**Topics: Descriptive Statistics and Probability**

1. Look at the data given below. Plot the data, find the outliers and find out

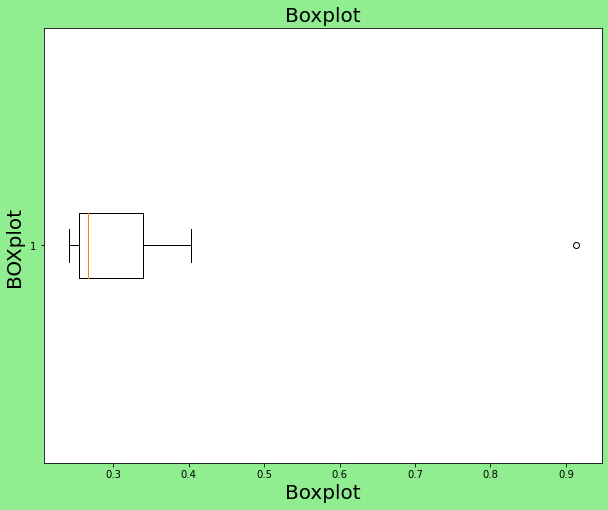
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft | 32.95% |
| Morgan Stanley | 91.36% |
| Sun Microsystems | 25.99% |
| Travelers | 39.42% |
| US Airways | 26.71% |
| Warner-Lambert | 35.00% |

Mean=33.2713

Standard deviation=16.9454

Variance=287.15

Outlier=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.
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)

Inter quartile range=q3-q1, i.e, somewhere around 12-5=7

This value implies 50% of data lies in this range.

(2)

Right skewed

(3)

If 25 is actually 2.5 than 25 will not considered as an outlier. And Box plot starts from 0 to 20.



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 data set lies between 4 and 8.

(2) right skewed

(3) In box plot 50% data lies between 5 and 12 and boxplot provide whisker length to find the outlier. Bout in histogram plot there is no way to find outlier. We can only guess by looking at the gap between 20 and 25, that 25 may be an outlier.

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.)

Solution: if 1 out of 200 phone calls is misdirected.

Probability of call misdirecting (p)= 1/200

Also, Probability of call not misdirecting(q)= 1-1/200=199/200

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

ncx P^x q^n-x p(x)

Therefore, P(x)=5c1 \*(1/200)\*(199/200)^4\*p(1)=0.0245

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?

Solution: The most likely monetary outcome of the business venture 2000$. As for 2000$ the probability is maximum.

1. Is the venture likely to be successful? Explain

Solution: Yes,

Because the probability of venture to make profit is(p(x>0)+p(x>1000)+p(x>2000)+p(x>3000))=0.8

Hence, there are 80% chances of Venture to make profit.

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

Solution: The long term average is expectation value=summation(x\*p(x))=800$

It means on an average it will return 800$

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

Solution: higher variance means more chances to of risk Var(X)=(E(X^2)-E(x)^2)=2160000