**Topics: Descriptive Statistics and Probability**

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

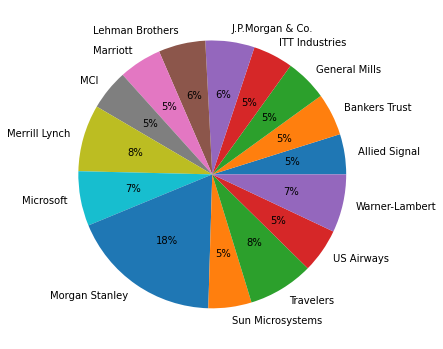
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft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

Ans: Morgan Stanley (91.36%) is an outlier, because it is different from other observation.

Mean = 33.271

Variance = 287.146

Standard Deviation = 16.945





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 = 12-5 = 7, this represents the range which contains 50% of the data points.

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

Ans: It 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: 2.5 will be not considered an outlier. The boxplot will start from 0 and send at 20 in representation.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: 4 to 8

1. Comment on the skewness of the dataset.

Ans: It is a Positive 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: Histogram provides the frequency distribution so we can see how many times each data point is occurring however boxplot provides the quantile distribution i.e. 50% data lies between 5 and 12. Boxplot provides whisker length to identify outliers, no information from histogram. We can only guess looking at the gap that 25 may be an outlier.

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 = 0.028 (Round 3%)

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: Max P = 0.3 for X (2000)

1. Is the venture likely to be successful? Explain

Ans: We can see that there is 60% (0.6) chance that the venture would get profit. So the venture is likely to be successful.

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

Ans: The average would be 800 (mean). This means the long-term average earning of business ventures is 800(including all loses and gains over period of time).

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

Ans: P(loss) =0.2. So, the risk associated with this venture is 20%.