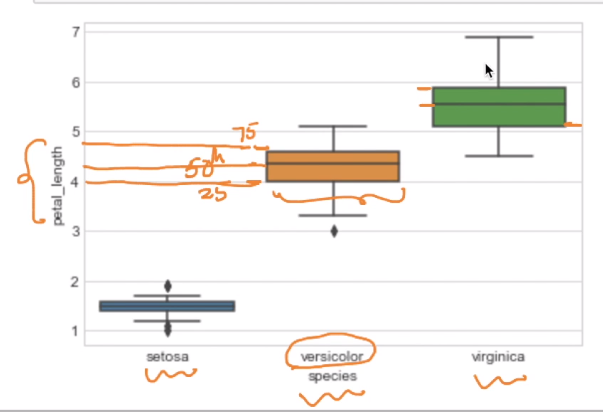
Histograms/PDF are good to know the density of points in a region, but they can’t tells us what is 25th, 50th and 75th percentile just by looking at them, so to overcome this Box plot are used, which clearly tells us the 25th, 50th and 75th percentile.

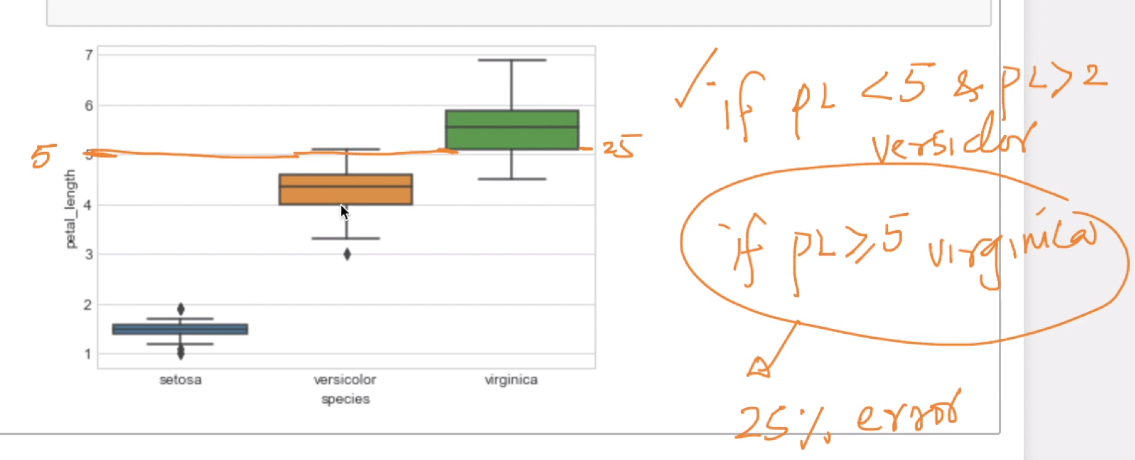
Below figure show box plot of setosa, versicolor and virginica, the rectangular box is basically a IQR.

* The top line of that box shows 75th percentile.
* The middle line shows 50th percentile.
* The lowest line shows 25th percentile.



Now seeing these 3 plots we can create a basic model as given in below figure.

But to classify virinica we are using petal length as 5 or more, so here box plot came into picture which tells how much % of virginica flowers are identified incorrectly, so in current case we have 25% error, because at length 5 there is 25%, that mean 25% flowers are there which virginica but identified incorrectly.



**Whiskers:**

Whiskers are the line drawn above and below rectangular box.

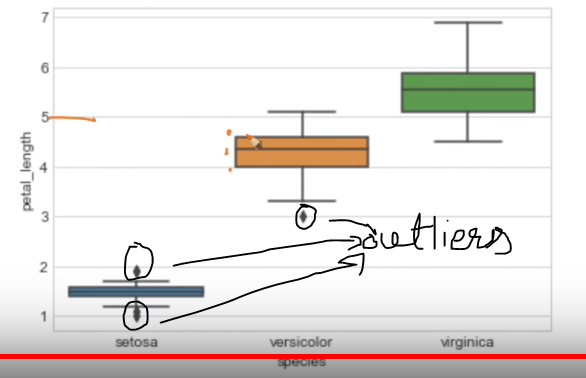
There are several rules for constructing Whiskers.

The value/Extent of whiskers are decided based on minimum value of the following two conditions:  
(1) + or - 1.5 \* IQR ( lowest whisker value  
(2) Maximum / Minimum value present in the data set.

Ie:

>The whiskers are extended-above up to   
minimum value of ( Max value present in data set AND Q3+1.5 IQR)  
>The whiskers are extended-below up to   
minimum value of ( Min value present in data set AND Q1-1.5 IQR)



The points outside the whiskers are sometimes considered to be outliers, not always. If there are a good number of such points, you tend to notice them like in the case of whiskers of Versicolor. 

**Why Box-plot can be visualized as a PDF on the side-ways? What does side-ways means?** <https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51>

