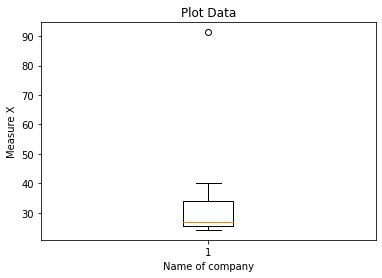
**Topics: Descriptive Statistics and Probability**

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 (μ) | 33.27133333333333 | |
| Variance(σ^2) | 287.1466123809524 | |
| Standard Deviation(σ) | 16.945400921222028 | |
| Outlier | Morgan Stanley | 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.

**Ans:- The First-Quartile Range (Q1) = 5**

**The Third-Quartile Range (Q3) = 12**

**The Second-Quartile Range(Q2) or Inter-Quartile Range (IQR) = Q3 – Q1 = 12 – 5 = 7**

**Inter-Quartile Range (IQR) is the Median Value**

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

**Ans:- The Data is Positively Skewed. The Tail is towards right side of the plot and the median is at the left. Right 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:- The Median Value remains same , where as the IQR may change. Mainly the outlier will be not will be not present and its will be Normally Distributed.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans:- **The Mode of the dataset are between 5 to 10 , approximately lies between 4 to 8. The mode is the data value that occurs the most in a dataset.**

1. Comment on the skewness of the dataset.

Ans:- **The dataset is Right Skewed , as the dataset lies on the right . which is Mean>Median>Mode.**

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:- **Both the Graphs are Right Skewed , and outliers are identified in both the graph. Box Plot median visualized easily where as in histogram Mode is Visualized easily .**

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:- **IF** 1 in 200 long-distance telephone calls are getting misdirected.

probability of call misdirecting   (p) = 1/200

Probability of call not Misdirecting (q) = 1-(1/200) = 199/200

**The** probability for at least one in five attempted telephone calls reaches the wrong number

Number of Calls (n) = 5

P(x) = at least one in five attempted telephone calls reaches the wrong number

P(x) = (nCx) (p^x) (q^n-x)

P(1) = (5C1) (1/200)^1 (199/200)^5-1

**P(1) = 0.0245037**

the probability that at least one in five attempted telephone calls reaches the wrong number = 0.0245037

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 |

**E (X) = Sum X \* P (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**

**E(X^2) = X^2 \* P(X)** => => ((-2000^2) \* 0.1) + ((-1000^2) \* 0.1) + ((0^2)\*0.2) + ((1000^2) \* 0.2) + ((2000^2) \* 0.3) + ((3000^2) \* 0.1) = (4000000\*0.1) + (1000000\*0.1) + (0 \*0.2) + (1000000\*0.2) +(4000000\*0.3) + (9000000\*0.1)

**= 400000 + 100000 + 0 + 200000 + 1200000 + 900000 = 2800000**

1. What is the most likely monetary outcome of the business venture?

Ans:- the most likely monetary outcome of the business venture is $ 2000 , As for $ 2000 the probability is 0.3 which is maximum as compared as to others.

1. Is the venture likely to be successful? Explain

Ans:- Yes , the Probability that the venture will make more than 0 or a profit

P(X>0)+P(X>1000)+P(X>2000)+P(X>3000) = 0.2+0.2+0.3+0.1 = 0.8

The States that there is a 80 % chance for the venture to make a profit

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

Ans:- The Long term average earning of Business venture is Expected value

E (X) = Sum X \* P (X) = 800 , Which means on an average the return will be $ 800.

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

Ans:- the good measure of the risk involved in a venture depend on Variability in the distribution , Higher the variance means more chances of risk

Var (X) = E(X^2) – (E(X))^2

= 2800000 – 800^2

= 2160000