# MAT 243 Project One Summary Report

Nicholas Cleveland

nicholas.cleveland1@snhu.edu

Southern New Hampshire University

## Introduction: Problem Statement

*Discuss the statement of the problem in terms of the statistical analyses that are being performed. In your response, you should address the following questions:*

* *What is the problem you are going to solve?*

*The problem is that we are a basketball data analyst and have been assigned to use descriptive statistics and data visualization techniques in order to study the distributions of key performance metrics of the basketball teams.*

* *What data set are you using?*

*For the data set, we have been given data with the name of the team, points scored by each team in a game, a measure of the relative skill level of the team in the league, and the year when the team played the games.*

* *What statistical methods will you be using to do the analysis for this project?*

*The statistical methods that I’ll be using to do an analysis will be calculating the mean the median, the variance, the standard deviation, and the confidence intervals.*

## Introduction: Your Team and the Assigned Team

*In this project, you picked a team and you were assigned a team to do comparative analysis.*

*See Steps 1 and 2 in the Python script to address the following items:*

* *What team did you pick and what years were picked to do the analysis?*

*The team that I was chose was the Cleveland Cavaliers from 2013 to 2015.*

* *What team and range of years were you assigned for the comparative study? (Hint: This is called the assigned team in the Python script.) Present this information in a formatted table as shown below.*

*For the comparative study, I was assigned the Chicago Bulls from 1996 to 1998.*

Table 1. Information on the Teams

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

## Data Visualization: Points Scored by Your Team

*In the Python script, you created a visualization for the distribution of points scored by your team.*

*See Step 3 in the Python script to address the following items in a paragraph response:*

* *In general, how is data visualization used to study data distributions and trends?*
* *In this activity, you were asked to pick one of the two plots that best describes the data distribution of the variable for your team. Include a screenshot of this plot in your report.*
* *Why did you pick this plot? Explain.*
* *What can you say about the distribution of the variable by visually inspecting this plot? What does this signify?*

* Answer the questions in a paragraph response. Remove all questions and this note before submitting! Do not include Python code in your report.*

## Data Visualization: Points Scored by the Assigned Team

*In the Python script, you created a visualization for the distribution of points scored by the assigned team.*

*See Step 4 in the Python script to address the following items in a paragraph response:*

* *In this activity, you were asked to pick one of the two plots that best describes the data distribution of the variable for the assigned team. Include this plot in your report.*
* *Why did you pick this plot? Explain.*
* *What can you say about the distribution of the variable by visually inspecting this plot? What does this signify?*

* Answer the questions in a paragraph response. Remove all questions and this note before submitting! Do not include Python code in your report.*

## Data Visualization: Comparing the Two Teams

*In the Python script, you created a visualization for the difference in the distributions of points scored by your team and the assigned team.*

*See Step 5 in the Python script to address the following items in a paragraph response:*

* *In general, how is data visualization used to compare two different data distributions?*
* *In this activity, you were asked to pick one of the two plots that best compares the data distributions of your team with the assigned team. Include a screenshot of this plot in your report.*
* *Why did you pick this plot? Explain.*
* *How do the two distributions compare to each other?*

* Answer the questions in a paragraph response. Remove all questions and this note before submitting! Do not include Python code in your report.*

## Descriptive Statistics: Relative Skill of Your Team

*In the Python script, you calculated descriptive statistics on the relative skill of your team. These included the mean, median, variance, and standard deviation for the relative skill of your team.*

*See Step 6 in the Python script to address the following items:*

* *Summarize* ***all*** *statistics in a formatted table as shown below. Use one row for each statistic. You will need to add rows to the table in order to include all of your statistics.*

Table 2. Descriptive Statistics for Relative Skill of Your Team

| **Statistic Name** | **Value** |
| --- | --- |
| Statistic  *(for example, Mean)* | X.XX  *\*Round off to 2 decimal places.* |

* *In general, how are the measures of central tendency and variability used to analyze a data distribution?*
* *Interpret each statistic in detail and explain what it represents in this scenario.*
* *Use the mean and the median to describe the distribution of relative skill of your team.*
  + *Describe the skew: Is it left, right, or bell-shaped?*
  + *Explain which measure of central tendency is best to use to represent the center of the distribution based on its skew.*

