**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**:- from the data the outlier is 91.36%

Mean:- 33.2713333

Variance:- 268.0035048888888

Standard deviation :-16.37081259097



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(Inter quartile range)=13**-**5 **=**8 This represents the range which contains 50% of the data points

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

**Ans :-** Fromtheboxplot data is shifted towards the left side hence it is positively skewed/right skewed distribution more than 50% of the data is between 5-13

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:-** ifthevalue25 is replaced by 2.5 then outlier will be removed median is shifts ,it will reduce right skewness if the data



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans**:- The mode can lie between 4 to 8 there are many values in this range but this oids an assumption. The two bars of the same height does not indicate mode

1. Comment on the skewness of the dataset.

**Ans** :- Most of the data lies on left side and tail is at right side therefore it is positively skewed distribution and also have outlier

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**:- From histogram and boxplot the value of the outlier is 25,both data set is right skewed or positively skewed distribution. median in boxplot and mode in histogram. Histogram Provides the frequency distribution so we can see how many times each data point is occurring however boxplot provides quantile distribution i.e 50% data lies between 5 and 12.

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 consider the probability of 1 call misdirected of 200 as event A .

Probability of occurring of a event A=1/200

P(A)=1/200 = 0.005

Probability of having at least one successful call will be ,1-P(A)=1-(1/200)=1-0.005=0.995

As every event is independent of other event the probability will be,

=1-(0.995)^5

=1-0.975

=0.025

=2.5% Chance

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 is high probability is 0.3

1. Is the venture likely to be successful? Explain

**Ans**:- venture is successful when x is positive therefore if x=1000,2000,3000 respective probability is 0.2+0.3+0.1=0.6 Hence 60% chance of venture is Successful.

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

**Ans:-** Long term average earning of business venture is =x\*p(x)

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

= $800

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

**Ans** :- A good measure to evaluate the risk involved in a variance and standard deviation of the variable x.

Var = 3500000

Std=1870.83

The large value of standard deviation of $1870 is considered along with the average returns of 800 % indicates that this venture is highly risky