**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 : refer to jupyter notebook Q1 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 (IQR) is calculated by the following equation**

**IQR = Q3-Q1**

**= 12-5 (approximate value)**

**= 7**

**The first quartile is the value below which 25% of the data falls, and the third quartile is the value below which 75% of the data falls.**

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

**Ans: We can see that one outlier is present at the upper extreme or right**

**side of the data so the data is rightly 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: The data point with value 25 is an outlier due to this the dataset is**

**rightly skewed if the value is 2.5 that means there is no outlier is the**

**dataset so we can say that the data is normally distributed**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans: Mode of the data set lie at the range of 5 and 6 because the bar has**

**the same frequency of 20**

1. Comment on the skewness of the dataset.

**Ans: The data is rightly skewed , most of the data point are clustered**

**towards the left of the dataset**

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: Both the dataset has the outlier which is present at the right side of**

**the dataset which means the data is rightly skewed most of the data point is present at the left side of dataset if the data single data point is not present we can say that the data is normally distributed.**

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 a call being misdirected is p=1/200

And the parameter n = 5

P(at least one misdirected call) = 1 - P(all calls correctly directed)

P(all calls correctly directed) = (199/200)^5

P(at least one misdirected call) = 1 - (199/200)^5

P(at least one misdirected call) ≈ 0.0244

The probability that at least one in five attempted telephone calls reaches the wrong number is approximately is 2.44 %

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: The most Likely monetary outcome of the business venture is $2000 which is**

**the maximum probability of 30 %**

1. Is the venture likely to be successful? Explain

**Ans: The positive return are (0.2+0.3+0.1) =0.6 i.e 60 % weather the negative return**

**are (0.1+0.1)=0.2 i.e 20% so the positive return are higher than the negative return.**

**Hence the Venture is successful**

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

**Ans: By x\*p(x)=(-2000\*0.1)+(-1000\*0.1)+ (1000\*0.2)+(2000\*0.3)+(3000\*0.1)**

**Long-term average earning of business venture is $ 800**

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

**Ans: Large value in the standard deviation of the variable x show that there is high**

**risk involved in the venture variance is 3.5 and Standard deviation is 1.87**