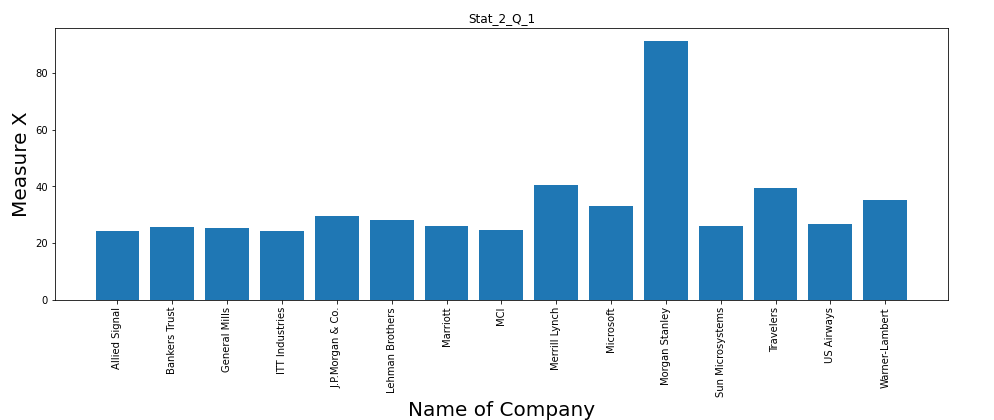
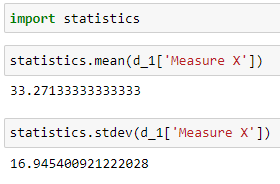
**Topics: Descriptive Statistics and Probability**

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 – After cleaning the data,





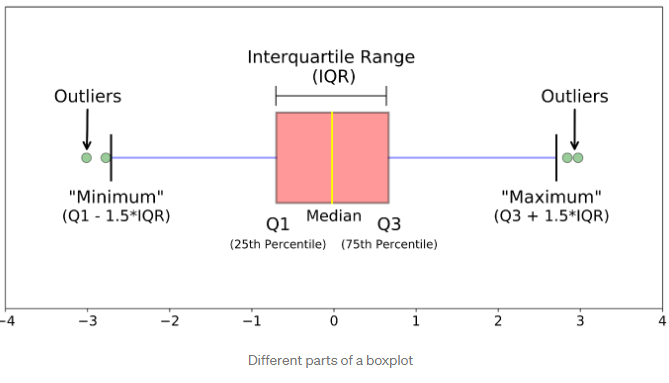
Therefore, Mean = 32.27 and Standard deviation = sigma\*sigma = 16.95

Thus,

Sigma = 4.18



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-** inter\_quartile range = 12 – 5 = 7 unit

Interquartile range shows the distribution on 50 percentile data.

That is data between 25th percentile and 75th percentile

Above case 50 percentile data lies in range 5 to 12.

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

**Ans –** median line is located at 7 on x axis, from boxplot we can observe that median is located near to left whisker of boxplot which indicates that boxplot 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 –** originally, the value at 25 was a outlier in the boxplot as it was located beyond the fourth quartile.

But in new case since the same value is located at 2.5 is no more outlier as it is located between 0th and 1st quartile.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**ANS-** Above data is bi modal, mode lie on 5 and 7 of y value located on x axis and equal to 20.

1. Comment on the skewness of the dataset.

**ANS -** For Median of the above data,

From range of values on x axis, (0+25)/2 = 12.5

Value corresponding to 12.5 on x axis is 8.

Therefore median = 8

From above y values on x axis corresponding to mode lies on the left side of the y values on x axis corresponding to median.

Thus, the above data is 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.

**ANS -** 1. From histogram and boxplot, we got same values for y value on x axis on fro median.

2. both histogram and boxplot shows the data is positively skew

3. also the information about the outliers, that is 2 outliers lies around the 25 on x axis.

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 -**

Probability of misdirecting the call = q = 1/200

Therefore p = 1 -1/200 = 199/200

To Find - the probability that at least one in five attempted telephone calls reaches the wrong number = ?

Solution –

P(x) = probability of at least one call to be misdirected

P(y) = Probability of no call to be misdirected

P(x) = 1 – P(y)

Thus to find the p(y), all should be not misdirected

P(y) = p\*p\*p\*p\*p

P(y) = 0.975

P(X) = 1 -0.975 = 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?

Ans – out of P(x) value 0.3 (highest) value is corresponding to $2000.

1. Is the venture likely to be successful? Explain

**Ans -** venture to be successful it should atleast have positive returns.

In above case chances of positive returns are 0.2 +0.3 +0.1 = 0.6.

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

**ANS -** (-2000)\*0.1+(-1000)-0.1+0\*0.2+1000\*0.2+2000\*0.3+3000\*0.1

-200-100+0+200+600+300

$800 is most likely monetary outcome

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

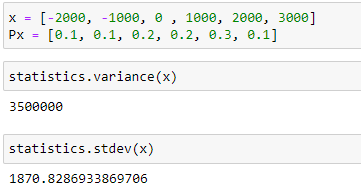
**Ans -**

To measure the risk involved in business, lets study the distribution pattern

Mean of x = 500

Var of x = 35000

Std of x = 1870.83



With mean of 500 and std of 1870.83, venture can be considered as risky.