

Exploratory Analysis of Shamiri Questionnaire Data

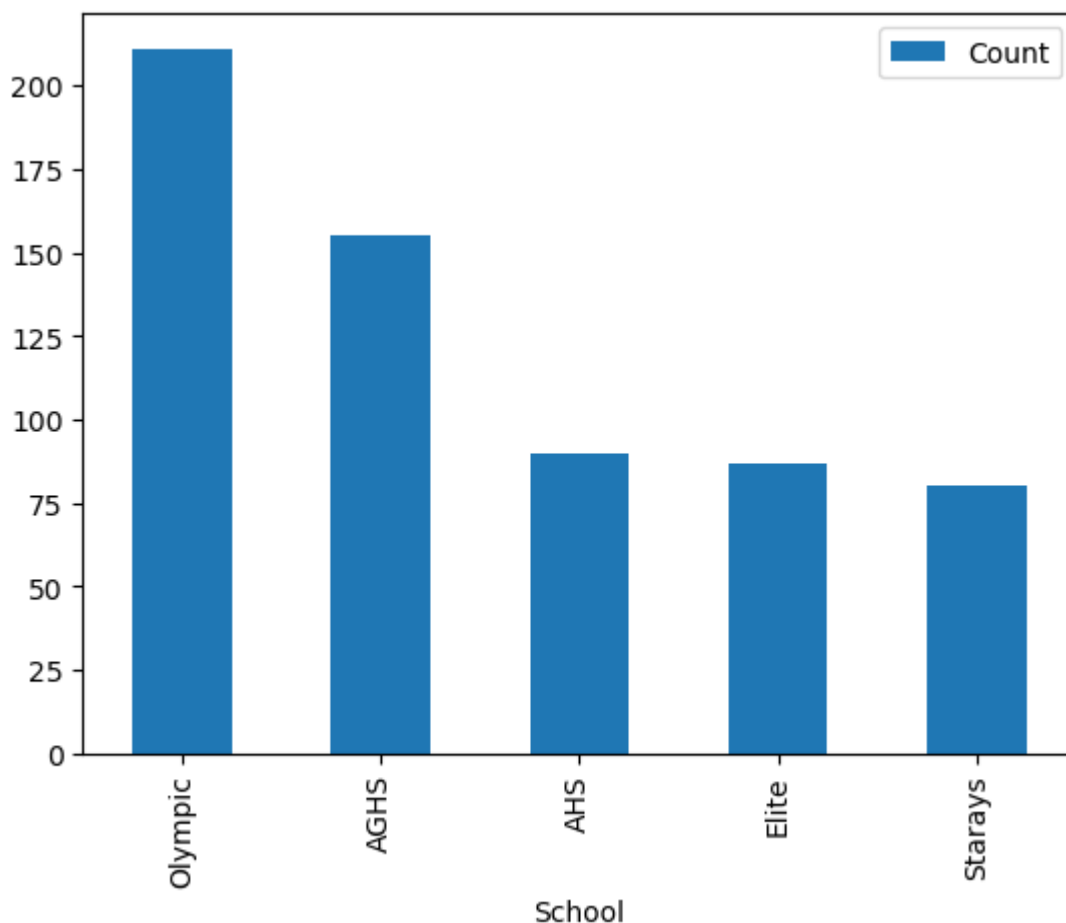
The data provided has both qualitative and quantitative qualities. Most columns have purely quantitative data while the rest have purely qualitative data. Most of the quantitative data is integers within a certain scale, for example 1 to 7 or 0 to 4. Only one column (age) has numbers of the float data type.

In total, the data has 33 columns and 658 rows. However, there are 32 rows that are a result of some participants filling the questionnaire twice. As such, they were dropped before any analysis was done. Three additional rows were dropped because they had values that were out of range given the questionnaire. For example, there was a '4' entered in a column that could only have values within the range of 0 to 3.

Fortunately, the data doesn't have any further problems like nulls. Ultimately, analysis was conducted on a dataset with 623 rows and 33 columns. The data can be segmented in a number of ways as outlined below.

