**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% |

**In Python**

import pandas as pd

df = pd.read\_excel('D:\\Study\\Assignments\\df.xlsx')

df['Measure X'].mean()

0.3327133333333333

df['Measure X'].var()

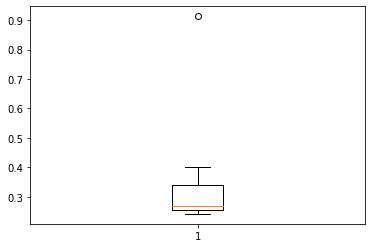
0.028714661238095233

df['Measure X'].std()

0.16945400921222029

import matplotlib.pyplot as plt

plt.boxplot(df['Measure X'])



|  |  |
| --- | --- |
| Morgan Stanley | 91.36% |

Outlier =



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-** IQR = Upper Quartile – Lower Quartile = 12 -5 = 7 (approx).

IQR implies that 50% of data lies inside this range.

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

**Ans-** The skewness of this dataset is right-skewed and the nature of skewness is positive.

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-** Then there will be no outliers found from the boxplot.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans-** The mode of this dataset lie between 4 to 8 values of ‘Y’.

1. Comment on the skewness of the dataset.

**Ans-** The skewness of the dataset is right-skewed. The nature of skewness is positive skewness.

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-** These two graphs confirms that there are outliers present in the dataset ie. 25 .

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:**As per data given 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) = ⁿCₓpˣqⁿ⁻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

**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?

**Ans: 2000**

1. Is the venture likely to be successful? Explain

**Ans:** Yes this venture is somehow likely to be successful as because the returns of positive probability is (0.2+0.3+0.1=0.6) or 60% which is more than 50%. So that we can say it as the venture is successful.

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

**Ans:**  To know the long-term average earning of business ventures, we need to calculate the expected value.

Average return (expected value)= (-2000 \* 0.1) + (-1000 \* 0.1) + (0 \* 0.2) + (1000 \* 0.2) + (2000 \* 0.3) + (3000 \* 0.1) = 800

So the long term average earning for these types of venture is around 800.

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

**Ans :** To know the risk possibility in this venture we need to calculate the mean and standard deviation. Here mean is 800 and standard deviation is 1870. So due to large difference this venture is risky in nature.