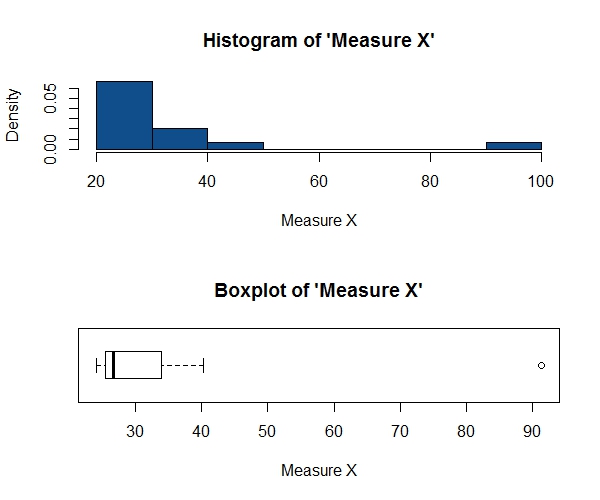
**CBA: Solutions to Practice Problem Set 1**

**Topics: Descriptive Statistics and Probability**

1. Here is a graphical representation of the data.



As we can clearly see, there is one outlier in the data. Observing the data table will be able to tell us that the value of ‘Measure X’ is particularly high for **Morgan Stanley.**

For this data,

|  |  |
| --- | --- |
|  | 33.27 |
|  | 16.95 |
|  | 287.15 |

1. (i) By looking at the box-plot we can approximate that the **First quartile (Q1)** is 5 and the **Third quartile (Q3)** is 12.

The Inter-quartile range indicates that 50% of the values of variable X lie within the range of 7 units i.e. in the interval 5 to 12.

(ii) By looking at the box-plot we can say that the distribution of X is **right-skewed** or **positively skewed.**

(iii) If it was found that the data value of 25 was actually supposed to be 2.5, the outlier in the box-plot will vanish.

Whether the median of the data shifts or not will depend on the size of the data, the frequency of the median and some other factors.

From visual observation we can say that it will reduce the skewness of the data.

1. (i) There are 2 bars in the histogram which are of the same height. But this does not indicate that the distribution has to be bimodal. We can only say that the mode lies between values 4 and 8 (both included). We will need to have the actual data in order to calculate the exact value of mode.

(ii) The distribution of variable Y is **right-skewed** or **positively skewed** as the histogram has a long tail on the right side. It indicates that there are many very high values with low frequency.

(iii) Box-plot readily tells us the values of minimum, maximum, first & third quartile (Q1 and Q3) and also the median. However, it is not possible to get information about the mode(s) of the data.

On the other hand, histogram readily gives information about mode(s). But it would not be easy to extract the values of above mentioned statistics from a histogram.

Both the plots clearly indicate the skewness of the dataset for which they are plotted.

Thus the two plots complement each other and are very useful when used in conjunction!

1. We are told that that,

We also know that whether a long-distance telephone call is misdirected or not is independent of any other long-distance call getting misdirected (or not).

We want to find,

Therefore, the probability that at least one in five attempted long-distance telephone calls reaches the wrong number is 0.0248.

1. (i) The most likely outcome of this business venture is a return of **$2000** as it has the highest probability of occurrence.

(ii) Success of the venture can of course be defined in multiple ways. But based on the data provided, we can look at **positive returns** as a measure of success.

The probability distribution gives us an idea about the long-term chances of earning given values of returns (indicated by **x**). Therefore, there is a *(0.2+0.3+0.1) = 0.6* probability that the venture would be successful.

In other words, the venture is 60% likely to succeed.

(iii) Now, we are not talking about only this venture. We are considering many ventures ‘of this kind.’ This means that we are looking at multiple ventures whose returns will have a distribution which is same as the one given for this venture.

In that case we say that the expected value of returns to this particular venture is the required average.

Therefore the long-term average earning for ventures of this kind is **$800**. This means, if we observe the returns to ventures of this kind (with the same probability distribution of returns) over a long period of time, the average returns would be $800.

(iv) Risk stems from the possible variability in the expected returns. Therefore a good measure to evaluate the risk for a venture of this kind would be the variance or the standard deviation of the variable X.

The standard deviation of the returns is **1469.69**. This large value of standard deviation considered along with the average returns of $800 indicates that this venture is highly risky.