**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:*** See in jupyter file.



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:*** Inter-quartile range of the given dataset is IQR=12 – 5 = 7.

From IQR it is observed that the datasets has less variability about its median. Its IQR is quit smaller. The middle 50% of dataset vary by somewhat 7 points.

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

***Ans:*** Dataset is Right 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:*** Spread of the graph will reduce and more accurate results will found.

Outlire will remove.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

***Ans:*** The distribution is right skew skewed distribution. The mode is the smallest value and it is located to the left on the distribution. The mode always occurs at the highest point of the peak. So the mode will lie between 4 and 8.

1. Comment on the skewness of the dataset.

***Ans:*** Dataset 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:*** 1. we can find the Median in box plot where as mode in histogram.

2.Histogram provides frequency of occurrence of data where outliers present in boxplot represent 50% of data lie between 5-12.

3. Identification of outlier point is possible in boxplot where no information about outliers in histrogram. We can only guess the outlier present at point 25.

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:***  Probability of call misdirecting  p = 1/200

      Probability of call not Misdirecting = 1 - 1/200 = 199/200

At least one in five attempted telephone calls reaches the wrong number

= 1 - none of the call reaches the wrong number

= 1 -P(0)

= 1 - ⁵C₀(1/200)⁰(199/200)⁵⁻⁰

= 1 - (199/200)⁵

= 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:*** From the data most likely monetary outcome of business venture is 2000 having a probability of 0.3.

1. Is the venture likely to be successful? Explain

***Ans:*** E(x) \* p(x) = -200 – 100 + 0 +200 +600 + 300 = $ 800

So venture is most likely to be successful as the expected value is positive 800.

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

***Ans:*** The long term average earning of the business venture is most likely 800 including all losses and gain.

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

***Ans:*** P(-2000) + P(-1000) =0.1 + 0.1= 0.2

20% risk involved in a business venture .