

DATA STORY

Project Title: PREDICTING ADMISSION DECISION FOR GRADUATE SCHOOL APPLICANTS

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

Wrangling:

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

There was no missing values in this dataset.

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

Data Story:

- The first feature of interest, Candidate's CGPA ranged from 2.42 to 3.97 on a 4.0 scale with an average of 3.43.
- Statement of purpose essays and letters of recommendation were graded on a 1 to 5 scale
- Most recorded GRE scores were in the 310 to 330 range, and most TOEFL scores were 100 to 115.
- The strongest parameter correlations were observed in the "CGPA", "GRE Score", "TOEFL Score" and the candidate ratings.

Hypothesis - Candidate rating and chance of admission may be estimated with limited parameters from the combinations of the identified test scores.