**Descriptive Statistics and Probability**

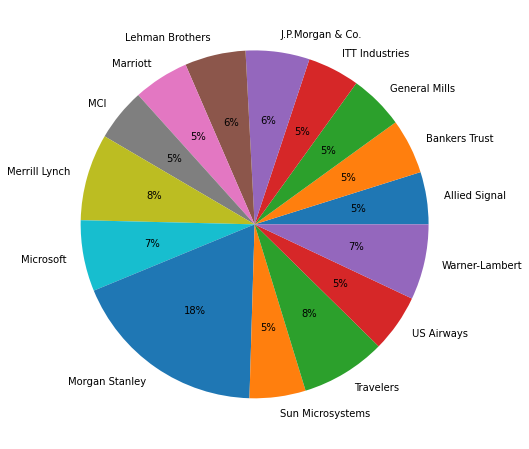
1. **Look at the data given below. Plot the data, find the outliers and find out .**

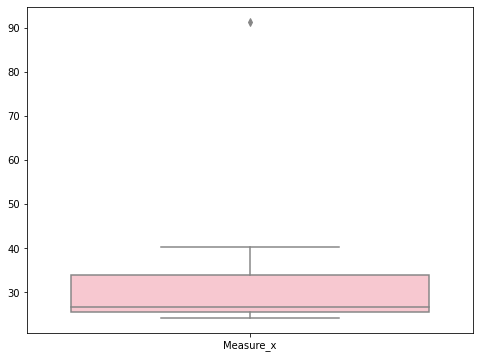
| **Name of company** | **Measure X** |
| --- | --- |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft |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

**Mean is 33.27**

**Variance is 287.14**

**Standard deviation is 16.945**

****

****

**From the boxplot it is clear that it has outliers.**

****

**Answer the following three questions based on the box-plot above.**

1. **What is interquartile range of this dataset? (please approximate the numbers) In one line, explain what this value implies.**
2. **What can we say about the skewness of this dataset?**
3. **If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?**

**IQR = R3 - R1 = 12 - 5 = 7**

**The data is Right skewed(+ve) as the box to the right is greater.**

**So there won’t be any affected only the values will get converted.**

****

**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.**

**The mode will lie in between 4 to 7.**

**The data is positively skewed as more data points are on the left side.**

**From the histogram and boxplot we can conclude that as the value of Y increases the frequency decreases, there is an outlier , most of the data points are in between 4 to 10.**

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.)**

**The probability of call misdirecting = 1/200**

**The probability of call not misdirecting = 199/200**

**Total number of calls = 5**

**LET:**

**at least one in five attempted telephone calls reaches the wrong number**

**= 1  -  none of the call reaches the wrong number**

**= 1  - P(0)**

**= 1   -  ⁵C₀(1/200)⁰(199/200)⁵⁻⁰**

**= 1  -  (199/200)⁵**

**= 0.02475**

**The probability that at least one in five attempted telephone calls reaches the wrong number is 2%.**

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?**
2. **Is the venture likely to be successful? Explain**
3. **What is the long-term average earning of business ventures of this kind? Explain**
4. **What is the good measure of the risk involved in a venture of this kind? Compute this measure**

**1) Most likely monetary outcome of the business venture  is 2000  $ as it has maximum probability = 0.3**

**2)Yes the venture is likely to be successful.**

**Because the expected value = ∑E(X)P(X)  = 800**

**3) long-term average earning of business ventures  = 800 $**

**4)The standard deviation for a probability distribution is**

**= ∑(x – μ)2 P(x)**

**For the given probability distribution the mean is 800**

**So ([-2000-800] ^2 [0.0] +…+[3000-800]^2 [0.1])**

**The Standard deviation is 0.08164966.**

**The risk involved in this venture is 8%.**