Ian Brandenburg

Dávid Szabados

Chosen city: Moscow

**Major decisions:**

**Keep missing rating values:**

Almost 25% of the observations have NaN values in the rating column, instead of completely dropping them out of the dataframe, we replaced the missing values with 0. The reason is, that missing values can represent very low numbers. As we are working with the highly\_rated binary variable, these observations automatically goes to have a FALSE value in that column, thus we can work with it.