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 Plot refer the attached python file.

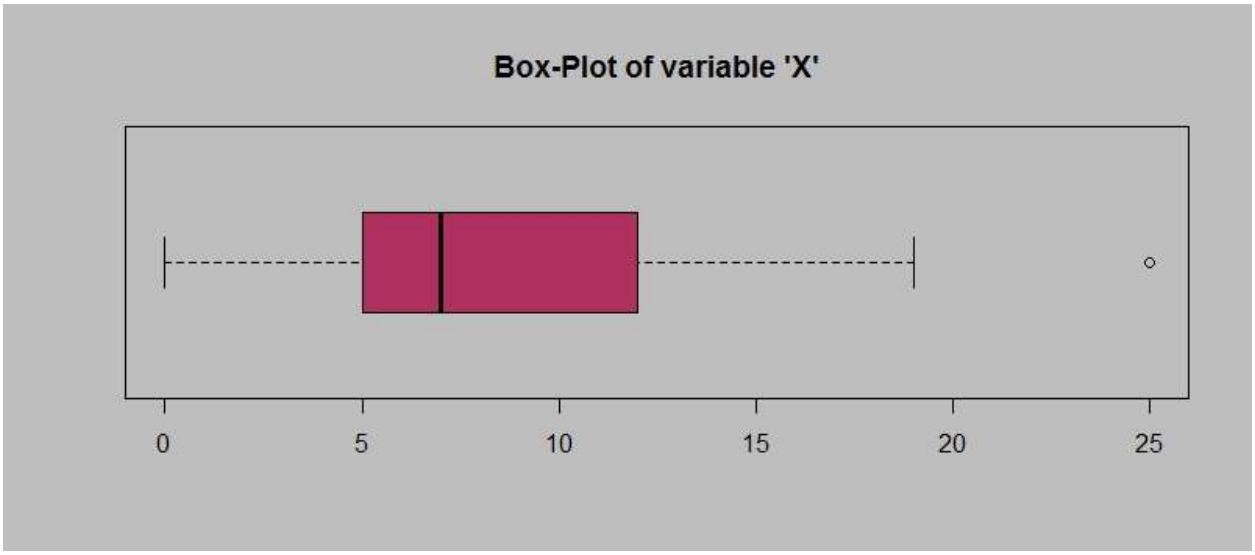
Mean – 33.27133333333333 %

Std Deviation - 16.945400921222028 %

Variance - 287.1466123809524 %

Questions referred to from *Aczel A., Sounderpandian J., Complete Business Statistics (7ed.)*

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: Inter Quartile Range = Q3-Q1 = 12-5 = 7

This value implies 50% of data is inside IQR. This shows variability inside the middle data.

