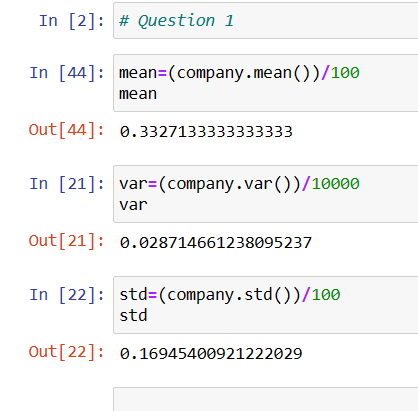
**Topics: Descriptive Statistics and Probability**

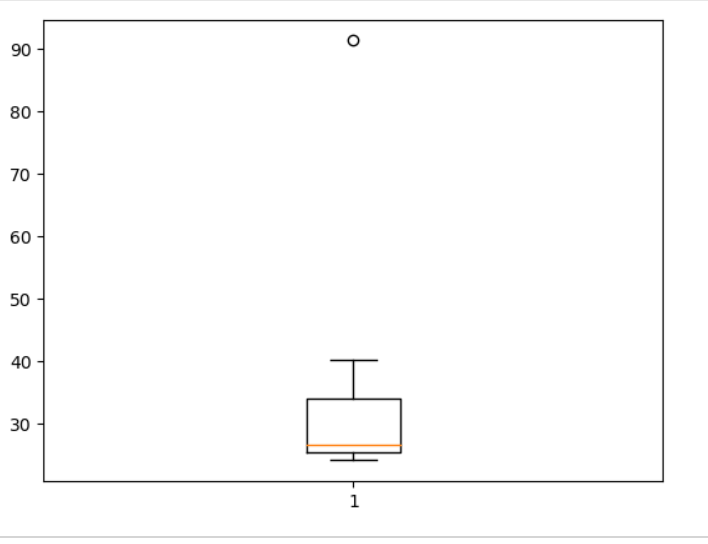
1. Look at the data given below. Plot the data, find the outliers and find out

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

From Python We get the values of as follows:



So from **Boxplot** it is clear that there is **outlier** on higher side which is **91.36%**





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 of the boxplot is approximately 5 to 12. This implies that the 50 % of the total data is lies between 5 to 12 datapoints.

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

Answer- from the above boxplot , as outlier is present at the upper extreme , it is clear that given dataset 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 - The new boxplot does not have outlier because value 2.5 is in lower Whisper.

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?

**Answer**- Mode of the dataset would be lies between points 4 to 8

1. Comment on the skewness of the dataset.

**Answer**- Dataset is clearly 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-** If histogram & boxplot are plotted then we can easily get the outliers, because in boxplot outliers can be seen and with the help of histogram we can understand it by frequency.

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-**

 if 1 in 200 long-distance telephone calls are getting misdirected.  
probability of call misdirecting = 1/200

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

The 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)

= (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?

Answer- 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

Answer- Venture is successful if X is + ve Hence if X is 1000 , 2000 or 3000 Probability is  0.2 + 0.3 + 0.1 = 0.6 as 0.6 > 0.5 Hence **venture likely to be successful**

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

Answer- Average Earning = Expected Value = ∑ Xi \* P(Xi) = 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 involved in a venture of this kind depends on the Variability in the distribution. Higher Variance means more chances of risk

Var (X) = E(X^2) –(E(X))^2 = 2800000 – 800^2 = 2160000

SD = √Var  ≈ **$ 1470** As **Variability is Quite high**  hence **Risk is high.**