DATA WRANGLING

Project Title: PREDICTING ADMISSION DECISION FOR GRADUATE SCHOOL APPLICANTS

Data source: https://www.kaggle.com/mohansacharya/graduate-admissions

What kind of cleaning steps did you perform?

The data cleaning steps involved
- correcting naming errors in the dataset e.g. "LOR " to "LOR"
- changed column names in some cases to more appropriate terms e.g. "Chance of Admit" to candidate "Rating" for easier data manipulation
- dataset reindexing with candidate serial numbers

How did you deal with missing values, if any?

There was no missing values in this dataset.

Were there outliers, and how did you handle them?

There were no outliers in this dataset, all candidate feature values had reasonable distribution.