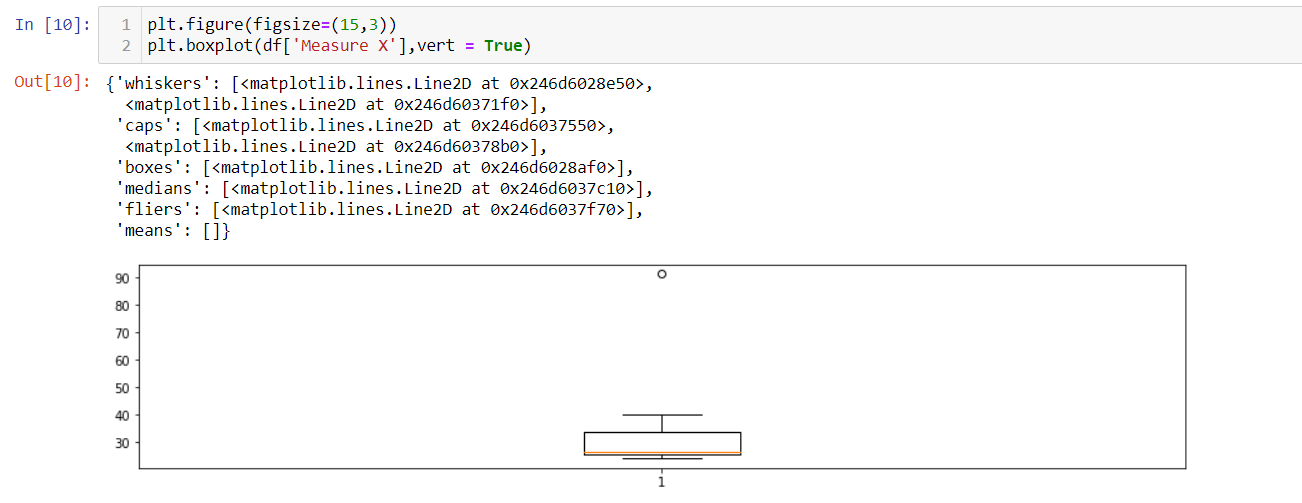
**Topics: Descriptive Statistics and Probability**

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% |

**Ans :**

| **Mean Σ X / N** | **33.27133333** |
| --- | --- |
| **Standard Deviation = √Σ(X-μ)^2 / N** | **16.37081259** |
| **varience Σ(X-μ)^2 / N** | **287.1466124** |



**Outlier - 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?

**Ans :**

**i - IQR (Inter quartile range) = Q3 - Q1, 25% to 75% (50%) data lies between Q3 to Q1**

**Here 12 - 5 = 7 = IQR**

**ii - Skewness tells about the asymmetry in data. In box plot when median is near to Min and whiskers are small at the lower end, then dataset is positively skewed (Mean > Median)**

**iii - Right most outlier is removed and placed inside between Min & Q1**



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.

**ANS:**

**i - Mode is most repeated value in dataset, here it is 20**

**ii - Skewness states the asymmetry in the dataset, 0 is the value for Normally distributed data. Direction of skewness can be seen with the direction of Tail, If the tail is towards right side (or) Mode is less than Mean then it is +vely skewed. I see the above Histogram is +vely skewed**

**iii - Both shows data spread in the dataset. Histogram is mainly used for continuous data type on X-axis, we can find skewness. Boxplot is used to find outliers and skewness of data. Through histogram we can see how data is grouped, skewed and by boxplot we can identify the outliers and skewness**

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 calls are misdirected**

**find the probability that atleast 1 in 5 attempted telephone calls reaches wrong number**

**Probability = Expected events / Total number of events**

**Total events = 200 P 1 (1 out of 200 reaches wrong number)**

**Expected events = 5 P 1 (1 reaching WRONG number of 5 calls)**

**Probability = 5 P 1/ 200 P 1 = 5 / 200 = 1/40 = 0.025**

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

ANS:

| x | P(x) | x \* P(x) |
| --- | --- | --- |
| -2,000 | 0.1 | -200 |
| -1,000 | 0.1 | -100 |
| 0 | 0.2 | 0 |
| 1000 | 0.2 | 200 |
| 2000 | 0.3 | 600 |
| 3000 | 0.1 | 300 |
| Total | | 800 |

i - Expected Monetary outcome = Σ (x \* P(x)) = 800

ii - YES , because prob(x>=0) is higher than prob(x< 0) , that means business is more likely to get profit.

iii - long-term avg. means to count for every every probability value, E(x) = $800

iv - prob(x>=0) = prob(0) + prob(1000) + prob(2000) + prob(3000)

= 0.2 + 0.2 + 0.3 + 0.1

= 0.8