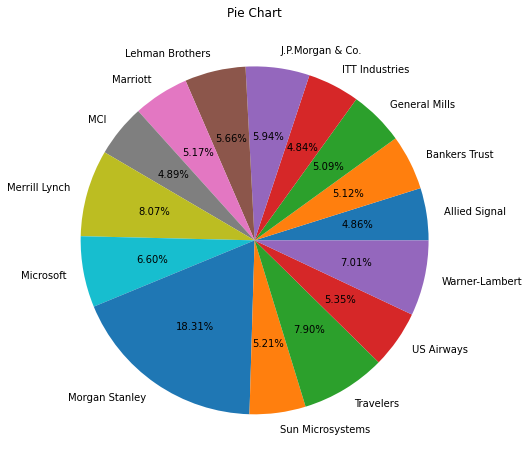
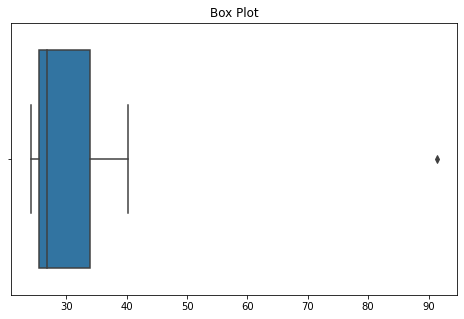
**Topics: Descriptive Statistics and Probability**

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

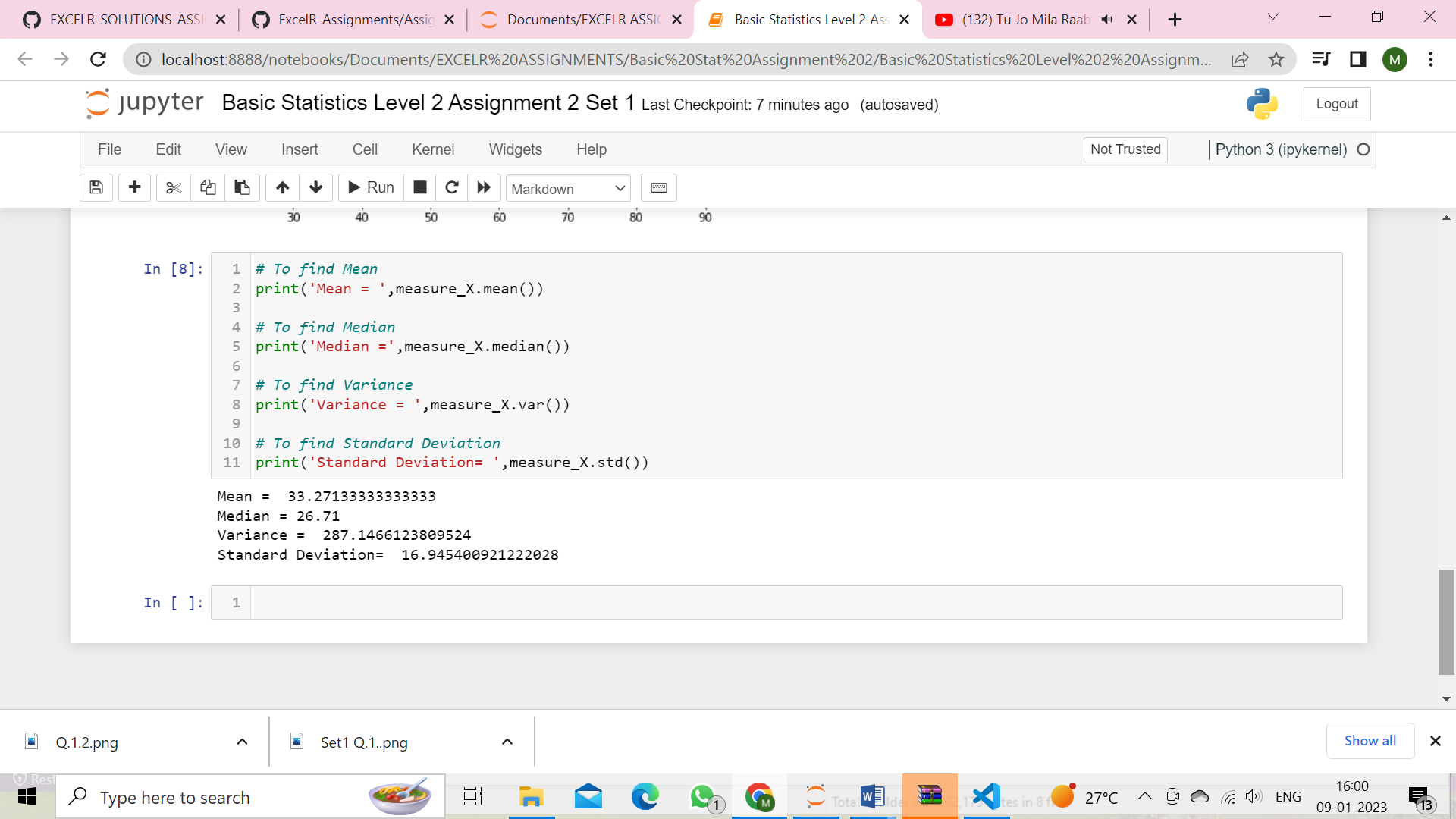
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft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

**Solution: #For code section refer Basic Statistics Level 2 Assignment 2 Set 1.ipynb file.**





There is Outlier in our data.





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.

**Answer:** Outlier=25

Median=7

First Quantile(Q1)= 5,

Third Quantile (Q3 )= 12

IQR = Q3-Q1

= 12 – 5 = **7**

IQR tells the range of the middle 50% values of the data.

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

**Answer:** From the above boxplot we can say that it is Positively 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?

**Answer:** If it was found that the data point with the value 25 is actually 2.5 then there will be no outliers, due to this the data will normal distributed and the new box-plot might have moved towards right i.e. positively skewed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Answer:** Themode of this dataset lie in between 5-7.

1. Comment on the skewness of the dataset.

**Answer:** The skewness of 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.

**Answer:** By comparing both of them, it is clear that the data is positively skewed and has outliers . The median can be easily visualized in box-plot where as mode is more visible in histogram.

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.)

**Answer:**

Probability of call misdirecting (p) = 1/200

Probability of call not Misdirecting( q )= 1-1/200

q = 199/200

Number of Calls (n)= 5

P(x) = ⁿCₓpˣqⁿ⁻ˣ

Probability that 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?

**Answer:** The most likely monetary outcome of the business venture is 2000 because it

has highest probability of 0.3.

1. Is the venture likely to be successful? Explain

**Answer:** The probability distribution gives us an idea about the long-term chances of

earning given values of returns (indicated by x) as Venture has positive profit.

P(x>1000)+P(x>2000)+P(x>3000) = 0.2+0.3+0.1 = **0.6**

Therefore, there is a 60% probability that the venture would be successful.

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

**Answer:** The long-term average earning of business ventures of this kind is

= ∑ ( x \* P(x)) = 800.

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

**Answer:** Risk stems from the possible variability in the expected returns. Therefore, good measure to evaluate the risk for a venture of this kind would be variance or standard deviation of the variable X.

Here,

SD= 1870 and Var=3500000

The large value of standard deviation of $1870 is considered along with the average returns of $800 indicates that this venture is highly risky.