In the last paper we examined the peculiar observation that for some reason homes with pools had a lower average price than homes without pools. We argued that since homes with pools are typically associated with luxury living and given the fact that there is the additional cost of a pool this strange difference in cost is worth further investigation. We noticed that the average price of homes with a pool is roughly 202.79 thousand dollars which is substantially less than the average price of all homes of 221.1 thousand dollars. The test conducted to check whether the two means are statistically different was relatively straightforward given that we assumed a normal distribution. However the challenge of substantiating or rejecting this observed behavior in the non-parametric test case has to be rather addressed differently.

In this study we will investigate the same anomaly in prices but without making an assumption about the distribution of prices (non-parametrically). There are several non-parametric tests that are applicable in this situation for example, Kolmogorov-Smirnov Test and Wilcoxon Signed Ranks Test. However we will perform a Wilcoxon Mann-Whitney (Rank-Sum) Test or the U-Test because of its simplicity and the fact that it applies well in our situation. Simply put the Rank-Sum test is a non-parametric test for assessing whether any two samples of observations come from the same distribution. In this case we would like to see whether the sample of homes with pools and sample of home without pools fall in the same distribution or not.

The procedure to conduct this test is not significantly complicated as it basically requires ranking of the sample data and summing up the ranks. The main idea of this test is that If two groups if data (pools/no pools) come from the same distribution, but we have just randomly assigned labels to them, then the values in the two different groups should have values somewhat equally distributed between the two labels. The test will as the previous one will be two sided and we will maintain the alpha value of the previous experiment/analysis (alpha=.05).