

Outcome-oriented educational tasks with flipbooks

For *Chance Magazine*'s 'Taking a chance in the classroom'

Evangeline Reynolds, June 2020

A goal of data science and statistics courses is giving students practice preparing and analyzing data. A new style of assignment presents outcome-oriented exercises in 'flipbooks' (evamaerey.github.io/flipbooks/about). Typically 'flipbooks' present code step-by-step and side-by-side with output. This gives you a movie-like experience, linking new code to the change that it triggers in output. But 'flipbooks' can also present *only output*. This mode can be used in data science and statistics classrooms to create engaging outcome-oriented assignments!

What's this new style of exercise?

A flipbook data manipulation assignment might look as follows. First, you have a reference section of functions that can be used to accomplish the required tasks. Then a target data frame is introduced. The target will be a final output of a data manipulation pipeline, where the input is a familiar data set (I use the 'gapminder' data set, which my students have worked with for data visualization). Then, a walk-through from input to target — including intermediary states — is displayed. Students are instructed to 'wrangle to it'; i.e., use the new data manipulation functions to get to the target output. A series of additional exercises are structured in the same way. A target is presented. Then a walk-through of states is shown. A collection of slides that constitutes one wrangling exercise follows, and a full set of exercises can be viewed here (<https://tinyurl.com/y6b9z5vx>).

What might the student/instructor experience look like?

Some students will be able to wrangle to accomplish the tasks just by inspecting the target. In that case, students can build the data manipulation pipeline without looking ahead. After producing the target, they are encouraged to compare their strategy to that presented in the walk-through.

Other students may initially be stumped about how to manipulate to the target. These students are encouraged to look ahead to the walk-through. They may be able to accomplish the task by using the intermediary outputs as a guide.

As an instructor, the task of assisting students who have trouble may be more manageable. Students are less likely to be overwhelmed by the data manipulation challenge, given the step-by-step presentations. Even if students don't know how to proceed, the task is logically broken up for the instructor or a fellow student to talk through the manipulation strategies.

Though not explicitly stated, hopefully students will take away the following idea: *you may have some big task ahead of you, but stay cool. Try to break up the task into pieces — it will probably feel more manageable.*

