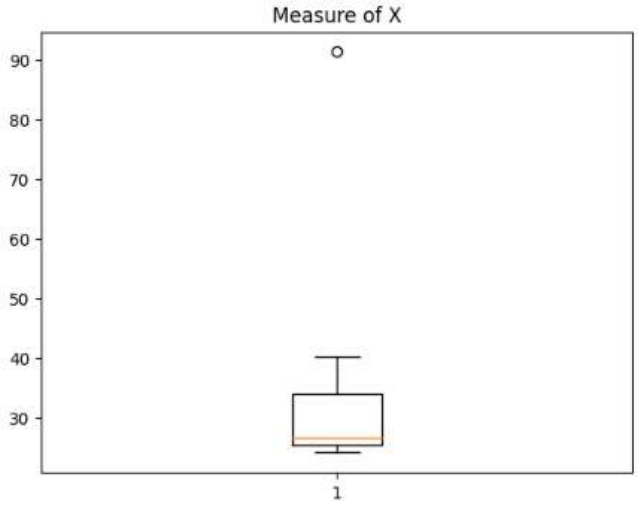
Topics: Descriptive Statistics and Probability

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:

OUTLIER:



There is one outlier in boxplot,that value is Margon Stanley(91.36%)

MEAN:

```
measures.mean()
```

```
33.27133333333333
```

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

VARIANCE:

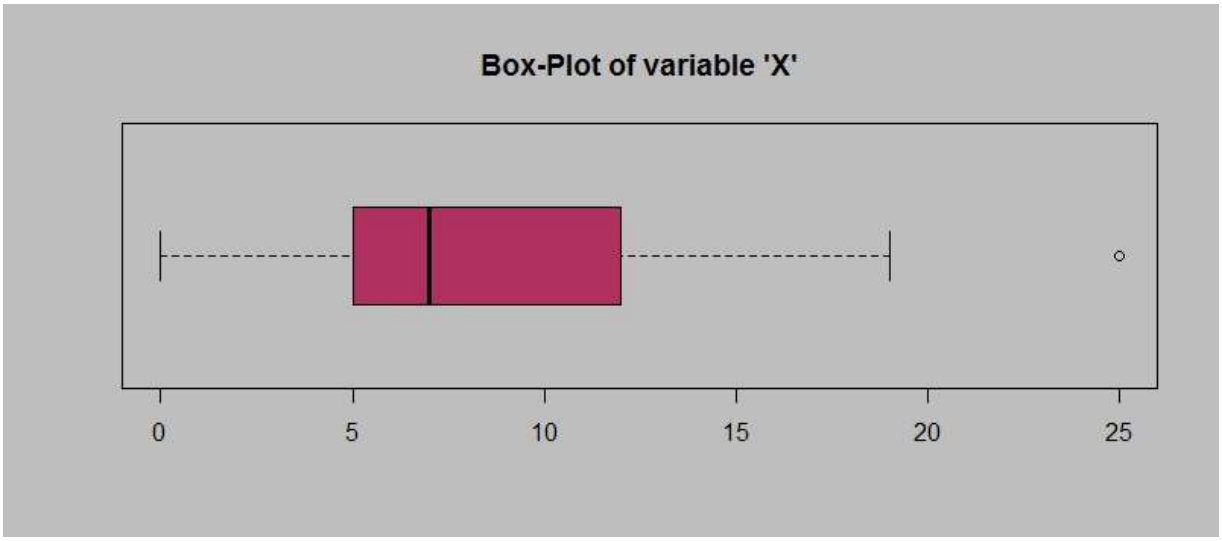
```
measures.std()
16.945400921222028
```

STANDARD DEVIATION:

```
[9] measures.var()
287.1466123809524
```

OUTLIER= 91.36
MEAN=33.27
VARIANCE=16.94
STANDARD DEVIATION=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:
IQR=INTER QUARTILE RANGE
IQR=Q3=Q1
Q3=12,Q1=5
Therefore,
IQR=12-5=7

