Baker Sample Lesson

David John Baker

20/03/2020

Lesson Outline

1. Introduce Problem (5 minutes)

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

2. Introduce Program (10 minutes)

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

3. Cooperative, Guided Practice

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

Learner Persona

General Background

Jason is a professor of social sociences at a state university. He is very knowledgeable about all things statistics and research and even completed a prestigious post-doc position before starting his current job. While he has mainly used the tools his discipline has used the past twenty years, his current graduate students want to learn R so that they can be competitive on any job market once they graduate. They have never programmed before.

Relevant Experience

Jason has a PhD in his discipline, has programmed before, and mastery of all of the statistical procedures that he has to do on on normal basis. He has not thought too much about the shape that data has to be in in order to run the models he does in programs he is familiar with like SPSS or SAS, as he normally has just gotten an undergraduate or grad student to reconfigure the data by hand. He has hundreds of hours experience working in SPSS, Excel, and SAS and considers himself and expert in these programs. He has played with a bit of R and likes how the pipe operator works because it reminds him of learning bash as a teenager. 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 has heard he needs to learn R and have to teach it in the coming year as his department is going to switch their curricula from SPSS to R. He has nine or so months to do this and figures what better way to start than to learn how to manipulate some of his own data as to help his own research and work in something more familiar. He figures, why not start by learning the tasks he often asks his students to do.

Special Considerations

While Jason does have a lot of time until he has to teach this course, he does not have that much time every day to devote to learning R. He is currently in the middle of negotiating how his department is trying to move some of their classes online due to a public health crisis and also needs to be home every day early in order to take care of his son.
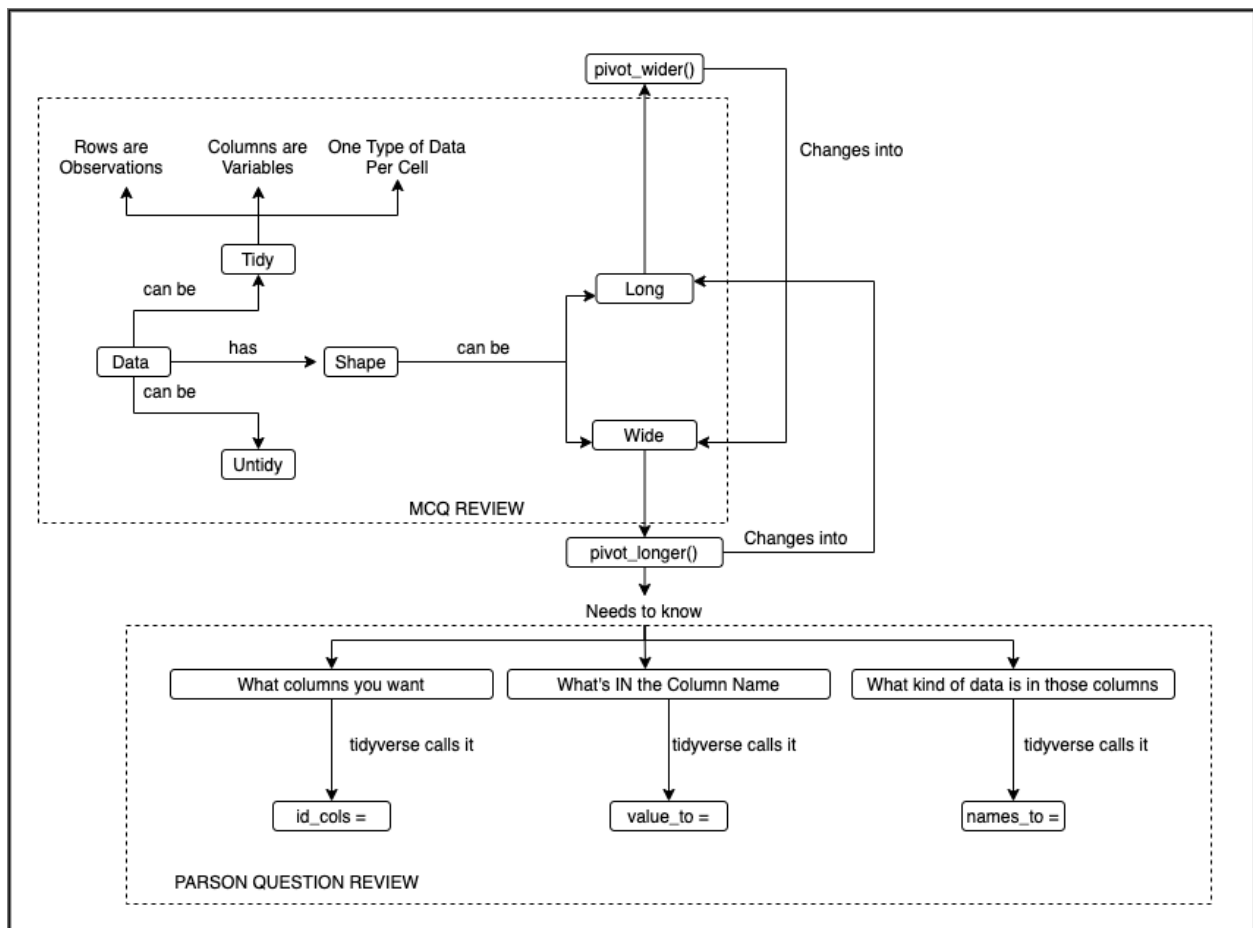
Concept Map

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.

    Find Dataset that is long and almost tidy

If the learner votes **A**, the learner is struggling with both the concept of wide and tidy data. The teacher should attempt to explain the concept using another approach. If the learner votes **B**, the learner understands what is meant by the shape of the data, but needs to be refreshed on what is meant by tidy so that they are not caught later by a problem they did not anticipate. If the learner votes **C**, the learner is still struggling with the concept of the shape of a dataset, but knows that the cells are not in the format they should be. If the learner votes **D**, the learner has correctly identified the answer.

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.

Slides and Notebooks