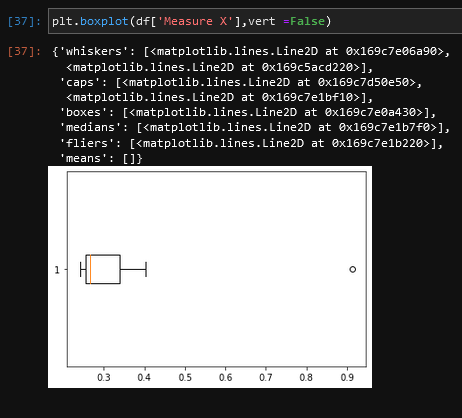
**Topics: Descriptive Statistics and Probability**

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% |



Mean(µ) = 33.27%(Including the outlier)

Mean(µ) = 29.12%(If the outlier 91.36 is excluded)

Variance(σ^2) = 0.2890

S.D(σ) = 5.376%



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.

IQR = Q3-Q1 = 12-5 = **7** (This number indicates how far is the spread of the middle of the data in 3 quartiles.

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

The distribution is slightly skewed right (Has a positive skewness)

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?

If the above scenario is considered, then the box would move bit towards the left and the value of the skewness would increase a bit more (making it more positive)



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie? Between 5 and 7
2. Comment on the skewness of the dataset. Positive Skewness
3. 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. The graphs show that the data is normally distributed, yet slightly right skewed. It clearly shows the outlier (in this case, 25). Also, it shows the kurtosis and frequency of the data.
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 1 wrong number out of 200 calls= P(Wrong) = 1/200 = 0.005

Probability of correctly connected calls= 1 - P(Wrong) =1- 0.005 = 0.995

Probability of at least one out of five calls is a wrong number

= 1 – Probability that all five calls are correctly connected

= 1 – (1 – P(Wrong)) ^5

= 1 – (1- 0.005) ^5

= 1 – 0.975

= 0.024 = **2.4%**

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?

Equals the highest probability = P(0.3) = $2000

1. Is the venture likely to be successful? Explain

Probability of Success = Probablity of Gain – Probability of Loss

= (0.2+0.3+0.1) – (0.1+0.1) = **0.4 = 40%**

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

Long-Term average earning is

= (-2000\*01)+(-1000\*01)+(0\*0.2)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)

= (-200)+(-100)+0+(200)+(600)+(300)

= **800**

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

Measure of risk can be calculated by the Standard Deviation(S.D).

Mean = 1333.33 and S.D = 2687.41

Hence this business venture is risky as the S.D value is high.