**Topics: Descriptive Statistics and Probability**

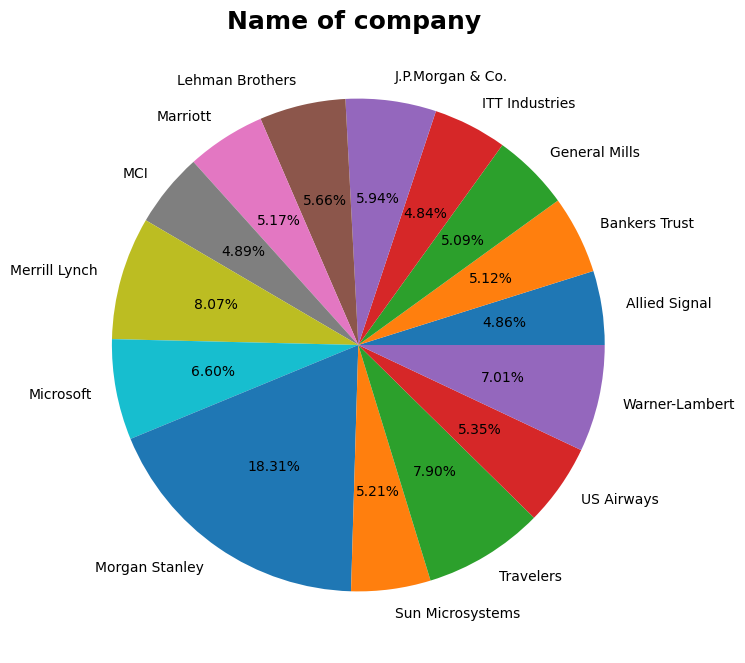
1. Look at the data given below. Plot the data, find the outliers and find out

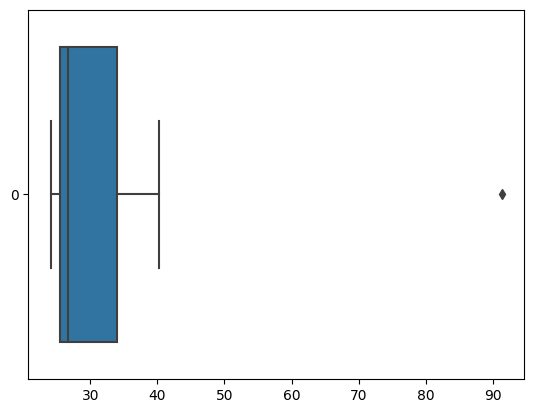
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft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

Ans: mean = 33.2713

Standard deviation = 16.9454

Variance = 287.1466







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: Q1(First Quantile Range) = 5

Q2 (Second Quartile Range) Median =7

Q3 (Third Quantile Range) = 12

IQR (Inter-Quartile Range) = Q3 – Q1

= 12 – 5

= 7

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

Ans: data is right 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?

Ans: 2.5 won't be regarded as an outlier. In representation, the boxplot will begin at 0 and end at 20.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: Mode lies between 4 and 8.

1. Comment on the skewness of the dataset.

Ans: Data set is Right-Skewed. (i.e 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: Mode in the histogram and Median in the boxplot The boxplot displays the quantile distribution, which indicates that 50% of the data falls between 5 and 12, while the histogram shows the frequency distribution, which allows us to determine how frequently each data point occurs. Whisker length is provided by a boxplot rather than information from a histogram to identify outliers. Considering the gap, our best guess is that number 25 might be an anomaly.

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: probability of call getting misdirected = 1/200

Hence probability of calls not getting misdirection =1 – 1/200 = 199/200

Number of phone calls not attempted = 5

So, probability that at least one in 5 attempted calls reaches the wrong number

=

=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?

Ans: Maximum P for P(2000) is 0.3. Thus, 2000 is the most likely result.

1. Is the venture likely to be successful? Explain

Ans: P(x>0) = 0.6 indicates that there is a 60% probability that the business will be profitable or generate higher than anticipated returns. Just 0.2 P(Incurring losses) exists. Thus, the project has a good chance of succeeding.

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

Ans: Weighted average = x\*P(x) = 800.

This indicates that, assuming all losses and gains over the course of the period, the average expected earnings over an extended period of time would be 800.

iv) 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. Thus, there is a 20% risk involved in this endeavor.