Applying Inferential Statistics:

Understanding what a normal NBA players Win Share contribution to an organization would be a huge benefit to an organization. What this could allow for is advanced insight towards expectations of your players as well as how to properly compensate these players for their on court presence. This should result in a better understanding of what average players in the NBA are contributing towards the teams’ success, then a better understanding of what the pay range for certain contributions should be and therefore help an organization that employs a maximum of 15 players make the best decisions when negotiating contracts.

To do this, we’re looking closely at the Win Share statistic from our data and learning more about the correlation to Salary. The hypothesis is as follows:

H0 : There is no correlation between Win Shares and Salary

H1 : Salary is correlated to Win Shares.

I conducted multiple correlation tests: Spearman’s, Pearson’s and Kendall’s. A Spearman Correlation is a nonparametric measure of rank correlation(statistical dependence between the rankings of two variables). It assesses how well the relationship between two variables can be described using a monotonic function.

The Spearman correlation between two variables is equal to the Pearson correlation between the rank values of those two variables; while Pearson's correlation assesses linear relationships, Spearman's correlation assesses monotonic relationships (whether linear or not).

Kendall's tau coefficient (after the Greek letter [τ](https://en.wikipedia.org/wiki/%CE%A4)), is a statistic used to measure the ordinal association between two measured quantities. A tau test is a non-parametric hypothesis test for statistical dependence based on the tau coefficient.

The results are as follows:

Spearman’s correlation: 0.51514

Samples are correlated (reject H0) p = 0.000

Pearson’s correlation: 0.53543

Samples are correlated (reject H0) p= 0.000

Kendall correlation coefficient: 0.362

Samples are correlated (reject H0) p= 0.000

The most important takeaway here is that in all three tests, we rejected the null hypothesis.

Summarizing:

After initial analysis, we can see that using ONLY the Win Shares as a feature for prediction yields a relatively low accuracy. A strong positive that came from this testing was statistically proving that there is a correlation between Win Shares and Salary. One thing to note that was extremely important is determining what value to use to represent the average of the data.

Initially, I used the mean, but this value was likely skewed because of the distribution we saw earlier. From there, I trimmed the data and removed the top and bottom 15% to analyze that new average value. The trimmed mean was reduced significantly, which was expected. The median is the value I chose to represent an average data point for both WS and Salary moving forward.

Once this was completed, I tested for correlation between WS and Salary using a variety of tests. In each case, we rejected the null hypothesis and can say with confidence that there is a correlation between WS and Salary. To be even more accurate, there is a positive correlation, and Pearson's test of correlation, a .53543 correlation between WS and 2018-19 Salary.

The goal moving forward is to utilize the data and Win Shares stat to predict the upcoming Salary of NBA players.

The goal moving forward is to utilize the data and Win Shares stat to predict the future salaries of NBA Players.