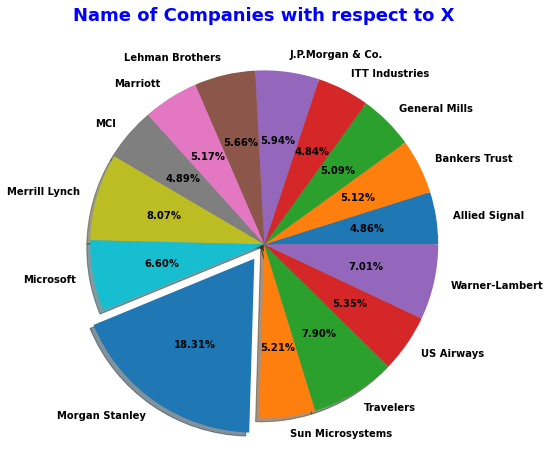
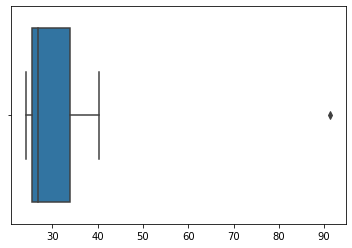
**Topics: Descriptive Statistics and Probability**

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% |





The following is the outlier in the boxplot: Morgan Stanley 91.36%

measure\_x.describe()

Mean = 33.271333

Standard deviation = 16.945401

measure\_x.var()

Variance = 287.1466123809524



1. What is inter-quartile range of this dataset? (please approximate the numbers) In one line, explain what this value implies.

Ans: Approximately (First Quantile Range) Q1 = 5 (Third Quantile Range) Q3 = 12, Median (Second Quartile Range) = 7

(Inter-Quartile Range) IQR = Q3 – Q1 = 12 – 5 = 7

Second Quartile Range is the Median Value

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

Ans: Right-Skewed median is towards the left side it is not normal distribution

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: In that case there would be no Outliers on the given dataset because of the outlier the data had positive skewness it will reduce and the data will normal distributed



1. Where would the mode of this dataset lie?

Ans: The mode of this data set lie in between 5 to 10 and approximately between 4 to 8 .

1. Comment on the skewness of the dataset.

Ans: Right-Skewed. Mean>Median>Mode

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: They both are right-skewed and both have outliers the median can be easily visualized

in box plot where as in histogram mode is more visible