

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%

In [24]:

```
# Mean  
x.mean()
```

Out [24]:

```
33.27133333333333
```

In [25]:

```
# Vairance  
x.var()
```

Out [25]:

```
287.1466123809524
```

In [26]:

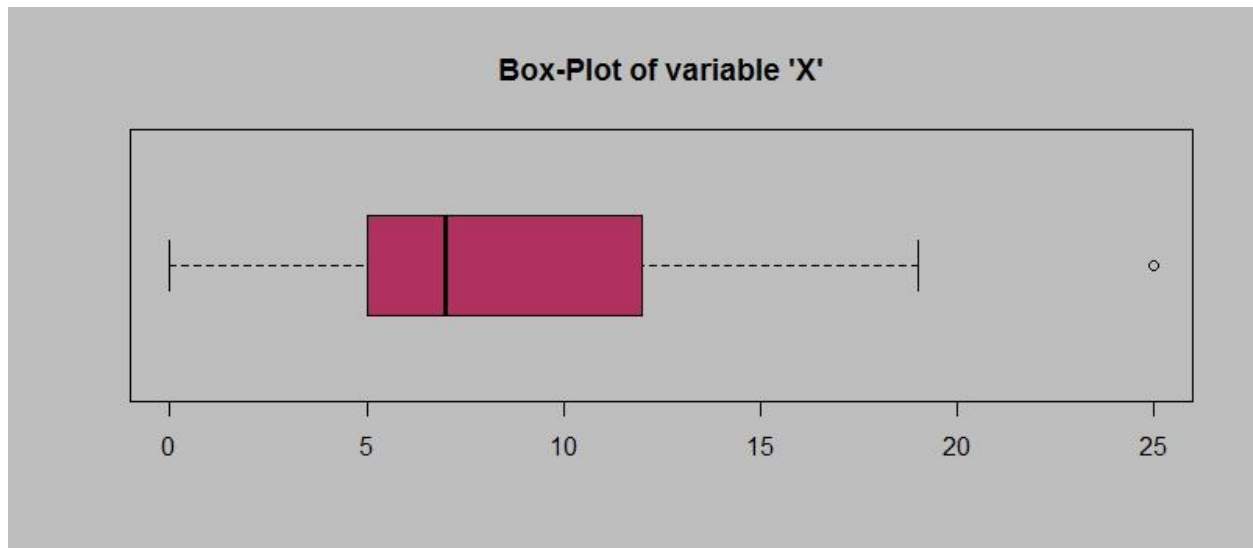
```
# Standard Deviation  
x.std()
```

Out [26]:

```
16.945400921222028
```

In []:

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 (First Quantile Range) $Q1 = 5$

