**Topics: Descriptive Statistics and Probability**

1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **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 : In jupyter notebook



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

1. What is inter-quartile range of this dataset? (please approximate the numbers) In one line, explain what this value implies.

Answer:

First quartile range Q1 = 5, Third quartile range Q3 = 12

So, Inter quartile range IQR = Q3 – Q1 = 12 – 5 = 7

That is second quartile range is the median value.

1. What can we say about the skewness of this dataset?

Answer:

It is right-skewed and median is towards the left side so it is not a Normal distribution.

1. 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:

In that case there would be no outliers and, on the given dataset because of the

Outliers there is positive skewness , it will reduce and data will be normally distributed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Answer:

The mode of this data set lie in between 5 to 10 and approximately between

4 to 8.

1. Comment on the skewness of the dataset.

Answer:

The distribution of the data is positively right skewed.

Mean>Median>Mode

1. 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:

Both histogram and boxplot are right-skewed also have outliers, the median can be

Easy to detect in boxplot whereas in histogram mode is more easy to detect.

1. 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:

Probability that at least one in five attempted telephone calls reaches the wrong number is 0.024.

1. 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 |

1. What is the most likely monetary outcome of the business venture?

Answer:

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

Probability is 0.3 which is maximum as compare to others.

1. Is the venture likely to be successful? Explain

Answer:

Yes, the venture likely to be successful, because the probability that this venture

Have that is more than 0 and talking about profit then,

P(x>0) + P(x>1000) + P(x>2000) + P(x>3000)

0.2 + 0.2 + 0.3 + 0.1 = 0.8

It means there is 80% chances for this venture for making profit.

1. What is the long-term average earning of business ventures of this kind? Explain

Answer:

The long-term average earning is

(0.1)(-2000) + (0.1)(-1000) + (0.2) (0) + (0.2) (1000) + (0.3) (2000) + (0.1) (3000)

= 800

So, the average will be 800$

1. What is the good measure of the risk involved in a venture of this kind? Compute this measure.

Answer:

**Mean = (-2000-1000+0+1000+2000+3000)/6**

**Mean = 500.**

**n-1 = 6-1 = 5**

**Std. dev = √ [(-2000-500)2+(-1000-500)2+(500-0)2+(1000-500)2+(2000-500)2+(3000-500)2]/5**

**Std.dev = 1870.8**