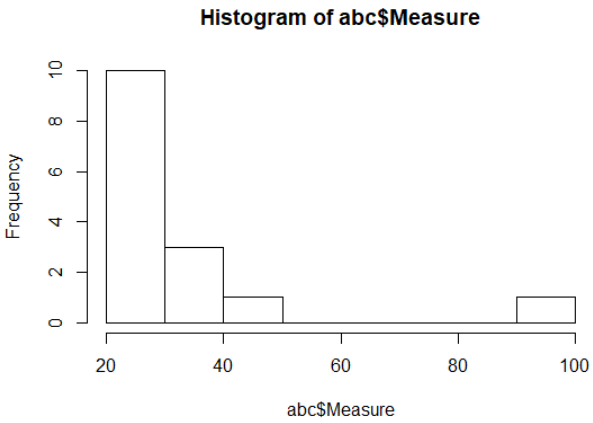
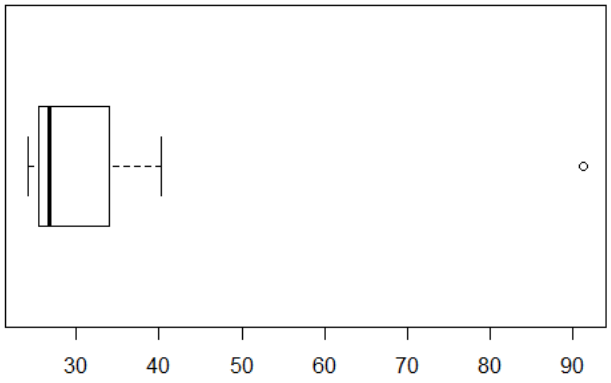
Topics: Descriptive Statistics and Probability

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%



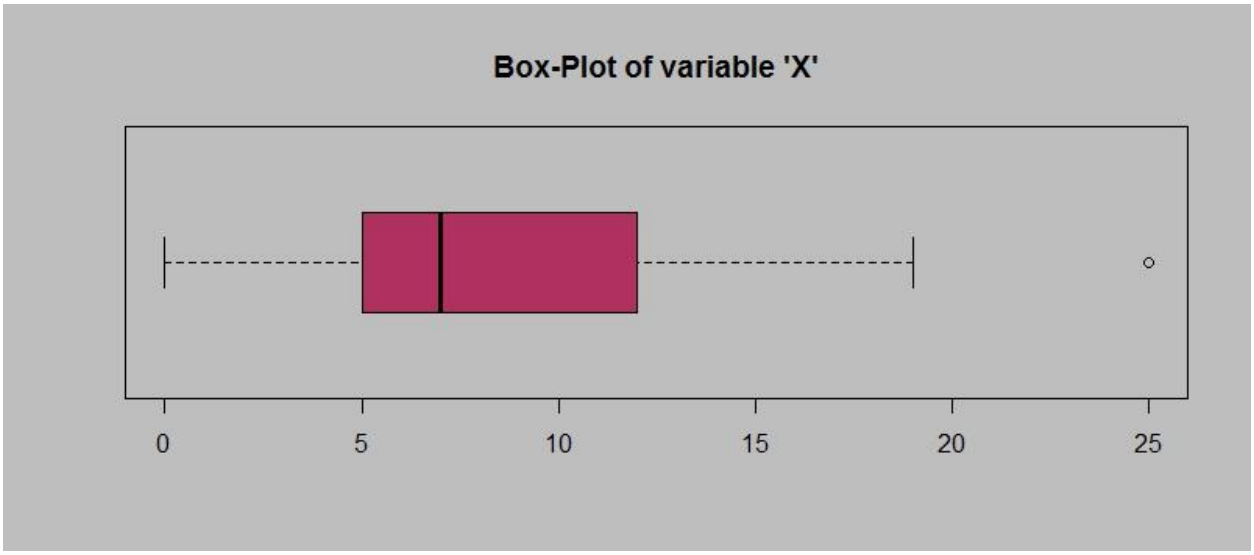


Answers are in Percentages

Mean 33.27133
variance 268.0035
SD 16.37081

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.

IQR = Q3 – Q1

12 – 5 = 7

Therefore, inter-quartile range is approximately around 7.

