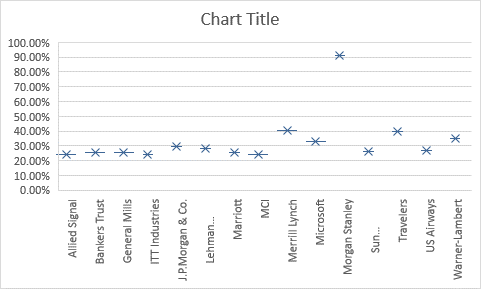
**Topics: Descriptive Statistics and Probability**

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:-**



Mean = ( ∑ Measure X ) / Count of Measure X data = ( 499.07 / 15 ) = 33.27%

Variance = ( (Xi−¯X)^2) / N = 0.028715

Std. Div. = √ (( (Xi−¯X)^2) / N) = 0.169454

Here, one outlier is there which is Morgan Stanley - 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?

**Ans :- (i)** IQR = Third Quantile Range - First Quantile Range

(First Quantile Range) Q1 = 5, (Third Quantile Range) Q3 = 12

= 12 – 5 = **7**

**(ii)** Because of the median is shifted towards left side of the data, it is **Right Skewed.**

**(iii)** If the data point 25 will become 2.5 then the boxplot will have no outliers, due to this the data become normally distribution, it will follow the Gaussian distribution.



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.

**Ans :- (i)** According to the histogram the mode lies between **4 to 8.**

**(ii)** The data set is slightly right skewed (positively skewed).

**(iii)** Both the data set is right skewed, both has outliers and median and outliers can be easily seen from the box plot but in histogram mode can be easily seen.

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 :-**  When 1 in 200 long-distance telephone calls are getting misdirected.  
The probability of the call misdirecting is = 1/200.

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

The probability for at least one in five attempted telephone calls reaches the wrong number of calls is = 5 n = 5 p = 1/200 q = 199/200.

P(x) = at least one in five attempted telephone calls reaches the wrong number.

P(x) = ⁿCₓ pˣ qⁿ⁻ˣ P(x) = (nCx) (p^x) (q^n-x) # nCr = n! / r! \* (n - r)! P(1) = (5C1) (1/200)^1 (199/200)^5-1 P(1) = **0.0245037.**

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

**Ans :- (i)** The most chances of the outcome of the business venture is 2000 because it has the probability of 0.3 which is the largest probability in the data set.

**(ii)** Yes, because the venture has more probability in positive values that is - 0.2+0.2+0.3+0.1 = 0.8, means 80%. So it is successful.

**(iii)** In the long term the average earning of business ventures of this kind is = ∑ ( x \* P(x) ) = 800. So it is profitable.

**(iv)** The good measure of the risk involved in a venture of this kind depends on Measure of spread (variance) which is equal to = Var (x) = E(X^2) –(E(X))^2 = 2800000 – 800^2 = 2160000.