**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** mean = 33.27133333333333 , std = 287.1466123809524, var = 16.945400921222028

there are no low value outliers but only one high valued outlier which is Morgan Stanley



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** 12 – 5 = 7
2. What can we say about the skewness of this dataset?

**Ans** 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:** there will be no outliers in the plot



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie? **Ans** around 5-7
2. Comment on the skewness of the dataset. **Ans** Right Skewed
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. **Ans** The median can be seen from the box-plot and mode can be seen from the histogram
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.) **Ans** 1-(1-(1/200))^5 = 0.0234751
5. 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** 2000$ because it has 0.3 as Prob

1. Is the venture likely to be successful? Explain **Yes** Profit is 3000$
2. What is the long-term average earning of business ventures of this kind? Explain 800$ as sum(xi\*p(xi)) = 800$
3. What is the good measure of the risk involved in a venture of this kind? Compute this measure **Ans** the variance is very high in the data hence the spread is too much which means this type of data is not good for analysis.