# MAT 243 Project Two Summary Report

Ryan Hatch

ryan.hatch@snhu

Southern New Hampshire University

**Notes:**

* Replace the bracketed text on page one (the cover page) with your personal information.
* You will use your selected team for all three projects

## Introduction: Problem Statement

In this project, I am trying to statistically validate the claims made on the performance of the Los Angeles Lakers basketball team in the period of 2013-2015. The dataset used contains point scoring records, relative skill ratings of the teams, or the Elo rating, and game outcomes for the teams and seasons, which basically is a historical record of NBA games. The data frame for the Bulls (assigned team,) is “assigned\_team\_df” and the data frame for the Lakers (my chosen team) is “your\_team\_df.” I will be using four variables in the data set to study, the points scored by the team in a game, the measure of relative skill level of the team in the league, the year when the team played the games, and the name of the NBA team.

In order to parse the data that I needed to solve my problem, I applied a number of statistical methods in the analysis, including the testing of the hypothesis. In the analysis, I used a one-sample t-tests to review whether the average relative level of the skills of the two years and the average points scored by the Lakers in the chosen years of 2013 – 2015, and how they differ from hypothesized values. I carried out a proportion z-test to test if the proportion of games won when the score was over a certain number of points is actually a match as it claimed. I then went to conduct a two-sample t-test to establish if the relative skill level mean difference of the Lakers and that of the Chicago Bulls was significant during the years that were used. The results of my analysis would have either confirmed the claims for the performance of the teams or it would have rejected them and help guide with making more informed and tailored decisions in order to improve the strategies used by the teams.

## Introduction: Your Team and the Assigned Team

*For my analysis I chose the Los Angeles Lakers for the years between 2013 – 2015. The assigned team I was given for comparative study was the Chicago Bulls, which covered the years between 1996 – 1998.*

Table 1. Information on the Teams

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

## Hypothesis Test for the Population Mean (I)

*In hypothesis testing, I used python in order to compute statistical methods to double verify claims about a population mean. This is considered a null and alternative hypothesis test, which ends up forming a claim that needs to be verified. It sets up the null hypothesis, which is representing the claim that one would want to test against an alternative hypothesis which would represent the outcome he or she would expect or would like to prove. The level of significance is then chosen to determine the threshold for rejecting the null hypothesis.*

*For this particular hypothesis test, the null and alternative hypotheses were:*

*Null Hypothesis (H0): The Lakers' average relative skill level for the years 2013-2015 is 1340 and below.*

*(H0: μ ≤ 1340)*

*Alternative Hypothesis (H1): The average number of relative skills levels of Lakers players across the span of the years 2013-2015 is not less than 1340.*

*(μ > 1340)*

*The degree of significance set for this test was at 5%, which meant that there was a 5% chance of wrongfully rejecting the null hypothesis.*

Table 2: Hypothesis Test for the Population Mean (I)

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | 2.94 |
| P-value | 0.0034 |

*The p-value, 0.0034, lies in the rejection region as it is less than the level of significance, 0.05; therefor, rejecting the null hypothesis. That is, for this specified level of significance, I would conclude that there is enough statistical evidence to say that the average relative skill of the Lakers team during 2013-2015 is more than 1340.  
  
The practical significance of this result is that it demonstrates the Lakers' performance in terms of relative skill level was above a critical low threshold during the specified years. This information can be valuable for the team's management when making decisions in terms of team improvement and strategic planning.*

## Hypothesis Test for the Population Mean (II)

*The null hypothesis tested the claim that the means of the average points of the Lakers from the year 2013 to 2015 had not been less than 106 points per game. I conducted this test at the 1% level of significance, assuming an unknown population standard deviation for the scores registered.*

*Null Hypothesis (H0): The Lakers have averaged 106 or more points for the seasons 2013-2015.*

*(H0: μ ≥ 106)*

*Alternative hypothesis (H1): The average of Lakers' points during 2013-2015 is less than 106.*

*(H1: μ < 106)*

*With the level of significance at 1% (or, α = 0.01) and since the P-value (or, 0.0000) is way less than the level of significance (0.01), I rejected the null hypothesis. This was because there was enough evidence to prove the claim of the coach, which was that the average number of points scored by the Lakers for 2013-2015 is less than 106. The practical implication of this is that it points out the poor scoring performance by the Lakers during the specified years, which is below the threshold associated with strong performance in the regular season. This was valuable information for the management team and coaching staff in order to properly identify and tailor what may need to be done in terms of offensive strategy and player development in order to get better results.*

Table 3: Hypothesis Test for the Population Mean (II)

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | -6.91 |
| P-value | 0.0000 |

## Hypothesis Test for the Population Proportion

*The claim that was made while analyzing the performance of the Los Angeles Lakers from 2013 to 2015 was the efficiency of this team in the games where they scored more than 102 points.*

*The management claimed that the Lakers would have an 0.90 probability of winning the game if they scored above 102 points per game. In order to check their hypothesis, I tested the population proportion was at the 5% level of significance.*

*Null Hypothesis (H0): The proportion of the Lakers won games, having scored more than 102 points, is equal to 0.90*

*(P=0.90).*

*Alternative Hypothesis (H1): The proportion of games won by the Lakers when they score over 102 points should not be 0.90.*

*(P ≠ 0.90)*

*This analysis was done at a level of significance at 5%, which meant that there was a 5% chance that if the null hypothesis was actually true, it would be rejected. And based on the test statistic and the P-value, the null hypothesis would be rejected. The P-value of 0.0000 is very small when comparing it to the level of significance being at 0.05. With this, it strongly suggests a rejection of the null hypothesis. This hypothesis test brought me to the conclusion that the claim made, which was that the Lakers won 90% of the games whenever they had scored over 102, is not possible based off of the data. This has practical significance because of the fact that the differences between these scores, both the perceived and true performance, become quite large in high scoring games.*

Table 4: Hypothesis Test for the Population Proportion

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | -2235.66 |
| P-value | 0.0000 |

## Hypothesis Test for the Difference Between Two Population Means

*You were asked to compare your team’s skill level (from its years) with the assigned team’s skill level (from the assigned time frame). You tested the claim that the skill level of your team is the same as the skill level of the assigned team, using a 1% level of significance.*

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

* In general, how is hypothesis testing used to test claims about the difference between two population means?
* Summarize all important steps of the hypothesis test. This includes:

1. Null Hypothesis (statistical notation and its description in words)
2. Alternative Hypothesis (statistical notation and its description in words)
3. Level of Significance
4. Report the Test Statistic and the P-value in a formatted table as shown below:

Table 5: Hypothesis Test for the Difference Between Two Population Means

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | 49.51 |
| P-value | 0.0000 |

1. Conclusion of the hypothesis test and its interpretation based on the P-value

* What are the implications of your findings from this hypothesis test? What is its practical significance?

* 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.