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

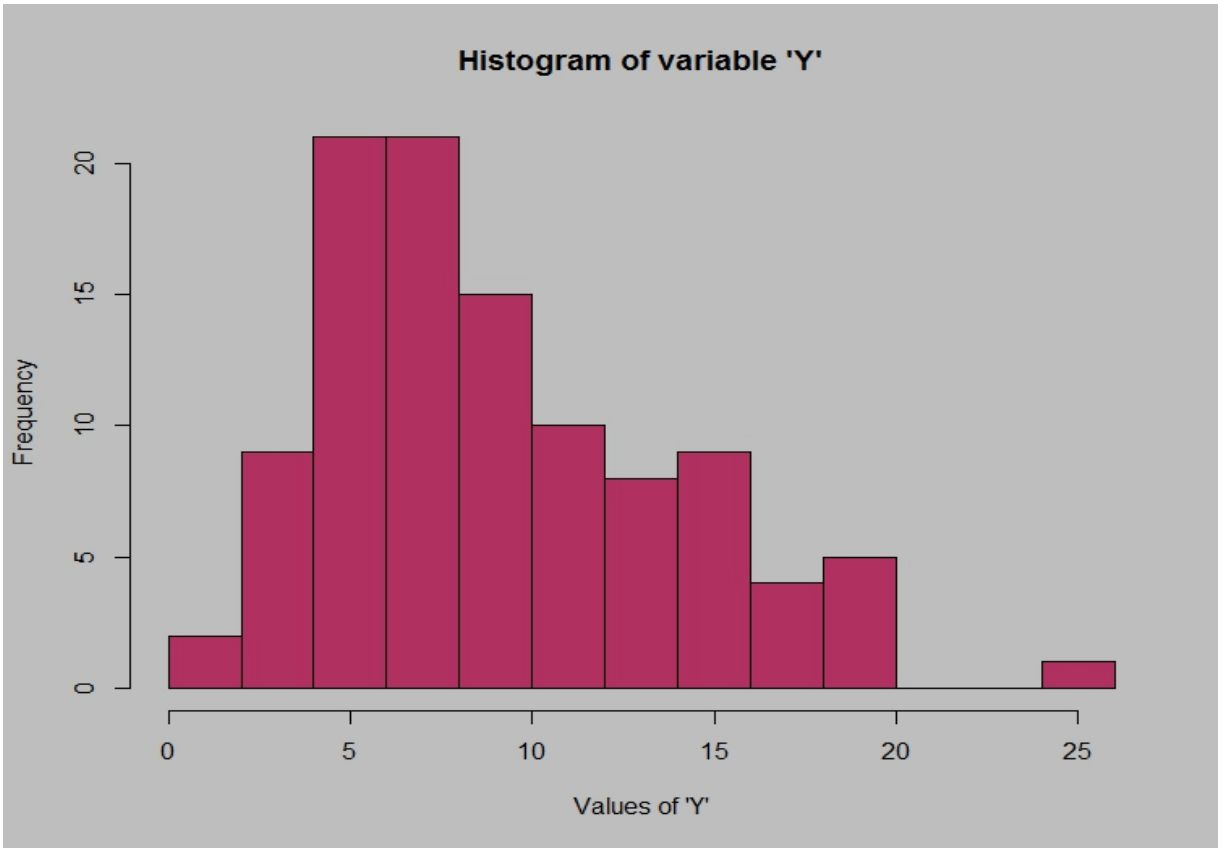
ANS:
In this boxplot,that vertical line much closer to first quartile than third quartile, which means that data is rightly skewed(positive skewness)

(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:

In that case there would have been no outliers, and it might have affected in the values of mean and median slightly. The boxplot might have been moved toward right slightly.

3.



Answer the following three questions based on the histogram above.

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

ANS:

The mode of this dataset lie between 4 to 8

(ii) Comment on the skewness of the dataset.

ANS:

The skewness of the dataset is positive skewness

(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 plots having outliers ,both are positively skewed.

In boxplot median is more visible,In histogram mode is more visible.

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

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:

   Let X be the probability of correctly call being a correctly, Y be the probability of call misdirected.

   Since in one in 200 calls is misdirected, Y=1/200 and X=1-Y

   lets find the probability atleast one in five attempted calls reaches the wrong number

   The probability of all five calls being correctly directed X^5

   Therefore, the probability that atleast one call is misdirected is 1-X^5

   Substitute the values

   X=1-(1-1/200)^5

   =1-0.9752

   =0.0248

   The probability of correctly call being a correctly is 0.0248

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 highest probability is associated with x=2000 where p(x)=0.3

   Therefore,

   The most likely monetary outcome of the business venture is $2000

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

   ANS:

   To determine if the venture is likely to be successful, We can consider the expected value of probability distribution

   The expected value is calculated by multiplying each possible outcome by its probability and summing up these values

   $E(X)=\Sigma(p(x)*x)$

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

Calculate the expected value for the distribution

E(x)=(-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

The expected value is $800.Since the expected value is positive ,it suggests that ,on average ,the business venture is likely to be successful.

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

ANS:

The long term average earning of a business venture of this kind can be determined by calculating the expected man of the probability distribution.

The expected value is calculated by multiplying each possible outcome by its probability and summing up these values

$$E(X)=\Sigma(p(x)*x)$$

Calculate the expected value for the distribution

E(x)=(-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

The long term average earning of a business venture of this kind,based on the probability distribution is $800,The business expected to earn $800 per venture in the long run.

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

ANS:

One common measure of risk in a probability is the standard deviation

Formula for the standard deviation is

$$\sigma=\sqrt{\Sigma(x-\mu)*p(x)}$$

Mean=800

Calculate the standard deviation

σ =√(-2000-800)^2*0.1+(-1000-800)^2*0.1+(0-800)^2*0.2+(1000-800)^2*0.2+(2000-800)^2*0.3+(3000-800)^2*0.1

=√784,000+324,000+128,000+8000+432,000+484,000

=√2,160,000

=1469.69

Therefore, the standard deviation is $1470.A higher standard deviation is highest variability or high risk.