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:-
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sn
%matplotlib inline
```

```
x=pd.Series([24.23,25.53,25.41,24.14,29.62,28.25,25.81,24.39,40.26,32.95,91.36,25.99,39.42,26.71,35.00])
```

A) Outliers=sns.boxplot(x)=91.36

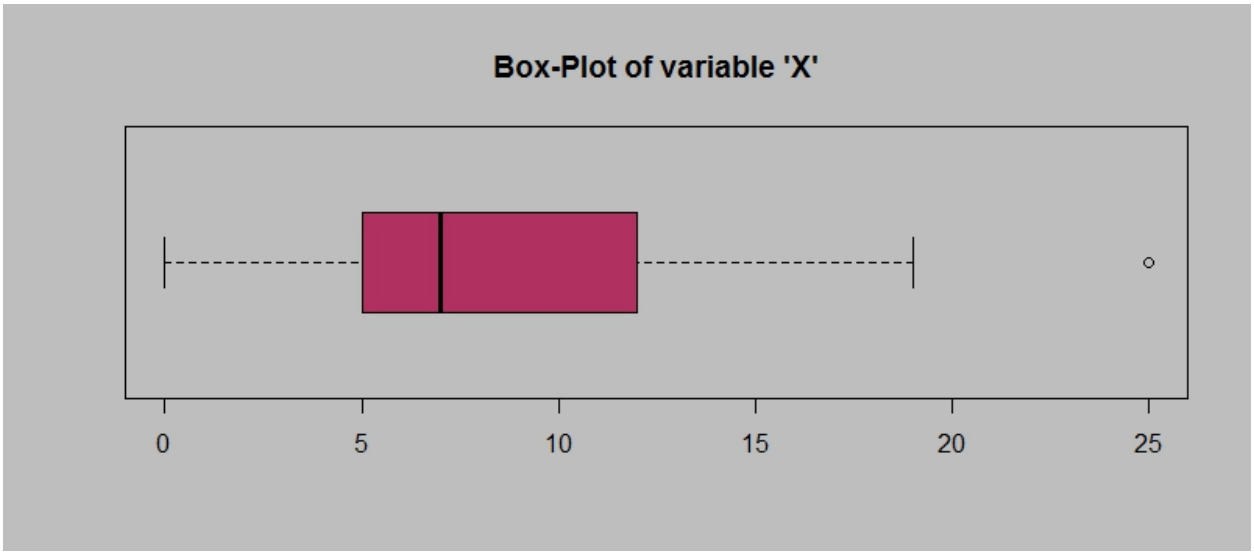
B)Mean= x.mean()=33.27133333333333

C)std= x.std()=16.945400921222028

D)(Std)^2=x.var()=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.

(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) (IQR)=Q3-Q2

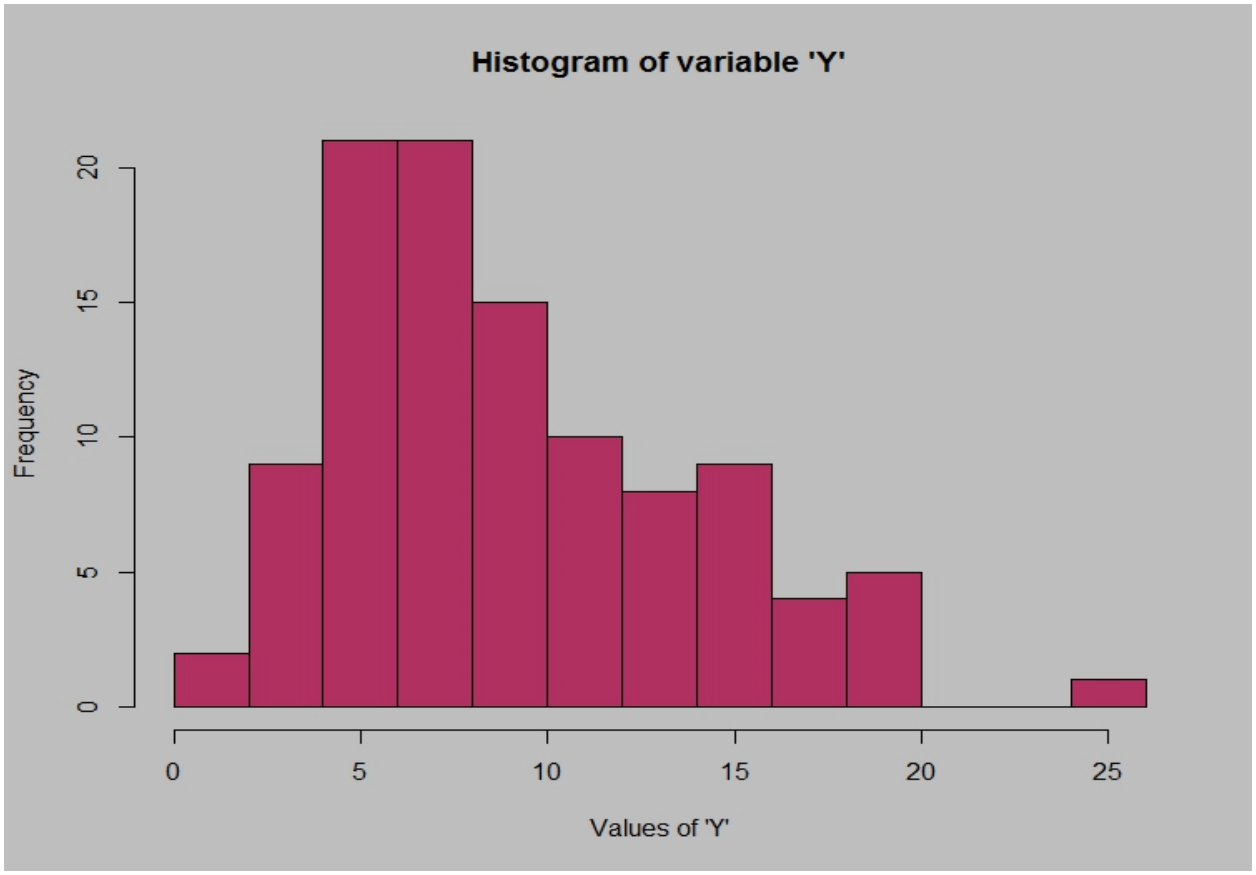
=12-5

=7

ii) we can say that it is positive skewness/Right skewed.

iii) It will affect to outlier 2.5 will not be considered as outlier.

3.



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) Mode will lie between 4 to 8.

ii) It is positive skewed.

iii) we can say that 50% of data lies in between 5 to 12.histogram provide frequency distribution and boxplot is providing most of the body lies in between 5 to 12. From the fiqures we can say that 25 is a outlier.

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;- out of 200 call one call is misdirecting

P (call is misdirecting) = 1/200

P (call is misdirecting) = (1-1/200)=199/200

We have formula p(x)=ncr*P(success)^r*P(failed)^(n-r)

Where, n=5,r=0

P (call is not misdirecting) =[(5*4*3*2*1)/(5*4*3*2*1)] x(1/200)*0 x(199/200)*5

= (199/200)*5

= 0.975

P (call is misdirecting) = 1-0.975

=0.025.

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,it has maximum probability of 0.3.

ii)-2000x0.1+(-1000x0.1)+(0x0.2)+1000x0.2+2000x0.3+3000x0.1=800$

As average is positive the venture is successful.

iii)AS we have already calculated above long term average earing =800$.

iv) we will take loss of(-2000)&(-1000)

The probability will be 0.1+0.1=0.2

The risk involved will be 20%