Dataset

We'll be using the same dataset as UCLA's Logistic Regression in R tutorial to explore logistic regression in Python, as explained in yhat's blog. This is an excellent resource for using logistic regression and summary statistics to explore a relevant dataset. Our goal will be to identify the various factors that may influence admission into graduate school. It contains four variables- admit, gre, gpa, rank.

- 'admit' is a binary variable. It indicates whether or not a candidate was admitted admit =1) our not (admit= 0)
- 'gre' is GRE score
- 'gpa' stands for Grade Point Average
- 'rank' is the rank of an applicant's undergraduate alma mater, with 1 being the highest and 4 as the lowest