Baker Sample Lesson

David John Baker

20/03/2020

Learner Persona

General Background

Jason is a undergradaute student in social sociences at a state university. He is very knowledgeable about all things statistics and research and even completed a summer research project in the lab of someone he might want to apply to do a PhD with. While he has mainly used the tools in his classes likes SPSS and Excel to do data analysis, he knows he wants to learn R to increase his chances of getting into a grad program he likes, but more importantly to make his job as a research assistant easier.

Relevant Experience

Jason is just finishing his undergraduate, has played with R before but always gives up, and has a good understanding of stats. He is often asked to change datasets by hand by the PI that runs his lab, but knows there must be a better way. He understands that data should one observation per row, one variable per column but doesn't refer to that as *tidy* data.

Percieved Needs

Jason wants to be able to apply what he is learning right away so that he can spend less time in the lab and more time with his friends in his last year of university. Ideally learning R will free up more time rather than make it a hobby or part of his identity.

Special Considerations

Jason is a smart student, but is easily distracted. He works best when lectures are short, to the point, and can try things out on his own. He'd rather do a lot of small examples to understand a concept, rather than one big one.
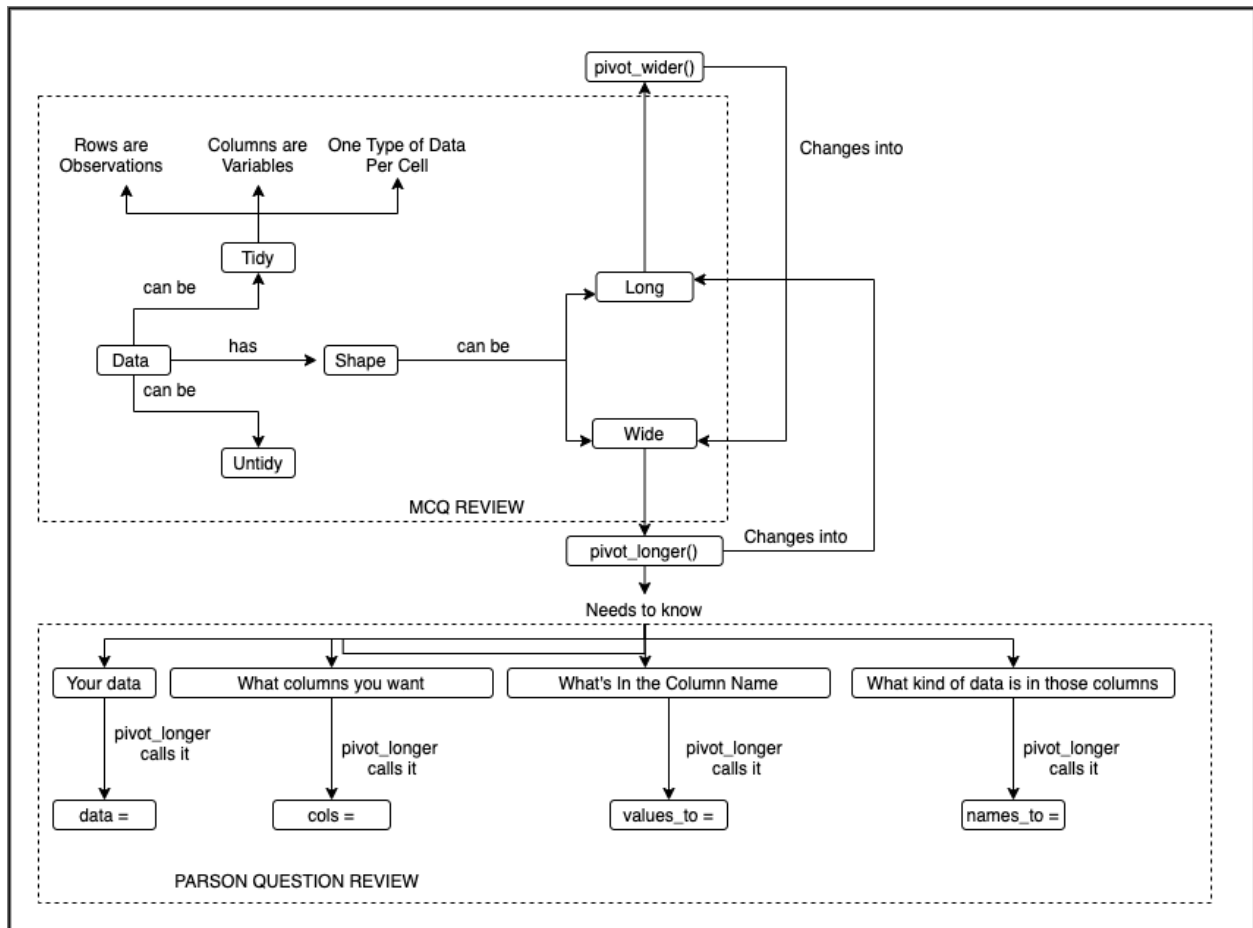
Concept Map

