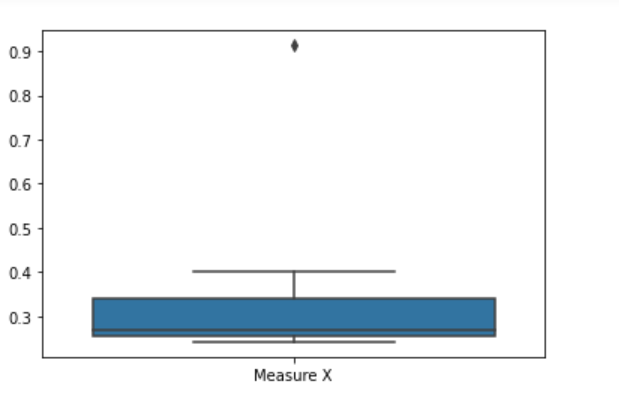
**Topics: Descriptive Statistics and Probability**

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% |



🡪Morgan Stanley is the outlier

Mean: 0.3327133333333333

variance: 0.028714661238095233

Standard Deviation: 0.16945400921222029



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.

Q3=12.5,Q1=5

IQR=q3-q1= 7.5

It is the range between q3 and q1. 50% values lie in this range hence we can infer that 50% of our values lie between 12.5 and 5 and area covered by it is 7.5

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

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?

It will lie in lower whisker and there will be no outliers



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie? 5-10
2. Comment on the skewness of the dataset.—Right skewed
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.

Boxplot provided information about outliers and median whereas histogram distributed data into 5 bins and clarified which range has maximum data and thus telling us the where our mode would lie.

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.)

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

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?

It can be seen from the above table that for x = 2000, the value of P(X) most. Hence, the most likely monetary outcome of the business venture is x = 2000.

1. Is the venture likely to be successful? Explain

Yes. Because, P(X>=1000) = 0.6 which is higher than P(X<1000) = 0.4.

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

Long term average earning of business venture = x.mean()= 500.

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

Risk = P(X<=0) = 0.2+0.1+0.1 = 0.4.