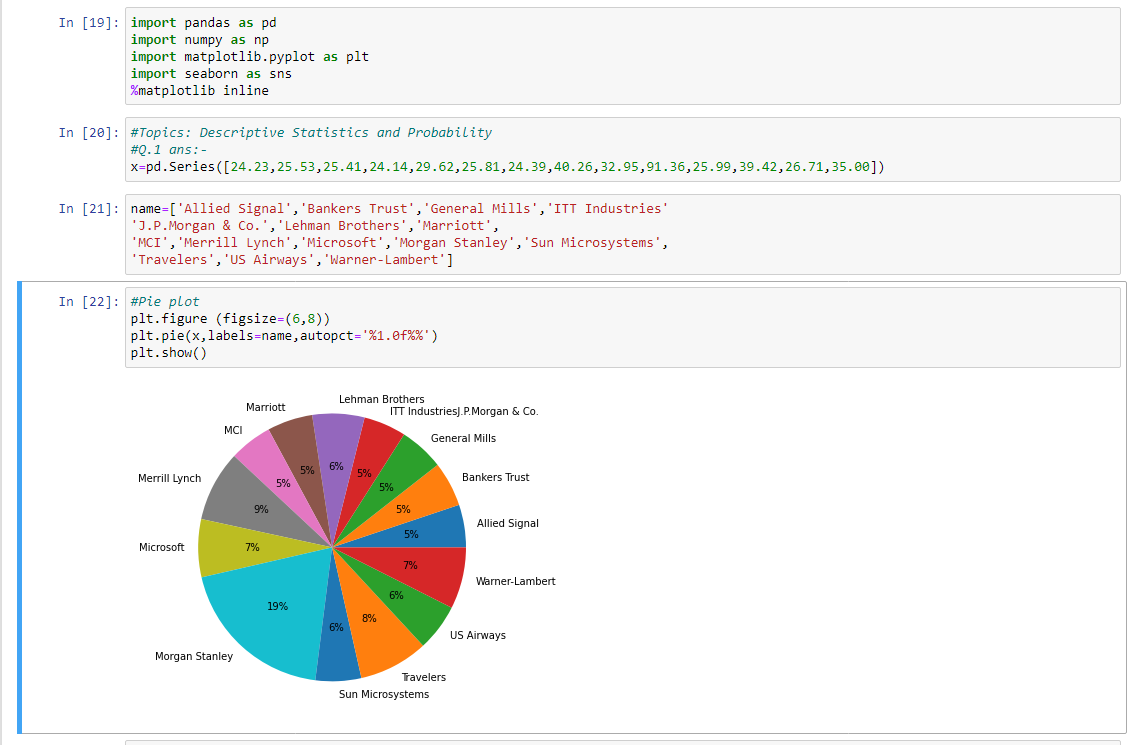
**Topics: Descriptive Statistics and Probability**

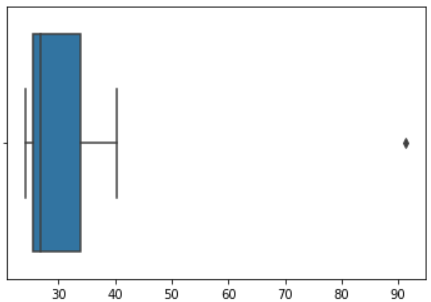
1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **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:-



Box plot to find outliers:-



Inference: There is One Outlier: Morgan Stanley at 91.36%

mean(x)= μ = 33.27133333333333

variance(x)= 𝞼 = 287.1466123809524

standard Deviation(x)= 𝞼2 =16.945400921222028



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:- :- IQR = 12-5 = 7, This represents the range which contains 50% of the data points.

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

Ans :- Right and positive 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 will be not considered an outlier. The boxplot will start from 0 and end at 20 in representation.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans:- Mode lies between 4 to 8

1. Comment on the skewness of the dataset.

Ans:- This data set is positively skewed.

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:- As show in Box plot, Median is shown in boxplot.

As show in Histogram, Mode is shown in Histogram.

Histogram provides the frequency distribution so we can see how many times each data point is occurring.

However, boxplot provides the quantile distribution. i.e. 50% data lies between 5 and 12.

Boxplot provide whisker length to identify outliers, no information from histogram. 25 may be an outlier.

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:- 1 in 200 long-distance telephone calls is misdirected.

Number of Calls(n) = 5, p = 1/200, q = 199/200

P(x) = at least one in five attempted telephone calls reaches the wrong number

P(x) = ⁿCₓ pˣ qⁿ⁻ ˣ

P(x) = (nCx) (p^x) (q^n-x)

P(x) = (5C1) (1/200)^1 (199/200)^5-1

P(x) = 0.0245037

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:- Max probability is 0.3 for x= 2000. So most likely outcome is 2000.

1. Is the venture likely to be successful? Explain

Ans:- P(x>0) = 0.6, implies there is a 60% chance that the venture would yield profits or greater than expected returns. P(Incurring losses) is only 0.2.So the venture is likely to be successful.

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(Incurring loss) = P(x= -2000)+P(x= -1000)=0.2.So the risk associated with this venture is 20