**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 pdf file assignment2,set 1- Q1**



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 value implies that 05 of data points lies between range of 5 to 12**

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

Ans 🡪 **The dataset is Positively skewed. Tail is extending on the right side of the curve.**

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 Median of the value will remain same but interquartile range will change to 25 .and there will not any outlier.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans 🡪 **Mode lies between 4 and 8.**

1. Comment on the skewness of the dataset.

Ans 🡪 **The dataset is positively skewed , the tail is extended towards Right side.**

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 🡪

**The histogram shows how much dataset is skewed and how higher the peak of the curve is.**

**The boxplot show the outlier of the dataset. It also shows where most of the data points lies,**

**We can understand the distribution of mean and median.**

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 wrong call = 1/200 = 0.005**

**Probability of not wrong call = 1-0.005 = 0.995**

**Probability of at least one out of five is a wrong number**

**= 1- Probability of all five calls are not wrong numbers**

**= 1- 0.995^5**

**= 0.0247**

**=2.5%**

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 🡪 **$2000 as it has the highest probability of occurrence**

1. Is the venture likely to be successful? Explain

Ans 🡪 **Yes, as positive values i.e. 0.3 +0.2 + 0.1= 0.6**

**0.6\*100 = 60%**

**So we can say there is 60% chance that venture will be**

**successful.**

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

Ans 🡪 **Long term average earning =**

**(-2000\*0.1)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000 \*0.3)+(3000\*0.1)=800**

**The long-term average earning for these type of ventures would be around = $800**

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

Ans 🡪 **A good measure to evaluate the risk would be variance and**

**standard deviation of the variable x**

**Var = 3500000**

**Sd = 1870.83**

**The large value of standard deviation of $1870 is considered along with the average returns of $800 indicates that this venture is highly risky**