

STAT 20 PS: TAXONOMY OF DATA

1. What type of variable is being measured by each of the following questions?

A.

How much coding experience do you have?

None.

Very little. I've just dipped my toe in.

A bit. I've used coding in a limited capacity, such as for a small part of a course

Some. I've had to write code for one or two classes and am comfortable with it.

A good deal. I have several years of experience writing code.

E.

How long have you been a student at Cal?

More than 4 years.

I'm in my fourth year.

I'm in my third year.

I'm in my second year.

I'm in my first year.

This is my first semester!

C.

⋮

Please also rate that experience on a scale from 1 to 10.

1 2 3 4 5 6 7 8 9 10

No experience Tons of experience

F.

Are hotdogs sandwiches?

Yes

No

D.

What zip code do you consider home? (If from abroad, use UC Berkeley's 94720)

2. Last class you collected data on six different variables on your classmates. List the data types of each of those variables, the units that they were measured in, and the range of possible values that you would expect them to take.

- Name:
- Contact info:
- Hometown:
- # Siblings + Pets:
- # Semesters in college:
- Farthest distance from campus during the break:

3. On the back of this page, sketch out a data frame that contains the data from your group mates on the questions above. Introduce yourself to other nearby groups to extend your data frame until it has 6 rows.

4. As discussed in the lecture notes, this is not a data frame:

Handed- ness Sex	Right-handed	Left-handed	Total
Male	43	9	52
Female	44	4	48
Total	87	13	100

It does, however, depict a data set, just in a different format. Sketch the data that is summarized here but structured as a data frame. Think through: what was the unit of observation? What were the variables? How many rows are there? How many columns? (you need not fill out the entire data frame; just a schematic)