Writing a Paper in Rmarkdown

Jack Blumenau   *University College London*

Everyone agrees that the issue in this paper is really important. But we do not know much about this specific question, although it matters a great deal, for these reasons. We approach the problem from this perspective. Our research design focuses on these cases and relies on these data, which we analyse using this method. Results show what we have learned about the questions. They have these broader implications.

Introduction

The important thing about academic papers is that you state your argument *very clearly* at the beginning of the introduction. Even if you think your argument is clear, there are certainly ways that you can make it **even clearer**. Often, the way to maximise clarity is to treat your audience as if they are really stupid, and to present your arguments in a list:

- You should be clear
- No, clearer than that
- Are you listening to me? Clarity is everything.

Or, even, as an *ordered* list:

1. One thing
2. Another thing
3. A third thing

If you are writing a paper, another important thing to clarify in your introduction is how you contribute to some important and extensive literature. For instance you might want to write about exciting work in electoral politics [Blumenau et al., 2017], legislative politics [Blumenau and Lauderdale, 2018]; or even politics and gender [Blumenau, 2019].[1] Alternatively, you might want to provide a link to a very interesting online resource that is relevant to your work.

Good papers have good sections

In addition, most good papers use a lot of math.

---

[1]I can use this footnote to tell you that other authors are available.
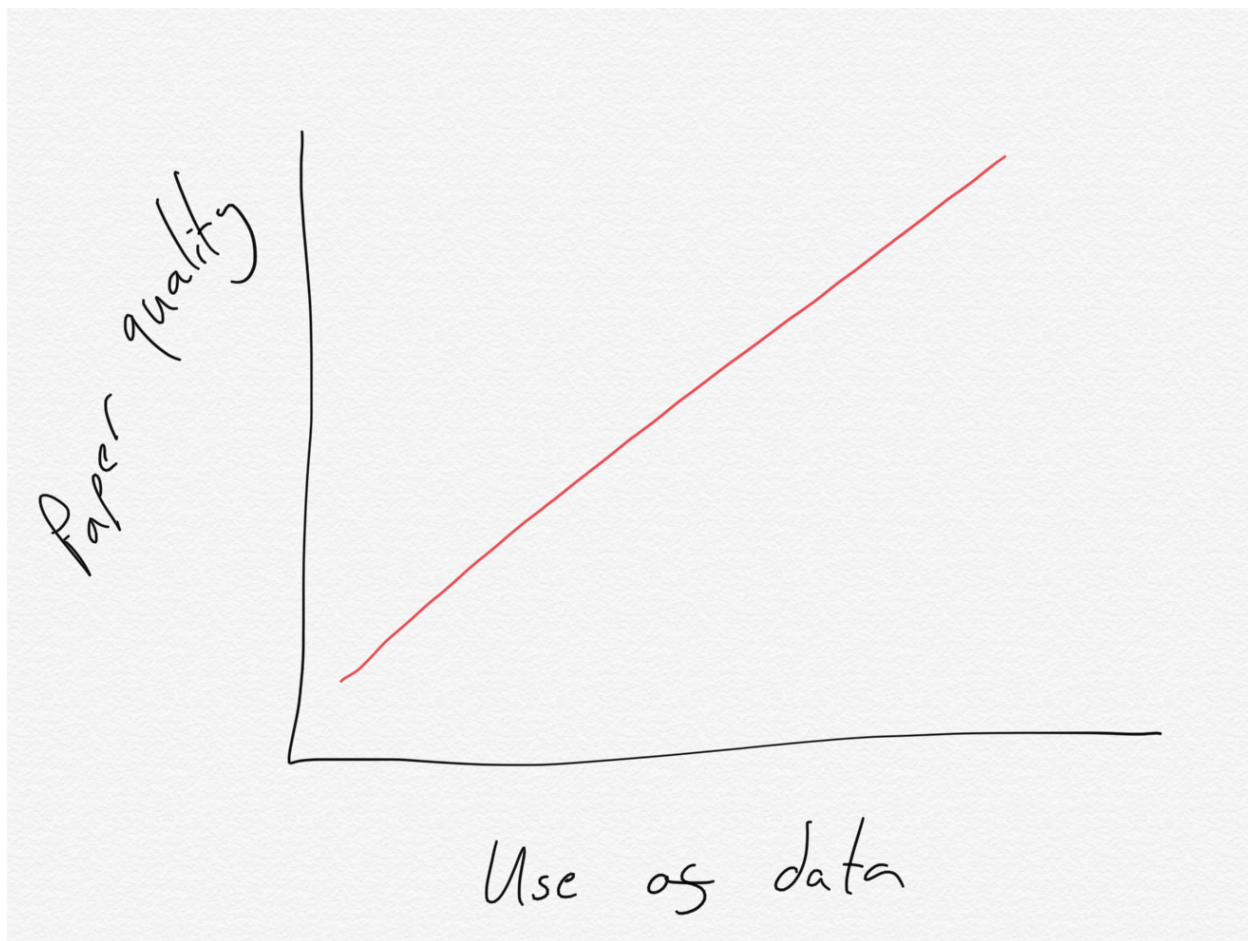
Figure 1: A very serious graph

$$Y_i = \alpha + \beta X_i + \epsilon_i$$

Sometimes you might not want to include a big equation like this, but instead just add a quick equation in-line with the text like this: $\frac{1}{2} = .5$. Sometimes you might not want to include math at all, but instead a more user-friendly graph. That's easy too. For instance, figure 1 shows a graph I created to explain what makes for a good paper.

