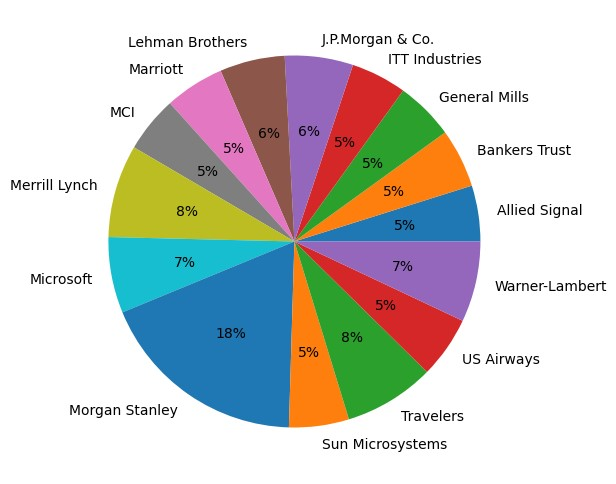
**Topics: Descriptive Statistics and Probability**

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

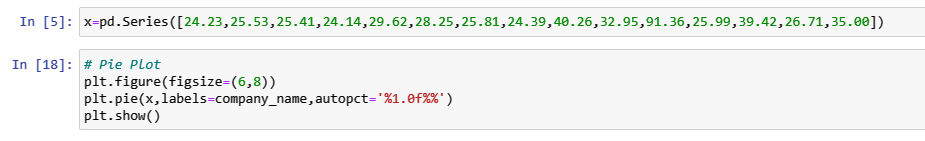
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft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

**Ans:**

**Piechart:**

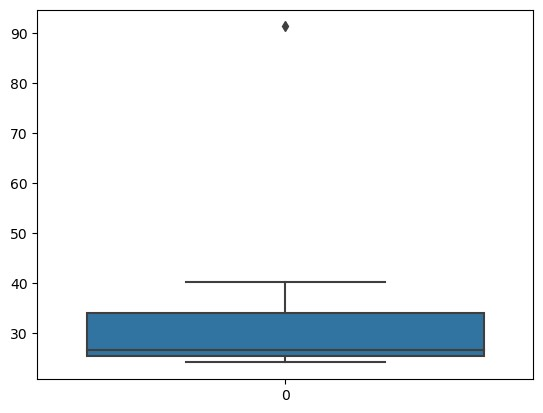






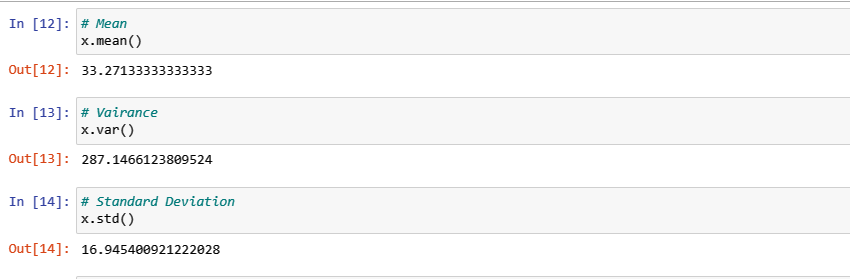
**Outliers:**





There is one outlier ‘Morgan Stanley’.

**:**



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:** Here from the above boxplot it is clearly seen 25 is the outlier.

Median = 7

Q1=1st quartile = 5

Q2=2nd quartile = 12

IQR = Q3-Q1=12-5 = 7

IQR gives the range of the middle half of the data.

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

**Ans:** Data set is **positively skewed** or right skewed since the median is closer to the bottom of

the box and the whisker is shorter on the lower end of the box.

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 have been no outliers and the values of mean and median

might change slightly.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans:** The mode of this dataset lie between values 4 – 8 since they are most frequently

observed values.

1. Comment on the skewness of the dataset.

**Ans:** The dataset 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:** Histogram as well as boxplot shows the dataset as positively skewed. Also they would

complement in finding approximate values of mean, median.

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:** Let the event of attempted call reaches the wrong number is considered a success.

Let X represent the number of calls reach the wrong number.

Probability of call getting misdirected = p=1/200 =0.005

Therefore probability of call not getting misdirected =q= 1- 0.005= 0.995.

Number of phone calls attempted = n=5

The probability of success is p = 0.005 and the probability of failure is q = 0.995

(where q = 1-p).

Here X follows Binomial Distribution.

Therefore, probability that at least one in 5 attempted call reaches the wrong

number is:

P(X>=1)= 1-P(X<1)

= 1-P(X=0)

= 1- (5C0)\*(0.005)^0\*(0.995)^(5-0)

= 1-0.9753

= 0.024

Therefore, the probability that at least one in five attempted telephone calls reaches the wrong number is 0.024.

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:** Here the highest probability is 0.3 and it corresponds to x= 2000. Hence, 2000 is the

most likely monetary outcome of the business venture.

1. Is the venture likely to be successful? Explain

**Ans:** Yes, because the total earnings of the venture is positive in value .

Here total negative returns are P(x=-2000)+P(x=-1000)+P(x=0) =0.1+0.1+0.2=0.4

and total positive returns are P(x=1000)+ P(x=2000)+ P(x=3000)=0.2+0.3+0.1 =0.6.

Here positive returns are are greater than negative returns. So venture is likely to be

successful.

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

**Ans:** the expected returns to the venture is considered as the required average.

Therefore Expected Returns= E(X)= ∑x\*P(x)

|  |  |  |
| --- | --- | --- |
| x | P(x) | X\*P(x) |
| -2,000 | 0.1 | -200 |
| -1,000 | 0.1 | -100 |
| 0 | 0.2 | 0 |
| 1000 | 0.2 | 200 |
| 2000 | 0.3 | 600 |
| 3000 | 0.1 | 300 |

E(X)= -200-100+0+200+600+300= 800

800 is the average outcome of business venture.

1. What is the good measure of the risk involved in a venture of this kind? Compute this measure

**Ans:** Standard deviation is the good measure of the risk involved in a venture of this kind.

Var (X) = E(X^2) –(E(X))^2 = 2800000 – 800^2 = 2160000

Standard Deviation= 1469.694