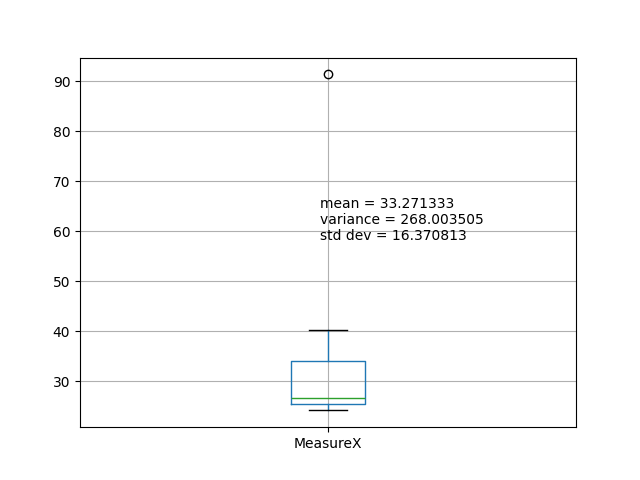
**Topics: Descriptive Statistics and Probability**

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% |



Mean = 0.3327

Variance = 0.0268

Std Dev = 0.1637

outlier = Morgan Stanley with 91.36% is the 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 = Q3-Q1 = 12-5 =7

50% of the data is spread from 5 to 12.

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

Ans: Distribution is Right skewed as median is less than mean.

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: 25 is an outlier. If it is actually 2.5, then falls in the first quartile of the distribution. So, there won’t be an outlier. IQR may change slightly and can be decreased. There will be a change in Mean, however it depends on the overall distribution. Median may or may not be changed. Skewness is positive with 25. With the correction to 2.5 might reduce the skewness slightly.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: mode lies between 6 to 8.

1. Comment on the skewness of the dataset.

Ans: Data is right skewed

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 shows that dataset has one outlier, median is 7 and positively skewed or right skewed.

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: Probability of call getting misdirected = 1/200

Probability of call getting rightly connected = 1-1/200 = 199/200

No.of attempts = 5

P(X) = 1- (199/200)^5 = 0.02475

P(X)= 2.475%

Hence, probability is 2.5% approx..

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: Probability of x, 2000 is 0.3 which is highest. It is the most likely monetary outcome.

1. Is the venture likely to be successful? Explain

Ans: P(X>0) = P(X=1000) +P(X=2000) +P(X=3000)

= 0.2+0.3+0.1

= 0.6

P(X<=0) = P(X=-2000) +P(X=-1000) + P(X=0)

= 0.1+0.1+0.2

= 0.4

P(X>0) > P(X<=0). Therefore, this business venture is likely to be successful.

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

Ans: Calculate the expected value to find the long-term average earning of business.

EV, E(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

E(X)= 800

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

Ans: Good measures are Mean, Variance and Standard Deviation

Mean or Expected value E(X)= 800

Variance =

= (-2000-800)^2\*(0.1) + (-1000-800)^2\*(0.1) + (0-800)^2\*(0.2) + (1000-800)^2\*(0.2) + (2000-800)^2\*(0.3) + (3000-800)^2\*(0.1)

= 216000

Standard Deviation = sqrt(variance)

= 1469.6938