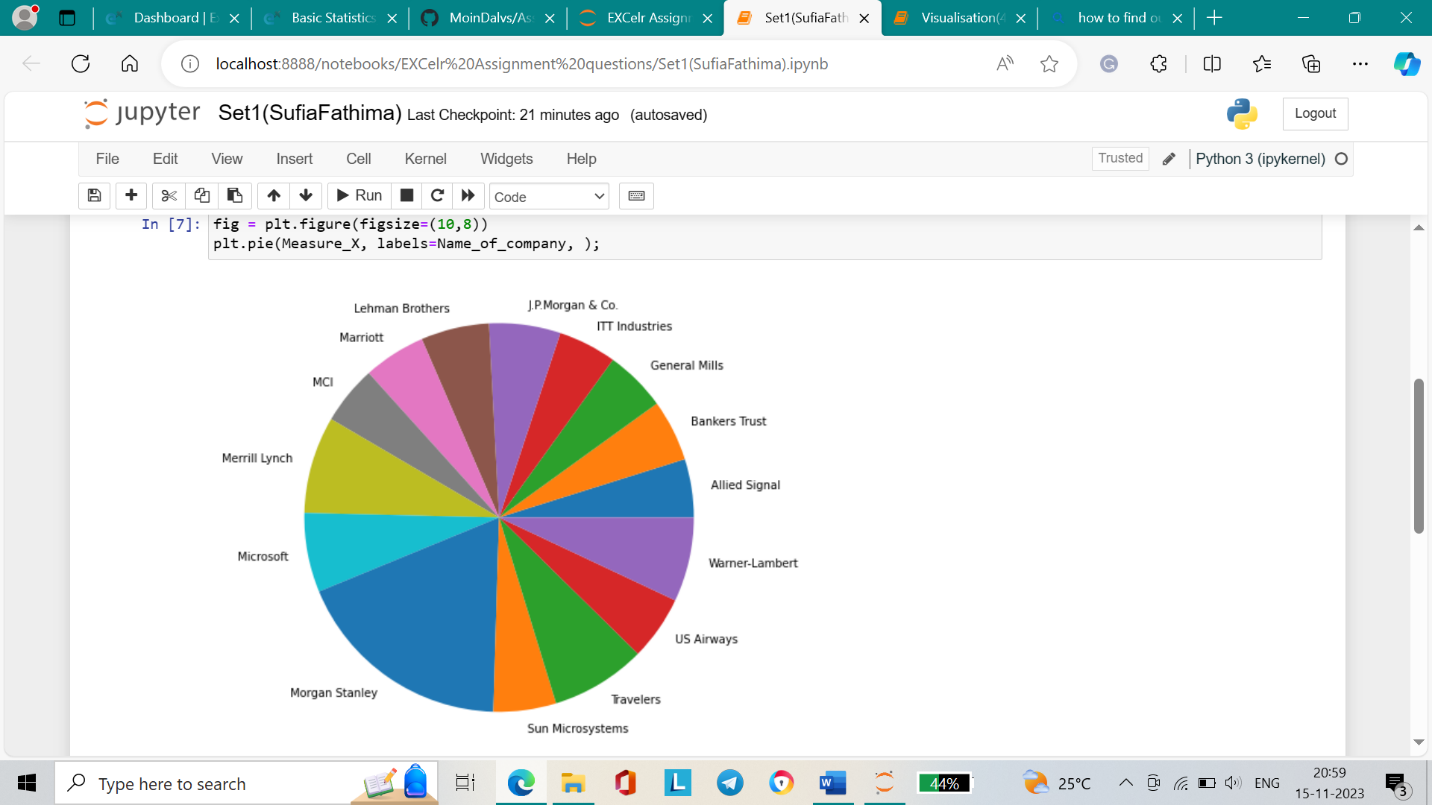
**Topics: Descriptive Statistics and Probability**

