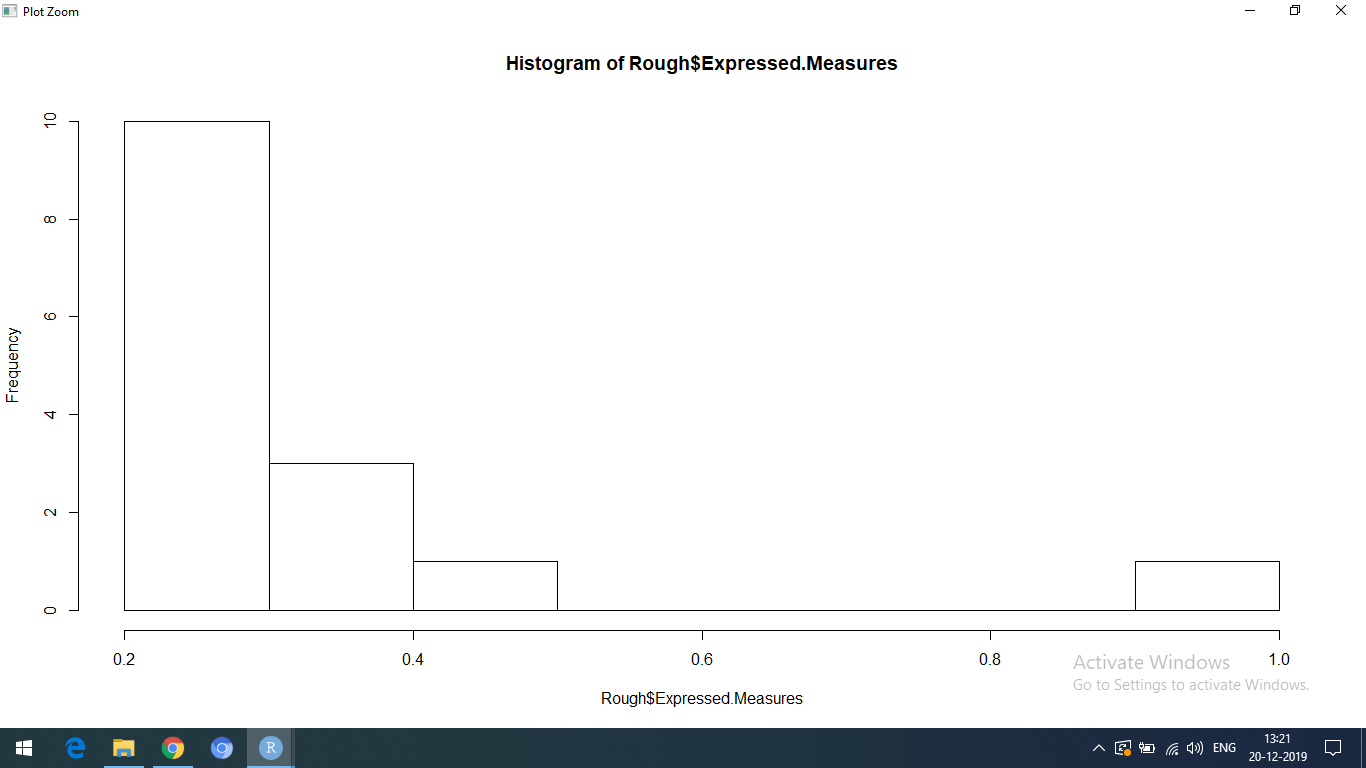
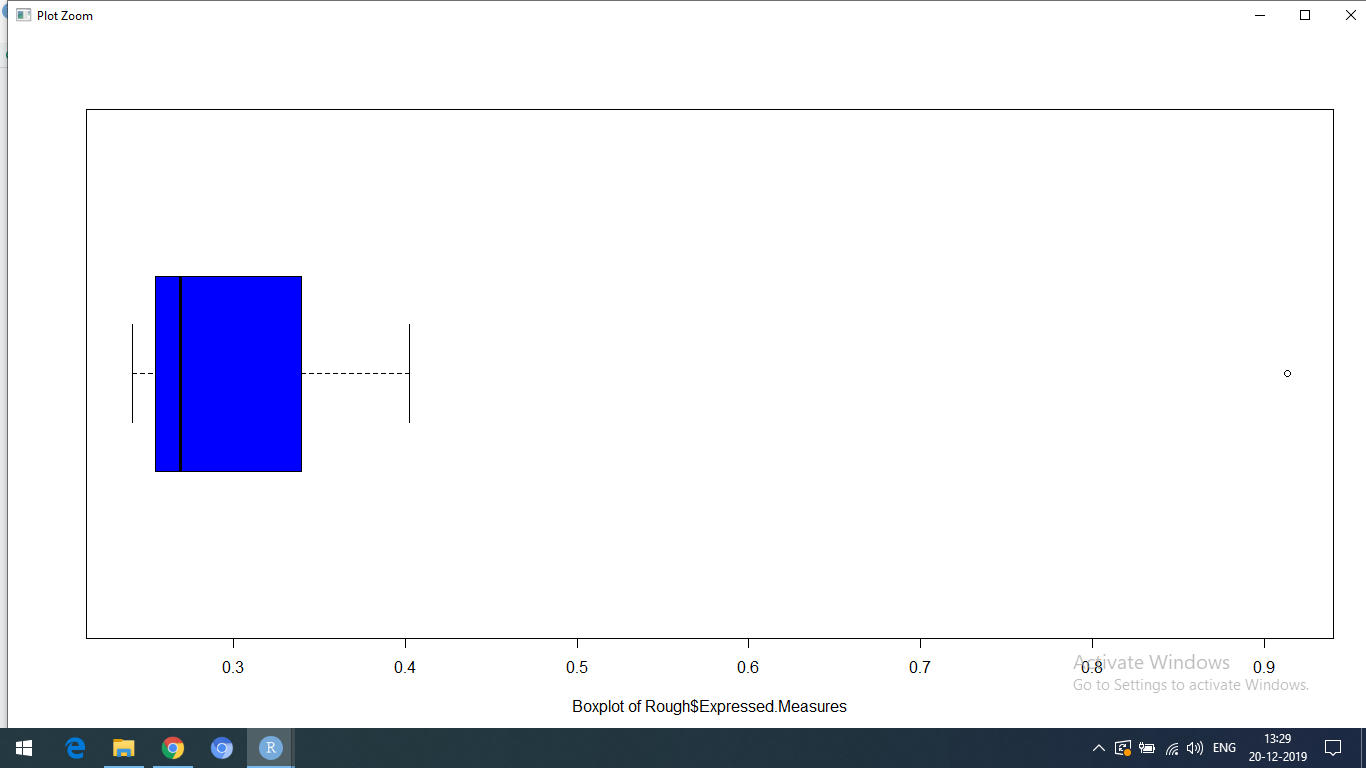
**Topics: Descriptive Statistics and Probability**

