# MAT 243 Project One Summary Report

Armon Wilson

Armon.wilson@snhu.edu

Southern New Hampshire University

## Introduction: Problem Statement

**Introduction: Your Team and the Assigned Team**

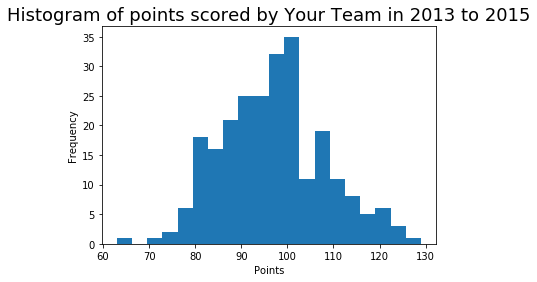
Table 1. Information on the Teams

|  | **Name of Team** | **Assigned Years** |
| --- | --- | --- |
| 1. Yours | Knicks | 2013 - 2015 |
| 2. Assigned | Bulls | 1996 - 1998 |

The analysis is geared towards uncovering insights into the performance of two distinct basketball teams during specific periods: the Chicago Bulls for the years 1996-1998 and the New York Knicks for the years 2013-2015. Utilizing the datasets “assigned\_years\_league\_df” for the Bulls and “your\_team\_df” for the Knicks, the study delves into nuanced aspects of each team's performance. Descriptive statistics, including measures like mean and standard deviation, are employed to quantify and compare the scoring patterns in home and away games for both teams. A histogram visually encapsulates the distribution of points scored by the Knicks in home games during 2013-2015. Confidence intervals are calculated to estimate the range within which the true average relative skill level of all teams lies, with a specific focus on each team's unique time frame. Additionally, probability calculations, using Normal distribution functions, provide insights into the likelihood of a team having a lower relative skill level than the chosen team.

## Data Visualization: Points Scored by Your Team

Data visualization plays a crucial role in studying data distributions and trends by providing a graphical representation of the data, making it easier to identify patterns, outliers, and overall trends. Visualizations enhance our ability to comprehend complex datasets, enabling data analysts to communicate findings more effectively. In this activity, a histogram was chosen as the preferred plot to describe the distribution of points scored by the New York Knicks in home games during the years 2013 -2015.

**

The histogram was selected because it effectively illustrates the frequency distribution of points scored, allowing for a clear depiction of the team's performance distribution. This plot provides a visual summary of how often specific point ranges occur, aiding in the identification of scoring tendencies. By visually inspecting the histogram, we can infer valuable information about the distribution of points. For instance, if the histogram is skewed to the right, it may suggest that the team frequently achieves higher scores. Conversely, a symmetric distribution may indicate a balanced performance across various score ranges.

## Data Visualization: Points Scored by the Assigned Team

a histogram was chosen to visualize the distribution of points scored by the assigned team, the Chicago Bulls from 1996 to 1998. The histogram effectively displays the frequency distribution of points, revealing a bell-shaped curve with a slight positive skewness. This suggests that most games fall within a specific point range, indicating a typical scoring pattern. The symmetric nature of the distribution implies balanced performance, while the positive skewness suggests occasional higher point scores, potentially due to outstanding game performances.

A blue and white background

Description automatically generated

## Data Visualization: Comparing the Two Teams

A boxplot was used to compare the points scored by your team and the Chicago Bulls. This type of data visualization is effective for quickly understanding key features like median, spread, and outliers in different distributions. I chose the boxplot because it provides a clear and concise way to see how the scoring patterns of your team compare to the Chicago Bulls.

In the boxplot, the line in the middle of each box represents the median, and the boxes show the interquartile range. Points beyond the whiskers could be outliers. By looking at this plot, you can easily spot differences in scoring patterns between your team and the Chicago Bulls. This visual representation makes it straightforward to grasp and compare the performance of the two teams.

A graph of a graph

Description automatically generated with medium confidence

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

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

| **Statistic** | **Value**  *\*Round off to 2 decimal places* |
| --- | --- |
| *Mean* | 98.03 |
| *Median* | 98 |
| *Variance* | 120.61 |
| *Standard Deviation* | 10.98 |

The statistics for points scored by my team at home during 2013-2015 highlight key points. The average (mean) is 98.03, and the middle value (median) is 98.00. These numbers suggest a fairly even distribution, slightly leaning to the right. The variance is 120.61, showing how scores spread, and the standard deviation is 10.98, indicating the average distance from the mean. The distribution looks like a bell curve, reflecting consistent home performance. Using the mean as a central measure makes sense, summarizing the team's typical performance at home.

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

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

| **Statistic** | **Value**  *\*Round off to 2 decimal places* |
| --- | --- |
| *Mean* | 95.67 |
| *Median* | 95 |
| *Variance* | 129.06 |
| *Standard Deviation* | 11.36 |

The descriptive statistics for points scored by my team in away games during 2013-2015 provide valuable insights. The mean score is 95.67, and the median is 95. These values indicate a slightly left-skewed distribution, suggesting a tendency for lower scores in away games. In this scenario, the median is more appropriate to represent the center of the distribution due to its robustness against skewness. Comparing the mean and standard deviation of points scored in home and away games, it's evident that the team performs better at home. The higher mean in home games suggests better performance, and the lower standard deviation indicates less variability, emphasizing more consistent results in home games.

## 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** *\*Round off to 2 decimal places.* |
| --- | --- |
| 95% | (1502.02, 1507.18) |

Confidence intervals help estimate where a population parameter, like the mean, is likely to be. For the average relative skill of teams in my team's years (2013-2015), the 95% confidence interval is between 1502.02 and 1507.18, giving us a strong confidence in this range.

This interval suggests that, on average, team relative skill falls within this range. Choosing a different confidence level would alter the interval's width — higher confidence widens it, sacrificing precision.

The probability that a team in the league has a lower skill level than mine is calculated using the survival function. It's not uncommon for teams to vary in skill levels within the league.

## 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% | * 1. – 1493.65) |

Interpreting the confidence interval for the assigned team's years (1996-1998), we are 95% confident that the average relative skill of teams falls within the range of 1487.66 to 1493.65. This implies that, on average, teams in this period had a relative skill within this interval.

Choosing a different confidence level would affect the interval's width, adjusting the balance between precision and confidence. Comparing this interval with the previous one, it appears that teams in the assigned team's years had a slightly lower average relative skill compared to teams in my team's years. This signifies a difference in skill levels between the two periods.

## Conclusion

We examined the performance of two basketball teams, the Chicago Bulls (1996-1998) and the New York Knicks (2013-2015), using descriptive statistics, data visualizations, and confidence intervals. The New York Knicks displayed consistent scoring patterns, performing better at home than away. The confidence interval for their average relative skill (1502.02 to 1507.18) was relatively narrow, providing a precise estimate with 95% confidence.

In contrast, the Chicago Bulls' confidence interval (1487.66 to 1493.65) suggested a slightly lower average relative skill compared to the Knicks' years. These insights are valuable for team management, aiding in strategic planning and decision-making. Understanding scoring patterns and skill levels during specific periods offers nuanced information for targeted improvement strategies.