**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 :-

Outlier = Morgan Stanley =91.36%

µ=33.271333

σ=16.945401

σ2=287.1466

[click here](http://localhost:8889/notebooks/Downloads/Data%20Science/Data%20Science%20Assignments/2.%20Basic%20statistics%20level%202/DONE%20ASSIGNMENT/Basic%20Statistic%20Level%202.ipynb)



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 :- Inter Quartile Range: Q3-Q1=12-5=7

This value implies median value of boxplot.

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

Ans :- dataset is positively 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 :- In that case there would be no outliers on the given dataset, because of the outlier of data has positive skewness it will get reduce and data will normally distributed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans :-Mode of this given dataset would lie between 4 to 8.

1. Comment on the skewness of the dataset.

Ans :- 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 :-Above Histogram and boxplot in Q.2 both are right skewed and both have outliers.

The median can be easily visible in the box plot where as mode is more

Visible in histogram.

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 :-

Let A be the event as,

A:The call is misdirected.

P(A)=1/200

Thus,

B: The call is getting connected correctly.

P(B)=199/200

Probability that at least one call in 5 attempts

reaches the wrong number = 1- probability that no attempted call reaches

the wrong number

= 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?

Ans :- the most likely monetary outcome of the business venture is 2000 because it’s probability is

high.

1. Is the venture likely to be successful? Explain

Ans :-The venture likely to be successful because

+p(x=1000)+p(x=2000)+p(x=3000) = 0.2+0.3+0.1 = 0.6

This states that there is a good 80% chances for this venture to be making a profit

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

Ans :-Long term average earning of business venture of given kind = = 800

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

Ans :- P(Loss) = P(X=-2000) + P(X=-1000)

= 0.1 + 0.1

= 0.2

So, the risk associated with this venue is 20%.