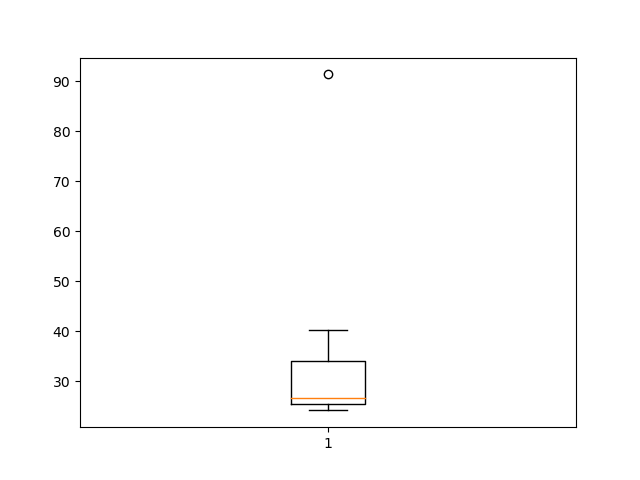
**Topics: Descriptive Statistics and Probability**

1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | *24.23%* |
| Bankers Trust | *25.53%* |
| General Mills | *25.41%* |
| ITT Industries | *24.14%* |
| J.P.Morgan & Co. | *29.62%* |
| Lehman Brothers | *28.25%* |
| Marriott | *25.81%* |
| MCI | *24.39%* |
| Merrill Lynch | *40.26%* |
| Microsoft | *32.95%* |
| Morgan Stanley | *91.36%* |
| Sun Microsystems | *25.99%* |
| Travelers | *39.42%* |
| US Airways | *26.71%* |
| Warner-Lambert | *35.00%* |
| \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* | *\*\*\*\*\*\*\*\** |
| Mean(mu) | *32.37* |
| Standard deviatiom | *16.37* |
| variance | *268.0035* |
| outlier | *91.36* |



[ point far away from boxplot represents outlier in our data]



Answer the following three questions based on the box-plot above.

1. What is inter-quartile range of this dataset? (please approximate the numbers) In one line, explain what this value implies.

Ans : inter-quartile range = [-5.5 , 22.5]

Inter-quartile range is use to find whether data has any outliers or not, if given data is within inter-quartile range then our data does not have any outliers.

1. What can we say about the skewness of this dataset?

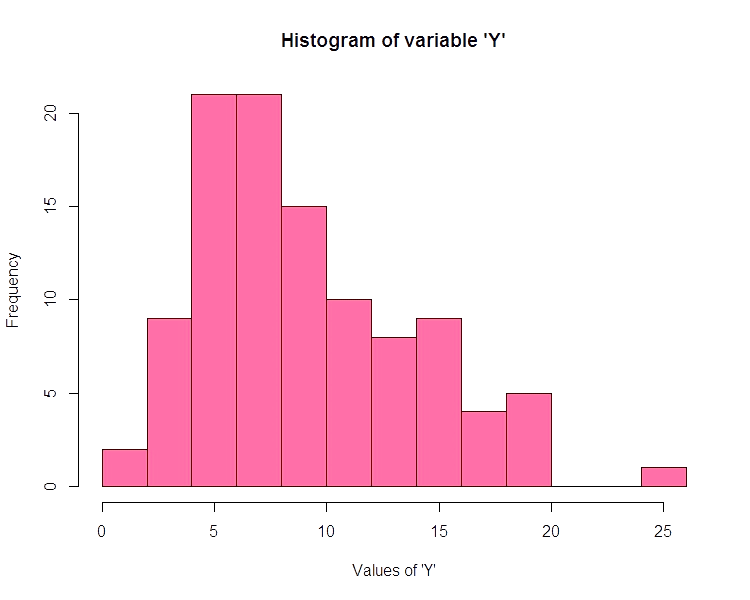
Ans : Here, (Q3-Q2) > (Q2-Q1) that means data is centered around Q2 quartile which is leftside in distribution and tail is at right side( POSITIVE – SKEWED )

1. If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?

Ans : if we remove outlier data point from box-plot then value of ( Q3-Q2) will dicrease and

(Q2-Q1) will increase after that if we add 2.5 value to box-plot then again Q2 will tends to

Move in Q3 direction which results in data tends to be normally distributed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans : above histogram has two bar at peak which means distribution is BIMODAL.

And mode lies between [4 – 8]

1. Comment on the skewness of the dataset.

Ans : POSITIVE SKEWED

1. Suppose that the above histogram and the box-plot in question 2 are plotted for the same dataset. Explain how these graphs complement each other in providing information about any dataset.

Ans : from both histrogram and box-plot we can say that both graph represents same nature of

Distribution , also from histogram we can say for box-plot that outlier at 25 affects graph to

become normally distributed because it pulls tail of graph towars right side.

1. AT&T was running commercials in 1990 aimed at luring back customers who had switched to one of the other long-distance phone service providers. One such commercial shows a businessman trying to reach Phoenix and mistakenly getting Fiji, where a half-naked native on a beach responds incomprehensibly in Polynesian. When asked about this advertisement, AT&T admitted that the portrayed incident did not actually take place but added that this was an enactment of something that “could happen.” Suppose that one in 200 long-distance telephone calls is misdirected. What is the probability that at least one in five attempted telephone calls reaches the wrong number? (Assume independence of attempts.)

Ans : probability = (1/200) \* 5 = 1/40

1. Returns on a certain business venture, to the nearest $1,000, are known to follow the following probability distribution

|  |  |
| --- | --- |
| x | P(x) |
| -2,000 | *0.1* |
| -1,000 | *0.1* |
| 0 | *0.2* |
| 1000 | *0.2* |
| 2000 | *0.3* |
| 3000 | *0.1* |

1. What Is The Most Likely Monetary Outcome Of The business venture?

*ans : most likely monetary outcome will be 2000 because it has high probability than others.*

1. is the venture likely to be successful? explain

ans *: yes , because prob(x>=0) is higher than prob(x< 0) , that means business is more likely to get profit.*

1. what is the long-term average earning of business ventures of this kind? explain

*ans : long-term avg. means to count for every every probability value, e(x) = $800*

1. what is the good measure of the risk involved in a venture of this kind? compute this measure

*ans : prob(x>=0) = prob(0) + prob(1000) + prob(2000) + prob(3000)*

*= 0.2 + 0.2 + 0.3 + 0.1*

*= 0.8*