**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 : please check in Jupiter notebook



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 : The IQR (inter-quartile range) describes the middle 50% of values when ordered from lowest to highest. To find the interquartile range (IQR), ​first find the median (middle value) of the lower and upper half of the data. These values are quartile 1 (Q1) and quartile 3 (Q3). The IQR is the difference between Q3 and Q1

IQR = Q3 – Q1 = 12 – 5 = 7

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

Ans : The skewness of the dataset is positive. Right skewed data and mean of the data is greater than the median.

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 : If the data point value 25 is actually 2.5, then there will be no outlier in the dataset boxplot. Skewness of the data will be reduced and data distribution will become symmetric.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans:** The mode of the dataset lie in between approximately 4.5 to 8

1. Comment on the skewness of the dataset.

**Ans :** Right skewed data and skewness is positive. Mean of the data is greater than the median.

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:** Data distribution can be seen more clearly in histogram. Outliers can be detected in both histogram and boxplot. Skewness can be identified in both histogram and boxplot.

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: one in 200 long-distance telephone calls is misdirected

Probability of call misdirecting = 1/200

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

Number of Calls = 5

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

= 1 - ((199/200)\*(199/200)\*(199/200)\*(199/200)\*(199/200))

= 1 – 0.975248

= 0.02475

The probability that at least one in five attempted telephone calls reaches the wrong number is 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 : The most likely monetary outcome of the business venture  is 2000  $

as it has maximum probability = 0.3

1. Is the venture likely to be successful? Explain

Ans: Yes.

P(X>0)= 0.2+0.3+0.1=0.6

It implies, 60 % chances that the venture yield profits.

P(loss) = 0.1+0.1=0.2

It implies, 20% chances that the venture get losses.

So overall, we can say the venture likely to be is successful.

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

Ans : =(-2000\*0.1)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)

= (-200-100 +0+200+600+300)

= 800$

The long-term average earning of business ventures is 800$

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

Ans : P(loss) = P(x= -2000)+P(x=-1000)=0.2

So, the risk involved in this venture is 20%.