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

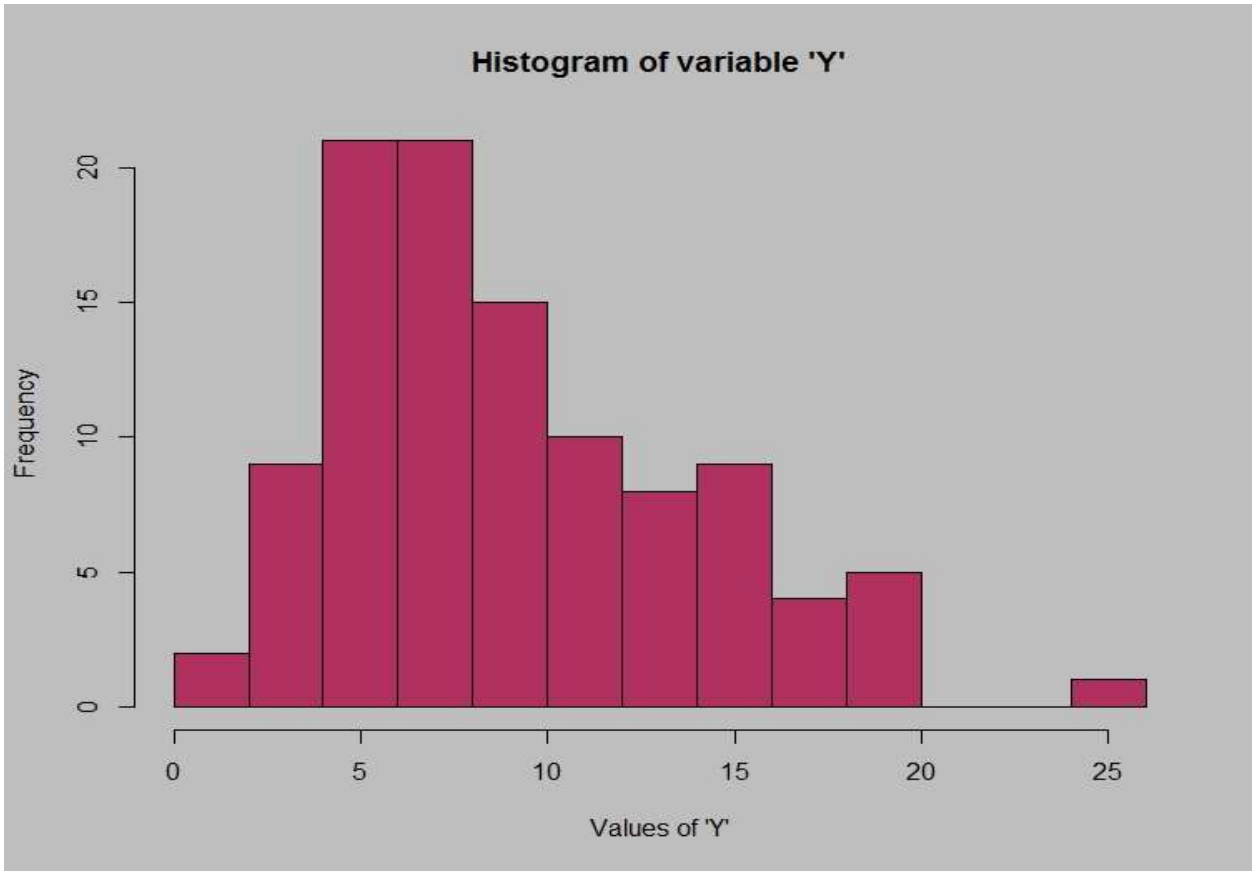
Ans: Since the median is slightly towards the left, this data has Positive Skewness and Right Skewed. Also we can see the outlier at right.

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: Then there will be no outlier. Median shifting will depend on size of data,since single data is not sufficient to pollute the median. It will reduce the right skewness.--------

Questions referred to from *Aczel A., Sounderpandian J., Complete Business Statistics (7ed.)*

3.



Answer the following three questions based on the histogram above.

i) Where would the mode of this dataset lie?

Ans: Values of Y – between 4 to 10 since the majority of data lie in this range.

ii) Comment on the skewness of the dataset.

Ans: Positive Skewness, Mean > Mode& Median.

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: Both the Plot showing Positive Skewness, Outliers as 25. From Box Plot we come to know about the median value i.e 6.5, in histogram we can find mode.

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 :

probability of misdirect = 1/200

probability of correct direct = 199/200

find At least 1 in five attempts?

This is example of binomial distribution with no of trials =5

Probability (x) = nCx . P^(x) . (1-P)^n-x

Probability of At least 1 in five attempts as misdirect = 1-P(all 5 are correct)

P(all 5 are correct) :

No of success(x) = 5 correct calls

Probability of success = 199/200

Probability of failure = 1/200

P(all 5 are correct) : (5!/(5-5)!*5!).(199/200)^5.(1/200)^0 = (0.995)^5

Probability of At least 1 in five attempts as misdirect = 1-P(all 5 are correct)

= 1-(0.995)^5

= 0.02475

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: most likely monetary outcome of the business venture is with highest probability i.e returns = 2000 with Probability of 0.3

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

Ans: Net returns = (-2000*0.1)+(-1000*0.1)+(0*0.2)+(1000*0.2)+(2000*0.3)+(3000*0.1)
= -200-100+0+200+600+300
= 800

Since the overall returns shows positive value(trend) we can say the venture likely to be successful.

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

Long term Avg= (-2000*0.1)+(1000*0.1)+(0*0.2)+(1000*0.2)+(2000*0.3) +(3000*0.1)
= -200-100+0+200+600+300
= 800

So long term avg earning is $ 800.

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

Ans: Std deviation can be good measure of the risk involved in a venture of this kind.

Std Dev is - 1870.8286933869706