

Dq op



**ASSIGN
BUSTER**

1) An outlier is a number or observation that lies outside the pattern of a distribution. Outliers suggest that model that we are studying has an error or due to erroneous observation, correct correlation between two variables cannot be determined. It is easier to identify outliers in the graphical representation such as histograms or line of best fits.

An example of outlier can be seen when drawing a line of best fit on scatter diagram. Any point which does not match the trend which is revealed by line of best fit can be classified as an outlier.

Let's look at another example of outlier in a graphical representation:

Y VARIABLE

.

X Variable

Suppose that above line of best fit correlates X and Y. A point which is above this line is an outlier in this example.

2) Yes. There is an outlier in data i.e Nigeria's literacy rate of 57%.

3) We have 46 countries for data analysis. In the first part we have taken out mean literacy rate which comes near to 90% or 89.55%. If we look at standard deviation of these countries, it's around 13.6% which shows the variation or literacy gap between different countries. The mode of this data is 99%, which means that it is the most repeated number in our data. Similarly, median of this data is 97.5%. Median value falls exactly at the center of the data and hence one can conclude that this is the normal literacy rate.

However, if we look at data from box plot, we will find out that data is skewed towards right side or towards higher values. This shows that more

countries have literacy rate above the average literacy rate. There's only one outlier in our data which is Nigeria's literacy rate of 47%.