* Answer the questions in a paragraph response. Remove all questions and this note (but not the table) before submitting! Do not include Python code in your report.*

## Descriptive Statistics: Relative Skill of the Assigned Team

*In the Python script, you calculated descriptive statistics on the relative skill of the assigned team. These included the mean, median, variance, and standard deviation for the relative skill of the assigned team.*

*See Step 7 in the Python script to address the following items:*

* *Summarize* ***all*** *statistics in a formatted table as shown below. Use one row for each statistic. You will need to add rows to the table in order to include all of your statistics.*

Table 3. Descriptive Statistics for Relative Skill of the Assigned Team

| **Statistic Name** | **Value** |
| --- | --- |
| Statistic  *(for example, Mean)* | X.XX  *\*Round off to 2 decimal places.* |

* *Interpret each statistic in detail and explain what it represents in this scenario.*
* *Use the mean and the median to describe the distribution of relative skill of the assigned team.*
  1. *Describe the skew: Is it left, right, or bell-shaped?*
  2. *Explain which measure of central tendency is best to use to represent the center of the distribution based on its skew.*
* *Use the variance and the standard deviation to compare the distributions of relative skill of your team and relative skill of the assigned team. Which of the two teams has a more consistent skill? How do you know?*

* Answer the questions in a paragraph response. Remove all questions and this note (but not the table) before submitting! Do not include Python code in your report.*

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

*In the Python script, you calculated a 95% confidence interval for the average relative skill of all teams in the league during the years of your team. Additionally, you calculated the probability that a given team in the league has a relative skill level less than that of the team that you picked.*

*See Step 8 in the Python script to address the following items:*

* *Report the confidence interval in a formatted table as shown below.*

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

| **Confidence Level (%)** | **Confidence Interval** |
| --- | --- |
| XX% (for example, 95%) | (X.XX, X.XX)  *\*Round off to 2 decimal places.* |

* *Describe how confidence intervals are generally used in estimating the measures of central tendency for a population.*
* *Provide a detailed interpretation of the confidence interval in terms of the average relative skill of teams in the range of years that you picked.*
* *How would your interval be different if you had used a different confidence level?*
* *What is the probability that a given team in the league has a relative skill level less than that of the team that you picked? Is it unusual that a team has a skill level less than your team?*

* Answer the questions in a paragraph response. Remove all questions and this note (but not the table) before submitting! Do not include Python code in your report.*

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

*In the Python script, you calculated a 95% confidence interval for the average relative skill of all teams in the league during the years of the assigned team. Additionally, you calculated the probability that a given team in the league has a relative skill level less than that of the assigned team.*

*See Step 9 in the Python script to address the following items:*

* *Report the confidence interval in a formatted table as shown below.*

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

| **Confidence Level (%)** | **Confidence Interval** |
| --- | --- |
| XX% (for example, 95%) | (X.XX, X.XX)  *\*Round off to 2 decimal places.* |

* *Provide a detailed interpretation of the confidence interval in terms of the average relative skill of teams in the assigned team’s range of years.*
* *Discuss how your interval would be different if you had used a different confidence level.*
* *How does this confidence interval compare with the previous one? What does this signify in terms of the average relative skill of teams in the range of years that you picked versus the average relative skill of teams in the assigned team’s range of years?*

* Answer the questions in a paragraph response. Remove all questions and this note (but not the table) before submitting! Do not include Python code in your report.*

## Conclusion

*Describe the results of your statistical analyses clearly, using proper descriptions of statistical terms and concepts.*

* *What is the practical importance of the analyses that were performed?*
* *Describe what these results mean for the scenario.*

* Answer the questions in a paragraph response. Remove all questions and this note before submitting! Do not include Python code in your report.*

## Citations

*You were* ***not*** *required to use external resources for this report. If you did not use any resources, you should remove this entire section. However, if you did use any resources to help you with your interpretation, you* ***must*** *cite them. Use proper APA format for citations.*

Insert references here in the following format:

Author's Last Name, First Initial. Middle Initial. (Year of Publication). Title of book: Subtitle of book, edition. Place of Publication: Publisher.