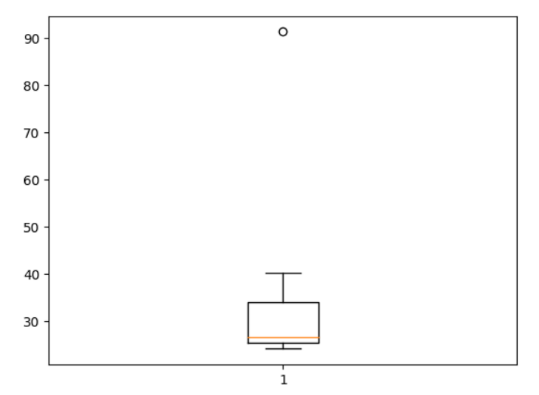
**Topics: Descriptive Statistics and Probability**

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% |

**We can see that we have only one outlier that is 91.36% (Morgan Stanley)**

**Mean = 33.27133333333333**

**Std. Dev = 16.945400921222028**

**Variance = 287.1466123809524**



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 = Q3-Q1 = 12-5 =7**

it implies that the range between the first quartile (Q1) and the third quartile (Q3) in a dataset is 7 units.

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

ANS: **Right-Skewed median is towards the left side it is not normal distribution**

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, Distribution could have changed to normal distribution from right skewed.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

ANS: **4 to 8**

1. Comment on the skewness of the dataset.

ANS: **It is positively skewed dataset**

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: 1. Both providing comprehensive information about the data.

2. Histograms are particularly useful for identifying patterns, skewness, or multimodality in the data distribution.

3. Boxplots are useful for identifying potential outliers or extreme values in the dataset.

4. By analyzing the patterns, skewness, outliers, and overall distribution of the data through histograms and box plots, you can gain a better understanding of its characteristics and make informed decisions in your analysis or modeling.

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: IF 1 in 200 long-distance telephone calls are getting misdirected.  
probability of call misdirecting = 1/200 Probability of call not Misdirecting = 1-1/200 = 199/200 The probability for at least one in five attempted telephone calls reaches the wrong number Number of Calls = 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?

ANS: **2000$ (30% probability)**

1. Is the venture likely to be successful? Explain

ANS: Yes, the probability that the venture will make more than 0 or a profit p(x>0)+p(x>1000)+p(x>2000)+p(x=3000) = 0.2+0.2+0.3+0.1 = 0.8 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: **The long-term average earnings of business ventures of this kind, based on the given probabilities, is $800.**

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

ANS: The good measure of the risk involved in a venture of this kind depends on the Variability in the distribution. Higher Variance means more chances of risk Var (X) = E(X^2) –(E(X))^2 = 2800000 – 800^2 = 2160000