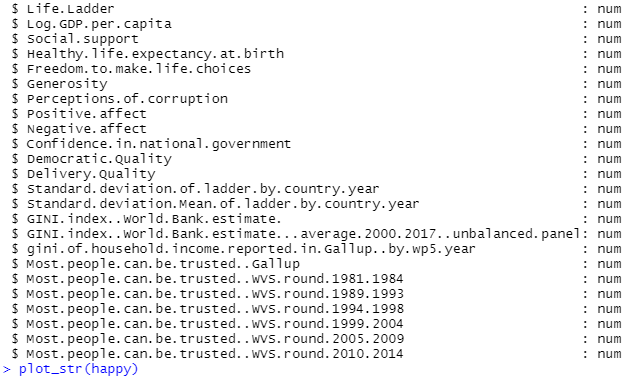
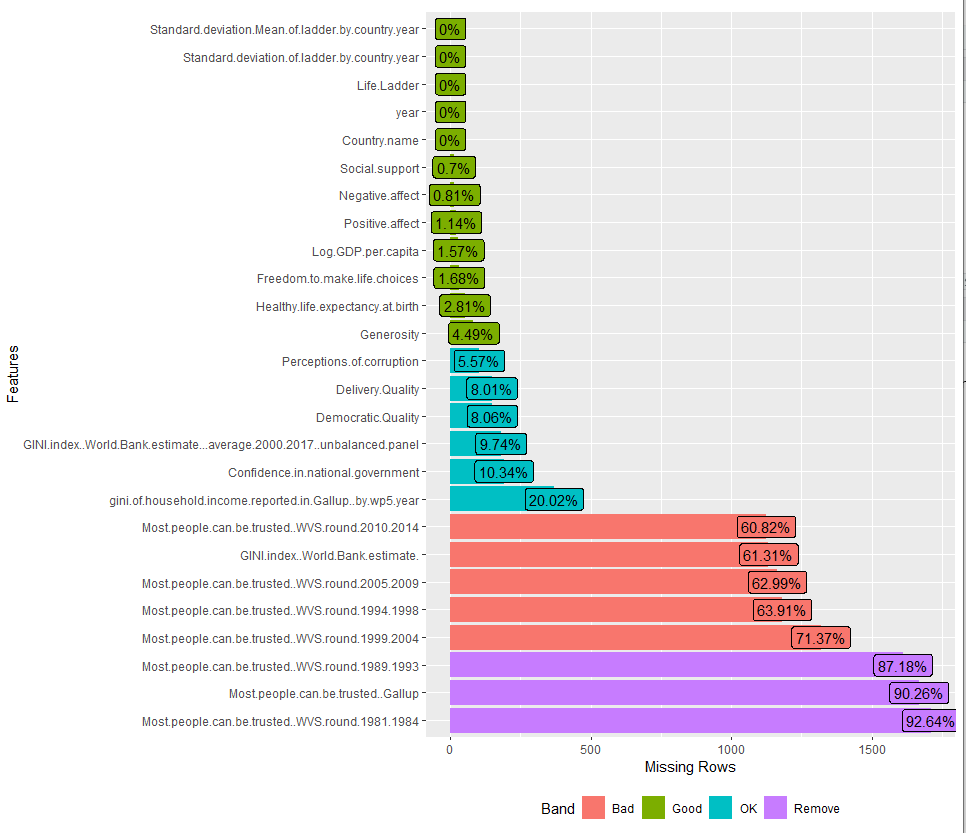
The data I have selected for the semester project is the happiness data. I believe it will be interesting to find more out about what countries are the happiest and the things they are doing in their daily lives to live happier and more fulfilling lives. This could potentially be used to support healthcare companies as a client in discovering what allows for the best mental health and physical health based on how long people are living and how happy they are each year. It could also be interesting to see how geography and recent event have changed the happiness levels of each country.

The data includes 26 variables, with 1848 observations. The variables are as follows. Country.name is a character variable, Year is an integer variable. The remining variables listed below are all classified as number variables. As there are no factor variables there are no factors here to explain.



The missingness of each variables are displayed in the following chart.



Given that the percentage of missing data for several of the variables is above 50% it might be a good idea to look into not using those certain variables, or finding out the best way to get the mean or median to fill in the missing data.

These variables will hopefully be used to create meaningful information about the overall happiness score for each country. I feel it will be particularly interesting to see how each country changes over time as well as how aspects such as freedom, democratic quality, confidence in national government, GDP per capita, and life expectancy impact an overall happiness score. Not only could this data be used in mental health and other healthcare aspects, but also for important social and political changes in many countries. I especially find this data so interesting because it takes something so subjective, a feeling of “happiness and well-being” and puts it into visible numbers and charts, thus making “the invisible, visible”.