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

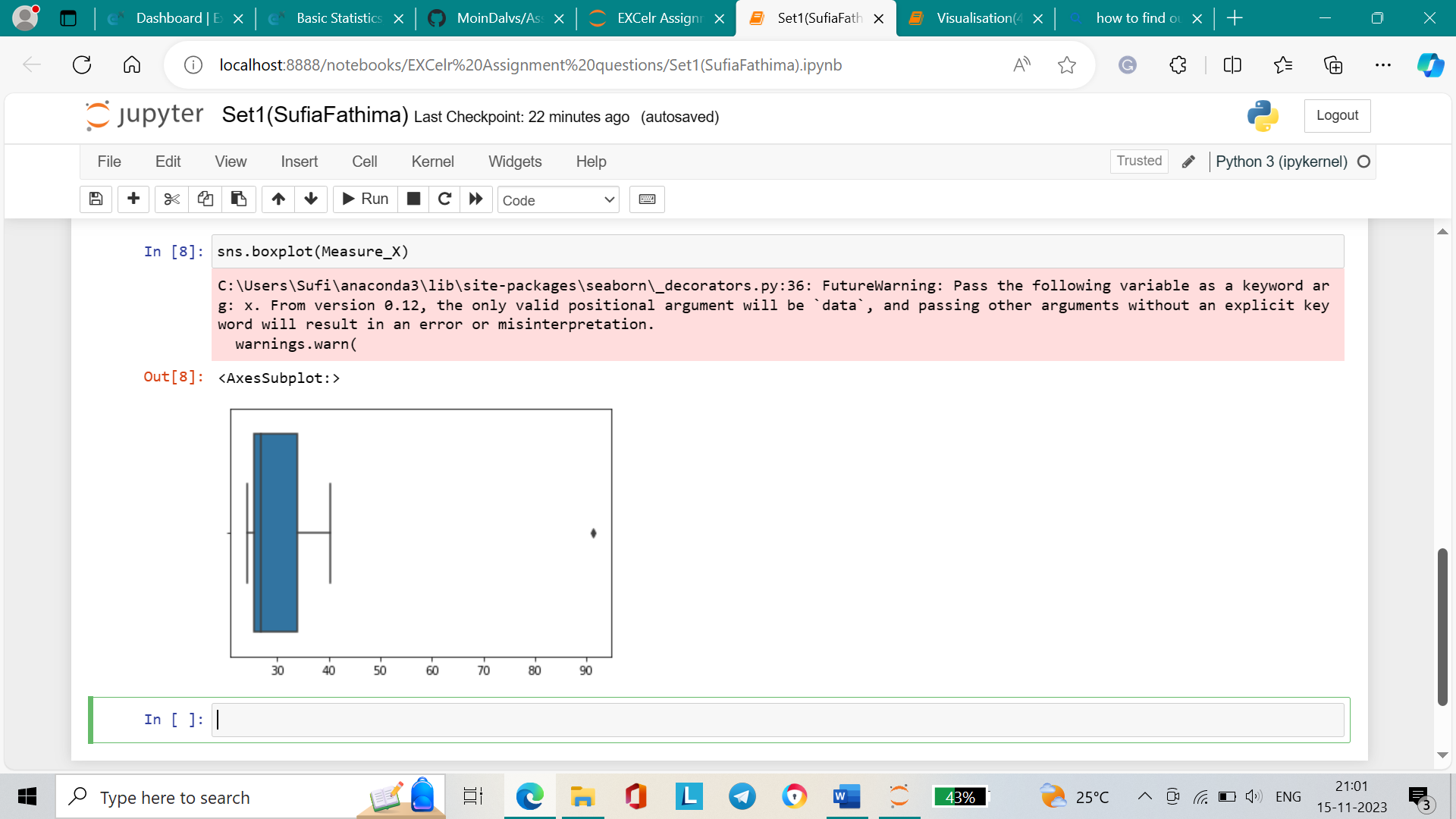
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft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

