Capstone 1: Inferential Statistics Steps

Overall statistic calculations:

* Calculated the means and standard deviations of squad values based on league. Not surprisingly, the Premier League has the highest mean.
* Calculated the means and standard deviations of goals in each league. Primera Division had the highest mean in goals scored. Serie A came closely in second, which is surprising since Italians have a fame for being more defensive.

Comparing Teams that qualified to the Champions League with those that do not.

* Mean squad value differences are vast
* Calculated misclassification rate and found that 14.2% would be misclassified if you tried to use squad value to guess if a team qualifies to CL.
* Calculated a Cohen’s D : 2.87
* Ran regression:
  + Dep variable: Wins
  + X = Squad Value in Millions
  + X coefficient: .0295
* Mean Difference confidence interval: 95 percent chance that mean difference in team values (Qualifies - Not Qualified ) is within: (222.651, 402.409) with Margin of error: 89.879
* Hypothesis Test: Null H: Testing a null hypothesis that clubs that qualify to CL and those that do not are equal, we get a t-statistic is 11.313 and the p-value is 0.0000000.