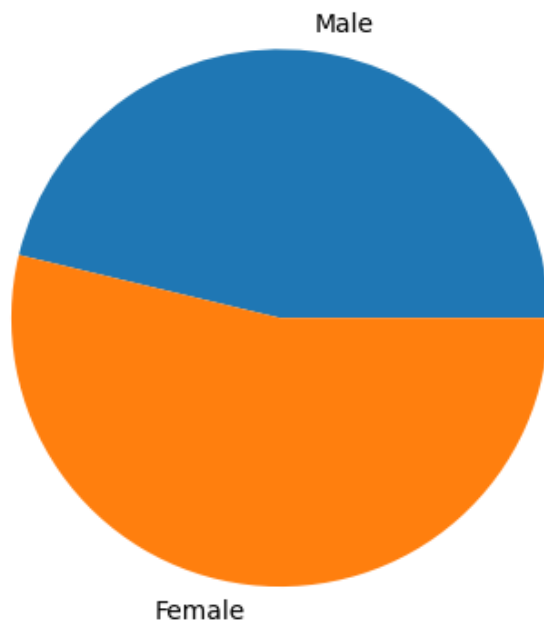
1. By School

The data was collected from 5 schools: Olympic, AGHS, AHS, Elite, and Starays. 211, 155, 90, 87, and 80 study participants come from each school respectively.



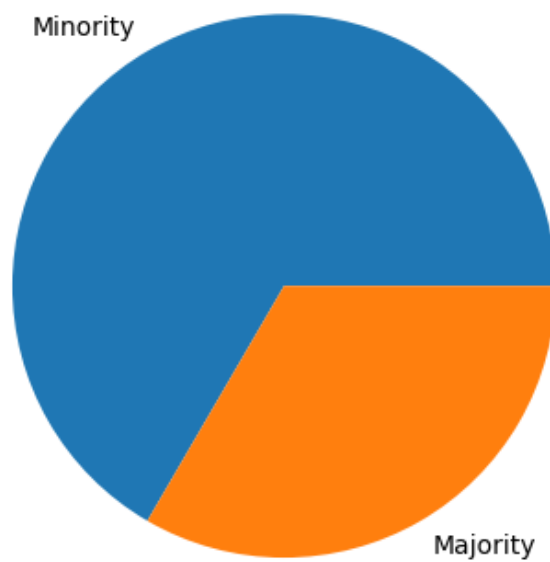
2. By Gender

Of all 623 study participants, 288 are male and 335 are female.