**Answer:**

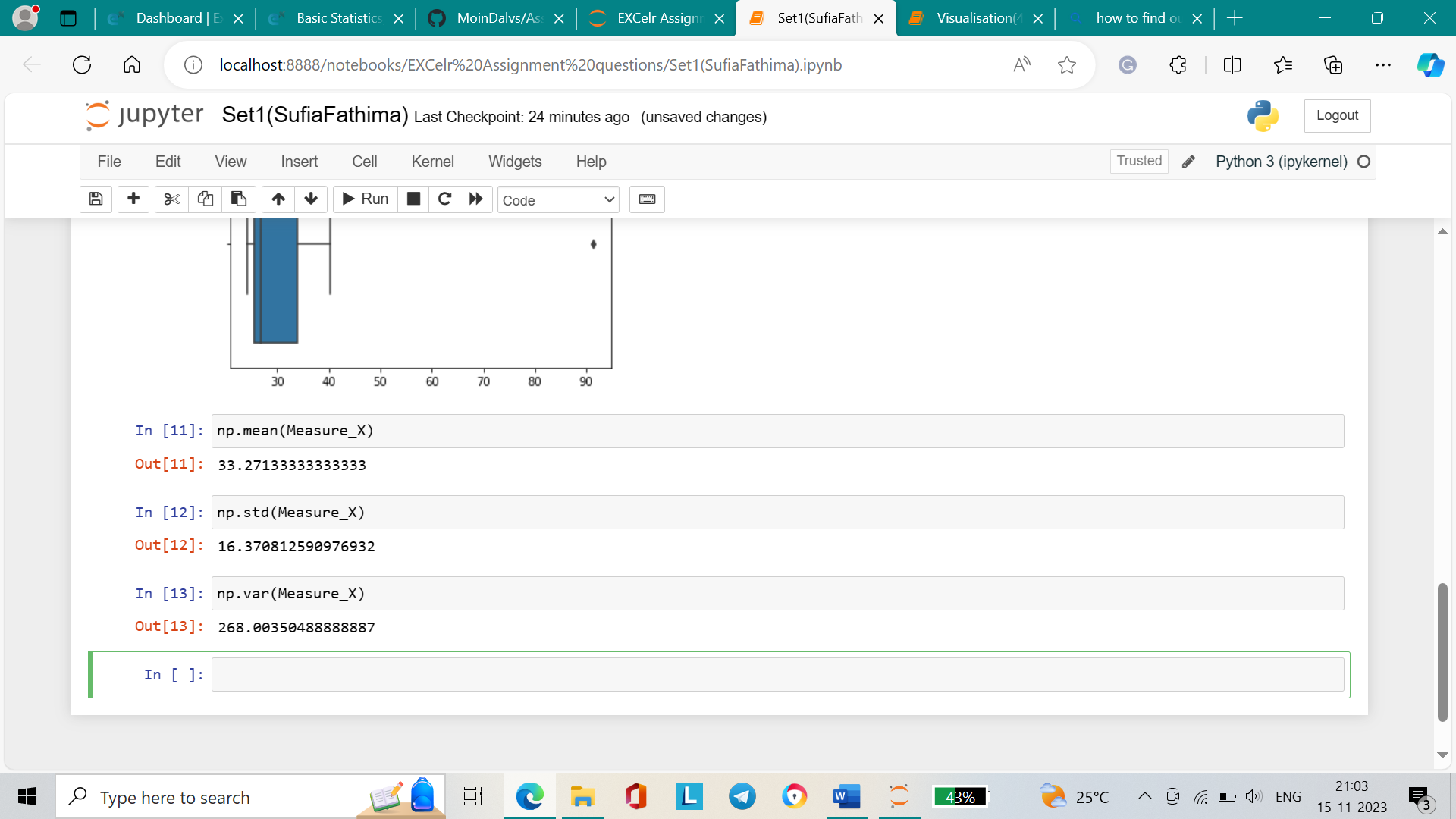
**Plot:**



**Outliers:**



**Mean, standard deviation, and variance**:



Mean = 33.27

Standard deviation = 16.37

Variance = 268.003



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**:

Interquartile range (IQR) = Q3-Q1 = 12-5 = 7

This means that 50% of the data is spread throughout this value.

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

**Answer**:

The data is right-skewed as the median is closer to the lower values.

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**:

There would be no indication of the outlier on the extreme right of the graph.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Answer**:

The mode of the dataset is the most frequently occurring value, hence it will lie approximately between 4-8.

1. Comment on the skewness of the dataset.

**Answer**:

The dataset is right-skewed or positively skewed as the tail is towards the right. Also, the median is greater than the mean.

1. Suppose 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:**

Both graphs provide information about the outlier and show that the dataset is positively skewed. Since the median is towards the lower values in both graphs, the skewness of the dataset is visible.

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**:

Misdirected call = p = 1/200

Correct Call = q = 199/200

n=5

P(1/5) = wrong number ?

P = nCr \* Px \* qn-x

P = 5C1 \* (1/200)1 \* (199/200)5-1

P = 5\*0.005\*0.98014

P=0.0245

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 |

**Answer**:

|  |  |  |  |
| --- | --- | --- | --- |
| x | P(x) | E(x) = X.P(X) | E(X2) =( X2 ).P(x) |
| -2,000 | 0.1 | -200 | 400000 |
| -1,000 | 0.1 | -100 | 100000 |
| 0 | 0.2 | 0 | 0 |
| 1000 | 0.2 | 200 | 200000 |
| 2000 | 0.3 | 600 | 1200000 |
| 3000 | 0.1 | 300 | 900000 |

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

**Answer**: The most likely monetary outcome of the business venture is 2000 as the probability is highest for 2000 at 0.3.

1. Is the venture likely to be successful? Explain

**Answer**: Yes, the probability that the venture will be successful is high. The probability that the venture will make more than 0 or a profit is given as p(x>1000)+p(x>2000)+p(x=3000) = 0.2+0.3+0.1 = 0.6.

Since 0.6>0.5, the venture is likely to be successful.

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

**Answer**: The long-term average earning is given by the expected value. Expected value = Sum (X \* P(X)) = 800. Which means the average return will be 800.

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

**Answer**: The good measure of the risk depends on the variability and standard deviation. **The higher the variance, the greater the risk.**

Expected Value (Mean) = weighted sum of all possible outcomes:

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

Variance **=** E(X2) – {E(X)}2 = 2800000 - 8002 = 2160000

Standard Deviation = sqrt(Var) = 1467