Problem 2

SPRING 24 FINAL

Michelle and Abel are each touring apartments for where they might live next year. Michelle wants to be close to UCSD so she can attend classes easily. Abel is graduating and wants to live close to the beach so he can surf. Each person makes their own DataFrame (called `michelle` and `abel` respectively), to keep track of all the apartments that they toured. Both `michelle` and `abel` came from querying `apts`, so both DataFrames have the same columns and structure as `apts`.

Here are some details about the apartments they toured.

- Michelle toured **one bedroom and studio** apartments at 12 different complexes, or 24 apartments total.
- Abel toured **one bedroom and two bedroom** apartments at 20 different complexes, or 40 apartments total.
- There are 8 complexes that are near both UCSD and the beach, and both Michelle and Abel toured these complexes.

We'll assume for this problem only that there is just one apartment of each size available at each complex, so that if they both tour a one bedroom apartment at the same complex, it is the exact same apartment with the same `"Apartment ID"`.

Problem 2.1

What does the following expression evaluate to?

```
michelle.merge(abel, left_index=True, right_index=True).shape[0]
```

8

> Click to view the solution.

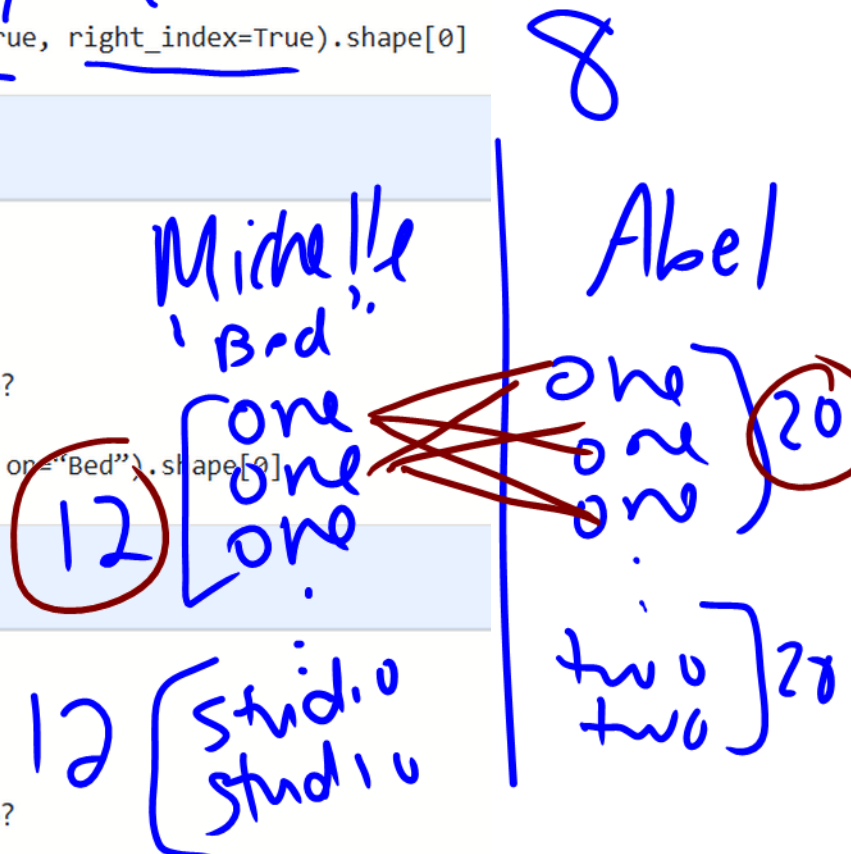## Problem 2.2

What does the following expression evaluate to?

*12*20*

```
michelle.merge(abel, on="Bed").shape[0]
```

Michelle
1 Bed

Abel

12

```
one
one
one
```
one
one
one
(20)

```
12 [ studio
     studio ]
```
two] 20
two

> Click to view the solution.

Problem 2.3

What does the following expression evaluate to?

Problem 2.3

What does the following expression evaluate to?

`michelle.merge(abel, on="Complex").shape[0]`

Click to view the solution.

Problem 2.4

What does the following expression evaluate to?

`abel.merge(abel, on="Bed").shape[0]`

Click to view the solution.

Handwritten annotations:

4 * 8

maldes complexes merged
per complex df has
four rows with A

Michelle "complex"   Abel "complex"