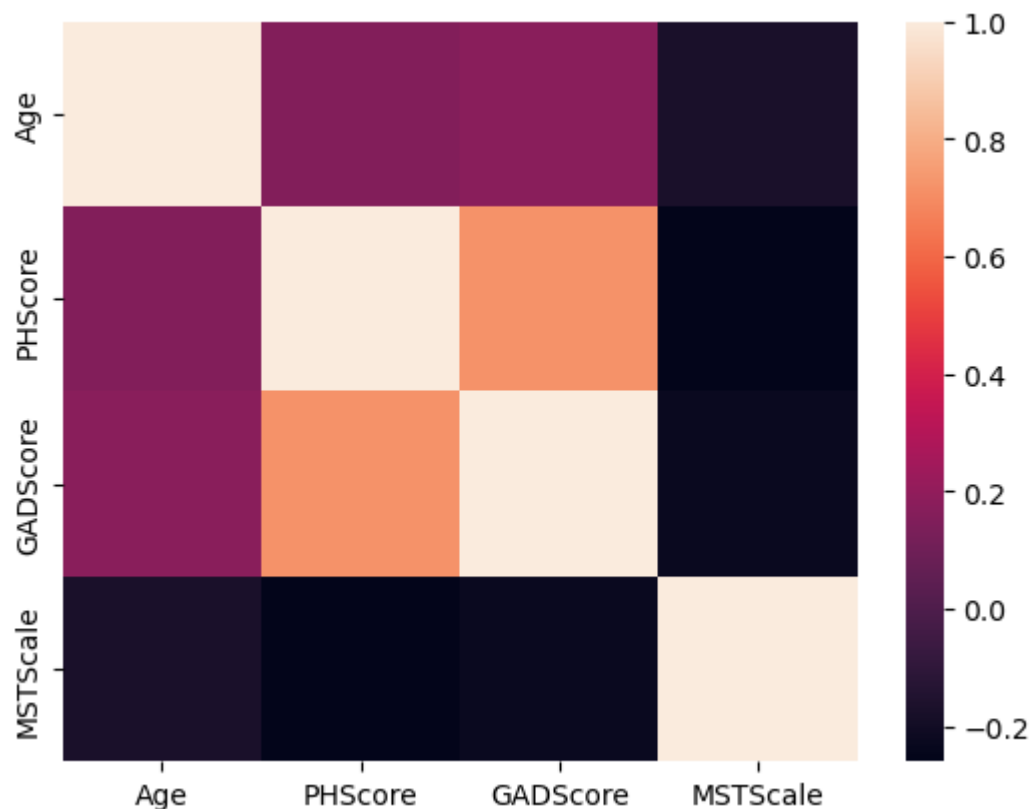
3. By Tribe

Of all study participants, 415 come from a minority tribe and 208 come from a majority tribe.

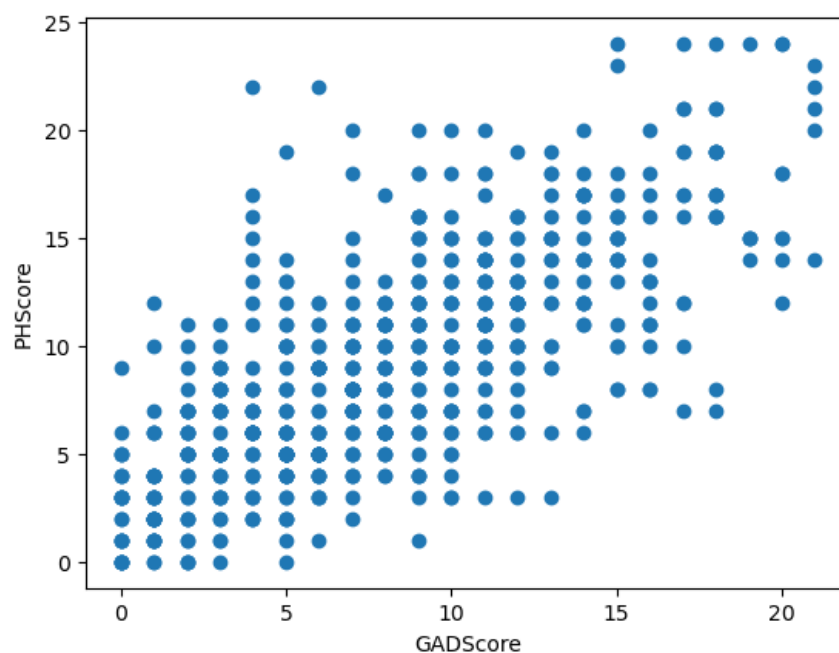


Important Trends

1. Strong correlation between PHQ-8 scores and GAD-7 scores

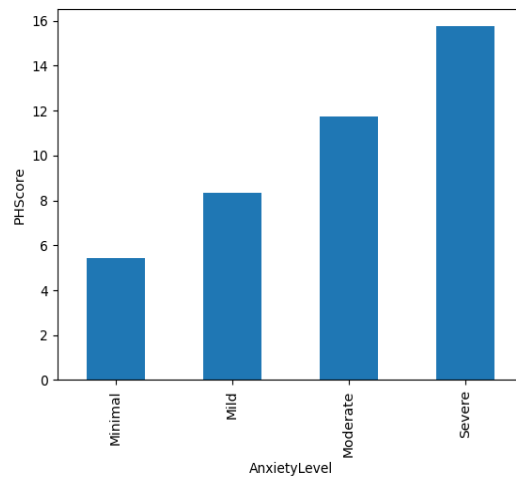
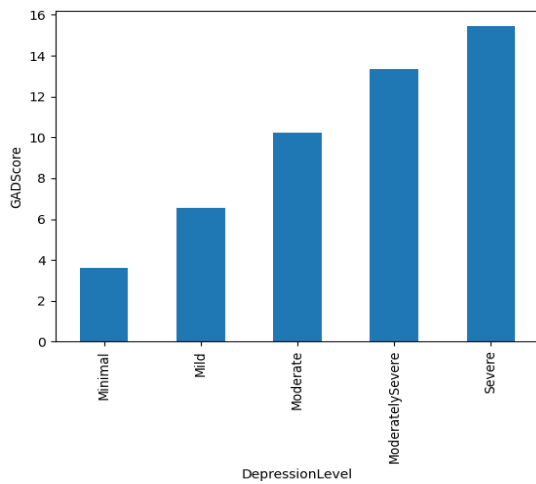


