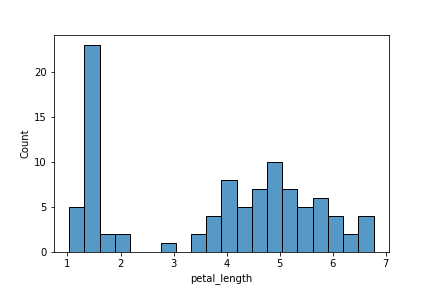
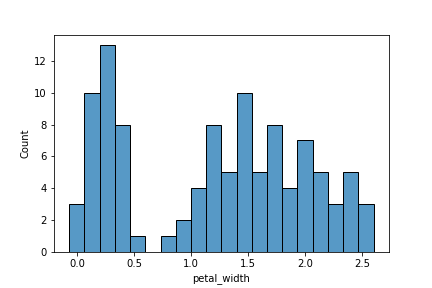
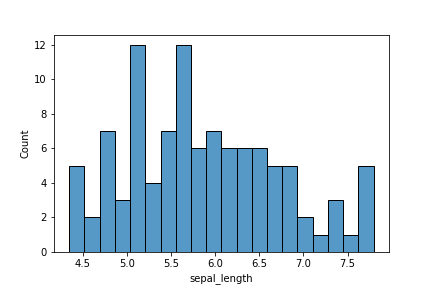
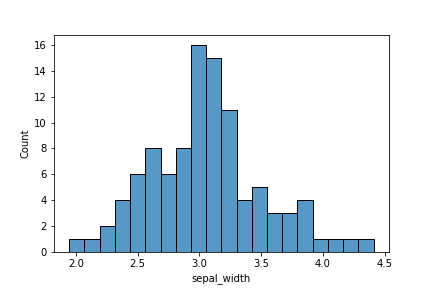
**Q 1.5**

**Iris Dataset**

**Inspecting outliers using histogram plot**:

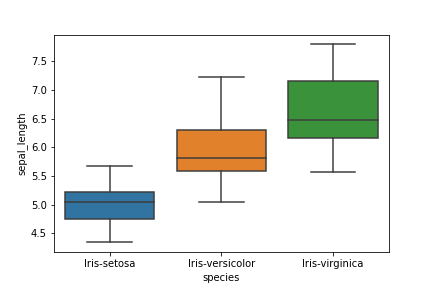
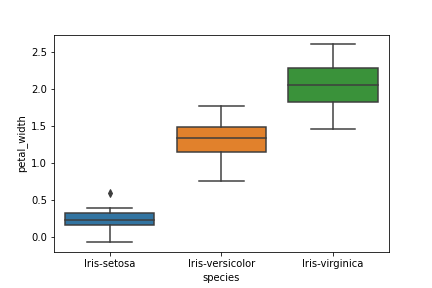
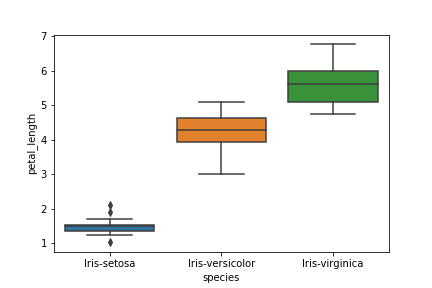
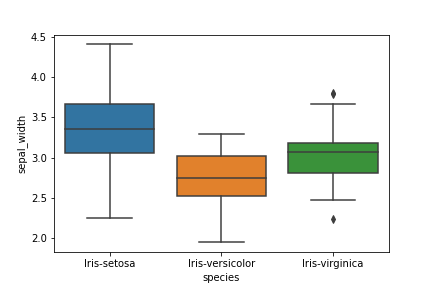


**Observations from the histogram plots:**

There are few outliers in petal width with values less than 0

There are also a few outliers in petal length

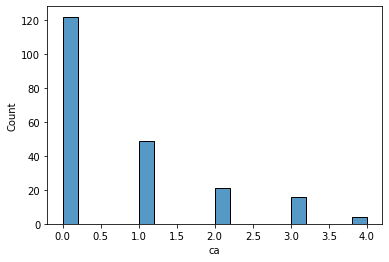
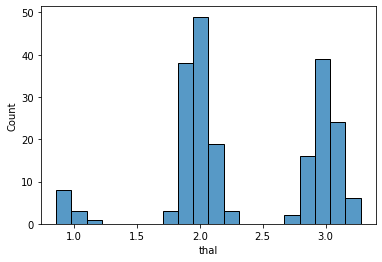
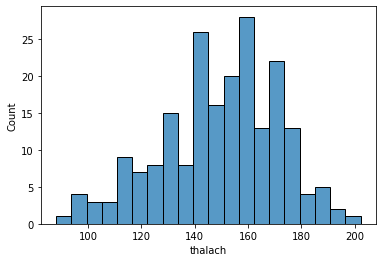
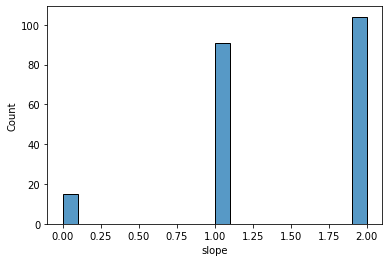
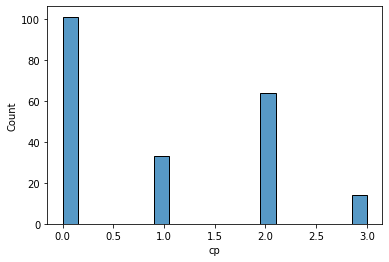
**Closer inspection of outliers using box plot**



**Observations from the boxplot**

* There are outliers in sepal width of the Iris-Virginica specie
* There are outliers in petal length and petal width of the Iris-Setosa specie
* It can be seen that the lower fence of the Iris-Setosa specie is below 0 indicating presence of negative values.

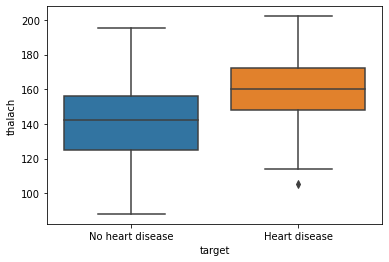
**Heart Disease Dataset**

****

**Comments about the plots**

* There is noise in thal feature as it is supposed to be categorical with unique values but it looks continuous (presence of Gaussian noise).
* There are outliers in ca because it is a categorical variable with range 0 to 3 and there are few observations with a value of 4.

**Closer inspection of outliers in numerical feature (thalach) using box plot**

****

**Observations from the boxplot**

* There is an outlier in thalach below the floor value for those with heart disease