Worksheet 16 - two sample permutation test

Monday, March 31, 2025

DS 002R - Jo Hardin

Name: _____

Names of people you worked with: _____

Where do you study? Do you have a favorite go-to place or are you trying different spots?

Task: To create a null sampling distribution, we repeatedly permute the variable (either one) and recalculate the statistic of interst (e.g., the difference in means across the two groups). Given the data below,

a. Which of the permute columns represents a valid permutation of the `math` variable? Why?
b. Add one additional column to the dataset which is a valid permutation of the `math` variable.

```
# A tibble: 10 x 6
   schtyp   math math_perm1 math_perm2 math_perm3 math_perm4
   <fct>   <int>      <int>      <int>      <int>      <int>
 1 public     49         63         49         41         49
 2 public     53         63         53         43         71
 3 public     41         71         41         52         63
 4 public     45         45         45         49         52
 5 public     71         43         71         43         43
 6 private    43         71         43         71         75
 7 private    75         52         75         75         63
 8 private    63         71         63         45         71
 9 private    43         53         43         63         63
10 private    52         71         52         53         41
```

Solution:

a. Both `math_perm2` and `math_perm3` are valid permutations.

b. Here is another one, `math_perm5`. Note that every single value in `math` must be represented exactly once in the permuted column.

```
# A tibble: 10 x 7
   schtyp   math math_perm1 math_perm2 math_perm3 math_perm4 math_perm5
   <fct>   <int>      <int>      <int>      <int>      <int>      <int>
 1 public     49         63         49         41         49         49
 2 public     53         63         53         43         71         63
 3 public     41         71         41         52         63         43
 4 public     45         45         45         49         52         53
 5 public     71         43         71         43         43         43
 6 private    43         71         43         71         75         71
 7 private    75         52         75         75         63         75
 8 private    63         71         63         45         71         45
 9 private    43         53         43         63         63         52
10 private    52         71         52         53         41         41
```