

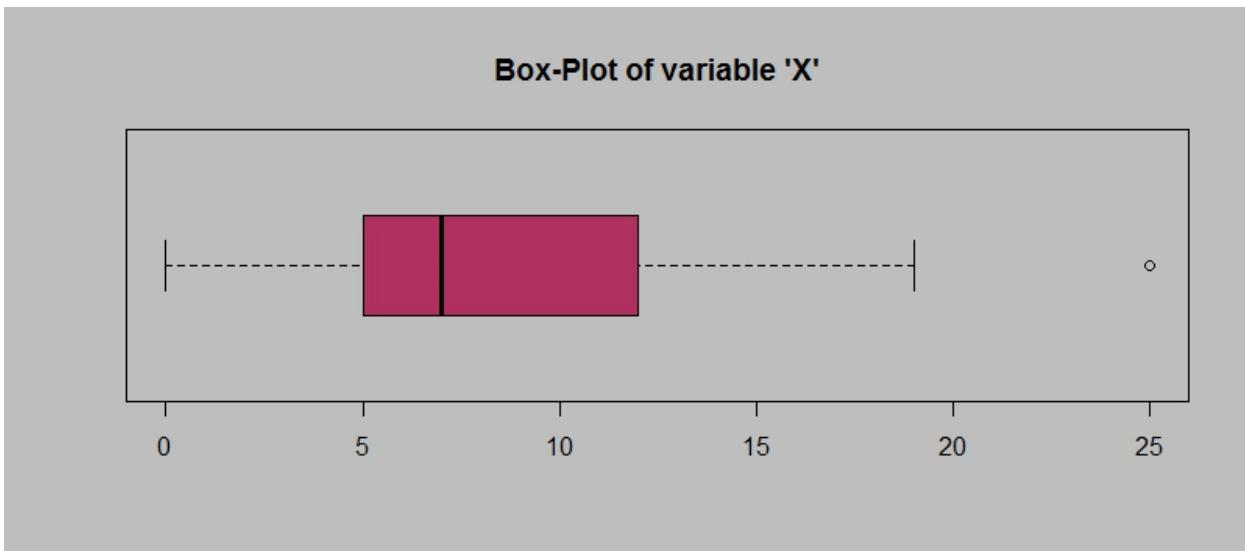
Topics: Descriptive Statistics and Probability

1. Look at the data given below. Plot the data, find the outliers and find out μ, σ, σ^2

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%

Refer to the jupyter notebook
Assignment 2 .

2.



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

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

IQR \rightarrow (5 to 13) approx.

It means 50% of the data lies between the range of 5 and 13 in the given dataset.