(Third Quantile Range) $Q3 = 12$

Inter-Quartile Range) $IQR = Q3 - Q1 = 12 - 5 = 7$

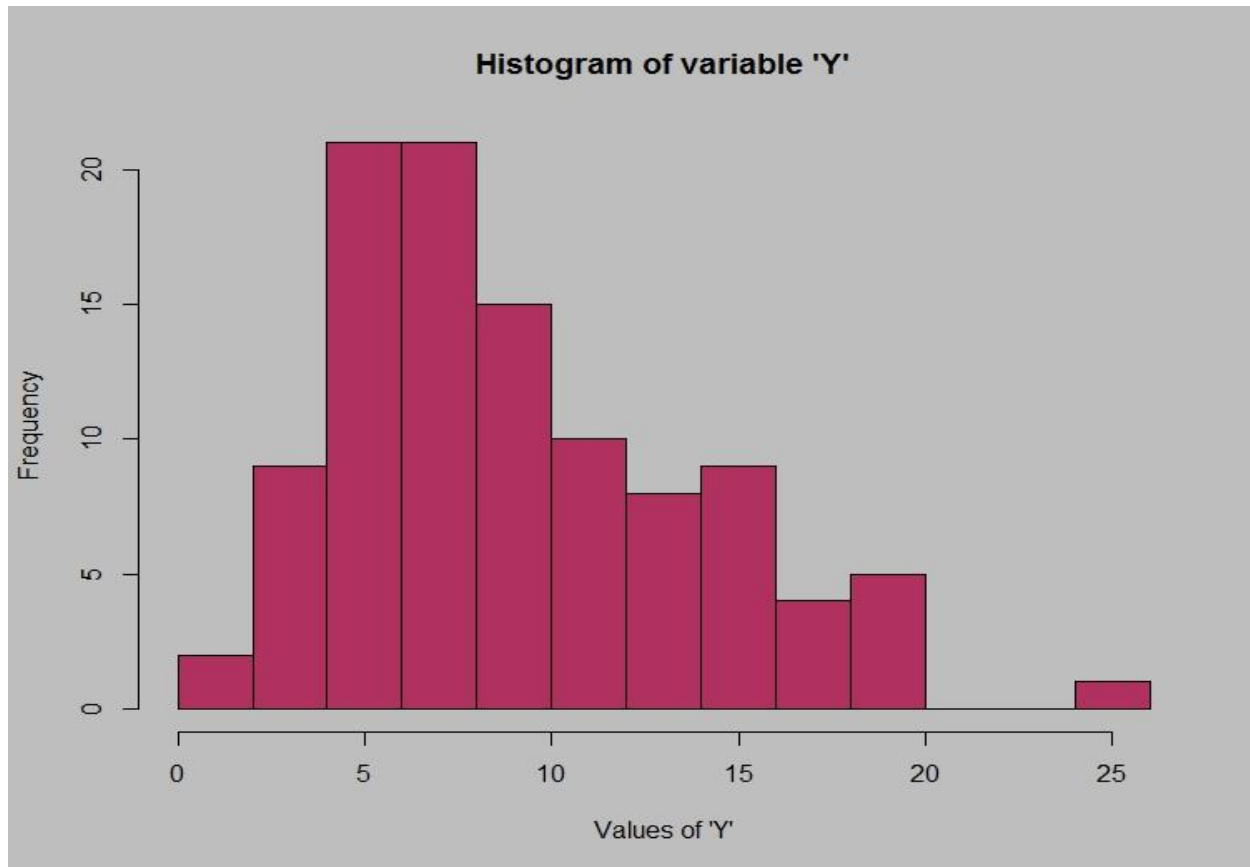
- (ii) What can we say about the skewness of this dataset?

Ans: Right-Skewed median is towards the left side

- (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 be no Outliers on the given dataset because of the outlier the data had positive skewness it will reduce, and the data will normal distributed

3.



Answer the following three questions based on the histogram above.

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

Ans The mode approx between 5 to 8

- (ii) Comment on the skewness of the dataset. [OBJ]

Ans: This is a positive skewed data

- (iii) Suppose that the above histogram and the boxplot in question 2 are plotted for the same dataset. Explain how these graphs complement each other in providing information about any dataset.

Ans: we can say that by checking the hist and boxplot that the data is positive skewed and has one outlier and also, would help us to find the mean mode value

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 call getting misdirected= (1/200)

Hence probability of call not getting misdirected=1-(1/200) = 199/200

Number of phone calls attempted= 5

Therefore, probability that at least one in 5 attempted call reaches the wrong number is
 $= 1 - (199/200)^5$
 $= 0.025$

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?

Ans: Here is the highest probability for 2000

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

Ans: Yes, the total earning is in positive that is 800 and 2000 is the highest probability

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

Ans: 800(explanation given below)

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

Standard Deviation = Square root of Variance = Square root $(((-2800^2) * 0.1) + ((-1800^2) * 0.1) + ((-800^2) * 0.2) + ((200^2) * 0.1) + ((1200^2) * 0.3) + ((2200^2) * 0.1)) = 1468$

x	P(x)	Income (x * P(X))
-2,000	0.1	-200
-1,000	0.1	-100
0	0.2	0
1000	0.2	200
2000	0.3	600
3000	0.1	300
Total		800