**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% |

Answer: From the given data, Morgan Stanley with 91.36% is the outlier.

Mean= 33.271

Standard deviation = 16.945

Variance = 287.14



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

Answer: From the above picture the inter-quartile range is ranging from 5 to 12. It tells how to spread out the middle values and find out the outliers from that.

.

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

Answer: From the figure it shows dataset is right skewed hence, it is positive 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?

Answer: Then, the new data point would be in between 0 to 5 which is lower quartile. And there would be no presence of outliers. And also there would be change in mean value.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Answer: The mode of this dataset would lie in between 4 and 7 data points.

1. Comment on the skewness of the dataset.

Answer: From the above picture, the right side distribution is longer. So, the skewness of the dataset is positive.

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.

Answer: Distribution of data set in both graphs would be good but the boxplot would show the distribution of data with more detailed information. In both graphs the skewness would be positive as the distribution would be right-skewed. And presence of outliers. Thus, histograms and boxplots are the graphical representations for the frequency of numeric data values which aim to describe the data and explore the central tendency and variability. Histograms are used to find the underlying probability too.

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.)

Answer: Probability that at least one in 5 attempted call reaches the wrong number is 0.025.

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?

Answer: The most likely monetary outcome of the business venture is 2000$.

1. Is the venture likely to be successful? Explain

Answer: Long term average = (-2000\*0.1) + (-1000\*0.1) + (0) + (1000\*0.2) + (2000\*0.3) + (3000\*0.1) = 800$

Thus, the venture is likely to be successful.

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

Answer: 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

Answer: Standard deviation is a good choice which is 1469.693 for the given data.