**Topics: Descriptive Statistics and Probability**

1. Look at the data given below. Plot the data, find the outliers and find out

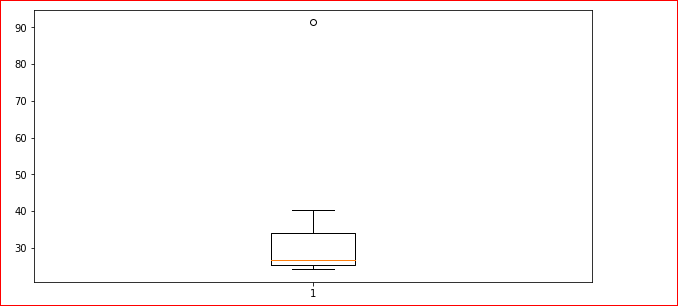
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft | 32.95% |
| Morgan Stanley | 91.36% |
| Sun Microsystems | 25.99% |
| Travelers | 39.42% |
| US Airways | 26.71% |
| Warner-Lambert | 35.00% |

Mean=33.271333

Standard Deviation= 16.945401

Variance= 268.00350488889

X= 91.36% is the outlier in the dataset corresponding to the Company name Morgan Stanley





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.

-(5-12) The Inter-quartile range always represents 50% of our data i.e. 50% of our data lie between 5-12

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

The dataset is Right Skewed or Positive Skewed since majority of data is concentrated between median and Quartile 3 out of 4 Quartiles.

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?

The new boxplot will not have any outliers and the range of values of data in Q1 quartile will increase causing the length of Whisker to increase a bit i.e. data is bit spread out in Q1 than it was earlier.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

-Somewhere between 4 and 8.

1. Comment on the skewness of the dataset.

-Right-Skewed Data or Positive Skewed 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.

-Both Histogram and Boxplot has its own Advantages and Disadvantages. If Histogram and Boxplot are drawn for the Same Dataset, then

1.Boxplot doesn't give any idea about the shape of symmetry while Histogram does.

2.Boxplot helps in identifying actual measures of spread and center directly from the boxplot, while histogram doesn't.

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.)

-0.02475 or 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?

-x=2000 cause P(x=2000)=0.3(Max.)

1. Is the venture likely to be successful? Explain

-Yes, Because the probability of this kind of Business venture to end up in profit is greater than 0.5 i.e. =0.6

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

-−2000 ∗ 0.10 − 1000 ∗ 0.10 + 0 \*0.2+ 1000 ∗ 0.20 + 2000 ∗ 0.30+ 3000 ∗ 0.10= 800

-$800

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

-Standard Deviation: $2160.The Standard Deviation value is much far from the mean value which warns that there’s a risk involved in this kind of venture