M1-07: DataFrame: Conditionals

Part of the "Basics of Data Science with Python" Learning Badge

Video Walkthrough: *https://discovery.cs.illinois.edu/m1-07/*

Conditionals and the Berkeley Admissions Dataset

Earlier, you learned about the Simpson's Paradox and learned about the admissions scandal at UC-Berkeley related to the graduate class of Fall 1973. The following dataset contains the data about all applicants who applied to UC-Berkeley's during that Fall 1973 cycle:

Year	Major	Gender	Admission
1973	C	F	Rejected
1973	B	M	Accepted
1973	Other	F	Accepted
...			

Dataset URL: https://waf.cs.illinois.edu/discovery/berkeley.csv

Puzzle #4: How do we find the number of records in a DataFrame?

Python:	
Description:	Returns the number of records (observations) in a DataFrame.

Puzzle #5: Determine the number of applicants by gender, and their accepted percentage:

(a):

Python:	
Description:	Selects all **female** applicants.

(b):

Python:	
Description:	Selects all **male** applicants.

(c):

Insight:	How do you modify your code in (a) and (b) to select only **admitted** females and males?

Analysis: Find the acceptance rate by using the code you wrote above:

Acceptance rate among females: **Acceptance rate among males:**

Puzzle #6: Determine the number of applicants by gender, acceptance status, and major:

Python:	
Description:	Selects all **female** applicants **accepted** into major "**A**".

Puzzle #7: Finally, continue to modify the code to complete the table below. Several cells are already filled in to help you check your answer and speed up the process:

	Female Applicants			Male Applicants		
	Accepted	Applied	%Accepted	Accepted	Applied	%Accepted
Major A		/		825	/ 1138	72%
Major B	17	/ 25	68%		/	
Major C		/	34%	120	/ 325	37%
Major D	131	/ 375			/	
Major E		/	24%	53	/ 191	28%
Major F	25	/ 341	7%	22	/ 273	6%

Analysis:
Looking back at the **overall data**, was a higher percentage of males or females accepted to UC-Berkeley in Fall 1973?

Analyzing the table above, **how many different majors** had a higher percentage of males accepted to UC-Berkeley than females?

Is this an example of Simpson's Paradox?