A   A
A ⤬ A
A   A

B   C
B ⤬ C
C   C
C

"Bed"   "Bed"
Abel   Abel

one ⎫   one ⎫
one ⎬ (20)   one ⎬ (20)
one ⎭   one ⎭

two ⎤ (20)   two ⎤ (20)
two ⎦   two ⎦

8 letters in both dfs

40 * 20
rows   matches per row

20*20 + 20*20
ones   twos

more/harder practice:
try Q9 on Wi24 Final

reject null vs. fail to reject null

obs value is like other from simulation

fail to reject null

dist generated from null through simulation
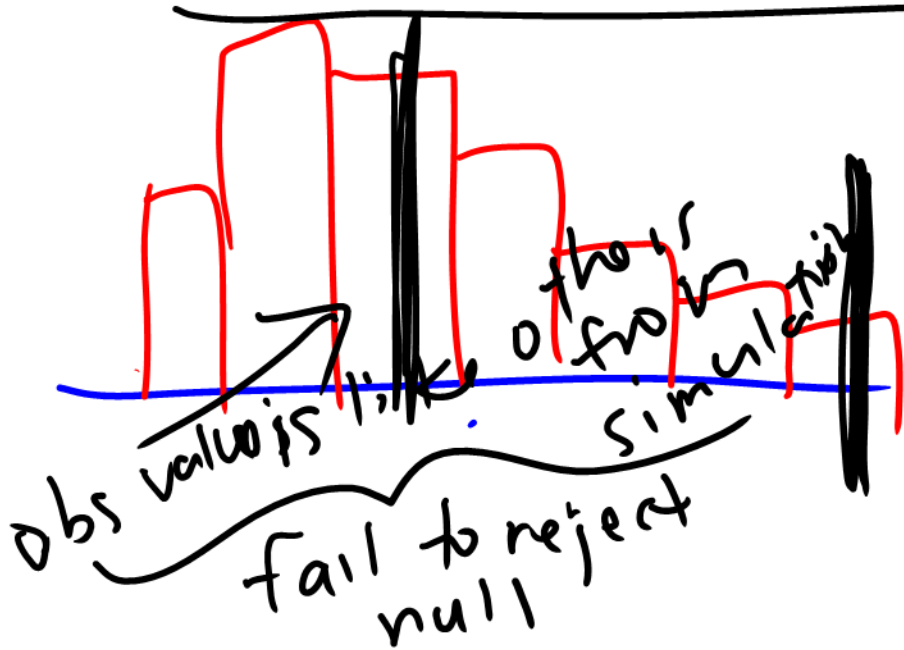
← observed value is unlike those from simulation
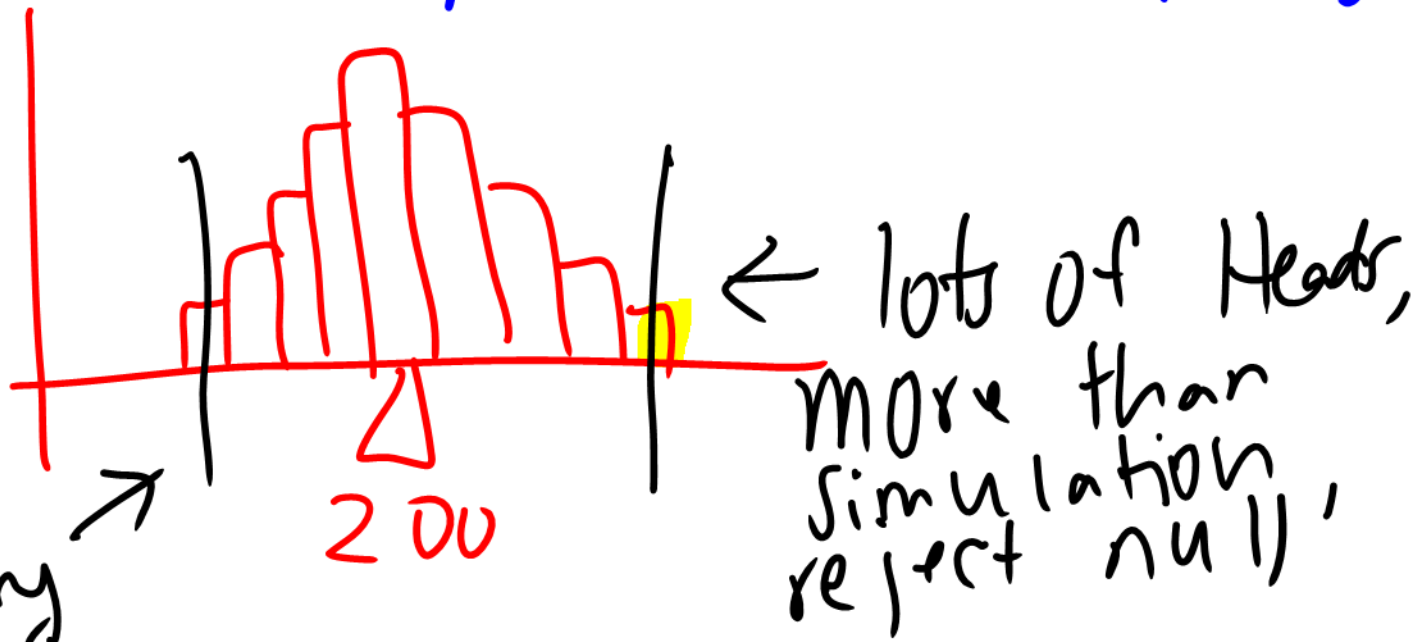
usually reject null

~ depends on direction- is it in the direction of alt?

Q: is our coin fair or biased towards Heads?

null

alt

Simulation: keep track of # Heads
flip coin 400

← lots of Heads,
more than
simulation,
reject null)

200

if very
few heads,
fail to reject null

extreme in
the direction
of alt ⇒ reject null, otherwise don't