Using T-tests to Compare Top Teams in March Madness

Daniel Zhang

January 11, 2023

Abstract

In Basketball many factors are used to measure team stats like Effective Field Goal Percentage, Offensive Rebound Rate, and Free Throw rate. Since the top teams that play in March Madness are already well above average in these factors, we are interested in the places where the finalists excel in order to see what gives these teams so much success during March Madness. In other words, we are interested in what makes great teams succeed against other top teams.

Introduction

What separates top basketball teams from the rest? Statistics like Effective Field Goal Percentage (EFG), Percentage of free throws (FTR), and Turnover Percentage Allowed (TOR) are all used to get a measure of the top teams in the country. The Teams that make it far in the NCAA are the best of the best, and their stats are at the very top of their division. This means that a lot of their stats are clustered closely together with less variation compared to the rest, making it harder to spot significant differences in top teams. In this analysis, we are interested in finding these significant differences by using T-tests in order to get an idea of what the top teams are doing differently.

Methodology

We use a T-test for differences in means. To use a T-test we must show that the data sets are normally distributed. We do this using normal Q-Q plots. These compare randomly generated, independent standard normal data (the line) to the data set we give it (the points). Looking at these plots we can say that most of the features are normally distributed. Some features like Power Rating (BARTHAG) definitely aren’t normally distributed seeing how the points clearly deviate from the normal line. Features like Turnover Percentage Committed (TORD) and Free Throw Rate Allowed (FTRD) are skewed right in some places.

Chart

Description automatically generatedChart, line chart

Description automatically generated

Chart

Description automatically generated

Next, we run the T-tests using the t.test function in R. For these tests, the null hypothesis is that the true difference in means is 0 and the alternative hypothesis is that the true difference is not 0. We will be using a significance level of 0.05, which means that the probability of rejecting the null hypothesis when it is true is 5%.

Data

We collected data from the ‘College Basketball Dataset’ by Andrew Sundberg on Kaggle. It contains data from the 2013-2021 Division I college basketball seasons. For our analysis, we looked at the 2013 to 2019 part of the dataset.

Results

Looking at boxplots there seems to be a trend where the teams that make it farther in the tournament have better stats on average. We are interested in which of these differences are significant and which could be there because of change so we look for a p-value under 0.05.

For Adjusted Offensive Efficiency (ADJOE) we got a p-value of 0.03066 and for Adjusted Defensive Efficiency (ADJDE) we got a p-value of 0.02256 when comparing the Champions to those who were knocked out in the Sweet Sixteen. This means that there is a difference in the Champions ADJOE and ADJDE compared to the people who only made the Sweet 16.

Chart, bar chart, histogram

Description automatically generatedChart, bar chart, histogram

Description automatically generated

For Effective Field Goal Percentage Shot (EFG\_O) and Effective Field Goal Percentage Allowed (EFG\_D) we got p-values of 0.06466 and 0.02134 when comparing the champions to those knocked out in the Round of 64. Although 0.06466 is not under 0.05 it is very close which means that it is still likely that there is a difference in means in EFG\_O, but we can’t say for certain as we can with EFG\_D.

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

For Turnover Percentage Allowed (TOR) and Turnover Percentage Committed (TORD), the bar graphs looked a lot flatter which means that it would be less likely to see a difference in mean. The T-test reflected this giving p-values of 0.01569 and 0.3846 respectively. This shows that there was a difference in TOR depending on if the team won or was eliminated in the round of 64, but there wasn’t a difference in TORD for those teams. The T-test for TORD wouldn’t have been considered anyways due to the data not being normal, resulting in the T-test working differently than intended.

Chart, bar chart, histogram

Description automatically generatedChart, bar chart, histogram

Description automatically generated

For Offensive Rebound Rate (ORB) and Offensive Rebound Rate Allowed (DRB), the bar graph looked more flat, so it would be unlikely that there would be a difference in mean. The P-values were 0.2543 and 0.9765 which means that we cannot say that there is a difference in mean ORB and DRB between the Champions and the teams that were eliminated in the round of 64.

Chart, bar chart

Description automatically generatedChart, bar chart, histogram

Description automatically generated

For Free Throw Rate (FTR) and Free Throw Rate Allowed (FTRD), we saw P-values of 0.1901 and 0.04755. Unfortunately the data for FTRD was seen to not be normal which makes this P-value more ambiguous. This means that we cannot rely on this P-value so we cant say that Free Throw Rate Allowed had a different mean in the teams that won and the teams that were knocked out in the round of 64.

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

For Two-Points Shooting Percentage (TwoP\_O) and Two-Points Shooting Percentage Allowed (TwoP\_D), we got P-values of 0.09911 and 0.0877. We cannot say that there is a difference in means of the Champions and teams eliminated in the round of 64 in the Two-Point statistics but seeing this P-value is still rare (under 0.1) which means that there still could be a difference.

Chart, bar chart, histogram

Description automatically generatedChart, bar chart, histogram

Description automatically generated

For Three-Point Shooting Percentage (ThreeP\_O) and Three-Points (ThreeP\_D) Shooting Percentage Allowed, we got P-values of 0.1219 and 0.07419. We cannot say that there is a difference in mean for the Champions and the teams that were eliminated in the round of 64. But there still may be a difference in Three-Point Shooting Percentage Allowed although less likely.

Chart, bar chart

Description automatically generatedChart, bar chart

Description automatically generated

Lastly, for Adjusted Tempo (ADJ\_T) the bar graph was rather flat, so there would most likely not be a difference. The P-value was 0.5794 which reflected this.

Chart

Description automatically generated

Discussion

A possible fault in the analysis is that there is less data for each round of the tournament because half of the get eliminated and there are only 7 data points for the Champions of the tournament. This means that it is harder to get an estimate of the true mean and variance for teams that make it farther in the tournament and as such the t-test could be a little less reliable.

In addition, most of the T-tests done compared the winners of the tournament to those who were eliminated in the round of 64. This was because for most of the features, the P-value when comparing teams that got farther in the tournament did not yield low P-values.

For Future work, we could take the means statistics for each of the groups and number each of the groups to later run a linear regression to see the change in each of the factors. We could maybe use this to create a model to predict how far a team will go during March Madness.

We used the ggplot library and other functions built into R. Their results can be found on GitHub