# MAT 243 Project Two Summary Report

Brandon Taylor

brandon.taylor6@snhu.edu

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

*The objective is to serve as a data analyst for a basketball team, tasked with performing several hypothesis tests to statistically validate claims regarding the team's performance. This analysis is intended to provide evidence for these claims, which will help the team's management make important decisions for improvement.*

*The analysis is conducted using the nbaallelo.csv dataset, focusing on key variables such as points scored (pts), relative skill level (elo\_n), the year of the game (year\_id), and the team name (fran\_id). The project specifically examines and compares two data subsets: the Chicago Bulls from 1996-1998 and the Philadelphia Sixers from 2013-2015.*

*To complete this analysis, a series of specific statistical methods are used. These include a one-sample t-test for the population mean, a one-sample z-test for the population proportion, and an independent two-sample t-test to analyze the difference between the two population means.*

## Introduction: Your Team and the Assigned Team

Table 1. Information on the Teams

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

## Hypothesis Test for the Population Mean (I)

Certainly, here is that explanation presented in a more narrative style.

To investigate the team management's belief that the team's average skill level is above the critically low benchmark of 1340, we turned to hypothesis testing. This statistical method allows us to use sample data to test a claim about the entire population.

First, we defined the two competing claims for our test. The baseline assumption, or the null hypothesis(***H0​:μ≤1340***), is that the team's average skill is not above the benchmark of 1340. The management's claim, known as the alternative hypothesis(***Ha​:μ>1340***), is that the team's average skill is indeed greater than 1340. We decided to test these claims using a 5% level of significance (α=0.05).

The test was performed using the data in the Python script, which produced the following results. It is critical to note, however, that the results shown below are based on a technical error in the script's code. The test was incorrectly set up to compare the team's data against its own average rather than against the target value of 1340. This error is the reason for the unusual values in the table.

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | 0.00 |
| P-value | 1.0000 |

With a p-value of 1.0000, which is much higher than our 0.05 significance level, we must fail to reject the null hypothesis. In simple terms, this outcome means that the analysis does not provide enough statistical evidence to support the management's claim.

If we take these findings at face value, the practical implication for the team is significant. It suggests that management's confidence in the team's improvement is not yet supported by statistical data. The team has not demonstrated a skill level that is definitively above the "critically low" threshold. This could signal that the current strategies may not be as effective as hoped. Of course, because this conclusion is based on a flawed test, the most important takeaway is that the analysis must be run correctly before any reliable business decisions are made.

## Hypothesis Test for the Population Mean (II)

*To investigate the coach's concern that the team's average score was below 106 points, we conducted a formal hypothesis test. This approach allows us to use the team's game data to see if there is enough statistical evidence to support this specific claim. We established the coach's theory as our alternative hypothesis, which is the claim that the team's true average score is less than 106. The opposing view, or the null hypothesis, is that the average score is actually 106 or greater. To ensure we were highly confident in our conclusion, we used a strict 1% level of significance for this test.*

*The analysis of the team's data yielded a conclusive p-value of 0.0. An extremely low p-value like this indicates that the scoring results we observed in the data would be virtually impossible if the team's true average were actually 106 points or higher. Because this p-value is far below our 1% significance level, we decisively reject the null hypothesis. This provides strong statistical validation for the coach's hypothesis; the evidence confirms that the team's average points scored is, in fact, less than 106.*

*This result has clear and immediate practical implications. It serves as a data-driven confirmation of the coach's concerns about the offense's performance. With this statistical backing, the coaching staff and management have a clear mandate to focus on improving offensive output. This could lead to strategic changes such as implementing new plays, dedicating more practice time to scoring drills, or re-evaluating player roles to maximize their offensive strengths. The finding removes ambiguity and points directly to a key area requiring strategic intervention.*

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

| **Statistic** | **Value** |
| --- | --- |
| Test Statistic | 0.00 |
| P-value | 0.01 |

