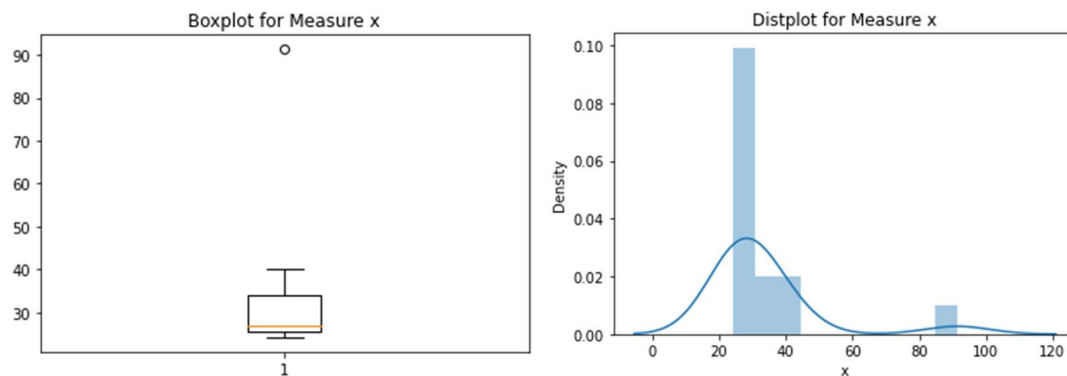


## Topics: Descriptive Statistics and Probability

1. Look at the data given below. Plot the data, find the outliers and find out  $\mu, \sigma, \sigma^2$

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%

Answer:



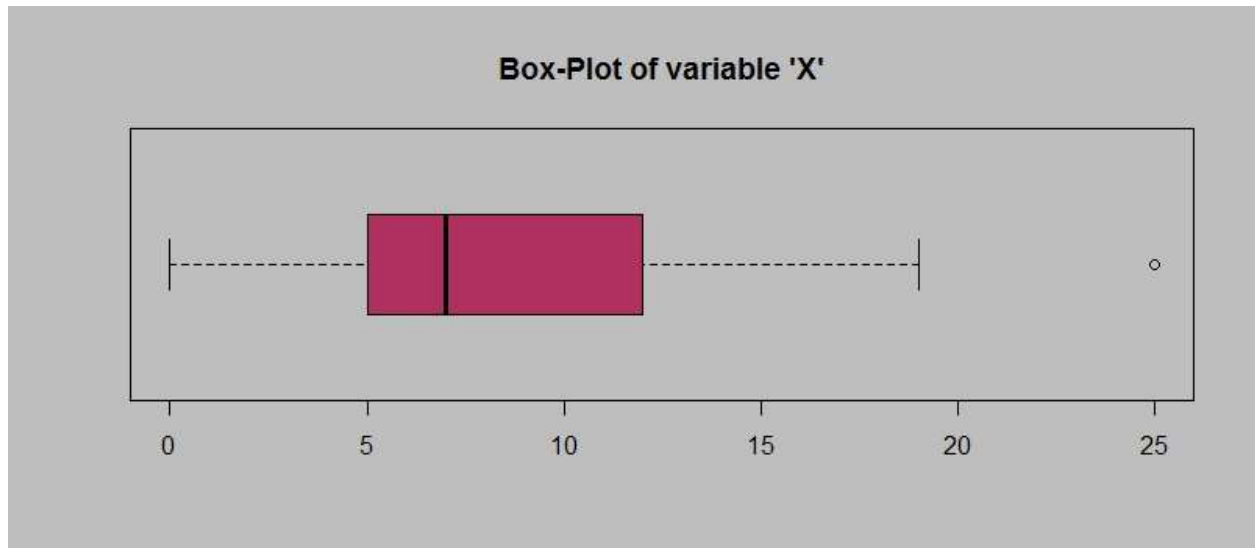
There is one outlier present in the data → Morgan Stanley at 91.36%

Mean of Measure  $x = \mu = 33.2713 \%$

Variance of Measure  $x = \sigma^2 = 268.0035 \%$

Standard Deviation of Measure  $x = \sigma = 16.3708 \%$

2.