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% |

Plotting of data using Histogram and Boxplot to have a clear inference regarding outliers.





Ans: In this we have a single outlier i.e., for Morgan Stanley and it is 91.36%.

|  |  |
| --- | --- |
| Mean | 33.28%  > mean(Rough$Expressed.Measures)  [1] 0.3328467 |
| Variance | 0.0286961812  > var(Rough$Expressed.Measures)  [1] 0.02869618 |
| Standard Dev | 0.1693994724  > sd(Rough$Expressed.Measures)  [1] 0.1693995 |



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: The IQR of this dataset ranges between 5 to 12. This value implies that the range between Upper quartile(Q3) and lower quartile(Q1).

IQR = Q3 - Q1 = 12 - 5 = 7

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

Ans: The skewness of this dataset implies that it is right skewed(Positive skewness).

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: The new boxplot would be affected by not having any outliers in the data and will have continuous distribution of data points.



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 lies between 4 to 8.

1. Comment on the skewness of the dataset.

Ans: Even the following Histogram represents a right skewed distribution.

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: Basically the Histogram represents the data in interval format and so it becomes hard to identify the outliers whereas in the Boxplot the outliers can be identified and even the median can be easily identified.

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 an event P(E) = 1 / 200 = 0.005

Probability of error, P(Ē) = 1 - P(E) = 1 - (1/200) = 199/200

Probability that atleast one in five attempted telephone calls reaches the wrong number will be

= 1 - (199/200)^5 = 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: The mostly likely monetary outcome of business venture will be when p(x) is 0.3 and x=2000.

1. Is the venture likely to be successful? Explain

Ans: Yes because the probabilties of positive sides of x is more than the negative sides of x and hence it is likely to be successful.

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

Ans: The long average earning of this business venture will be equal to expected value of x and p(x)

= -2000(0.1) - 1000(0.1) + 0(0.2) + 1000(0.2) + 2000(0.3) + 3000(0.1)

= 800

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

Ans: We can measure mean using the random variables(x) and other thing which we can find and which is a good measure involved in such a venture would be finding out the Variance.