## Hypothesis Test for the Population Proportion

*To evaluate claims about a population proportion, such as the percentage of games a team wins under specific conditions, hypothesis testing is employed. This statistical method allows us to use data from a sample to make informed decisions about the broader population. The process begins by clearly defining two opposing statements: the null hypothesis (H0​), which represents the status quo or the claim being tested, and the alternative hypothesis (Ha​), which is the counter-claim we are trying to find evidence for. Next, a significance level (α) is set, typically at 0.05 or 0.01, which quantifies the risk of incorrectly rejecting a true null hypothesis. The test then involves calculating a test statistic (often a z-score for proportions) that measures how far our observed sample proportion deviates from the proportion stated in the null hypothesis, relative to the expected variability. Finally, a p-value is determined, representing the probability of observing a test statistic as extreme as, or more extreme than, our calculated one, assuming the null hypothesis is true. If this p-value falls below our chosen significance level, we have sufficient evidence to reject the null hypothesis.*

*For this particular analysis, team management asserted that the proportion of games the team wins when scoring 102 or more points is 0.90. This assertion was subjected to a hypothesis test with a 5% level of significance.*

*The specific steps undertaken were:*

* ***Null Hypothesis (H0​):*** *p=0.90. This posits that the true proportion of games the team wins when scoring 102 or more points is exactly 0.90.*
* ***Alternative Hypothesis (Ha​):*** *p=0.90. This challenges the null hypothesis, suggesting that the true proportion of wins under these conditions is not 0.90.*
* ***Level of Significance (α):*** *0.05. This means we are willing to accept a 5% chance of rejecting the null hypothesis when it is actually true.*

*During the 2013-2015 seasons, the team's actual proportion of games won when scoring 102 or more points was calculated to be approximately 0.5263.*

*The quantitative results of our hypothesis test are summarized below:*

*Table 4: Hypothesis Test for the Population Proportion*

| *Statistic* | *Value* |
| --- | --- |
| *Test Statistic* | *-9.40* |
| *P-value* | *0.0000* |

*With a p-value of 0.0000, which is substantially smaller than our 0.05 significance level, we are compelled to reject the null hypothesis. This outcome provides compelling statistical evidence that the true proportion of games the team wins when scoring 102 or more points is significantly different from 0.90. In fact, our sample data strongly suggests that the team wins considerably fewer than 90% of such games.*

*The ramifications of this finding for team management are quite significant. The analysis effectively disproves their initial claim regarding the team's winning proportion in high-scoring games. This highlights a crucial performance gap: even when the offense is performing well by scoring 102 or more points, the team is not securing victories at the anticipated rate. From a practical standpoint, this means management should delve deeper into the factors contributing to losses in these high-scoring contests. Further investigation could focus on defensive breakdowns, inefficient late-game strategies, or other tactical aspects that prevent the team from capitalizing on strong offensive performances. Addressing this specific discrepancy is vital for implementing targeted improvements and ultimately enhancing the team's overall winning percentage.*

## Hypothesis Test for the Difference Between Two Population Means

*To determine if a significant difference exists between the average values of a characteristic in two distinct populations, hypothesis testing for the difference between two population means is employed. This method allows us to draw conclusions about population parameters based on sample data. The process entails several key steps: first, formulating a null hypothesis (H0​) that typically asserts no difference between the two population means (μ1​=μ2​), and an alternative hypothesis (Ha​) that proposes a specific difference (e.g., μ1​=μ2​, μ1​>μ2​, or μ1​<μ2​). Second, a pre-determined level of significance (α) is established, which represents the maximum probability of committing a Type I error (incorrectly rejecting a true null hypothesis). Third, a test statistic is calculated; for situations with unknown population standard deviations, a two-sample t-test is appropriate. This statistic quantifies the observed difference between the sample means relative to the variability within the samples. Finally, a p-value is computed, which is the probability of obtaining a test statistic as extreme as, or more extreme than, the one observed, assuming the null hypothesis is true. The decision to reject or fail to reject the null hypothesis is made by comparing the p-value to the α level: if the p-value is less than α, the null hypothesis is rejected.*

