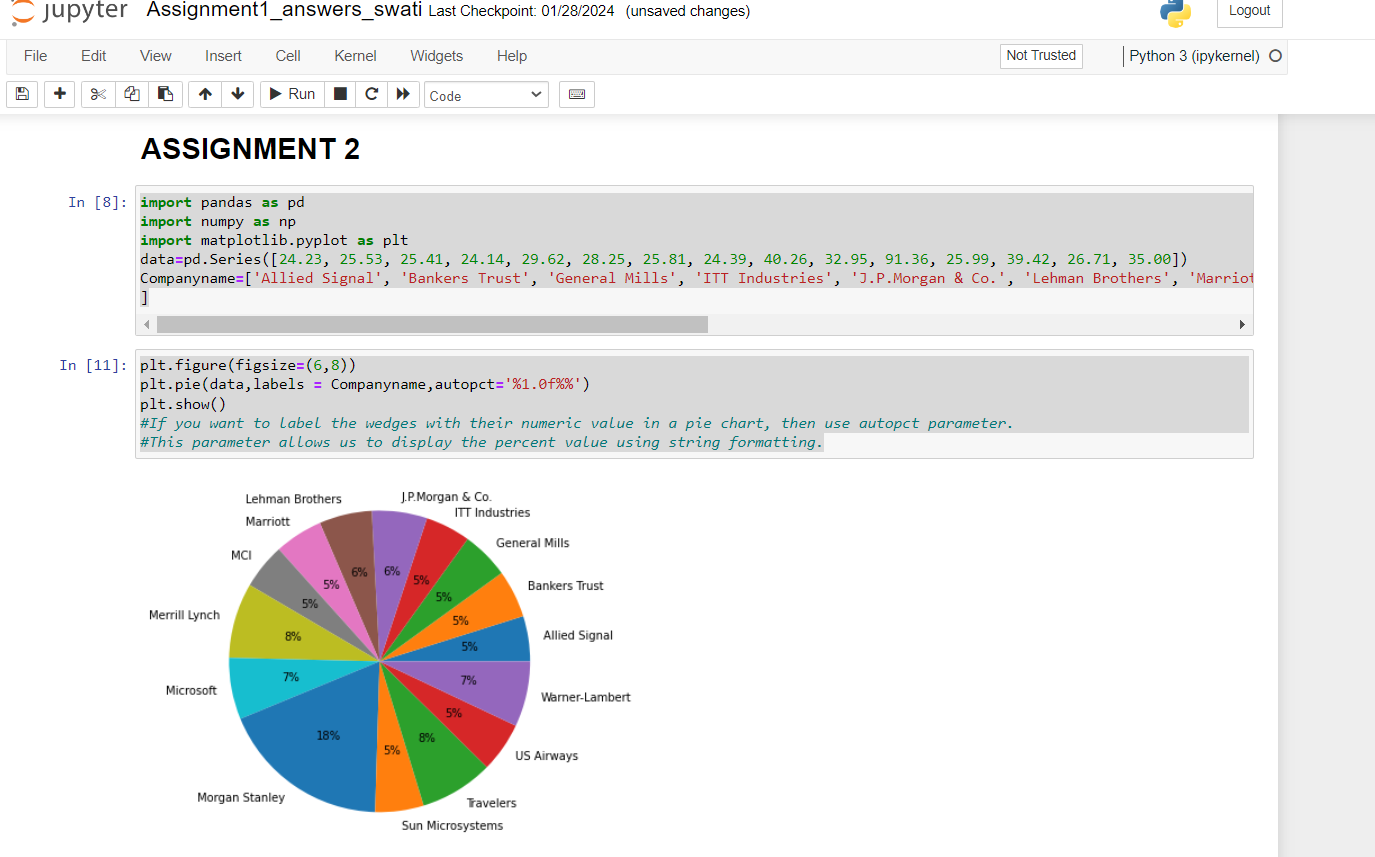
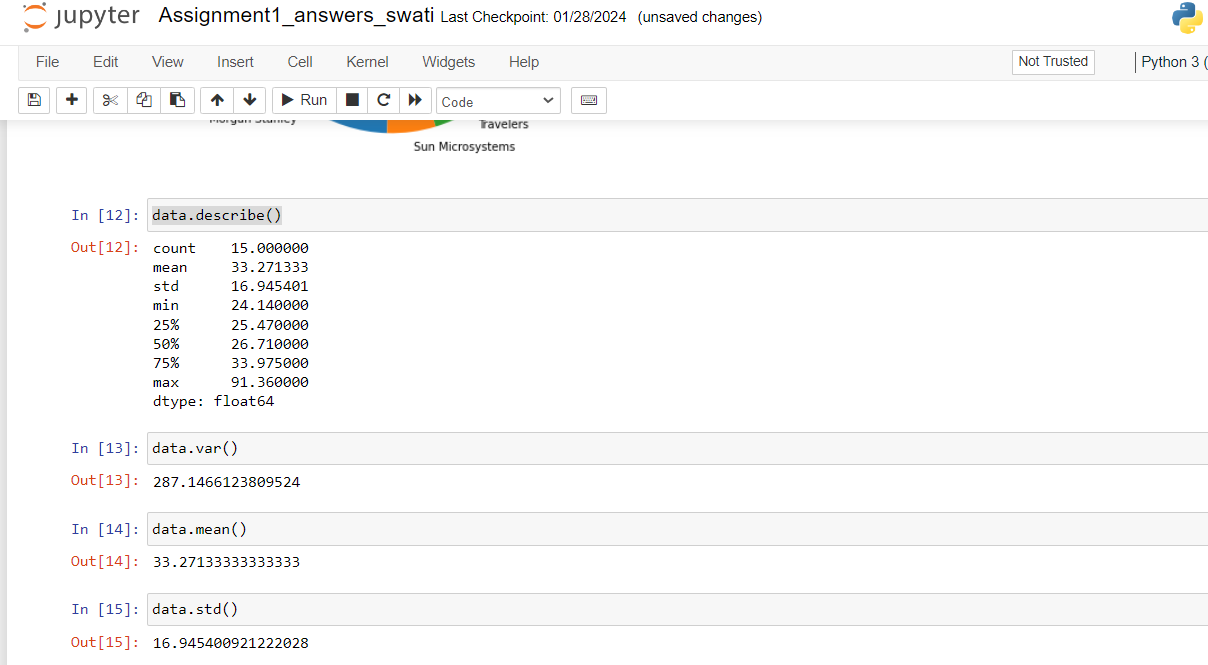
**Topics: Descriptive Statistics and Probability**

