# MAT 243 Project One Summary Report

Jon Hoover

Jonathan.Hoover1@snhu.edu

Southern New Hampshire University

## Introduction: Problem Statement

In this project, I will be using statistical analysis on a dataset to answer some key questions about a couple NBA teams. The data come from a large historical set of data--the FiveThirtyEight NBA Elo dataset found on Kaggle. The question I am answering is how well my NBA team, the Detroit Pistons, did during the years 2013 to 2015, compared to how well the Chicago Bulls did during the years 1996 to 1998.

The statistical methods I am using are:

* Mean
* Media
* Variability
* Standard Deviation
* Cumulative Distribution Function (CDF)

## Introduction: Your Team and the Assigned Team

Table 1. Information on the Teams

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

## Data Visualization: Points Scored by Your Team

Visualization helps people see and interpret the data quickly. Charts come in several different forms. Some have bars to help visualize how the data is distributed. Others show scatter plots where people can see which way the data is trending over time. The chart below is a histogram, and I chose this chart over the scatter plot, because it shows how often my team scores points ranging from 70 points up to 130 points. It also shows that it is close to a normal distribution, which we could not see in the scatter plot.

Chart, histogram

Description automatically generated

## Data Visualization: Points Scored by the Assigned Team

Below is the chart I chose, a histogram, to represent the Bulls performance for the years 1996 to 1998. The histogram makes it easier to see how well the Bulls scored during these years, and it shows how it resembles a normal distribution as well.

Chart, histogram

Description automatically generated

## Data Visualization: Comparing the Two Teams

When comparing the two teams using an overlapping histogram below, viewers can clearly see how well the Pistons and the Bulls compared in how many points they both scored and how often. This chart works better than scatter plots because the bars are easier to discern than random overlapping dots. And it helps show that both teams had a normal distribution of points scored.

Chart, histogram

Description automatically generated

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

The measures of central tendency, mean, median, variance and standard deviation gives us an idea of what the center of the data and the spread look like. The mean and the median show the center of the data. Below you can see the center is about 99 points and is close to bell-shaped. The variance and standard deviation show the spread. Variance is the squared difference between the mean and the outer edges of the data. Standard Deviation is the square root of the variance and is expressed in the same units as the data. My results show that the standard deviation is 10.89 points. The Empirical Rule states that in a normal distribution, 99.7% of our data falls within three standard deviations of the mean. (Menon, 2022). Therefore, we can conclude that 99.7% of points by the Pistons in home games range between 67 and 132 points.

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

| **Statistic Name** | **Value** |
| --- | --- |
| Mean | 99.48 |
| Median | 99.0 |
| Variance | 118.55 |
| Standard Deviation | 10.89 |

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

The statistics for the Pistons’ away games are very similar to their home games. However, the data is slightly skewed to the left, with a mean of 96.89 points and a median of 98 points. The variance and standard deviation are about the same, where 99.7% of their points were in the range of 66 to 128 points. Therefore, the Pistons played better in their home games than their away games.

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

| **Statistic Name** | **Value** |
| --- | --- |
| Mean | 96.89 |
| Median | 98.0 |
| Variance | 107.38 |
| Standard Deviation | 10.36 |

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

The confidence interval shows us what range we are confident our population mean falls in between, for a particular confidence level. In this case, we are 95% confident that the average skill of all teams during 2013 to 2015 falls between the range of 1283 and 1726. The 95% interval shows the margin of error for two standard deviations from the mean. For less than 95%, it would include a smaller margin of error. And a larger margin of error for greater than 95% confidence level. There is a 19% probability that a given team in the league as a relative skill less than the Pistons for years 2013 to 2015.

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

| **Confidence Level (%)** | **Confidence Interval** |
| --- | --- |
| 95% | (1283.05, 1726.16) |

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

For the years 1996 to 1998, we are 95% confident that the mean skill of all teams is between 1238 and 1744. If the confidence level were lower, the interval range would be narrower. A higher level would make the range wider. Overall, this confidence interval is very similar to the one for the years 2013 to 2015. Therefore, the skill level is about the same for all teams in both sets of years. At about 97%, there is a significantly higher probability that teams during the years 1996 to 1998 have a lower skill level than the Bulls, versus the probability teams had a lower skill level than the Pistons during 2013 to 2015.

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

| **Confidence Level (%)** | **Confidence Interval** |
| --- | --- |
| 95% | (1237.69, 1743.62) |

## Conclusion

In conclusion, The Pistons had about the same average points per game from 2013 to 2015 as the Bulls did from 1996 to 1998. However, when we look at the probabilities that teams ranked lower than each team, we see that the Bulls ranked much higher at 97% than the Pistons did at 19%. This tells me that teams score much higher points on average from 2013 to 2015 than they did from 1996 to 1998.

## Citations

FiveThirtyEight. (April 26, 2019). FiveThirtyEight NBA Elo dataset. Kaggle. Retrieved from https://www.kaggle.com/fivethirtyeight/fivethirtyeight-nba-elo-dataset/

Menon, K. (2022, September 27). *All about the empirical rule in statistics: Simplilearn*. Simplilearn.com. Retrieved January 29, 2023, from https://www.simplilearn.com/tutorials/statistics-tutorial/all-about-the-empirical-rule-in-statistics

*Confidence intervals and margin of error (video)*. (n.d.). Khan Academy. https://www.khanacademy.org/math/ap-statistics/xfb5d8e68:inference-categorical-proportions/introduction-confidence-intervals/v/confidence-intervals-and-margin-of-error