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: Attached Jupyter Notebook Python File

2.

Box-Plot of variable 'X'
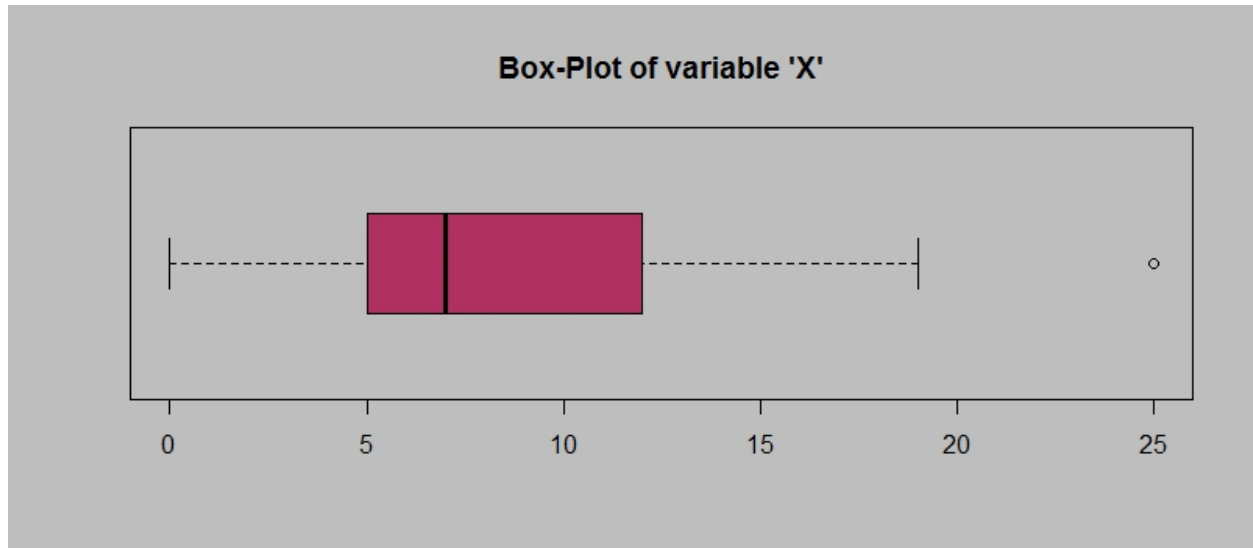
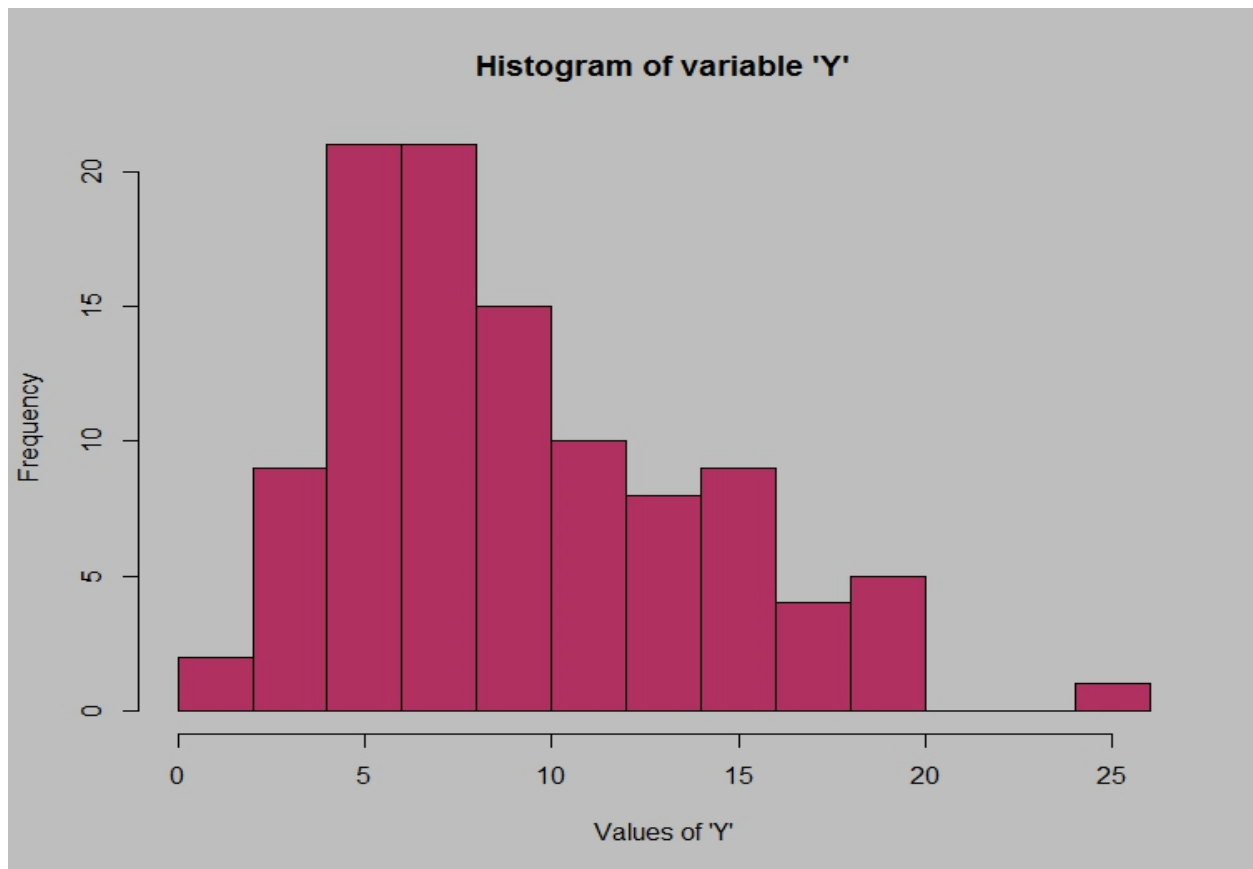Answer the following three questions based on the box-plot above.

