

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

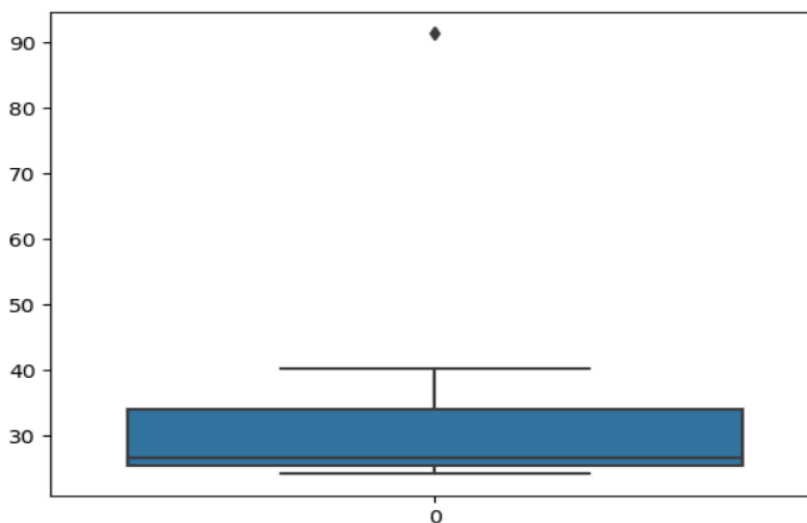
Name of company	Measure X
Allied Signal	24.23%
Bankers Trust	25.53%
General Mills	25.41%
ITT Industries	24.14%
J.P.Morgan & Co.	29.62%
Lehman Brothers	28.25%
Marriott	25.81%
MCI	24.39%
Merrill Lynch	40.26%
Microsoft	32.95%
Morgan Stanley	91.36%
Sun Microsystems	25.99%
Travelers	39.42%
US Airways	26.71%
Warner-Lambert	35.00%

ANS-For the given data the outlier is 90.

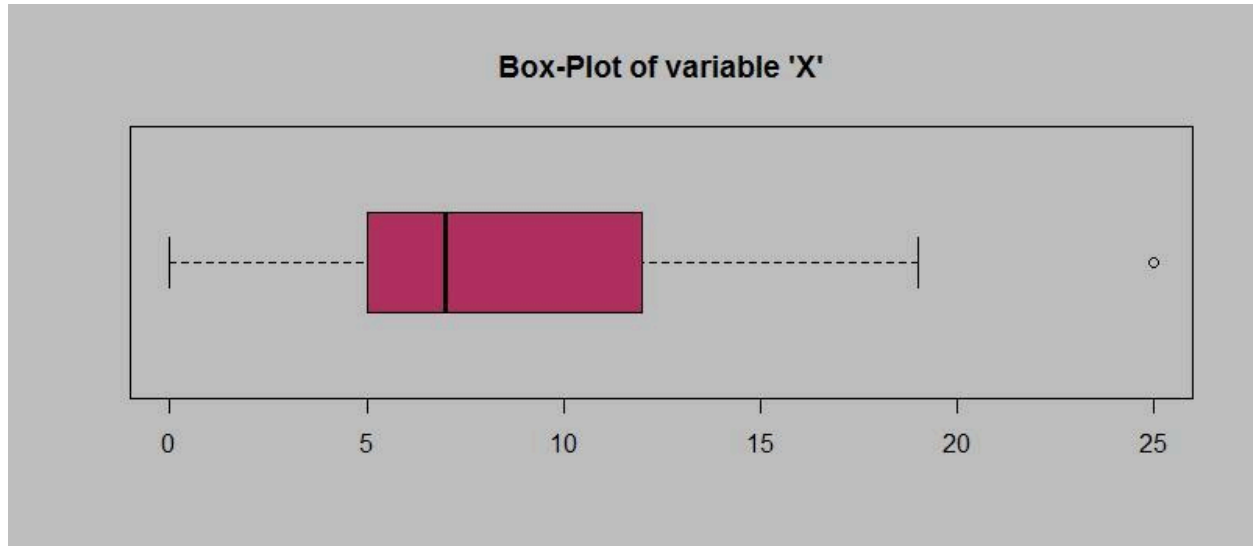
Mean = 33.271

Standard deviation =16.945

Variance = 287.14



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.

