The Data: The Use of Reformulations in Human-Computer Interaction

male	female		
2	0	1	5
2	13	0	2
2	7	4	1
5	3	1	2
1	2	0	1
0	0	1	3
2	1	2	2
0	1	0	
2	5	4	
3	13	2	
0	1	2	
4	4		
3	4		
4	0		
4	5		
6	1		
1	0		
0	5		

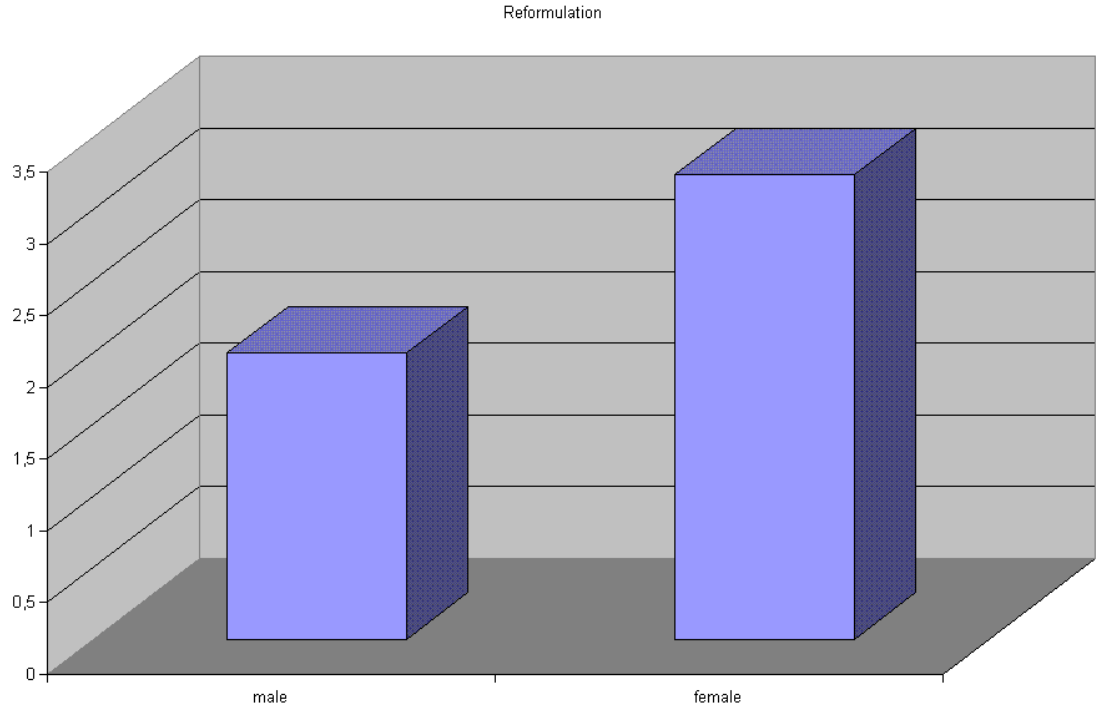
The first thing we can do:

- calculate the mean

in this case:

- females: 3.24
- males 2.00

reformulations per 20 turns



Does that mean that males use fewer reformulations than females?

- we can't decide that yet, because we don't know whether our data are just due to chance – maybe we have just by accident found all non-reformulation-using males and all reformulation-using females
- we need to decide when we'll be satisfied
  - are we convinced when there is a 5% chance that our data came about accidentally? Or rather when the chance level is just 1% or even less?
    - we need to choose a level of significance
      - e.g. $p<0.05$, $p<0.01$, $p<0.001$
    - and whether we want to expect differences between our groups in both directions or just in one
      - i.e. whether we want a one- or a two-tailed test

The most common tests:

- t-test
- ANOVA (Analysis of Variance)

in our case: the difference between the use of reformulations by males and females is not significant!

Now *male* and *female* are discrete categories

What if we want to find out whether people who use many reformulations use fewer repetitions – or the other way around?
That is, what is if we want to learn about the relationship between two non-discrete categories?

The test we make is one of correlation: How are the two variables *reformulation* and *repetition* correlated?

- Tests of correlation

The more X, the more Y: positive correlation
The more X, the less Y: negative correlation