**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 (1): Please refer to the attached notebook “Basic\_Statistics\_Level\_2\_Set1.ipynb”



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 (i): The interquartile range (IQR) of this dataset is approximately 7. This means that the middle 50% of the data points have values between 5 and 12.

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

Answer (ii): The box plot is slightly skewed to the right, which means that there are more data points above the median than below it.

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 (iii): If the data point with the value 25 is actually 2.5, the new boxplot would be slightly shifted to the left and there will be no outliers.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Answer (i): The mode of this dataset can be anywhere from 4 upto 9.

1. Comment on the skewness of the dataset.

Answer (ii): The dataset is slightly skewed to the right, which means that there are more data points above the median than below it. This is evident from the histogram, which shows a longer tail 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.

Answer (iii): Boxplots displays central tendency and outliers, while histograms highlight distribution shape, mode, skewness, and kurtosis.

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 (4): Please refer to the attached notebook “Basic\_Statistics\_Level\_2\_Set1.ipynb”

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 (i): 0.3 is the most likely monetary outcome of the business venture.

1. Is the venture likely to be successful? Explain

Answer (ii): Yes, the venture is likely to be successful. We can see that there is a 60% probability of success based on the expected outcome.

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

Answer (iii): Average Earning (AE) = (-2000 \* 0.1) + (-1000 \* 0.1) + (0 \* 0.2) + (1000 \* 0.2) + (2000 \* 0.3) + (3000 \* 0.1) i.e. AE = -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

Answer (iv): Please refer to the attached notebook “Basic\_Statistics\_Level\_2\_Set1.ipynb”