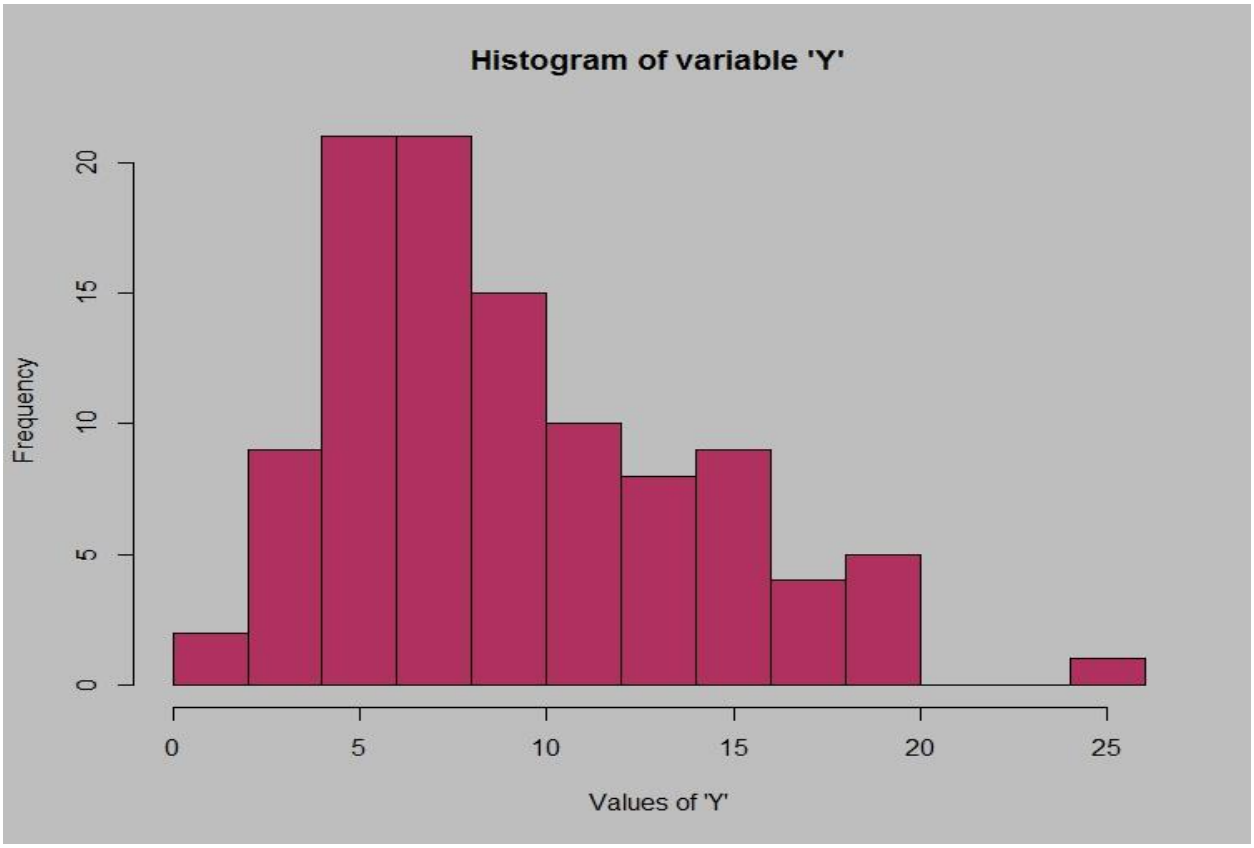
(ii) What can we say about the skewness of this dataset?

The Distribution is Right Skewed or Positive Skewed.

(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?

There would be no outlier in the above box-plot figure if 25 was actually 2.5.

3.



Answer the following three questions based on the histogram above.

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

Mode of the data set lies around 5.

(ii) Comment on the skewness of the dataset.

The Distribution is Right Skewed or Positive Skewed.

(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.

From the above histogram and box-plot figures we could say that distribution is right or positive skewed and outlier lies at 25.

Box-Plot helps us to determine Median, Inter Quantile Range, smallest value, Range, Outlier and Skewness in the datasets.

Whereas Histogram represents Mean, Median, Mode and Skewness.

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.)

Probability of Misdirected class A = $\frac{1}{200}$

$P(\overline{A})$ = 1 - $\frac{1}{200}$ = $\frac{199}{200}$

Probability that at least one in 5 call reaches wrong number.

= 1 – (Probability that calls reaches to attempted person)^5

= 1 – ($\frac{199}{200}$) ^5 = 1-(0.995) ^5 = 1 – 0.975248

=0.0247 or 2%

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?
2000 is the most favorable outcome since it has the highest probability.

(ii) Is the venture likely to be successful? Explain
For the above probability distribution, we could say that business venture is likely to be successful with 60% or .6 probability. 0.2+0.3+0.1 = .6 or 60%.

(iii) What is the long-term average earning of business ventures of this kind? Explain
(-2000*0.1) + (-1000*0.1) + (0*0.2) + (1000*0.2) + (2000*0.3) + (3000*0.1) = 800
Therefore, long term average of this venture is $800.

(iv) What is the good measure of the risk involved in a venture of this kind? Compute this measure.
By measuring the variance and standard deviation from the above data set we could say that this venture is very risky.
Variance = 3500000
Standard Deviation = 1870.829