**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% |
| 14661 | 32.95% |
| Morgan Stanley | 91.36% |
| Sun Microsystems | 25.99% |
| Travelers | 39.42% |
| US Airways | 26.71% |
| Warner-Lambert | 35.00% |

Measure = [24.23, 25.53, 25.41, 24.14, 29.62, 28.25, 25.81, 24.39, 40.26, 32.95, 91.36, 25.99, 39.42, 26.71, 35]

Name of company = [ Allied Signal, Bankers Trust, General Mills, ITT Industries, J.P.Morgan&Co., Lehman Brothers, Marriott MCI, Merrill Lynch, Microsoft, Morgan Stanley, Sun Microsystems, Travelers, US Airways, Warner-Lambert]

**=** [24.23+ 25.53+ 25.41+ 24.14+ 29.62+ 28.25+ 25.81+ 24.39+ 40.26+ 32.95+ 91.36+ 25.99+ 39.42+ 26.71+ 35]

**=** 33.2713333333333

**STANDARD DEVIATION (**) **=** **16.370812590976932**

**VARIANCE (** **= 268.00350488888887**



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 of the dataset is 5-12 aprrox. It will represent you 50% data of the data set.**

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

**ANS = The data is right skewed or 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 = If it was found that the data point with the value 25 is actually 2.5, then there will be no outliers in the data sets.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**ANS = The mode of this dataset lie between 5-7**

1. Comment on the skewness of the dataset.

**ANS = It is right skewed or positively skewed data**

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 plots gives idea about skewness of the dataset which is right skewed.**

**From above box plot and histogram we can confirm an outlier at 25 in Y value.**

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 us define an event**

**E: The call is misdirected**

**then probability of the event E is**

**P(E) = 1/200**

**Therefore,**

**P(E) = 1-P(E) = 1 - (1/200) = 199/200**

**Probability that at least one in 5 attempted call reaches the wrong number**

**= 1 - Probability that no attempted call reaches the wrong number**

**= 1 – (199/200)\*(199/200)\*(199/200)\*(199/200)\*(199/200)**

**= 1-0.975248753**

**= 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?

**ANS = 2000 is the most likely monetary outcome of the business venture because it has highest probability.**

1. Is the venture likely to be successful? Explain

**ANS = YES venture is likely to be successful because is gives good returns.**

**E(x) = (-2000\*.01)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)**

**E(x) = $800**

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

**ANS = $800 is the long-term average earning of business ventures, because**

**E(x) = (-2000\*.01)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)**

**E(x) = $800**

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

**ANS =  Risk involved in a venture**

**Var (X) = E(X²)  - { E(X) }²**

=   **2800000 -   800²**

= **2160000**

**SD = sqrt(Var)  =** **$ 1470**

**Hence**,**Variability is Quite high** **so** **Risk is high.**