**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% |

* Refer ‘A 2] set 1.ipynb’ for calculations.



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.

* IQR -> [5 , 12]
* It implies 50% of data lie in range 5 to 12

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

* Data is right skewed (positively 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?

* Box-plot will have no fliers.
* Median will not change significantly
* Since mean is susceptible to outlier , mean will change sightly



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

* Mode of this dataset will lie between range [4,8]

1. Comment on the skewness of the dataset.

* 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.

* boxplot gives quartiles data points and also maximum/minimum values using which outliers can be found
* histogram gives more detailed underlying distribution of data points

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.)

* probability that at least one in five attempted telephone calls reaches the wrong number is 2.45%
* Refer ‘A 2] set 1.ipynb’ for calculations

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?

* Since x = 2000 have highest probability of 0.3 therefore
* $2000 is the most likely monetary outcome of business venture

1. Is the venture likely to be successful? Explain

* Since expected monetary outcome of business is positive , it can be concluded that

Venture is likely to be successful

* Refer ‘A 2] set 1.ipynb’ for calculations.

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

* Long-term average earning of venture = $800
* Long-term average is nothing but expected value of probability distribution.
* Refer ‘A 2] set 1.ipynb’ for calculations.

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

* Refer ‘A 2] set 1.ipynb’ for calculations.