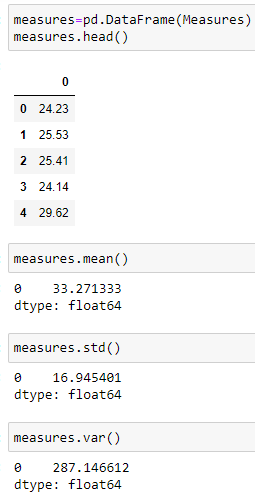
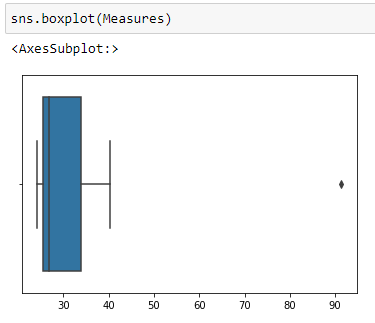
**Topics: Descriptive Statistics and Probability**

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% |





* The above boxplot shows that there is one outlier = **91.36%**
* Mean(µ) = **33.27%.**
* Standard deviation () = **16.94%.**
* Variance () = **287.14%.**



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.

IQR = upper quartile - lower quartile

= 12-5 = 7.

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

* In the above boxplot, the density of data toward left side is maximum, hence data is right side. (i.e 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?

* If the value is 2.5, then there is no outliers present in the boxplot.
* The median will shift to the very next positive value, since box plot is plotted by using median.
* Range will differ and the value 2.5 will be considered as the lowest value and it will lie in 1st quartile that is among the 25% of our data. And also its mean, median value might change.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

* The mode of this dataset line lies between 4 to 8

1. Comment on the skewness of the dataset.

* As most of the data lies towards the left side, the dataset is right 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.

* Boxplot helps to find the outliers, median,interquartile range, and maximum and minimum values
* Histogram helps to find the frequency of occurrence of data
* Compared to boxplot, the data distribution is shown good in Histogram.
* Skewness can be determined easily by using boxplot than histogram
* Both histogram and boxplot show that the data are right skewed

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.)

* Probability of getting wrong number (p)= 1/200
* Probability of not getting wrong number(q)=199/200
* By using Binominal distribution equation,
* When X=0, p(X=0) = 0.975
* When x>=1, P= 1-0.975 = 0.0247
* **So, probability that at least one in five attempted telephone calls reaches the wrong number is 0.0247.**

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?

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

=800

1. Is the venture likely to be successful? Explain

* From the above problem, the most likely outcome is 800, hence the venture is successful.

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

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

=800

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

* V(x) = 2160000
* Sd(x) = 1469.69.