

# Advanced Statistics

## Simpsons Paradox:

A phenomenon in probability and statistics where a **trend** appears in **several data groups** but **disappears or reverses** when the **groups are combined**.

(It shows how different data groupings can lead to very different conclusions)

EX:

DATA:

	MALE			FEMALE		
	APPLIED	ADMITTED	RATE	APPLIED	ADMITTED	RATE
MAJOR A	900	450	50%	100	80	80%
MAJOR B	100	10	10%	900	180	20%

IS THERE A GENDER BIAS ?

~~YES~~ = NO

A couple years ago UC Berkeley did a test to see if there was any gender bias in their acceptance rates. (the example above follows what UC Berkeley did but with made-up numbers).

In the example if we look at the **admittance rates** for both males and females **separated by major**

we find out that the **males have a lower admittance rate** instead of the **females** in both majors

we reach a conclusion that : **there is female bias in admittance rates!**

however if we **look at both majors combined** we'll find that for **males**, out of the **applied 1000**, **460 got accepted** leading to an **admittance rate of 46%**.

as for **females**, out of the **1000 applied** only **260 got accepted** leading to an **admittance rate of only 26%**.

by looking at such data we conclude that there is **severe male bias** in overall acceptance.