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

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

(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?

Ans: i. Approximately (First Quantile Range) Q1 = 5 (Third Quantile Range) Q3 = 12, Median (Second Quartile Range) = 7 (Inter-Quartile Range) IQR = Q3 − Q1 = 12 − 5 = 7 Second Quartile Range is the Median Value

ii. Right-Skewed. Median is towards the left side it is not normal distributed.

iii. In that case there would be no Outliers on the given dataset because of the outlier the data had positive skewness it will reduce and the data will normal distributed

3.

Histogram of variable 'Y'

Answer the following three questions based on the histogram above.

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

(ii)   Comment on the skewness of the dataset.

(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.

Ans: i. The mode of this data set lie in between 5 to 10 and approximately between 4 to 8 .

ii. Right-Skewed. Mean>Median>Mode

iii. They both are right-skewed and both have outliers the median can be easily visualized in box plot where as in histogram mode is more visible.

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. To solve this problem, we first need to calculate the probability that a given telephone call will be misdirected. Since 1 in 200 long-distance telephone calls is misdirected, this probability is $1/200 = 0.005$.

    Next, we need to find the probability that at least one in five attempted telephone calls reaches the wrong number. To do this, we can use the complementary probability formula: P(at least one success) = 1 - P(no successes).

    Since the probability of a given call being misdirected is 0.005, the probability of a given call not being misdirected is $1 - 0.005 = 0.995$. Since these events are independent, the probability that none of the five attempted calls will be misdirected is $(0.995)^5 = 0.976$. Therefore, the probability that at least one of the five attempted calls will be misdirected is $1 - 0.976 = 0.024$.

    Thus, the probability that at least one in five attempted telephone calls reaches the wrong number is 0.024.

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

Ans: i. The most likely monetary outcome of the business venture is 2000$ As for 2000$ the probability is 0.3 which is maximum as compared to others

ii. Yes, the probability that the venture will make more than 0 or a profit

p(x>0) +p(x>1000) +p(x>2000) +p(x=3000)

= 0.2+0.2+0.3+0.1

= 0.8 this states that there is a good 80% chances for this venture to be making a profit.

iii. The long-term average is Expected value = Sum (X * P(X)) = 800$ which means on an average the returns will be = 800$

iv. 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 Var $(X) = E(X^2) - (E(X))^2$

= 2800000 − 800^2 = 2160000