PHQ-8 scores and GAD-7 scores have a correlation coefficient of around 0.715, pointing to a strong positive linear relationship between the anxiety and depression levels in the study participants. This relationship can clearly be seen in the scatter plot below.

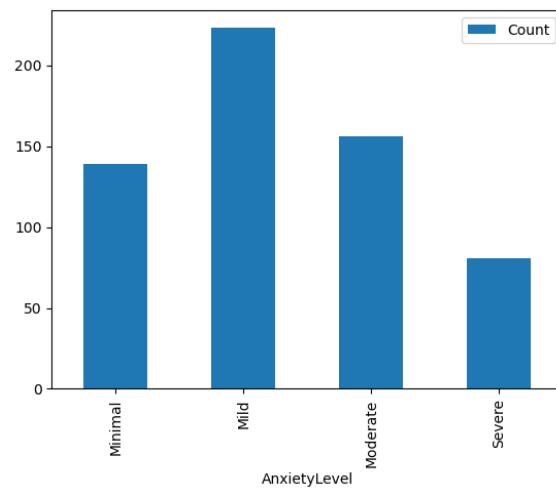
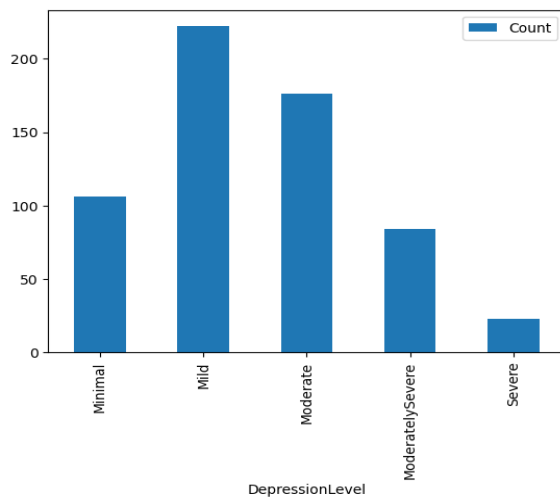


On taking a closer look, we discovered that severely depressed participants had an average GAD-7 score of 15.4, indicating severe anxiety. On the other hand, severely anxious

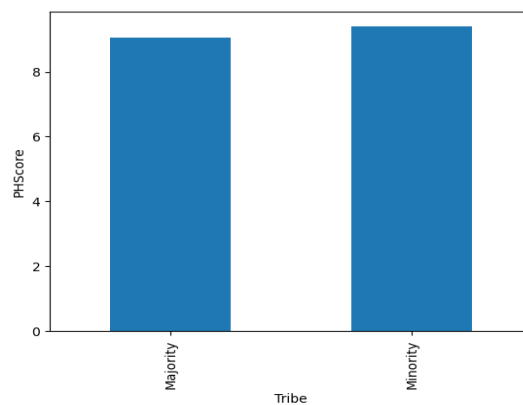
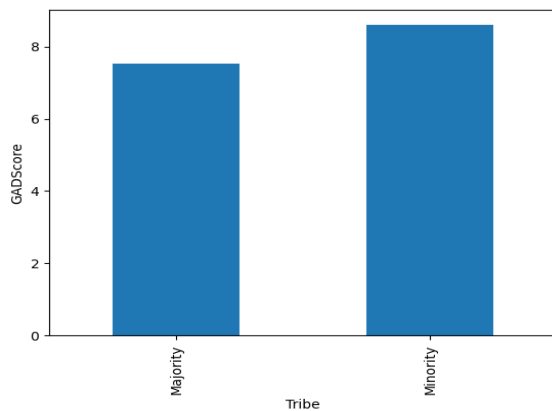
participants had an average PHQ-8 score of 15.7, indicating moderately severe depression.