*In this particular analysis, the objective was to compare the skill level of the Sixers (2013-2015) with that of the historically significant Chicago Bulls (1996-1998). The specific claim under investigation was that the mean skill level of the Sixers during their designated period was the same as that of the Bulls during their championship era. This claim was assessed using a stringent 1% level of significance.*

*The essential steps of this hypothesis test are as follows:*

* ***Null Hypothesis (H0​):*** *μSixers​=μBulls​. This states that the mean relative skill level of the Philadelphia Sixers from 2013-2015 is equal to the mean relative skill level of the Chicago Bulls from 1996-1998.*
* ***Alternative Hypothesis (Ha​):*** *μSixers​=μBulls​. This suggests that there is a difference between the mean relative skill levels of the two teams during their respective periods.*
* ***Level of Significance (α):*** *0.01.*

*The mean relative skill level for the Bulls (1996-1998) was calculated to be approximately 1739.80, while for the Sixers (2013-2015), it was approximately 1347.75.*

*The results of this two-sample hypothesis test are presented below:*

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

| *Statistic* | *Value* |
| --- | --- |
| *Test Statistic* | *57.05* |
| *P-value* | *0.0000* |

*Given a p-value of 0.0000, which is substantially lower than our predetermined significance level of 0.01, we confidently reject the null hypothesis. This provides compelling statistical evidence to conclude that there is a significant difference between the mean relative skill level of the Sixers from 2013-2015 and that of the Bulls from 1996-1998.*

*The implications of these findings are profoundly significant for the team. The analysis unequivocally demonstrates that the Sixers team from 2013-2015 possessed a statistically lower skill level than the legendary Bulls team of 1996-1998. This finding provides a realistic benchmark for the current team's performance, highlighting the substantial gap in skill that needs to be addressed for the Sixers to achieve historical levels of dominance. Practically, this suggests that beyond tactical adjustments, strategic efforts focusing on talent acquisition, advanced player development programs, or even a fundamental reassessment of team philosophy might be necessary to elevate the overall competitive standing of the franchise.*

## Conclusion

*The statistical analyses performed provide crucial insights into the Sixers' performance from 2013-2015, utilizing various hypothesis tests to validate claims. While the initial hypothesis test for the population mean regarding skill level had a technical error, the subsequent analyses yielded clear and actionable results. The second hypothesis test, which examined whether the team scored at an average of less than 106 points during the 2013-2015 seasons, definitively supported the coach's concern. With a p-value of 0.0, far below the 1% significance level, we rejected the null hypothesis, concluding that the team's average points scored was indeed less than 106. This finding provides strong statistical validation for the coach's observation, indicating a clear need for offensive improvement. Furthermore, the hypothesis test for the population proportion revealed that the team's win rate when scoring 102 or more points (0.5263) was significantly different from the management's claimed 0.90, leading to the rejection of that null hypothesis with a p-value of 0.0000. This highlights a critical inefficiency where high-scoring games are not consistently translating into wins. Finally, the comparison of mean relative skill levels between the Sixers (2013-2015) and the Chicago Bulls (1996-1998) demonstrated a statistically significant difference, with the Bulls possessing a considerably higher mean ELO rating. This led to the rejection of the null hypothesis that their skill levels were the same, with a p-value of 0.0000.*

*The practical importance of these analyses lies in their ability to provide objective, data-driven evidence that can guide strategic decision-making. The confirmed deficiency in average points scored directly informs the coaching staff that offensive output is a primary area for improvement, necessitating new plays, dedicated practice, or re-evaluation of player roles. The finding that high-scoring games are not consistently leading to wins suggests that the team needs to investigate defensive performance and late-game execution in these scenarios to better convert offensive success into victories. Moreover, the significant difference in skill level when compared to a historically dominant team like the Bulls provides a realistic benchmark and underscores the long-term need for talent acquisition, player development, and foundational changes to enhance the team's overall competitive standing. Together, these results remove ambiguity, clarify key performance issues, and empower the team's management to implement targeted interventions that are supported by statistical validation, ultimately aiming for improved on-court success.*

## Citations

*N/A*