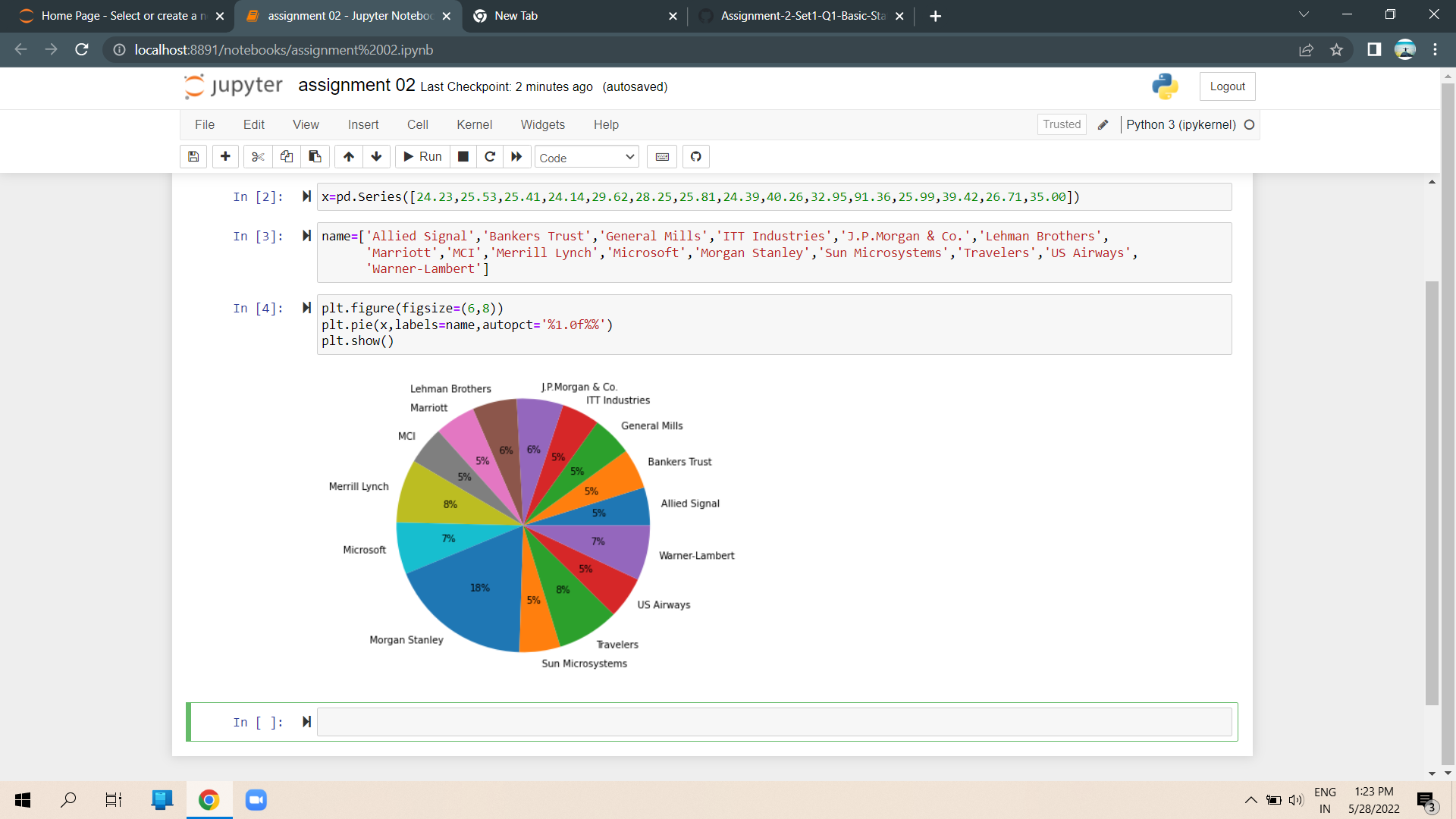
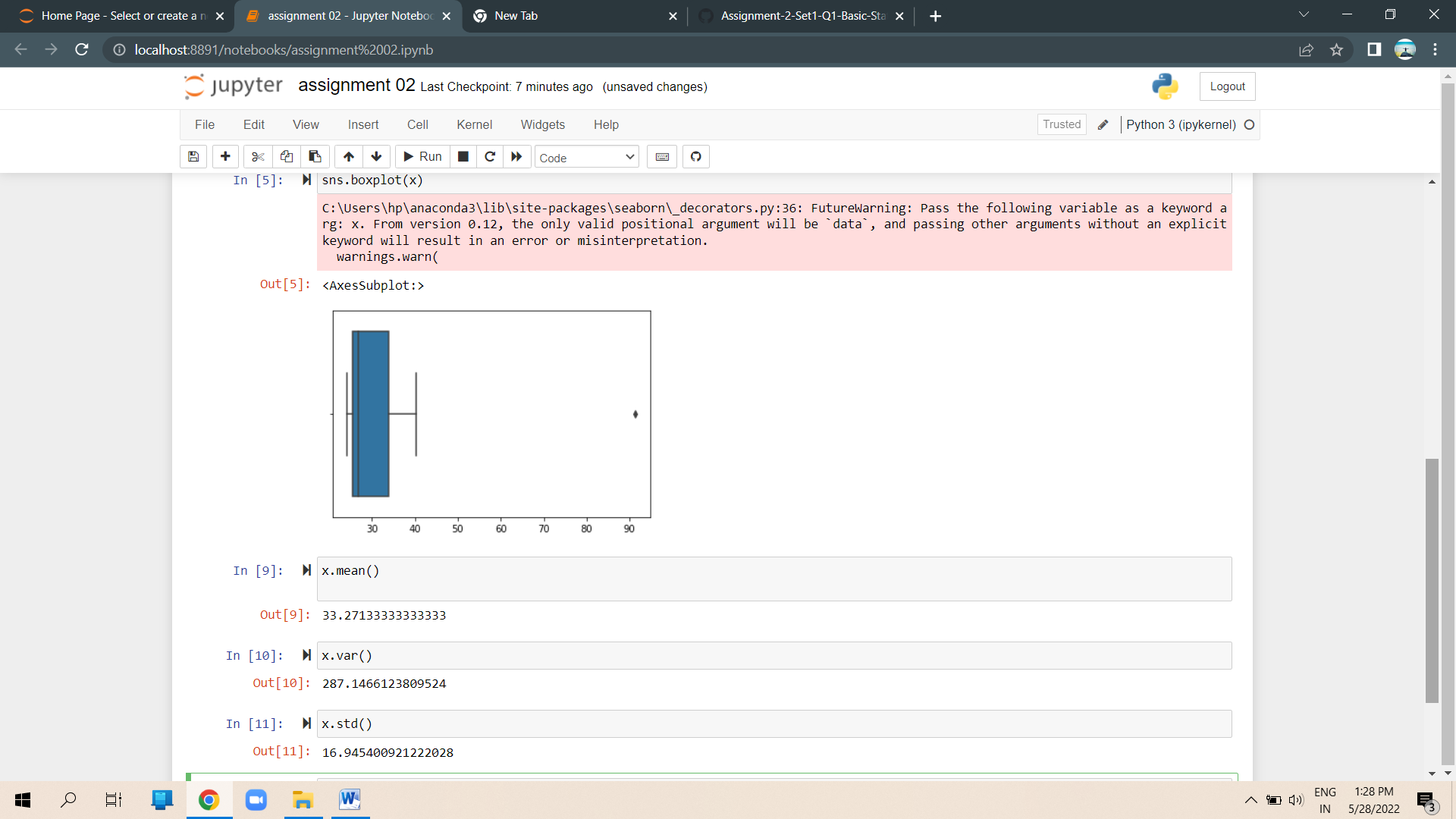
Topics: Descriptive Statistics and Probability

