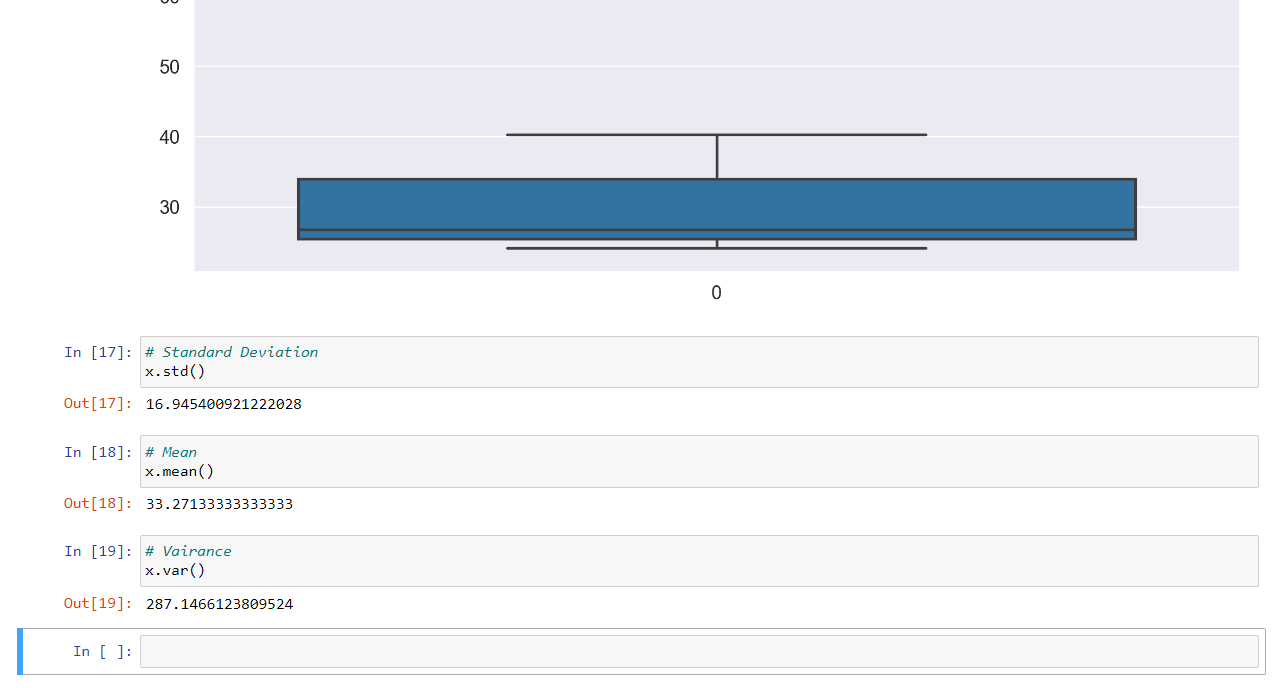
**Topics: Descriptive Statistics and Probability**

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% |







1. Answer the following three questions based on the box-plot above.
2. 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, Median (Second Quartile Range) = 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 :** 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 :** In that case there would be no Outliers on the given dataset because of the outlier the data had positive skewness it will reduce and the data will normal distributed



1. Answer the following three questions based on the histogram above.
2. Where would the mode of this dataset lie?

**ANS :** Between 5 – 8

1. Comment on the skewness of the dataset.

**ANS :** Positively Skewed.

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 :** By comparing both of them it is very clear that the data would be positively skewed. Also, would help us finding mean, mode 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 :** Probability of call geting misdirected = (1/200)

Hence probability of call not getng misdirected = 1-(1/200) = 199/200Number of phone calls attempted = 5

Therefore, probability that at least one in 5 attempted call reaches the wrong number is:=1-(199/200) ^5= 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 :** highest probability is for 2000.

1. Is the venture likely to be successful? Explain

**ANS :** Yes, because the total earnings of the venture is positive in value i.e 800 and highest probability of earning is 2000

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

**ANS :** The long-term average is Expected value = Sum (X \* P(X)) = 800$ which means on an average the returns will be + 800$

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

**ANS :** # Given probability distribution

returns <- c(-2000, -1000, 0, 1000, 2000, 3000)

probabilities <- c(0.1, 0.1, 0.2, 0.2, 0.3, 0.1)

# Calculate expected return

expected\_return <- sum(returns \* probabilities)

# Calculate variance

variance <- sum((returns - expected\_return)^2 \* probabilities)

# Calculate standard deviation

std\_deviation <- sqrt(variance)

# Display results

cat("Expected Return (mu): $", expected\_return, "\n")

cat("Variance (sigma^2): $", variance, "\n")

cat("Standard Deviation (sigma): $", std\_deviation, "\n")

# Assess risk

risk\_assessment <- ifelse(std\_deviation > expected\_return, "High Risk", "Low Risk")

cat("Risk Assessment:", risk\_assessment, "\n")

Risk involved in a venture

**SD** = √Var ≈ $ 1470

**Var (X)** = E(X²) - { E(X) }²

= 2800000 - 800²

= 2160000

Risk is high