**Topics: Descriptive Statistics and Probability**

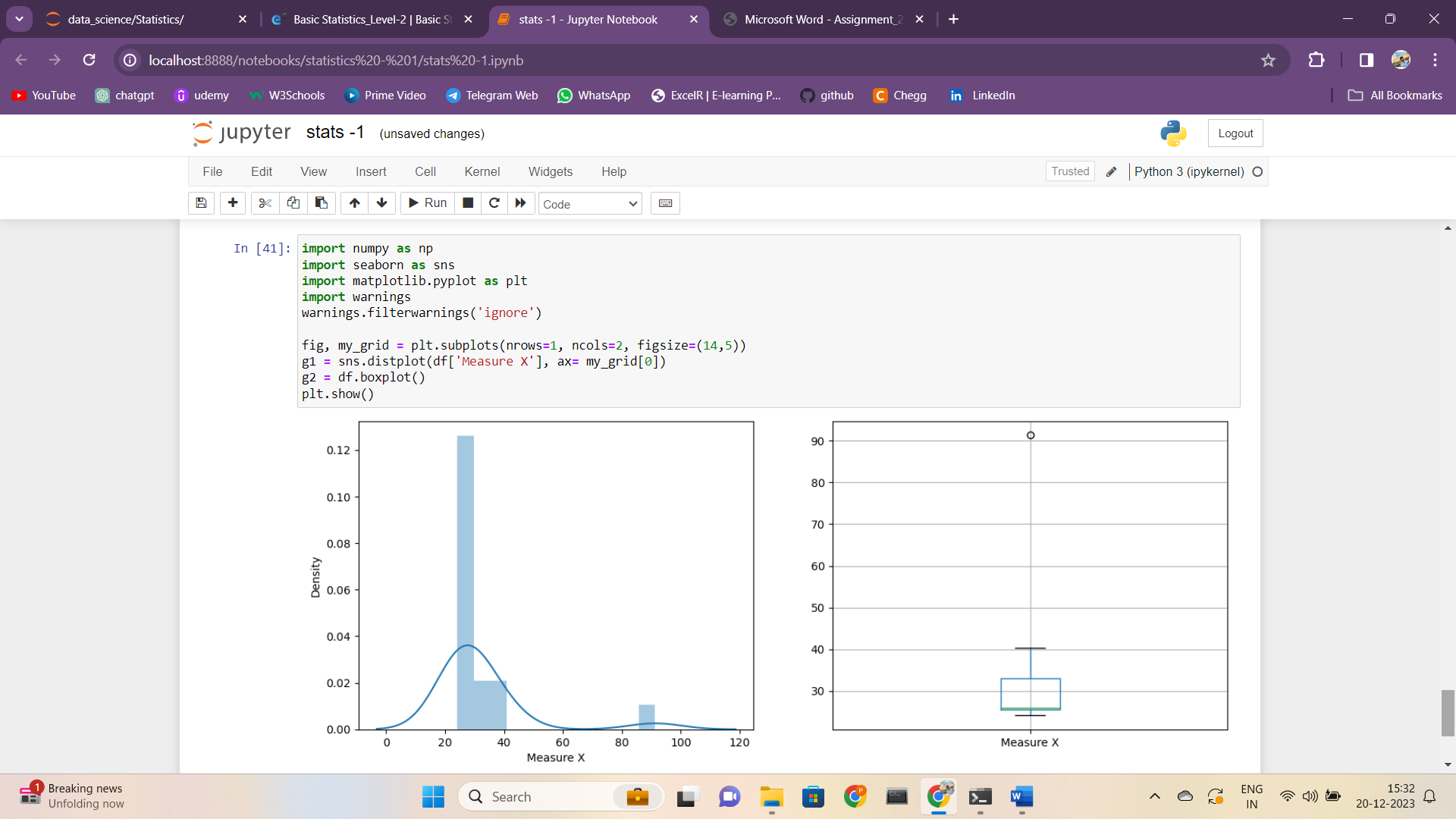
1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **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% |

**Sol:-**

A screenshot of a computer

Description automatically generated



From the above distribution plot and Boxplot, we can easily detect the

outliers.

Morgan Stanley-91.36 is the outlier.

Mean, Standard Deviation and Variance are as follows:

A screenshot of a computer

Description automatically generated



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.

**Solution**: IQR = Upper Quartile-Lower Quartile=12-5=7

This value implies that 50% of the data lies in this IQR , and 1.5 IQR from the upper and lower quartile denotes upper and lower extreme points.

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

**Solution**: This given dataset is Right Skewed as the size of the plot towards the

right side of the mean is more.

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?

**Solution**: This change in value will affect the mean in a way that it reduces , and the median also reduces.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Solution**: This dataset is multimodal, mode is at 4 to 6 values of Y

1. Comment on the skewness of the dataset.

**Solution:** This is Right 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.

**Solution:** We can find the outlier and the kind of skewness caused in the dataset from both histogram and boxplot.

• Also the histogram gives the frequency of the distribution of values of Y and the box plot provides the IQR and the Whiskers.

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.)

**Solution:**

The total no long-distance calls = 200

Probability of call being misdirected= 1/200

Sample calls = 5

Probability of at least one in five attempted calls reaches wrong number =

5\*(1/200) = 0.025 = 2.5%

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?

**Solution:** The value of x with highest probability= 2000

1. Is the venture likely to be successful? Explain

**Solution:** Probability of the venture to fail = 0.1+0.1 = 0.2

Probability of the venture to neither fail nor be successful = 0.2

Probability of the venture to be successful= 1-(0.2+0.2) = 0.6=60%

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

Solution:- Expected Mean = Σ X\* p(X)

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

Average Earnings = 800

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

**Solution:**

Risk involved in a venture = E(X^2) - [E(X)]^2

Var = 2800000 - 8002 = 2160000

Standard Deviation = √var=1470

Since the deviation from the expected value is higher the risk involved

is also very high.