**Topics: Descriptive Statistics and Probability**

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

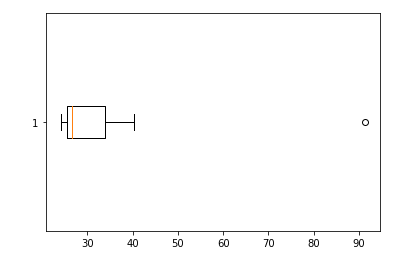
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft | 32.95% |
| Morgan Stanley | 91.36% |
| Sun Microsystems | 25.99% |
| Travelers | 39.42% |
| US Airways | 26.71% |
| Warner-Lambert | 35.00% |

Ans: Mean = 33.27

Stanfard deviation = 16.94

Variance = 287.146

Outliers = 91.36





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.

Ans: IQR = Q3 – Q1 = 12 – 5 = 7

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

Ans: From the above Boxplot, we can understand that the data is centered around Q2.

Here, (Q3-Q2) > (Q2-Q1) which implies that the dataset is Positive Skewed

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?

Ans: the data 2.5 will be added to the boxplot, which will increase the value of (Q2 – Q1).

This will reduce the Positive skewness of the data.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: The mode of the data lies between (4 – 8) of y values

1. Comment on the skewness of the dataset.

Ans: The given histogram shows that the data is Right Skewed

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.

Ans: From the above Histogram and Boxplot, we can understand that both graphs shows nature about the dataset. In both the graph we can see that the outlier 25 affects both the graph noticeably. Both representation implies the data is Right Skewed.

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.)

Ans:

P (misdirected call) = 1/200

P ( success call) = 1-(1/200) = 0.995

Given that the attempts are independent

So, P (Success call for 5 attempts) = (1/200)^5

Ie , p( misdirected call) = 1 – (1/200)^5 = 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?

Ans: The most likely monetary outcome of the business venture is $2000 because it has the highest probability

1. Is the venture likely to be successful? Explain

Ans: Yes , because P (x>=0) is higher than P (x< 0) , that means business is more likely to get profit and successful

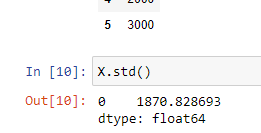
1. What is the long-term average earning of business ventures of this kind? Explain

Ans: The long term avg earning of business ventures of this kind is $800.

Ie, (-2000\*0.1)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1) = 800

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

Ans: If we can analyse the deviation of value from the avg expected value, then we can understand the amount of risk involved in such venture.



From the above calculation of std, we can understand the venture is in highly risk. (note: mean = $800)