# MAT 243 Project One Summary Report

[Marc Anthony Aradillas]

[marc.aradillas@snhu.edu]

Southern New Hampshire University

## Introduction: Problem Statement

The goal of this project is to analyze historical basketball data using statistical methods to understand team performance and patterns. The data set used for this analysis is the NBA Elo dataset obtained from FiveThirtyEight. The statistical methods employed include descriptive statistics, confidence intervals, and data visualizations to study distributions, trends, and team performance.

## Introduction: Your Team and the Assigned Team

This project involves a comparative analysis of two basketball teams. The selected team for analysis is Dallas Mavericks for the years 2013–2015. The assigned team for comparison is the Chicago Bulls for the years 1996–1998.

**Table 1. Information on the Teams**

| **Name of Team** | **Assigned Years** |
| --- | --- |
| Dallas Mavericks | 2013 - 2015 |
| Chicago Bulls | 1996 - 1998 |

## Data Visualization: Points Scored by Your Team

Data visualization is a powerful tool for studying data distributions and identifying trends. It provides a graphical representation of the data, making it easier to spot patterns, anomalies, and general distributions.

A graph of points scored by a team

Description automatically generatedThe plot selected to represent the distribution of points scored by the Mavericks is a histogram.

The histogram was chosen because it effectively shows the frequency distribution of points scored by the Dallas Mavericks from 2013 to 2015, offering insights into the central tendency and spread of the data. The distribution appears bell-shaped, which suggests a relatively normal distribution of points scored. This indicates consistency in the team’s scoring performance during the specified period, with most games falling within a central range of scores. The histogram provides a clear visualization that supports identifying trends and patterns in the data.

## Data Visualization: Points Scored by the Assigned Team

A graph of points scored by the bulls

Description automatically generatedThe plot selected to represent the distribution of points scored by the assigned team (Chicago Bulls) is a histogram.

The histogram was chosen for its ability to clearly display the frequency distribution of points scored by the Chicago Bulls. The distribution appears bell-shaped, suggesting a relatively consistent scoring performance during the years analyzed.

## Data Visualization: Comparing the Two Teams

*A diagram of a box plot

Description automatically generated*Data visualization is also used to compare two different data distributions by overlaying or juxtaposing their visual representations. The plot selected to compare the points scored by The Dallas Mavericks and the Chicago Bulls is a side-by-side boxplot.



The side-by-side boxplot was chosen as it provides a clear comparison of the central tendencies and variability for both teams. It reveals that The Dallas Mavericks had slightly higher median scores compared to the Chicago Bulls, with less variability, as indicated by the narrower interquartile range.

## Descriptive Statistics: Points Scored By Your Team in Home Games

**Table 2. Descriptive Statistics for Points Scored by Your Team in Home Games**

| **Statistic Name** | **Value** |
| --- | --- |
| Mean | 105.68 |
| Median | 105.0 |
| Variance | 123.32 |
| Standard Deviation | 11.1 |

Measures of central tendency (mean and median) provide insights into the typical points scored by the team in home games. The mean and median are close, suggesting a symmetric distribution. The variability, as indicated by the standard deviation and variance, shows moderate consistency in the points scored by the team at home.

The distribution of points scored in home games appears bell-shaped. Based on this skewness, the mean is an appropriate measure of central tendency to represent the data.

## Descriptive Statistics: Points Scored By Your Team in Away Games

**Table 3. Descriptive Statistics for Points Scored by Your Team in Away Games**

| **Statistic Name** | **Value** |
| --- | --- |
| Mean | 103.52 |
| Median | 103.0 |
| Variance | 131.76 |
| Standard Deviation | 11.48 |

The mean and median for away games indicate a similar central tendency, suggesting a symmetric distribution. The distribution appears bell-shaped, and the mean is an appropriate measure of central tendency. Comparing home and away games, the team performs better at home, as evidenced by the higher mean and slightly lower standard deviation, indicating more consistent scores at home.

## Confidence Intervals for the Average Relative Skill of All Teams in Your Team’s Years

**Table 4. Confidence Interval for Average Relative Skill of Teams in Your Team’s Years**

| **Confidence Level (%)** | **Confidence Interval** |
| --- | --- |
| 95% | (1502.02, 1507.18) |

Confidence intervals estimate the range within which the true population parameter is likely to fall. For the years 2013–2015, we are 95% confident that the average relative skill of all teams lies between 1502.02 and 1507.18.

If a different confidence level were used, such as 99%, the interval would be wider, reflecting increased uncertainty. The probability that a team has a skill level less than the Dallas Mavericks is 0.45, indicating that nearly half the teams in the league had a relative skill level below the Mavericks during this period.

## Confidence Intervals for the Average Relative Skill of All Teams in the Assigned Team’s Years

**Table 5. Confidence Interval for Average Relative Skill of Teams in Assigned Team’s Years**

| **Confidence Level (%)** | **Confidence Interval** |
| --- | --- |
| 95% | (1487.66, 1493.65) |

For the years 1996–1998, we are 95% confident that the average relative skill of all teams lies between 1487.66 and 1493.65. Compared to the confidence interval for 2013–2015, the earlier years show a slightly lower average relative skill. This suggests that, on average, teams in the later years demonstrate slightly improved skill levels, reflecting potential advancements in strategies or player performance.

## Conclusion

The statistical analyses conducted in this project provide valuable insights into team performance and league trends. Descriptive statistics highlighted consistent scoring patterns for the Dallas Mavericks, with better performance at home compared to away games. Confidence intervals allowed us to estimate average relative skill levels, showing slight improvements in team skills in later years compared to earlier periods.

These results are significant as they provide a data-driven basis for understanding performance and making informed decisions to improve the team’s strategies. Visualizations effectively highlighted trends and distributions, aiding in the interpretation of the data.