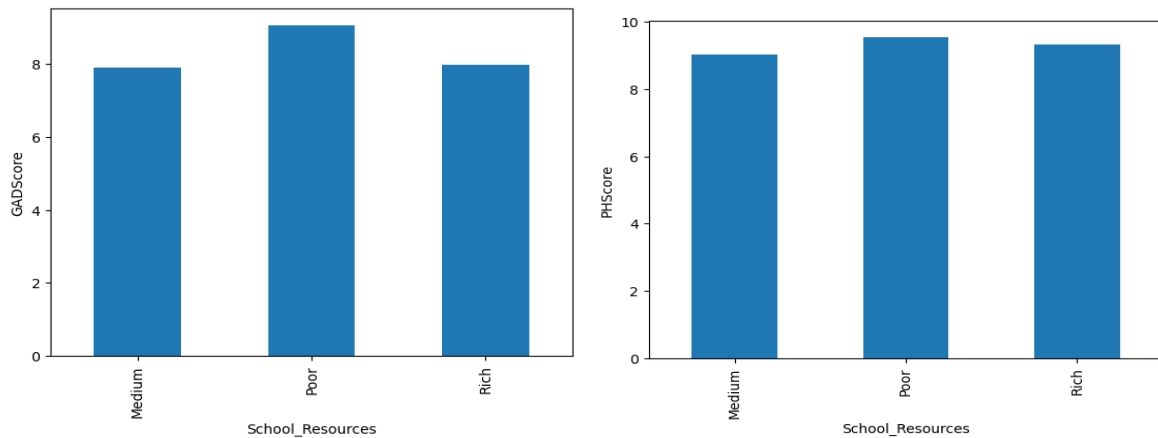
2. Most participants have mild levels of anxiety/depression



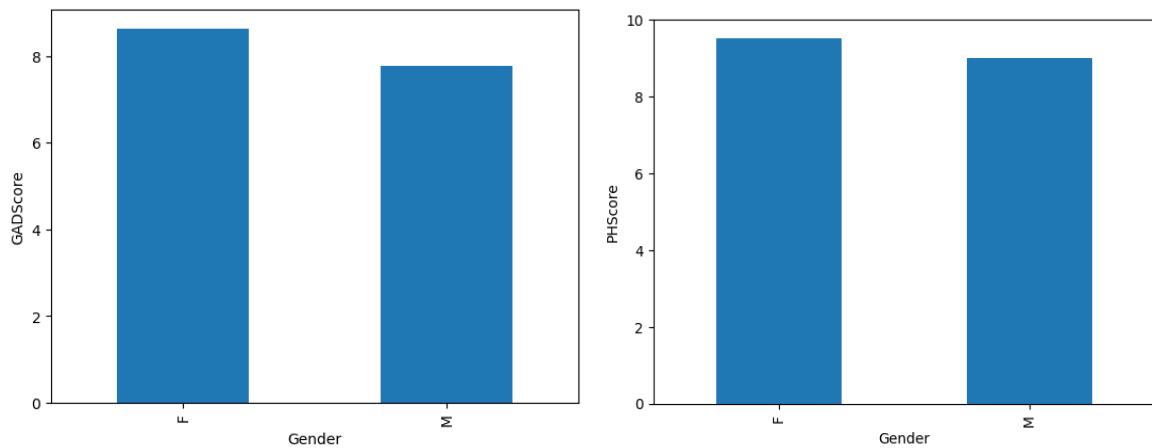
3. Participants that come from a minority tribe have a higher average PHQ-8 and GAD-7 score



4. Participants that come from schools with poor resources have a higher average PHQ-8 and GAD-7 score



5. Female participants have a higher average PHQ-8 and GAD-7 score



6. Severely depressed/anxious participants have a lower average total MSSS score than those with minimal depression/anxiety

