My first replicable Paper

MyFirstName MyLastName
Evans School of Public Policy and Governance
University of Washington
Seattle, WA 98115, United States
greatguy@uw.edu

January 8, 2018

Abstract

This is an example on how to make a reproducible paper. We are using R from Rstudio, creating an RSweave document. This is a nice start to create a nice paper and get an A+. The next sections will show the steps taken.

1 Introduction

This is my intro to my great paper, I will explain the cool things I can do with my new 'computational thinking' powers combined with some Latex.

This is my nice intro to my great paper, I will explain the cool things I can do with my new 'computational thinking' powers combined with some Latex.

This is my nice intro to my great paper, I will explain the cool things I can do with my new 'computational thinking' powers combined with some Latex.

2 Explaining Labels

Sections may use a label[1]. This label is needed for referencing. For example the next section has label *datas*, so you can reference it by writing: As we see in section 3.

---

[1]In fact, you can have a label wherever you think a future reference to that content might be needed.

3 Data analysis

Here you can explain how to get the data:

```
> states=read.csv("https://goo.gl/So48s5")
```

3.1 Exploration

Here, I start exploring the data. The first step is to know what variables I have, and in what scale they are:

```
'data.frame':          51 obs. of  8 variables:
 $ state               : Factor w/ 51 levels "Alabama",""..
 $ satMean             : int  991 920 932 1005 897 959 89..
 $ satDemand           : num  0.08 0.41 0.26 0.06 0.47 0...
 $ k12ExpenditurePupil : int  3627 8330 4309 3700 4491 50..
 $ incomeHouseholsMedian: num  27.5 48.3 32.1 24.6 41.7 ...
 $ diplomaHsAdults     : num  0.669 0.866 0.787 0.663 0.7..
 $ collegeDegreeAdults : num  0.157 0.23 0.203 0.133 0.23..
 $ region              : Factor w/ 4 levels "Midwest","N"..
```

A next step demands:

- Knowing the *central* and *dispersion* values.

- Visualizing the variables of interest.

Except for the column *state*, we can compute the centrality and spread measures for the other variables in the data. I will do that in Table 1 in the next page.

Table 1: Mean and Spread values

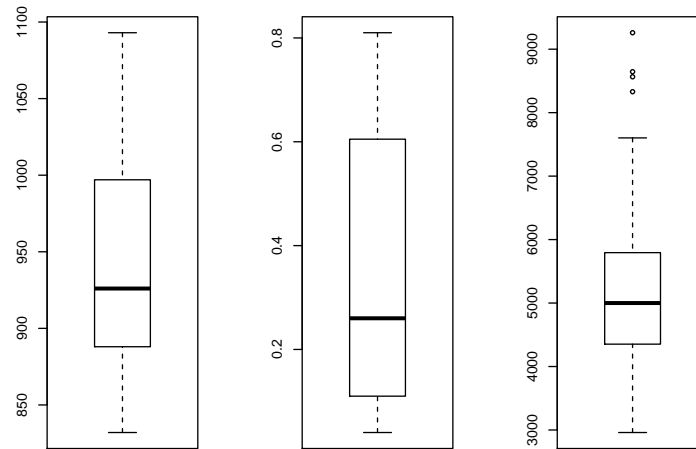| Statistic	N	Mean	St. Dev.	Min	Max
satMean	51	944.098	66.935	832	1,093
satDemand	51	0.358	0.262	0.040	0.810
k12ExpenditurePupil	51	5,235.961	1,401.155	2,960	9,259
incomeHouseholsMedian	51	33.957	6.423	23.465	48.618
diplomaHsAdults	51	0.763	0.056	0.643	0.866
collegeDegreeAdults	51	0.200	0.042	0.123	0.333

Figure 1: Location of values

As you saw, my Table 1 is nice. As you, saw the mean of the variable *satMean* is 944.098039215686. Now let's use a boxplot to explore location:
As we have a categorical variable, we could create a frequency table:

Table 2: Distribution of Region

Region	Frequency
Midwest	12
N. East	9
South	17
West	13

3.2 Modeling

Here, I propose that the amount of money spent for child per state in the US has an effect on the mean average pupils in a state get in SAT:

```
> reg1=lm(satMean~k12ExpenditurePupil, data = states)
```

Here, I modify the previous model; while I insist that the amount of money spent for child per state in the US has an effect on the mean average pupils in a state get in SAT; I will control the effect the demand per state (as demand were equal accross states). Then,

Model 2:

```
> reg2=lm(satMean~k12ExpenditurePupil+satDemand, data = states)
```

I have the results, but have not display them, let's do it in the coming subsection

3.3 Modeling nicely

What about this:

Table 3: Regression Models

	Dependent variable:	
	satMean	
	(1)	(2)
Dollars per Student	−0.022***	0.009**
	(0.006)	(0.004)
Share taking SAT		−253.770***
		(22.491)
Constant	1,060.732***	989.807***
	(32.701)	(18.396)
Observations	51	51
R^2	0.217	0.786
Adjusted R^2	0.201	0.777
Residual Std. Error	59.814 (df = 49)	31.623 (df = 48)
F Statistic	13.615*** (df = 1; 49)	88.009*** (df = 2; 48)
Note:	*p<0.1; **p<0.05; ***p<0.01	

4 Explaining Citations

Citing requires a *bib* file with all the books. You can create it from Zotero, and then add it here with the command (Heath & Heath, 2010). For example, open the file named 'GovernanceAnalytics' and write the name of the author here (Magallanes Reyes, 2017).

References

Heath, C., & Heath, D. (2010). *Switch: how to change things when change is hard* (1st ed ed.). New York: Broadway Books.

Magallanes Reyes, J. M. (2017). *Introducing to data science for social and policy research: collecting to organizing data with R and Python.* Cambridge, United Kingdom New York, USA Port Melbourne, Australia Delhi, India Singapore: Cambridge University Press. (OCLC: 998518603)