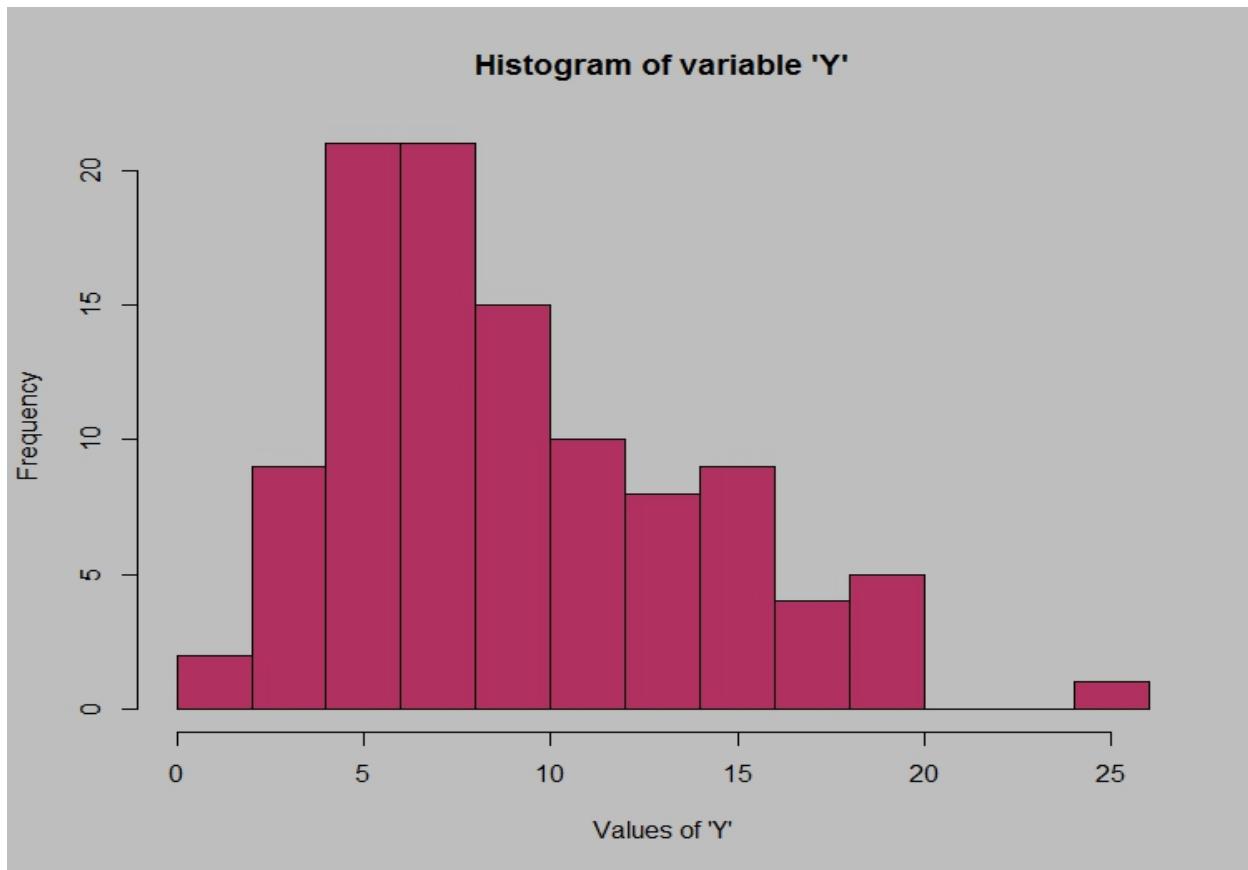
- (ii) What can we say about the skewness of this dataset?

Since the upper whisker is greater than the lower whisker, we have more data points on the upper range. Thus, we can say the dataset is positively skewed.

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

There will be no outlier in the boxplot.

3.



Answer the following three questions based on the histogram above.

- (i) Where would the mode of this dataset lie?

The mode will be between 4 to 8.

- (ii) Comment on the skewness of the dataset.

It is positively skewed.

- (iii) 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.

From both the plots we see that values are more dense within (5-12). The long whisker at the right side of boxplot and the occurrence of more bars in the right side of the histogram shows the positive skewness. Both show the outlier at 25.

4. 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.)

$$\text{prob. of calls misdirected, } P(M) = \frac{1}{200}.$$

$$\text{u n u not n, } P(NM) = 1 - \frac{1}{200} = \frac{199}{200}$$

we can use the formula for binomial distribution -

no. of calls done, $n = 5$

prob. that one of the calls misdirected

$$= 1 - P(0) \xrightarrow{\text{no. calls got misdirected}}.$$

$$= 1 - {}^n C_0 p^x q^{1-x} = 1 - {}^5 C_0 \left(\frac{1}{200}\right)^0 \left(\frac{199}{200}\right)^5$$

$$= 1 - \left(\frac{199}{200}\right)^5 = 0.02475 \approx 2.45\%.$$

5.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

- (i) What is the most likely monetary outcome of the business venture?

The most likely monetary income is 2000 with maximum probability of 0.3

- (ii) Is the venture likely to be successful? Explain

Chances of getting more than or equal to 1000 = $(0.2+0.3+0.1)=0.6$

Therefore, it is likely to be successful.

- (iii) What is the long-term average earning of business ventures of this kind? Explain

- (iv) What is the good measure of the risk involved in a venture of this kind? Compute this measure
- We can use standard deviation to understand the spread or deviation from 1000.
Refer to the notebook