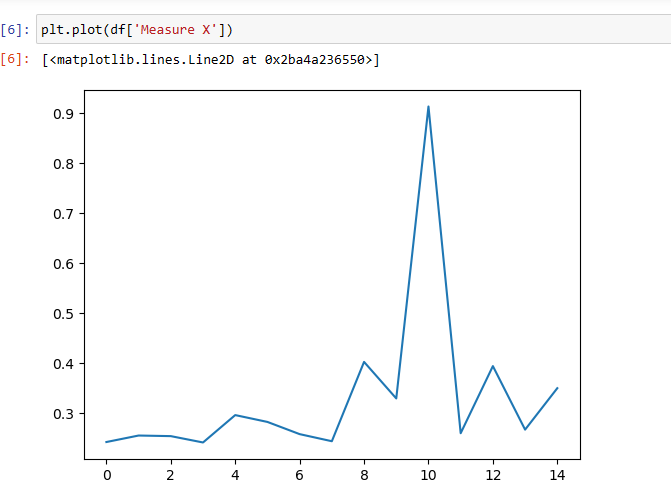
**Topics: Descriptive Statistics and Probability**

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% |

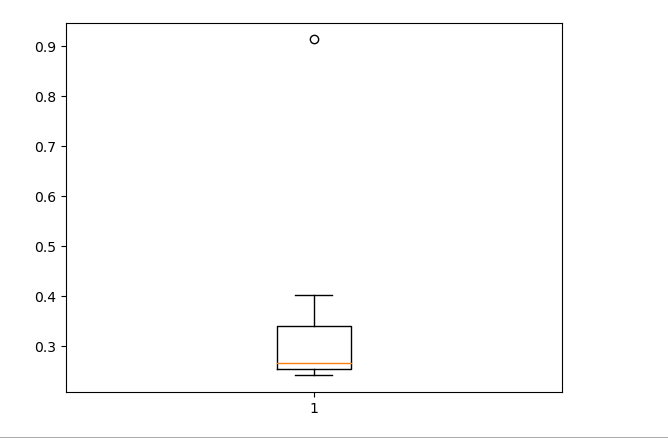
ANS:

**Plotting line graph** -

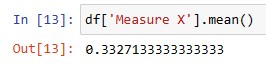


**Outliers:**

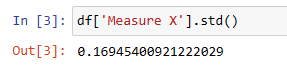
The below boxplot shows the outlier, only one outlier present at the upper extreme.



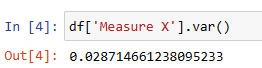
Mean = (24.23+25.53+25.41+24.14+29.62+28.25+25.81+24.39+40.26+32.95+91.36+25.99+39.42+26.71+35.00)/3 = 33.27



Std =



Variance =



1. 

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.
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?

ANS:

1. Q1 (lower quartile) =5 and Q3(Upper quartile) =12

Inter-quartile range = Q3 – Q1 = 12 – 5 = 8

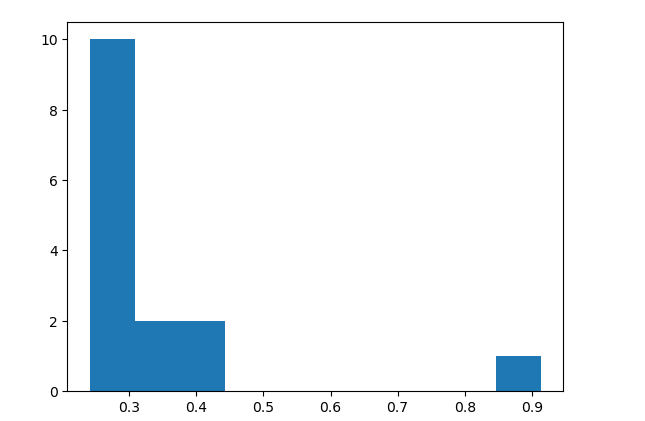
Inter-quartile range measures the spread of the middle half of your data.

1. The skewness is positive(right) skewed, the outlier present in the upper extreme.
2. If it was found that the data point with the value 25 is actually 2.5, then the data point will be present below the lower extreme value, it will be considered as outlier.
3. 

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. Mode value is most frequent values present in the particular column (Y), according to the above graph, the mode will be 5.
2. Skewness is positive or right skew.
3. question 2 are plotted for the same dataset: 

From the above graph we can say that it is positively skewed.

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:

One in 200 long-distance telephone calls is misdirect probability of call misdirecting  p = 1/200

      Probability of call not Misdirecting = 1 - 1/200 = 199/200

Number of Calls = 5

P(x) = ⁿCₓpˣqⁿ⁻ˣ , n = 5,p = 1/200,q = 199/200

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

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

Ans:

1.What is the most likely monetary outcome of the business venture?

As the probability (0.3) is more for 2000 $ as compared to others,

Therefore, most likely monetary outcome of the business venture = 2000$

2. Is the venture likely to be successful? Explain

Long term average = \sum{P(xi)\*Xi} = (-2000\*0.1) +(-1000\*0.1) +(0) +(1000\*0.2) +(2000\*0.3) +(3000\*0.1) = 800$

As the long-term average gives positive numbers the Business venture likely to be successful.

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

Long term average = \sum{P(xi)\*Xi} = (-2000\*0.1) +(-1000\*0.1) +(0) +(1000\*0.2) +(2000\*0.3) +(3000\*0.1) = 800$

Means on an average Return will be 800 $

4. Risk involved in a venture

Var (X) = E(X²)  - { E(X) }²

=   2800000 -   800²

= **2160000**  ( Quite High)

SD = √Var  ≈ **$ 1470**

As **Variability is Quite high**  hence **Risk is high.**