**Topics: Descriptive Statistics and Probability**

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

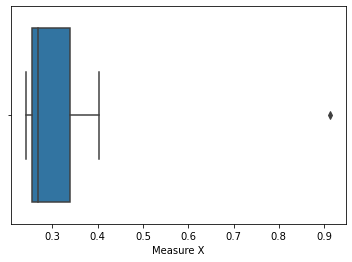
|  |  |
| --- | --- |
| **Name of company** | **Measure X** |
| Allied Signal | 24.23% |
| Bankers Trust | 25.53% |
| General Mills | 25.41% |
| ITT Industries | 24.14% |
| J.P.Morgan & Co. | 29.62% |
| Lehman Brothers | 28.25% |
| Marriott | 25.81% |
| MCI | 24.39% |
| Merrill Lynch | 40.26% |
| Microsoft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

A:

\* mean of the Measure X 0.3327133333333333

\* standard deviation of the Measure X: 0.16945400921222029

\* variance of the Measure X: 0.028714661238095233



Here there is only one point as the oulier i.e, at 0.9 and the name of the company is Morgan Stanley. It is having the highest measure of X when compared to all the other companies



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?

A :

* min value=0
* max value=19,
* quartile1=5
* median=7
* quartile3=12
* outlier=25

1. The interquartile range of this data set is (Q3-Q1)=12-5=7.

It implies the value of the median.

1. The data is right skewed data or positive skewed data as more values are towards right.
2. As 25 is an outlier ,and the value is changed to 2.5 the will be much variation in the boxplot and the median value also decreases.



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.

A:

1. The mode of the dataset lies between 4 to 8
2. The data is right skewed or positive skewed
3. Both the graphs describe the same information such as mean , median,mode,25%quartile,75%quartile,and outliers. Prefering a boxplot is better when compared to histogram.
4. 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.)

A:

One in 200 long distance telephone calls is misdirected.

i.e, The probability of misdirecting is 1/200

The probability of not misdirecting is 1-1/200=199/200

Number of Calls =5

PROBABILITY CALUCULATION:

1 0.995 3.10938E-09

2 0.995 1.23753E-06

3 0.995 0.000246269

4 0.995 0.024503738

5 0.995 0.975248753

Here, as we need to find the probability that at least one in five attempted telephone calls reaches the wrong number

i.e, 1-0.9752=0.024751247.

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?
2. Is the venture likely to be successful? Explain
3. What is the long-term average earning of business ventures of this kind? Explain
4. What is the good measure of the risk involved in a venture of this kind? Compute this measure

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| A: |  |  |  |  |  |  |
|  |  | i)The most likely monetary outcome of the business ventureis 2000 as the probability of **2000 is max i.e, 0.3.** | | | | |
|  |  |  |  |  |  |  |
|  |  | ii)As we need to check whether the venture is likely to be succesful we need to sum up all the postive values of x | | | | |
|  |  | considering 1000,2000,3000 we add 0.2+0.3+0.1 = 0.6 | | |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  | iii)Now, by comparing 0.6>0.5 .hence the chance of venture is to be likely sucessful . | | | | |
|  |  |  |  |  |  |  |
|  |  | The long-term average earning of business ventures | | |  |  |
|  |  |  |  |  |  |  |
|  |  | x | P(x) | E(X)=x\*P(x) | |  |
|  |  | -2,000 | 0.1 | -200 |  |  |
|  |  | -1,000 | 0.1 | -100 |  |  |
|  |  | 0 | 0.2 | 0 |  |  |
|  |  | 1000 | 0.2 | 200 |  |  |
|  |  | 2000 | 0.3 | 600 |  |  |
|  |  | 3000 | 0.1 | 300 |  |  |
|  |  |  |  | 800 |  |  |
|  |  |  |  |  |  |  |
|  |  | On an average the return will be 800 | |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  | iv) | the good measure of the risk involved in a venture | | |  |  |
|  |  | E(X) = ∑X . P(X) |  |  |  |  |
|  |  | E(X²) = ∑X² . P(X) |  |  |  |  |
|  |  | Var (X) = E(X²)  - { E(X) }² |  |  |  |  |
|  |  | SD = √Var |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  | Here we need to find variance .So, |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  | E(X²) = ∑X² . P(X) |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  | x | P(x) | X^2 | E(X^2) |  |
|  |  | -2,000 | 0.1 | 4000000 | 400000 |  |
|  |  | -1,000 | 0.1 | 1000000 | 100000 |  |
|  |  | 0 | 0.2 | 0 | 0 |  |
|  |  | 1000 | 0.2 | 1000000 | 200000 |  |
|  |  | 2000 | 0.3 | 4000000 | 1200000 |  |
|  |  | 3000 | 0.1 | 9000000 | 900000 |  |
|  |  |  |  |  | 2800000 |  |
|  |  |  |  |  |  |  |
|  |  | E(X)^2=800^2 | 640000 |  |  |  |
|  |  |  |  |  |  |  |
|  |  | HENCE THE E(X^2)=2800000 AND E(X)^2=800^2=640000 | | | |  |
|  |  |  |  |  |  |  |
|  |  | VAR= | 2160000 | (THE VARIATION IS VERY HIGH) | | |
|  |  | STD= | 1469.694 | (THE STD IS ALSO VERY HIGH) | | |
|  |  |  |  |  |  |  |
|  |  | Hence we can say that the measure of risk involved in the venture is very high. | | | | |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |