UC Berkeley Admissions

Queries to extract knowledge from the data set:

- Is there a gender bias in the graduate admissions process?

Output:

First looking at total percentages of admitted males vs. females, it looks unfair (44.5% for males vs. 30.4% for female). Second consider these numbers by departments. Departments A and F have much lower acceptance rates for males, B and D slightly lower for males, and C and E slightly higher for males. Third, look at the overall acceptance rates per department. A and B ave higher acceptance rates, F has a very low acceptance rate and C, D, and E have moderate acceptance rate. Finally, look at applicants by department. Notice that the departments A and B with high acceptance rates have predominantly male applicants. C and E (departments with middle value acceptance rates) have an applicant pool with roughly two female applicants for every male applicant. The genders are roughly balanced in the applicant pools of Departments D and F. In summary, “the apparent association between admission and sex stems from differences in the tendency of males and females to apply to the individual departments (females used to apply more to departments with higher rejection rates).” [data set help manual in R]
This plot shows that a smaller % of females were admitted. Of the male applicants 44.5% were admitted, whereas for the female applicants, only 30.4% were admitted. As Page 74 of the R book notes, “barplots can be used effectively to show the data in a two-way table. One variable is chosen to be the categories of the barplot. Then, the bars for each level of the category are segmented to indicate the proportions of the other variable.” Also side-by-side plots are possible. In the above example, gender category is the chosen variable and proportions of the admit category (admitted/rejected) are shown on the y-axis.
Thus females are admitted at a higher rate in Dept A, B, D, F and at a lower rate in Dept C and E. The preference for males is quite slight in C and E whereas the preference for females is quite large in A and F. In B and D, females are admitted at a slightly higher rate than males. For each possible value of components 2 (Gender) and 3 (Dept), form proportions over the remaining variables (Admit).
The proportions are plotted in this next figure.
Below is a mosaic plot of the admit and gender variables summed across all departments.
Per department mosaic plots:

Student admissions at UC Berkeley

Dept A

Dept B

Dept C

Dept D

Dept E

Dept F
Transposed per-department mosaic plots:

**Student admissions at UC Berkeley**

Dept A

Dept B

Dept C

Dept D

Dept E

Dept F