**Workshop5B**

**Ex1: What is Box Plot? Explain it in detail with a SAS example (code + graph).**

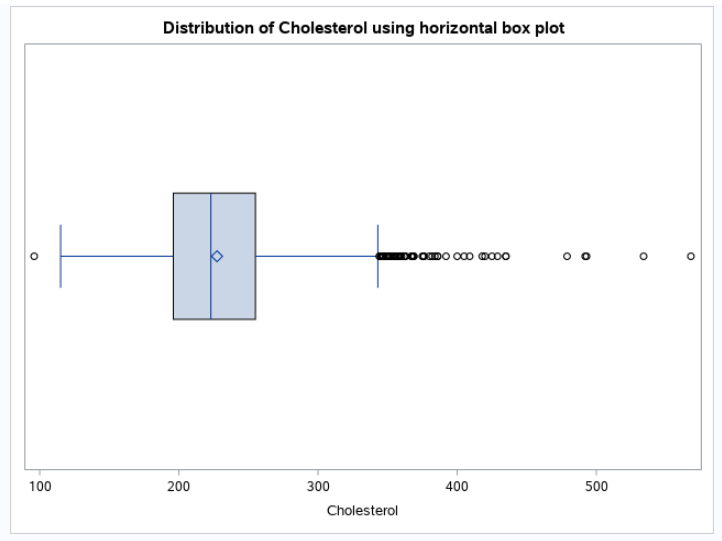
Ans**.** Box Plot is used in statistics for exploratory data analysis. It shows critical values like median, first and third quartiles and data points called outliers.

Below is an example of Box plot using code:

**CODE:**

title "Distribution of Cholesterol using horizontal box plot";  
proc sgplot data=SASHelp.Heart;  
 hbox Cholesterol;  
run;

**OUTPUT:**



* The vertical line inside the box represents median, while left and right sides of the box represent 1st and 3rd quartiles.
* The lines extending from both sides of box represent a distance of 1.5 interquartile ranges on both sides of box.
* Diamond inside the box represents mean of cholesterol.
* Circles falling outside the 1.5 interquartile ranges are data points which are known as outliers.

**Section: 20.12**

**Exercises: 9 to 12**

**Question 1**

\* 20.12-9 Using the SASHelp data set, Heart, generate a histogram for the variable, Cholesterol.;

title "Histogram for Cholesterol from Heart dataset";

proc sgplot data=SASHelp.Heart;

histogram Cholesterol;

run;

**Question 2**

\* 20.12-10 Using the SASHelp data set, Heart, generate a histogram with a curve for the variable, Cholesterol.;

title "Plotting curve on histogram for Cholesterol";

proc sgplot data=SASHELP.Heart;

histogram cholesterol;

density cholesterol;

run;

**Question 3**

\* 20.12-11 Using the SASHelp data set, Heart, generate two horizontal box plots that show the distribution of Cholesterol for men and women (variable Sex).;

title "Distribution of Cholesterol for men and women using horizontal box plots";

proc sgplot data=SASHelp.Heart;

hbox Cholesterol / group=Sex;

run;

**Question 4**

\* 20.12-12 Using the SASHelp data set, Heart, generate a plot showing the mean height for men and women. On the same plot, show the mean weight for men and women. For the latter display, set transparency to .2, and make the bar width 25% of the full width.;

title "Mean height and weight for men and women";

proc sgplot data=SASHELP.Heart;

vbar Sex / Response=Height stat=mean barwidth=0.75;

vbar Sex / Response=Weight stat=mean barwidth=0.25 transparency=0.2;

run;