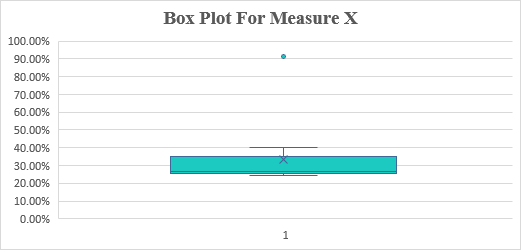
**Topics: Descriptive Statistics and Probability**

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% |

**Solution:**



By analyzing the above box plot we can conclude following points:

1. We can figure out that the given dataset has one outlier i.e., 91.36%.
2. Majority of the data lies approximately between 24.14% to 40.26% values.
3. Where the median is 26.71%.
4. The inter quartile range IQR = Q3 - Q1 = 35.00 - 25.41 = 9.59.

For the above data the solution for mean, variance and standard deviation is as follows:

1. Mean () = 33.27%
2. Variance = 0.0268
3. Standard Deviation = 0.163708126

[Solution Q1](Solution%20Q1.xlsx)

|  |  |  |  |
| --- | --- | --- | --- |
| Solution | | | |
|  | Mean | Variance | Standard Deviation |
| Measure X | 33.27% | 0.0268 | 0.163708126 |



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?

**Solution:**

1. To find inter-quartile range we assume from the above graph the values of Q3 and Q1 as 12 and 5. So,

Inter-quartile range (IQR) = Q3 – Q1 = 12 – 5 = 7

The inter-quartile defines middle half of your data set i.e., 50% of my data lies in that range.

1. The data set is right skewed, or we can say it has positive skewness.
2. If we found out that the value 25 which is and outlier is actually 2.5 then, the median and the range of the IQR will change with a minor difference.



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.

**Solution:**

1. The mode of the data set will lie between value 5 – 10 since the peak of the histogram lies between these values.
2. The given graph is right skewed or we can say it has positive skewness. Since majority of the data is towards left.
3. Here, if we compare the both the graphs of Q 2 and Q 3, we can see that both the graphs have positive skewness or the graph is right skewed. Since in both graph majority of the data is on the left side. By analyzing both the graphs has an outlier at 25. So we can conclude that the both the graphs are derived from same data set.
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.)

**Solution:**

One in 200 long-distance telephone calls is misdirected

Probability of call misdirecting p = 1/200

Probability of call not Misdirecting = 1 - 1/200 = 199/200

Number of Calls = 5 P(x) = ⁿCₓ pˣ qⁿ⁻ˣ

n = 5

p = 1/200

q = 199/200

at least one in five attempted telephone calls reaches the wrong number

= 1 - none of the call reaches the wrong number

= 1 – P (0)

= 1 - ⁵C₀ (1/200)⁰ (199/200)⁵⁻⁰

= 1- (199/200)⁵

= 0.02475

**Answer:** probability that at least one in five attempted telephone calls reaches the wrong number = **0.02475**

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

**Solution:**

1. x = 2000 with highest probability of 0.3 is the most likely monetary outcome of the business venture.
2. Since the probability of non-negative returns is more than 0.5 which is 50%, the venture will be successful if these rates are maintained. 0.2+0.3+0.1=0.6
3. p(x)\*x

= (-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

long-term average earning of business ventures = $ 800

1. venture is likely to be successful as Expected value is positive i.e., $ 800