Data and analysis

Data

Of course, while good papers might use math, the best papers use *data*. You can load the data that I provided by writing the following in R:

```
## Load data
covid_data <- read.csv("data/covid_survey.csv")
```

This data includes responses from from a survey of 3049 individuals which was fielded in March 2020. We will focus on information on 8 questions posed to those individuals. Our primary interest is in the responses provided to the following questions:
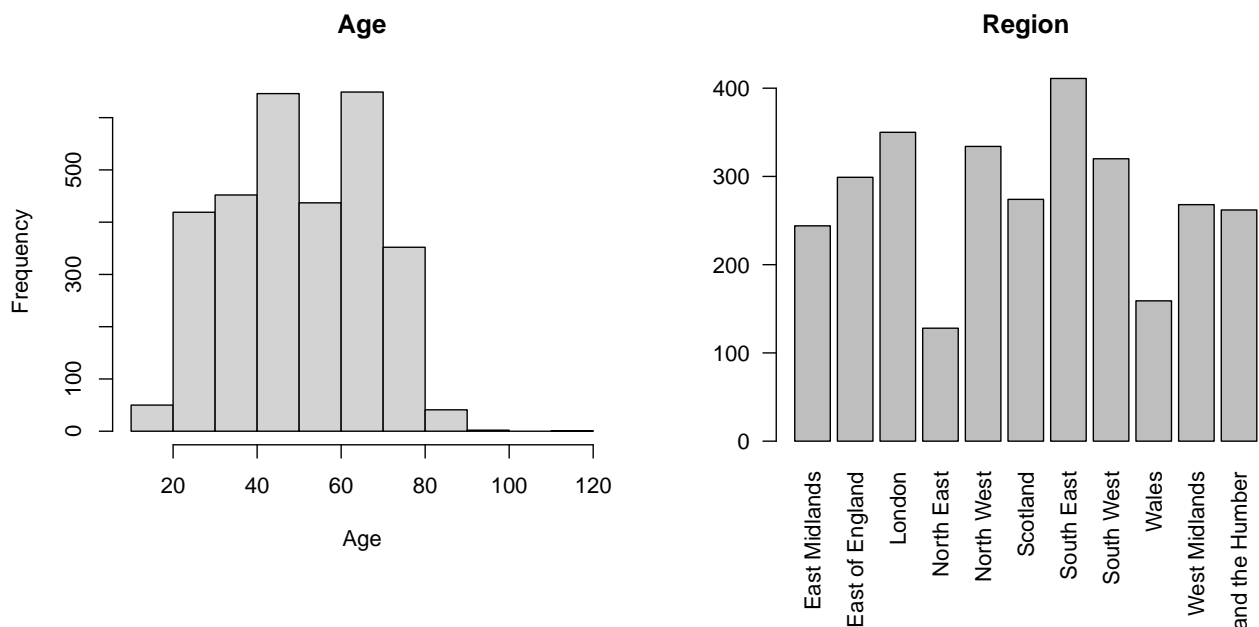
1. `incomeChangeHH`

   - Has your household monthly income changed since the coronavirus outbreak?

2. `approveUKGovt`

   - Do you approve or disapprove of the job that the UK government is doing?

Analysis

Given this data, we might want to conduct some cutting-edge analysis which will produce some awesome findings. For instance, you might want to create plots like this:

Note that by including `echo = TRUE` in the chunk options, you can see the code that I used to produce such beautiful graphs

Or you could run a regression and present the results in a table by using the `stargazer` package (let's hide the relevant code in the document this time by setting `echo = FALSE`):

Here, we use `results = 'asis'` in the chunk options to ensure that the table is rendered properly in the PDF.

	Table 1:
	Dependent variable:
	age
profile_genderMale	2.151***
	(0.612)
Constant	49.588***
	(0.412)
Observations	3,049
R^2	0.004
Note:	*p<0.1; **p<0.05; ***p<0.01

If we care about reporting some quantity of interest from this model in the text, then that is trivial to do using in-line code chunks. For instance, the coefficient on the gender variable is equal to 2.1514788.

Your work goes here!

In this section you should include some R code that conducts some additional analyses of the `covid_data`. You might, for instance, do the following:

1. Create a histogram of incomeChangeHH
2. Create a histogram of approveUKGovt
3. Recode incomeChangeHH to be a binary variable, where 1 is whether an individual's household has lost income as a result of COVID-19 and else the variable is 0
4. Recode approveUKGovt to be a numeric variable
5. Implement one regression predicting approveUKGovt as a function of the binary income variable you have used
6. Implement one regression predicting approveUKGovt as a function of the binary income variable you have used + demographics
7. Write a one paragraph summary of your results which includes describing a) the sample size used in the regressions; b) the size of the coefficient on approveUKGovt in the two models; and c) the r-squared of the two models.

In your document you should make sure that:

1. We cannot see the code you use to do your recoding and analysis
2. We can see both the graphs and tables you produce
3. Any quantities that you report in the paragraph of text you write should be included as in-line code chunks.

Helpful resources

1. Rmarkdown cheat sheet
2. Bibdesk
3. Github and Rstudio
4. Writing reproducible research papers with R Markdown

References

Jack Blumenau. The effects of female leadership on women's voice in political debate. *British Journal of Political Science*, pages 1–22, 2019.

Jack Blumenau and Benjamin E Lauderdale. Never let a good crisis go to waste: Agenda setting and legislative voting in response to the eu crisis. *The Journal of Politics*, 80(2):462–478, 2018.

Jack Blumenau, Andrew C Eggers, Dominik Hangartner, and Simon Hix. Open/closed list and party choice: Experimental evidence from the uk. *British Journal of Political Science*, 47(4):809–827, 2017.