ANS-The first quartile (Q1) is 5, the third quartile (Q3) is 12, and the median (second quartile) is 7. The interquartile range (IQR) is calculated as $Q3 - Q1$, which equals 7

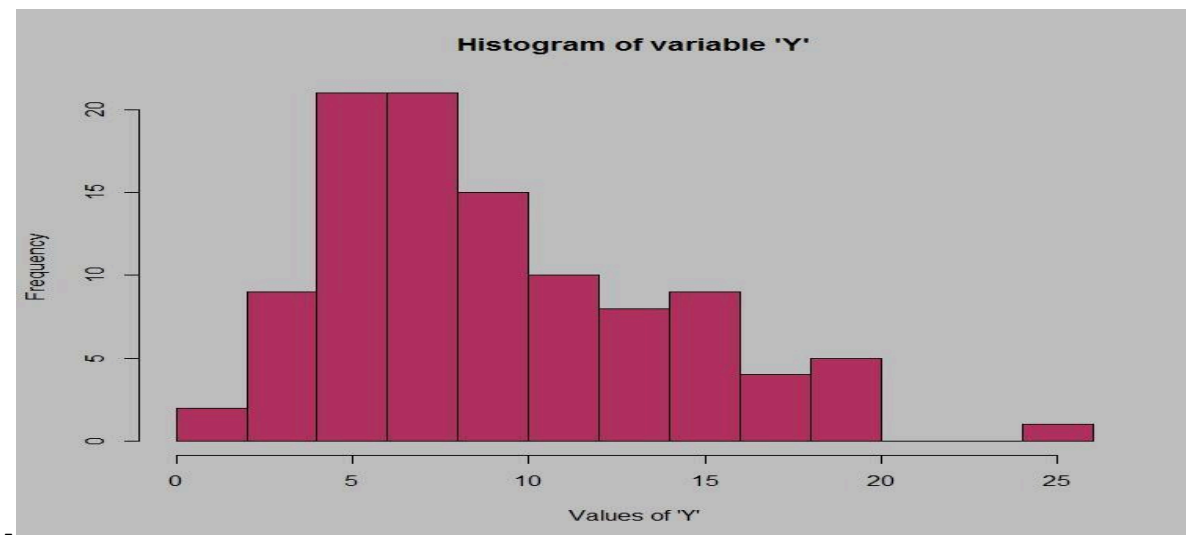
What can we say about the skewness of this dataset?

ANS-The skewness of the data is Right skewed.

- (ii) 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- In this scenario, no outliers are present in the above boxplot. The presence of outliers was causing positive skewness in the data; their removal results in a reduction of skewness, leading the data to exhibit a more normal distribution.

3.



Answer the following three questions based on the histogram above.

(i) Where would the mode of this data set lie?

ANS- The mode of this data set is estimated to be approximately within the range of 4 to 8

(ii) Comment on the skewness of the data set.

ANS- The data is positively skewed.

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

ANS- Both distributions exhibit positive skewness and contain outliers. The median is easily discernible in the box plot, while the mode is more prominently visible in the histogram.

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.)

ANS- The chance of encountering at least one wrong number among five attempted telephone calls is around 99.5%, not a perfect 1%. This means there is a high likelihood of experiencing an instance of dialing a wrong number within this set of calls

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?

ANS- The most likely monetary outcome for the business venture is \$2000, with a probability of 0.3, the highest among all possibilities.

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

ANS- Yes the probability that the venture will make more than 0 or a profit is

$P(x > 0) = p(x > 1000) + p(x > 2000) + p(x = 3) = 0.2 + 0.2 + 0.3 + 0.1 = 0.8$. This suggests an 80% chance of success for the venture.

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

ANS- The long-term average is the Expected value = $\sum(X * P(X))$

$$\begin{aligned}\text{Expected value} &= (-2000 * 0.1) + (-1000 * 0.1) + (0) + (1000 * 0.2) + (2000 * 0.3) + (3000 * 0.1) \\ &= -200 - 100 + 0 + 200 + 600 + 300 \\ &= 800\end{aligned}$$

Therefore the long-term average earnings of business ventures of this kind is \$800.

- (ii) What is the good measure of the risk involved in a venture of this kind? Compute this measure.

ANS-The good measure of the risk involved in a venture of this kind depends on the variability in the distribution.

Higher Variance means more chances of risk,

$$\begin{aligned}\text{Var}(X) &= E(X^2) - (E(X))^2 \\ &= 2800000 - 800^2 \\ &= 2160000\end{aligned}$$