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 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 = IQR is Q3-Q1**

**= 12 – 5**

**= 7 ( The value implies that is Mean > median )**

**.**

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

**Ans=The data 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?

**Ans=if it was found that the data point is actually 2.5 instead of 25, the outlier in theboxplot will be removed.Whether the median shifts or not depends on the size of the data.It will reduce the right skewness of the data.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans = The mode can lie between 4 and 10. The 2 bars of the same height doesn’t indicate mode every time.**

1. Comment on the skewness of the dataset..

**Ans = The data is Right side 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.

**Ans= from the above histogram and barplot we can confirm an outlier at 25 in Y value. Both the plots indicate the +ve skewness of the dataset.**

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.)

one in 200 long-distance telephone calls is misdirected

=> 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?

**Ans: The most likely monetary outcome of the business venture is 2000$ As for 2000$ the probability is 0.3 which is maximum as compared to others**

1. Is the venture likely to be successful? Explain

**Ans: Yes, the probability that the venture will make more than 0 or a profit p(x>1000)+p(x>2000)+p(x=3000) = 0.2+0.3+0.1 = 0.8 this states that there is a good 60% chances for this venture to be making a profit**

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

**Ans: The long-term average is Expected value = Sum (X \* P(X)) = 800**

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

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

**df[“x”].std()1870.829**

**df[“x”].var()**

**3500000**