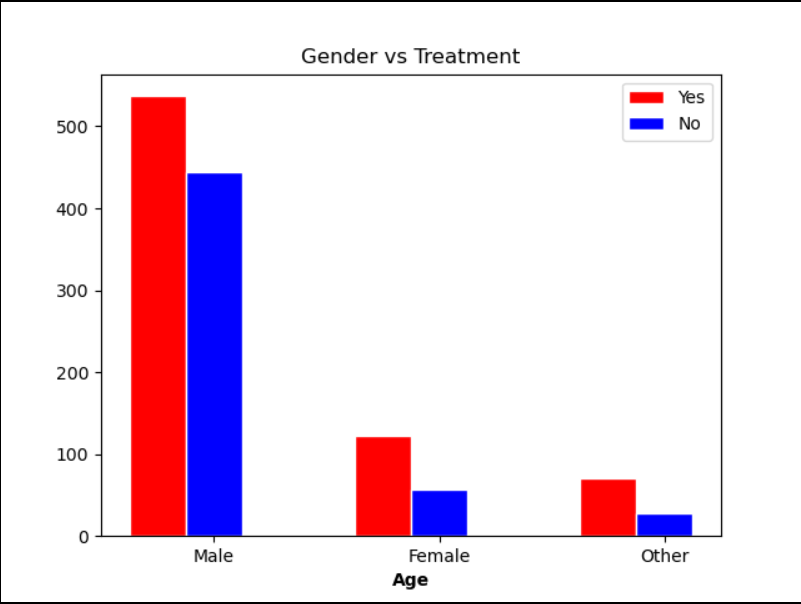




Independent Variable: Gender

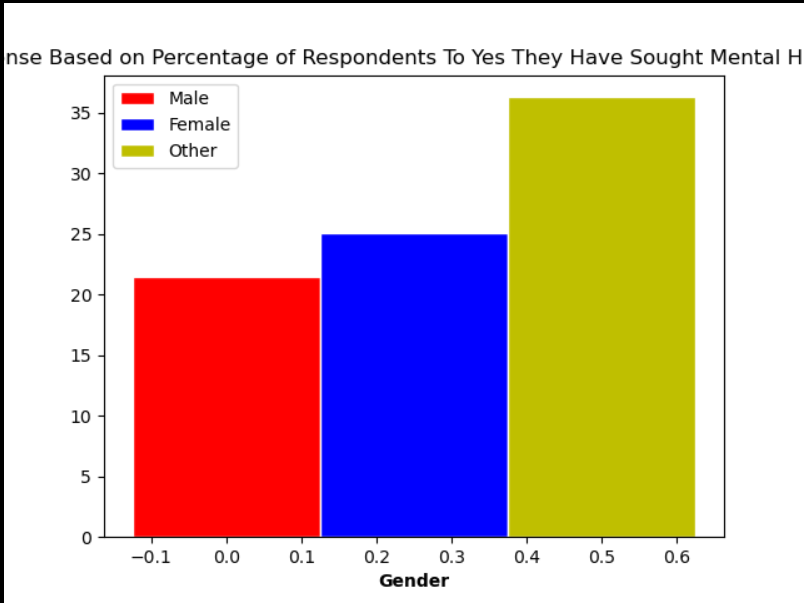






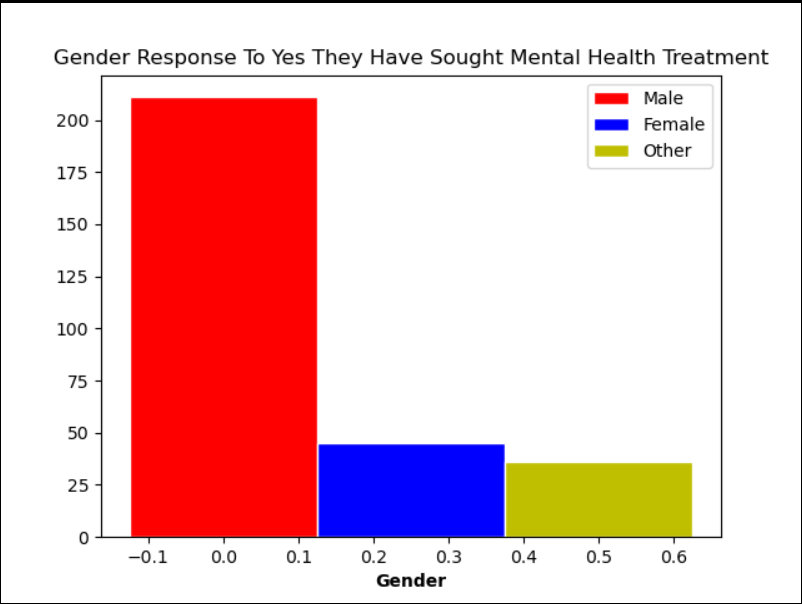
**Fairness Graph:**

The fairness graph, shows based on a categorical percentage of responses that the respondents who replied “Yes” to having sought mental health treatments is much more even than what is conveyed in the bias graph. Now there is still bias within this graph due to it only being shown off 1 independent variable, but it is much better than the below in its representation.

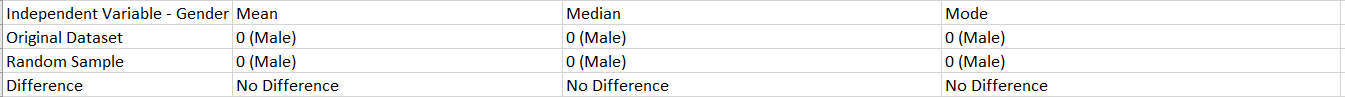
4

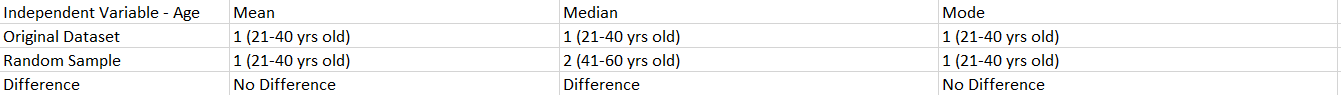
**Bias Graph:**

The bias graph makes it appear that Males are much more likely to have answered yes, to seeking mental health treatment. But when the data is normalized those ranges look much more even.



**Sampling & Statistics**





**Random Sampling Frequency/Histograms**

Based on step 4, with our protected class (gender), being compared against whether they sought treatment or not, the random sampling does not appear to have benefited/hurt any of the 3 categories. At least not by any large measure.