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=Q3-Q1=12-5=7.

25 is an outlier

1st quartile Q1=5

Q2=7

Q3=third quartile=12

The interquartile range (IQR) is the difference between the 75th and 25th percentile of the data.

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

Ans: It is right skewed and positive

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: Length of left whisker: Q1-1.5IQR=5-1.5\*7=-5.5

Length of right whisker: Q3+1.5IQR =12+1.5\*7=22.5

As per my above calculation, There would be no outlier then.

2.5 value is within range and it could have slightly affected mean and median. Box plot might have moved towards right slightly



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

The most repeating values are data 5,6,7 and 8.

1. Comment on the skewness of the dataset.

Right and positive 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.

Both tells data is positive skewed and would help us to find mean, median and mode values.

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: **:** 1 in 200 long-distance telephone calls are getting misdirected.

probability of call misdirecting ,p   = 1/200

Probability of call not Misdirecting,q = 1-1/200 = 199/200

probability for at least one in five attempted telephone calls reaches the wrong number

Number of Calls = 5

n = 5

p = 1/200

q = 199/200

P(x) = at least one in five attempted telephone calls reaches the wrong number

P(x) = ⁿCₓ pˣ qⁿ⁻ˣ

P(x) = (nCx) (p^x) (q^n-x)

P(1) = (5C1) (1/200)^1 (199/200)^5-1

P(1) = 0.0245037

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?

It is for 2000 as it’s having highest probability

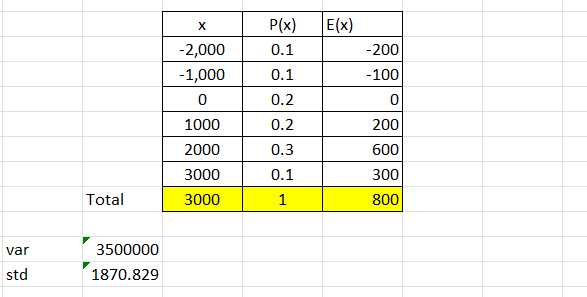
1. Is the venture likely to be successful? Explain

Ans: Yes, avg earnings would be around 800 and highest probability for earning is for 2000. So business venture likely to be successful.

E(X)=-2000\*0.1+(-1000)\*0.1+0\*0.2+1000\*0.2+2000\*0.3+3000\*0.1=800

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

It’s 800



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

Answer: Here the standard deviation is quite large which tells values are far away from the avg earnings/mean earnings. It indicates too much variability in data which could be positive as well as negative.