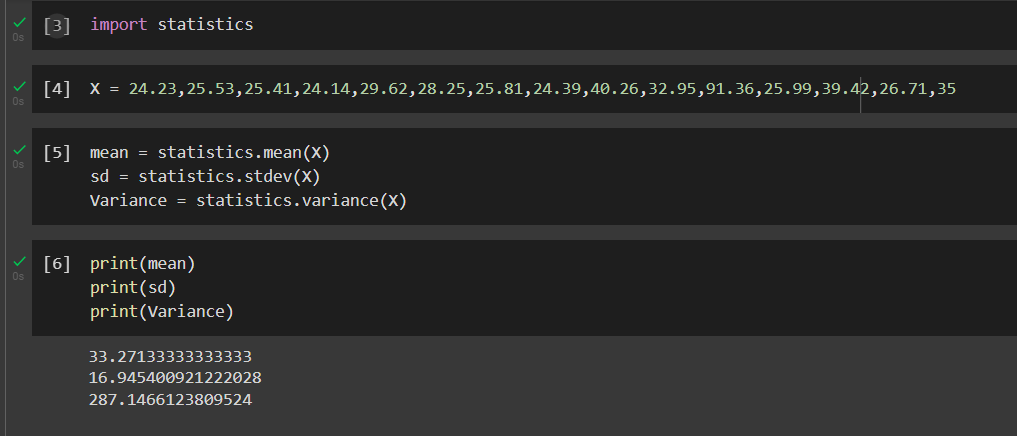
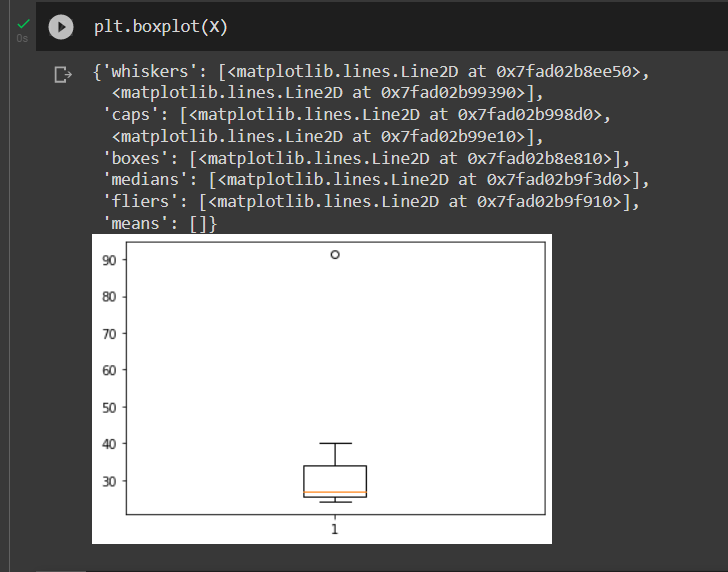
**Topics: Descriptive Statistics and Probability**

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% |
| JPMorgan & 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-**





Outlier is lying at X=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.
2. What can we say about the skewness of this dataset?
3. 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-**

1. Inter-quartile range is (12-5=7). IQR represents 50% of data when ordered from lower extreme to higher extreme.
2. Since tail is on right side, above plot is positively skewed.
3. If data point is 2.5 instead of 25, then their will not be any outlier as all the values are lying between and left extreme and right extreme.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?
2. Comment on the skewness of the dataset.
3. 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.

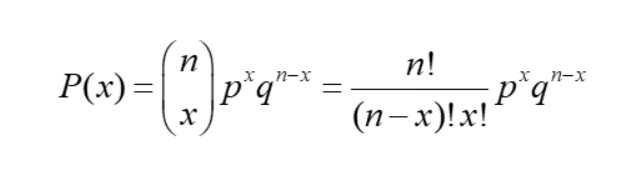
**Answer-**

1. Mode is the peak value of dataset. So, in a given dataset peak value is lying between 5-7. Therefore, mode is 5 and 7 for given dataset.
2. Above dataset shows positive skewness since tail is on right side.
3. Histogram and Boxplot or Whisker plots are methods used to find outlier. From both the plots we can conclude that 25 is an outlier.
4. 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-**

It is the case of binomial distribution because there are only two possible outcomes and distribution is discrete.

Binomial distribution can be calculated using formula,



Where,

n= Number of trails

x=number of successes

p= Probability of getting a success

q= 1-p= probability of getting a failure.

In binomial distribution we find probability of single variable, we will not find range. Hence “probability that at least one in five attempted telephone calls reaches the wrong number’” can be Rewritten as “No call reaches the wrong number”.

Given:

n=5

x=0

p=1/200

q=199/200

Above formula can be given as,

P(x)=1-no calls reach the wrong number

Therefore,

P(x)=1-(199/200)5=0.25

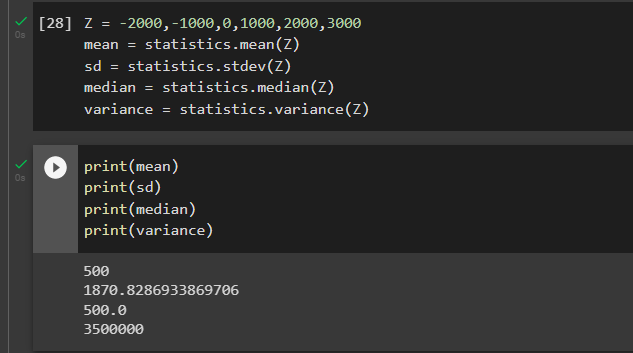
*Probability of at least one in five attempted telephone calls reached wrong number = 0.25*

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?
2. Is the venture likely to be successful? Explain
3. What is the long-term average earning of business ventures of this kind? Explain
4. What is the good measure of the risk involved in a venture of this kind? Compute this measure

**Answer-**

****

1. Highest probability is 0.3. Therefore, most monetary outcome is $2000.
2. This business venture most likely to be successful as profit is more than loss. ($800).

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

1. Long term average earning of business venture of above will be Expected value i.e., $800
2. Finding standard deviation and variance is good measure to understand the risk involved in the business venture as these parameters gives us information about variability of returns on business venture. More variability indicates more risk. Standard Deviation of this venture is $1870 and average return is $800 which indicates business venture is risky.