**MAT 243 Project One Summary Report**

Kayla Casper-Dixon

Kayla.Casper-Dixon@snhu.edu

Southern New Hampshire University

With this study the intent is to use key performance metrics to look for areas of improvement in the overall team performance. Looking over key data points from the years 2013 – 2015 we can have a large sample of data to look for common trends in the team’s performance and comparatively how they matched up in skill to various opposing teams. We will be looking into the central tendencies i.e., the mean, median, mode, variance, and standard deviation for the data overall. Confidence intervals will also be found for every team in the specified years to build a confidence to how accurate the data is. We will also be using the cumulative density function to find the probability of a team having a lower average relative skill in comparison to our team.

For our team the Timberwolves we will be accessing data from between the years 2013- 2015. For the team used to show a comparison the Bulls data was selected and was pulled from the years 1996-1998.

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

For an easier way to clearly see the data that shows the frequency of points scored for each game a histogram was chosen. This chart allows us to see which point score we tend to trend to with our highest frequencies falling between 90 to 100 points per game. We do have an outlier with the game scoring over 140 points which will pull our mean further to the right meaning the data is skewed right. This data also shows us which frequencies have the highest Chart, histogram

Description automatically generatedmode or the most frequently scored number of points.

For our comparison team we have another histogram to clearly show the frequency of points scored during the period of 1996 -1998. We can see that they more consistently score between 100 and 110 points. The data for this chart resembles close to a normal distribution with the data not quiet skewing left or right. The higher frequency scores trending closes to the Chart, histogram

Description automatically generatedmedian of the graph.

To compare the data samples between the two teams a box and whisker was chosen as an overlapping histogram was not easy to read. With this graph you can clearly compare the two ranges of data and see where the median falls between the two. In comparison we see our team does in fact have a median lies at around 100 points and the Bulls fall above that. The Q1 and Q3 for our data is a bit lower than the Bulls meaning that compared to their data median between our lowest score and the median is lower as well as our upper bound as well. The Bulls do not have an outlier that would throw the data from their range off, but our team has one on the upper side of the scoring that would pull our overall mean slightly higher because of it and may throw our Chart, box and whisker chart

Description automatically generateddata off if not taken into consideration.

Our central tendencies show us the overall trend of the data as seen in the chart below. The mean gives us the overall average of the team’s relative skill. The median gives us the true middle of the data. Our mean and median are close together meaning that the bulk of the data is close to the middle and would lead us to believe this is a normal distribution, but due to our outlier game from before and the high variance and large standard deviation confirms that we have a high possibility of skewed data. Due to the large amount of variance in the data the median would be best for the representation of our true center of the data.

| **Statistic Name** | **Value** |
| --- | --- |
| Mean | 1448.02 |
| Median | 1449.49 |
| Variance | 7013.67 |
| Standard Deviation | 83.75 |

For our comparison team we see a much different set of information as seen in the chart below. The mean and median fall close to one another well within one standard deviation of the data. This confirms that our data is closer to that of a normal distribution or a bell-shaped curve. As the mean is close to the median, we can confirm that the Bulls during this time range did not have any abnormal games that created any outliers that would skew the data. Because of the closeness in data either the mean or the median could be used. Given that the Bulls have a smaller variance and smaller standard deviation they are much more constant skill wise than our team.

| **Statistic Name** | **Value** |
| --- | --- |
| Mean | 1739.8 |
| Median | 1751.23 |
| Variance | 2651.55 |
| Standard Deviation | 51.49 |

Our confidence level at 95% shows our confidence interval that the population mean lies within the range of 1502.02 and 1507.18. Which is to say that we are 95% certain the population mean is within this range. Given that most times the true population mean is unknown we can use confidence intervals to give a range that mean is most likely to be in with 95% probability. Meaning that there is a 5% chance the value is not in this range. This means that we can say with 95% confidence that the average of all the team’s average skill level is between 1502.02 and 1507.18 with the data between 2013-2015.

If we were to use a confidence level of 99% our range would increase as to increase our confidence, we must have a wider spread of the data. If we reduced the confidence level our confidence interval would get tighter as we are more laxed with how certain we are in the range. When finding the probability that a team would have a lower level of skill in comparison to our team, we used a cumulative density function to look for the probability of being lower or equal to our own score. The result was that it is 31% probable that the team we are facing will have a lower relative skill level to our team. Which means that a team facing a team with a higher skill level is much more likely and facing someone with a lower level will be a bit unusual.

| **Confidence Level (%)** | **Confidence Interval** |
| --- | --- |
| 95% | (1502.02, 1507.18) |

For the data ranging from 1996-1998 the confidence interval showed that the population average of all the teams skill level was between 1487.66 and 1493.65. This means within 95% confidence we can say that the overall level of the competing teams in 1996-1998 is lower than the skill level for the teams in 2013-2015 range. If we were to increase our level over confidence, then the interval would have to become a wider range and tight if we were to go for a lower level of confidence. With the Bulls consistent playing and higher mean skill level the probability of a team they faced in the time having a lower level was 97%. Meaning they were almost always at a higher skill level during this time.

| **Confidence Level (%)** | **Confidence Interval** |
| --- | --- |
| 95% | (1487.66, 1493.65) |

For our analysis we compared our team’s data for 2013 – 2015 to data from the Bull’s from 1996-1998. We have found that as a team we are much less consistent with the level of skill per game. With our high variance in our skill level per game shows that we are rather inconsistent with our skill level. Given that the mean for the overall skill level of all teams in the years 2013-2015 are slightly higher than they were in 1996-1998 means that the skill level across the teams has gotten better, meaning that teams are on average playing better than they were in 1996-1998, but that our skill level is on the lower side of the average. This tells us that we are playing with lower skill level than the average relative skill level of the entire range of teams. We will be outmatched close to 69% of the time.