Figure 1: Concept Map

Formative Assessments

Formative Assessment One: Multiple Choice Question

This multiple choice qustion will be given at the start of the lesson, after the initial terms are introduced to make sure the learners are all comfortable with the new terminology. If done in larger class, students could vote publicly on what they think the answer is, then discuss with a neighbor in order to clear up any misconceptions, then vote again. The question was designed so that the learner is not able to use any sort of process of elimination, but rather needs to rely on their knowledge of the content to answer the question.

Given the dataset below, which of the following choices best describes the current state of the data using terminology we just learned.

- A. The dataset is wide and tidy
- B. The dataset is long and tidy
- C. The dataset is wide and not tidy
- D. The dataset is long and not tidy.

	product	November	December	January	February	March	Department	Location
1	hammer	25	23	56	7	46	Home	Milwaukee
2	nail	105	452	454	563	222	Home	Milwaukee
3	screwdriver	34	67	43	32	52	Home	Chicago
4	pliars	24	23	23	24	39	Home	Milwaukee
5	drywall	58	78	56	34	64	Home	Chicago
6	stapler	21	11	18	31	33	Office	Chicago
7	printer ink	42	43	64	53	55	Office	Chicago
8	posit	97	75	37	97	64	Office	Milwaukee

Figure 2: MCQ

If the learner votes **A**, the learner correctly identifies the shape of the data, but does not realise that the `Month` columns are actually names of a varible that they could perform some sort of calculation on. If the learner votes **B**, the learner does not understand what it means for data to be wide/long. If the learner votes **C**, the learner has correctly identified the answer. If the learner votes **D**, the learner does not understand both the shape and concept of tidy.

Formative Assessment Two: Parson Problem

This Parson Problem will be given after the `pivot_longer()` function is introduced. At this point they were introduced to the **four** arguments needed to take a wide dataframe and make it long. Here the learner needs to transfer the model the just saw in the example and apply it exactly as they just learned to a new set of data.

Based on what we learned previously, take the next 3-5 minutes to look at this new dataset and fill in the blanks provided so that when you run this chunk of code, the shape of this new dataset matches the shape from our previous example.

```
#student_data

# pivot_longer(data = _____ , cols = _____, names_to = _____, values_to = _____)
```

Outline

1. Introduce Problem (~5 minutes)

   - Tidy
   - Define Long
   - Define Wide
   - Formative on Definitions (Short)

2. Introduce Program (~ 7 minutes)

   - Introduce Data
   - Ask if Long or Wide and Why
   - Introdce `pivot_longer()`
   - Introduce arguments

3. Cooperative, Guided Practice (~ 2 Min, if time)

   - Guess what output is going to be.
   - Try same thing, but new data frame (Parson Problem, Longer)
   - If time, third dataset