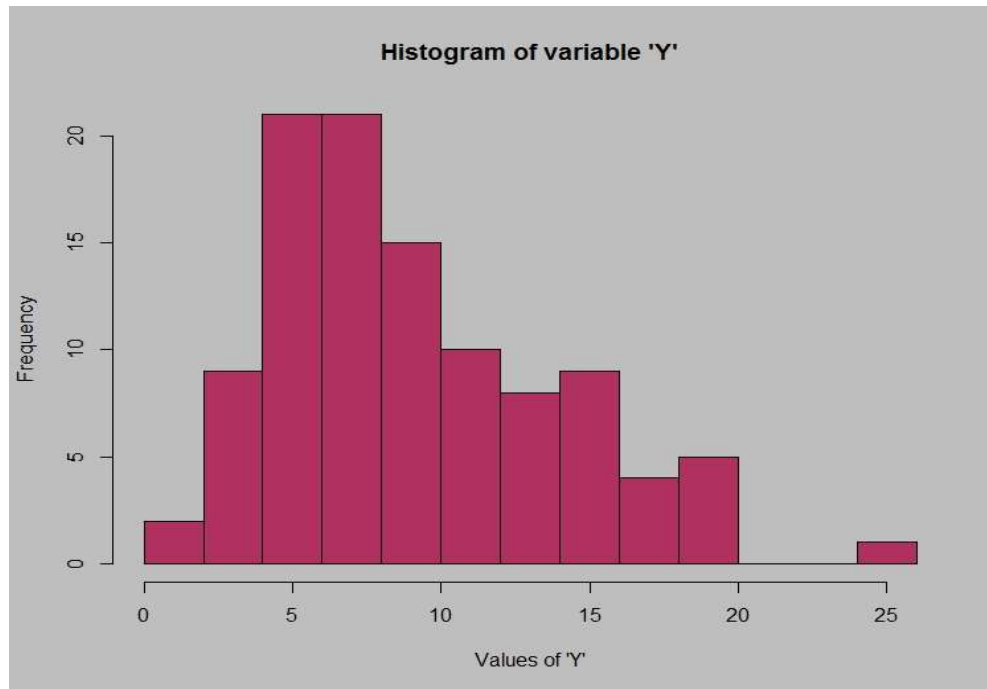
Answer the following three questions based on the box-plot above.

- (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?

Answer:

- (i) Inter-quartile range of this dataset is  $12.5 - 5 = 7.5$  approximately. This value implies that the majority of values falls within this range except one outlier at 25.
- (ii) The data is **right skewed** or **positively skewed**.
- (iii) If it was found that the data point with the value 25 is actually 2.5, then the inter-quartile range will be changed to  $12.5 - 2.5 = 10$ . And there will be no outliers in the data also the distribution of the data will be perfectly normally distributed.

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.

Answer:

- (i) Mode of the dataset would be at the peak point of the histogram i.e. between 4 to 8.
- (ii) The data is **right skewed or positively skewed**.
- (iii)

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.)

Answer:

Given data – one in 200 long-distance telephone calls is misdirected

No. of attempts =  $n = 5$

Probability of calls misdirected =  $P(m) = 1/200$

Probability of calls not misdirected =  $1 - P(nm) = 1 - (1/200) = 199/200$

$$P(x) = {}^nC_x P(m)^x P(nm)^{n-x}$$

at least one in five attempted telephone calls reaches the wrong number =  $1 - P(0)$

$$\begin{aligned} P(0) &= \text{none of the call reaches the wrong number} = {}^5C_0 * (1/200)^0 * (199/200)^{5-0} \\ &= (199/200)^5 \\ &= 0.97525 \end{aligned}$$

$$1 - P(0) = 1 - 0.97525 \rightarrow 0.02475$$

So, the probability that at least one in five attempted telephone calls reaches the wrong number is **24.75%**

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

Answer:

- (i) Most likely outcome of the business venture = \$2000 which has highest probability value 0.3.
- (ii) 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 net returns is 800 (which is positive) we can say that the venture is successful.
- (iii) Long-term earning of business ventures of this kind =  $\sum x * P(x)$   
 $= (-2000)(0.1) + (-1000)(0.1) + (0)(0.2) + (1000)(0.2) + (2000)(0.3) + (3000)(0.1)$   
 $= (-200) + (-100) + (0) + (200) + (600) + (300)$   
 $\sum x * P(x) = 800$   
 Hence, long-term earning of business ventures is \$800.
- (iv) Standard deviation is the good measure of the risk involved in this kind of a venture.  

$$SD \text{ of } x = \sigma = \sqrt{\frac{\sum |x - \mu|^2}{n}} = 1707.8251$$