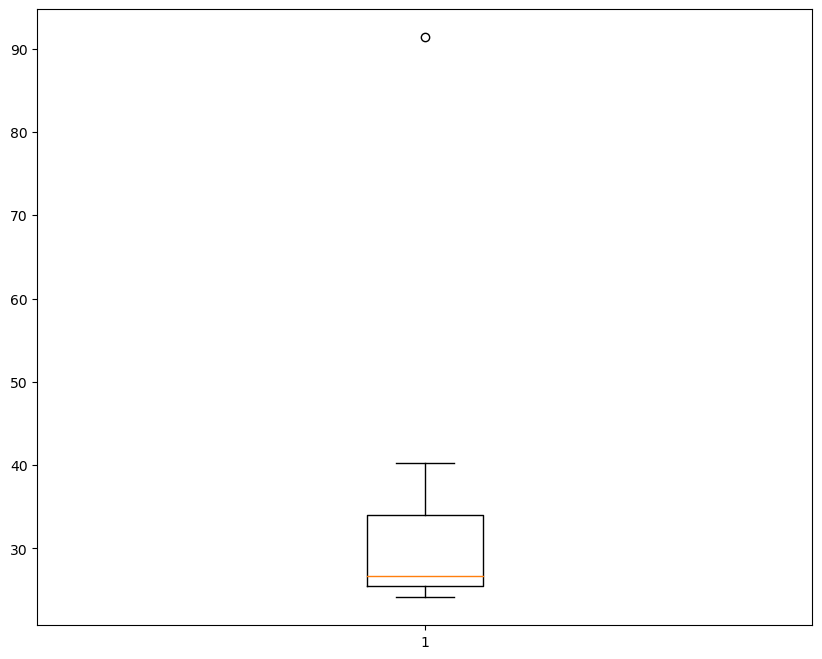
**Topics: Descriptive Statistics and Probability**

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:** Outlier = 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.

**Ans:**  Approximately,

First Quantile Range (Q1) = 5 ; Third Quantile Range (Q3) = 12 ; Second Quartile Range (Median) = 7 ; Inter-Quartile Range (IQR) = (Q3 – Q1) = (12 – 5) = 7

Second Quartile Range is the Median Value.

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

**Ans:** According to the given boxplot we can say that the data is Right-Skewed and left side is not normally distributed.

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 data point 25 is actually 2.5 then, then the outliers will not exist in the dataset and there might be a chance that data may follow normal distribution.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans:** The mode in the dataset approximately lie in between 4 and 8.

1. Comment on the skewness of the dataset.

**Ans:** It is a right-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:** From the histogram and the boxplot, we can say that the data is right-skewed and it has only one outlier. Whereas boxplot gives a clear visualization of median and histogram of the mode.

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:** Probability of misdirecting atleast 1 call (p) = 1/200;

Probability of not misdirecting any call (q) = 199/200;

n=200 , x=1

 Since, P(x) = ⁿCₓ pˣ qⁿ⁻ˣ

Then, P(1) = (200)C(1) \* (1/200)^1 \* (199/200)^(5-1)

P(1) = 0.025

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$. Because at 2000$ the probability is 0.3, which is maximum as compared to others.

1. Is the venture likely to be successful? Explain

**Ans:** Yes, the venture is likely to be successful.

 p(x>0)+p(x>1000)+p(x>2000)+p(x=3000) = 0.2+0.2+0.3+0.1 = 0.8.

It states that the probability of success is 80%.

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

**Ans:** The long term average earning is,

Expected Value = (x\*P(x)) = [(-2000\*0.1)+

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

Expected Value = 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 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

Var(X) = = 2160000.