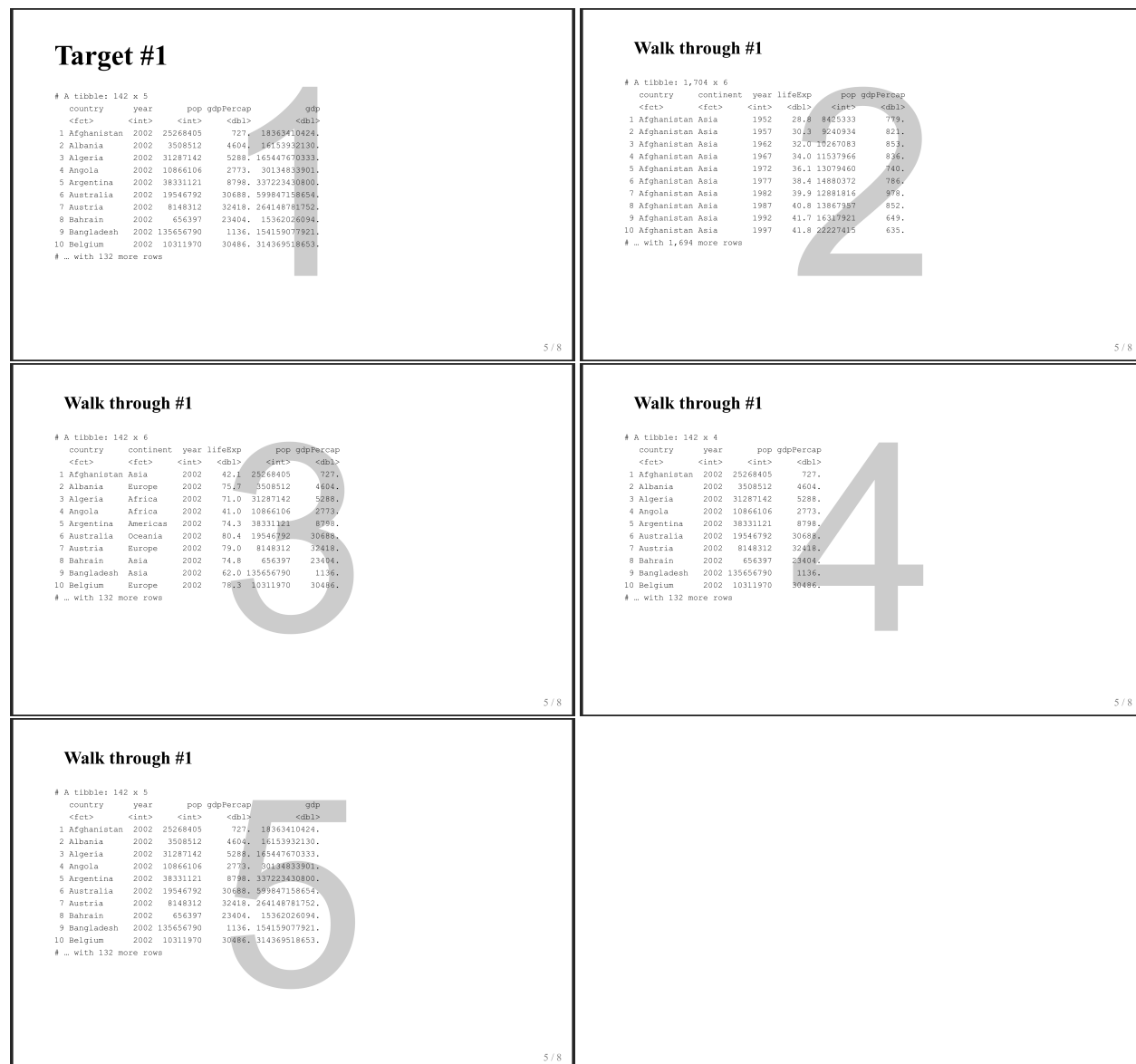


Figure 1: Slides constituting one data wrangling task.

Data manipulation as a general problem

I think an advantage of this outcome-oriented exercise is that it de-emphasizes specific tools. Sometimes the focus of data manipulation exercises is syntactical. There is a lot of focus on the tools being used. But this exercise puts the data transformation itself center stage (subsetting, dropping columns, summarizing, etc.). In fact, a large number of statistical software or syntactical paradigms could be used to accomplish the manipulation tasks presented here.

I personally teach and have built this exercise using ‘the tidyverse,’ which does have some signatures in the output (such as a separate `group_by` step when compared to `data.table`). But the manipulation tasks could also be solved with Stata, SQL, SPSS, SAS, R tidyverse, base R, `data.table`, Pandas, – or something yet to be invented! This shifts the emphasis from a *how* question (specific implementation and syntax) to a more general *what* question. The question “what data manipulation moves can I make?” is primary, and the how-to syntax is secondary.

How to build your own exercise

Educators should feel free to use this exercise as-is or can create their own outcome-oriented tasks using the new `{flipbookr}` package (<https://github.com/EvaMaeRey/flipbookr>)!

If you are familiar with R and Rmarkdown, you’re in a good place to start building these outcome-oriented educational tasks. Still, some things will be new, and I’ll discuss them here briefly. For more detailed guidance, it will also be useful to look at a basic tutorial (https://evamaerey.github.io/flipbooks/flipbook_recipes) on creating your first flipbooks and a short example (<https://tinyurl.com/y25d2nhz>) of exactly how to build an outcome-oriented exercise.

To get the movie-like experience of flipbooks we need a slide show tool. The `flipbookr` package is designed to be used with the Rmarkdown package `xaringan` to deliver material to .html slides. Your Rmarkdown file will need to have the appropriate meta data setup so that these slides are produced.

Once you have this in place, you can then write a code chunk with steps of a data manipulation task. The code chunk options should be set to `echo = FALSE` and `eval = TRUE` so that the output displays but the code is hidden. After a slide break, you’ll then use the `flipbookr` function `chunk_reveal` to refer to the prepared chunk. You should set `display_type` to “output” so that the output is shown, but the code used to accomplish the manipulation is hidden.

```
---

# Target #1

```{r task1, eval = FALSE}
mtcars %>%
 select(mpg:wt) %>%
 arrange(-mpg) %>%
 slice(1:5)
```

---

`r chunk_reveal(chunk_name = "task1", display_type = "output", title = "## Walk-through #1" )`
```

You’ll probably want to host the resulting html and associated files online for students. Github Pages has worked well in my experience.

Going further

A bit more may be asked of students to further engage them with the pipelines they have built. At the end of data manipulation pipelines, you might ask students to create a descriptive name for an object that will hold the manipulated data or to describe the data manipulation steps in plain language. Also, educators might think about creating outcome-oriented tasks like these for data visualization or data analysis.