**Topics: Descriptive Statistics and Probability**

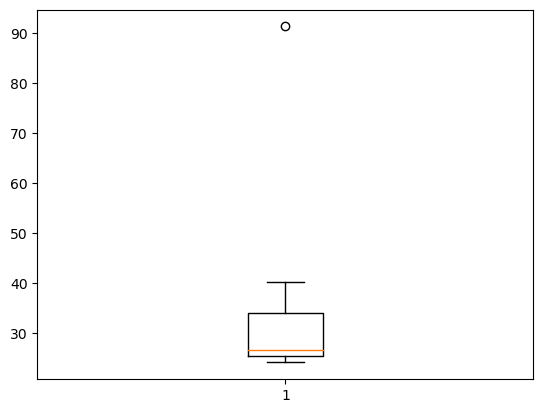
1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **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: - Population mean of the data: 33.2713.

Standard deviation of the data: 16.3708

variance of the data: 268.0035





Answer the following three questions based on the boxplot above.

1. What is the 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(approx.)

It tells about middle 50% of the data distribution and this data not affected by outliers.

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

**Ans: -** Positive Skewness, due median is closer to lower quartile.

1. If it was found that the data point with the value 25 is actually 2.5, how would the new boxplot be affected?

**Ans: -** There will be no outlier in the data and the point will be included in the main body or box of boxplot. The whisker on the right side of the boxplot would likely extend further, and the overall range of data represented by the boxplot would decrease.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans: -** The mode of this dataset would lie around the value of **5** on the X-axis as that is where the highest frequency (around 20) occurs.

1. Comment on the skewness of the dataset.

Ans: - The dataset is **positively skewed**, meaning that there are a number of smaller observations stretching the distribution to the right.

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: - Histogram can provide insights into the distribution and skewness of data, while box-plots can offer additional information about data’s central tendency, variability, and presence of outliers. Box-plots can complement histograms by providing a visual representation of the median, quartiles, and outliers of the data.

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 (directed correctly) = 1 – 1/200 = 199/200

The probability for at least one in five attempted telephone calls reaching the wrong number is:

P (at least one in five is wrong number) = 1 – P (none in five is wrong) = 1 – (199/200)^5 ≈ 0.02475

Therefore, the probability that at least one in five attempted telephone calls reaches the wrong number is approximately **0.02475**.

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 **$2,000** as it has the highest probability of 0.3.

1. Is the venture likely to be successful? Explain

**Ans: -** The venture has a 0.6 probability of either breaking even or making a profit, and a 0.2 probability of making a loss. So, it is more likely to be successful than unsuccessful.

1. What is the long-term average earning of business ventures of this kind? Explain
2. **Ans: -**  The long-term average earning can be calculated by multiplying each outcome by its probability and summing those products:

x = 800

Therefore, the average earning is equals to $800

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

**Ans: -** The expected value is $800, the standard deviation, which measures the risk involved in this venture, is approximately $1280.62. The higher the standard deviation, the higher the risk. In this case, the risk is relatively **high.**