2a. The skewness of each variable in the dataframe was measured where 8 of the 14 variables have their data skewed to the right due to the skewness being positive. Of the 8 variables, 4 have a skewness greater than 1. The remaining 6 variables are skewed to the left due to the negative skewness, though none are less than -1. The normal distribution graphs visually show the distribution of each variable.

2b.
1. The 'sex' variable can be shown as a barplot where '1' denotes male and '0' denotes female. There is at least double the male patients over the female patients.
2. The 'cp' variable denotes the type of chest pain. The barplot of 'cp' shows that most patients experienced chest pain type '0'. Followed by chest pain type '2', '1', and '3'.
3. The 'fbs' variable can be shown as a barplot where '1' denotes true and '0' denotes false. The vast majority of the fbs entries are false.
4. The 'restecg' variable denotes the resting electrocardiographic results. The barplot of 'restecg' shows both '0' and '1' had a near equal amount and took up the vast majority of entries combined. The remaining small minority of entries is '2'.
5. The 'exang' variable can be shown as a barplot where '1' denotes yes and '0' denotes no. There are at least double the no entries over the yes entries.
6. The 'slope' variable denotes the slope of the peak exercise ST segment. The barplot of 'slope' shows both '1' and '2' had a near equal amount and took up the vast majority of entries combined. The remaining small minority of entries is '0'.
7. The barplot of the 'thal' variable shows that most entries were '2'. Followed by thal type '3', '1', and '0'.
8. The 'target' variable can be shown as a barplot where '1' denotes the patient has CVD and '0' denotes no CVD. There are slightly more patients who have CVD over patients that don't.

2c. A countplot of the occurrence of CVD across the Age category shows that CVD can happen between the ages of 41-45 and 51-54. The ages of 57-62 is where CVD doesn't occur.

2d. When generating a heatmap of the dataframe, the Sex variable does not have either strong negative or positive correlations with the other variables.

2e. A box plot graph of the 'trestbps' variable shows that one can't detect heart attacks based on anomalies. Both categories of the CVD variable have outliers in a similar range.

2f. A box plot was graphed to show the relationship between cholesterol and CVD. The interquartile range for patients with CVD is smaller than those who don't, but there are more and greater outliers of patients who have CVD over those who don't.

2g. A countplot graph between CVD and peak exercise shows there is an increase of patients with CVD when peak exercising levels rise.

2h. A countplot graph between CVD and thal shows that higher 'thal' states will correlate with an increase of CVD patients. Thal state '2' has the vast majority of CVD patients while that state '3' is a distant second with thal states '1' and '0' trailing the furthest behind.

2i.
1. A countplot graph between CVD and cp shows that an increase in chest pain correlates to a rise in CVD.
2. A countplot graph between CVD and fbs shows that either state of fbs doesn't affect the occurrence of CVD.

3. A countplot graph between CVD and restecg shows that state '1' is most likely to determine occurrence of CVD.
4. A box plot graph of the 'thalach' variable shows that occurrence of CVD is correlated with a higher thalach level as the median and interquartile range for patients with CVD is higher than patients without CVD.
5. A countplot graph between CVD and exang shows that patients without exang are more likely to have CVD. Patients without CVD are not affected by whether or not they have exang.
6. A box plot graph of the 'oldpeak' variable shows that the oldpeak variable is less likely to affect patients with CVD due to having a lower median, interquartile range, and outliers than patients with CVD.
7. A countplot graph between CVD and ca shows that patients with a 'ca' state of '0' are the most likely to have CVD. However for patients with/without CVD, there are increasingly less entries for 'ca' states above '0'.

2j. A pairplot graphed between all the variables shows various correlations, dependencies, and associations. The pairplot shows that the variables 'trestbps', 'chol', and 'oldpeak' are correlated with the 'age' variable. All three variables are skewed to the left when compared to the 'age' variable. This is also true of the 'thalach' variable. For categorical variables, chi-square tests show that the 'target' variable is most dependent on the 'cp' and 'exang' variables while p-values show that 'target' is not dependent on the 'fbs' variable.