**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% |

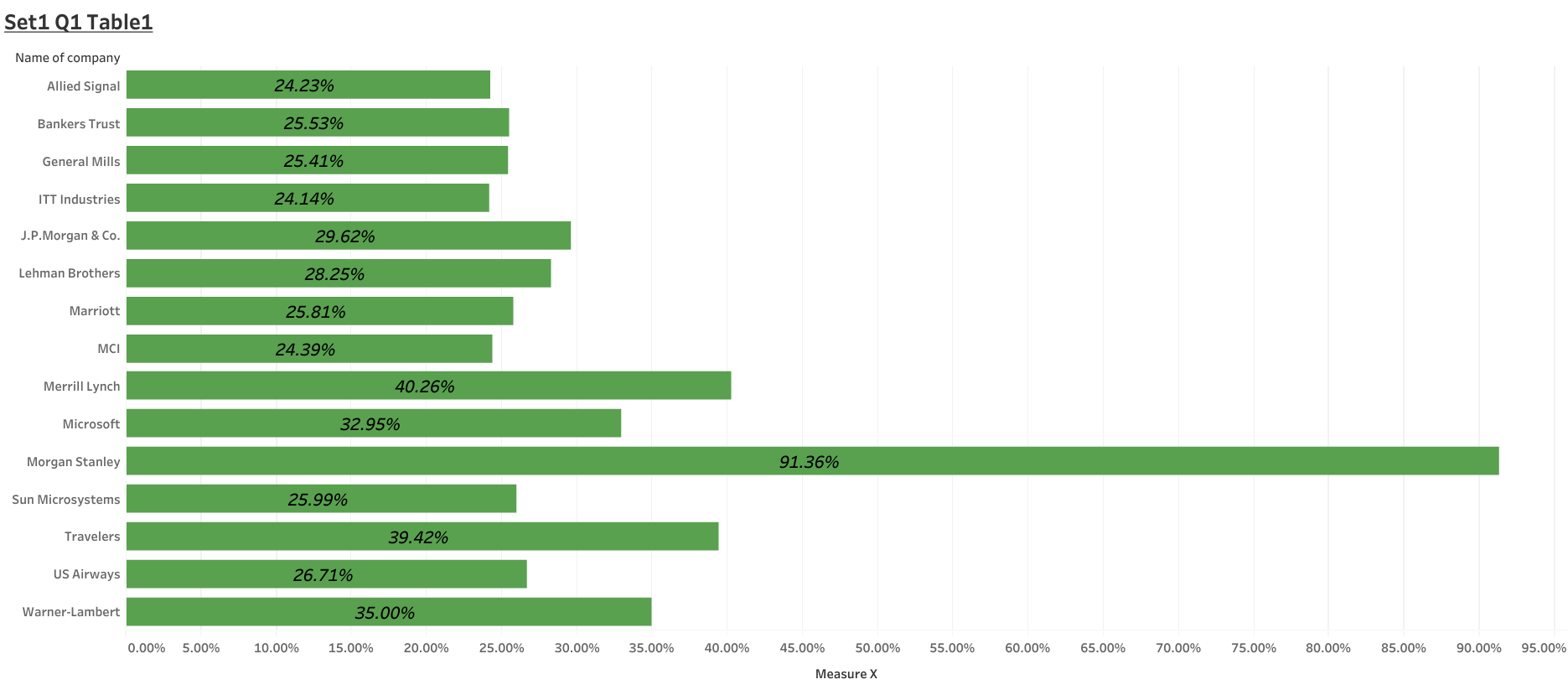
**Ans :- Outliers ( 91.36 )**

**Mean ( 33.271 )**

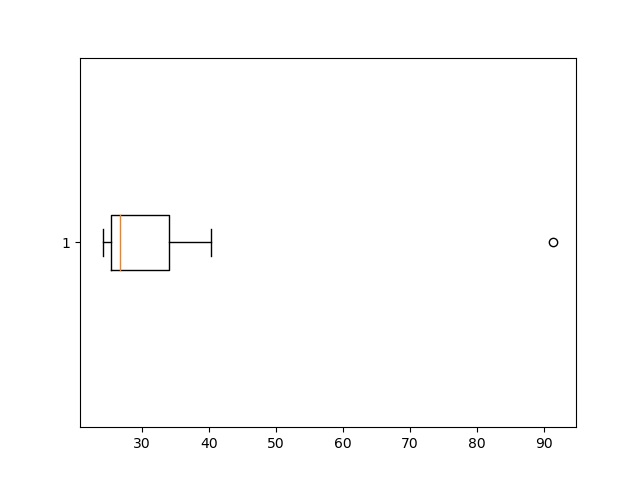
**Standard Deviation ( 16.945 )**

**Variance ( 287.147 )**

**Bar Chart :-**



**Boxplot :-**





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 :- Interquartile range (IQR) = 12-5 = 7**

**This is implies that the middle 50% of the data lies between 5 and 12.**

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

**Ans :- The dataset is right skewed, data is more concentrate towards the left and right tail is**

**longer.**

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 :- 2.5 will be not considered an outlier. The boxplot will start from 0 and send at 20 in representation.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?
2. Comment on the skewness of the dataset.
3. 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 :-**

**1. The mode of this dataset lie will be 4 to 8.**

**2 . The data is right skewed ,which means that the data is more concentrated towards the left**

**and right tail is longer.**

**3 . Median in boxplot and mode in histogram.**

**Histogram provides the frequency of distribution ,so we can identify the how many times data**

**point occurring however boxplot provides quantile distribution.**

**Boxplot provides whisker length to identify outliers ,no information from histogram.**

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 :-**

**The probability of call being misdirected = 1/200 = 0.005**

**The probability of call not being misdirected = 1 - 0.005 = 0.995**

**The probability that at least one in five attempted telephone calls reaches the wrong number = 1 - Probability of all calls reaching the correct number**

**= 1 - ((0.995)^5)**

**= 0.0248**

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 :- Maximum Probability = 0.3 for P(2000). So most likely outcome is 2000**

1. Is the venture likely to be successful? Explain

**Ans :- The venture is likely to be successful because the probability of the venture having a return on investment is P = 0.6 (0.2+0.3+0.1).**

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

**Ans :- Weighted average = x\*P(x) = 800. This means the average expected earnings over a long period of time would be 800 (including all losses and gains over the period of time)**

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 associated with this venture is 20%.**

|  |  |  |
| --- | --- | --- |
| **x** | **P(x)** | **E(x)** |
| -2000 | 0.1 | -200.0 |
| -1000 | 0.1 | -100.0 |
| 0 | 0.2 | 0.0 |
| 1000 | 0.2 | 200.0 |
| 2000 | 0.3 | 600.0 |
| 3000 | 0.1 | 300.0 |