Comparing Rnw to Rmd

An example

Robert McDonald[*]
Joe Smith[†]

August 7, 2018

Abstract

It's possible to create a document using either LaTeXcoupled with R code, or using Rmarkdown coupled with R code. How hard is it to do one vs the other? This is the Rmarkdown version.

Contents

[*]Kellogg School of Management, Northwestern University
[†]Someother University

Table 1: Top 10 names, 1880-2015

Name	Total	Year most popular	Proportion that year
James	5,144,205	1944	0.0554
John	5,117,331	1880	0.0815
Robert	4,823,167	1931	0.0566
Michael	4,345,569	1969	0.0466
Mary	4,133,216	1880	0.0724
William	4,087,556	1880	0.0805
David	3,602,623	1955	0.0413
Joseph	2,592,388	1914	0.0276
Richard	2,567,700	1946	0.0357
Charles	2,383,998	1880	0.0452

1 Introduction

This is a test document to illustrate the difference between Rmd and Rnw files.

1.1 Auto-references and equations

It is possible to include equations (such as Equation 1 below) and references (such as to Wickham and Grolemund (2017) or Schmuller (2017))

$$\alpha + \beta/\Delta + \int_0^\infty f(x)dx \tag{1}$$

2 Babynames

2.1 Top 10 names historically

Table 1 shows the top names historically, from 1880 to 2015.

2.2 Popularity of Mary and James

Figure 1 shows the popularity of the names Mary and James over 136 years. The placement of the figure is tricky. Latex permits the placement options “hptb”.

When placing the figure, knitr has to decide whether to write a figure environment or a markdown figure specification (`![]()`). The way to force a figure environment (from this post on stackoverflow) is to use a figure caption and at least one of fig.align, out.width, or out.extra. It works to use `out.extra=''`, which is what I do in this document.

Appendices

A Code

A.1 Table code

Here is the code used to produce Table 1:

```
topnames <- babynames %>%
    group_by(name) %>%
    summarize(Total=sum(n)) %>%
    arrange(-Total) %>%
    filter(row_number() <= 10)

tbl <- left_join(topnames, babynames) %>%
    group_by(name, Total) %>%
    summarize('Year most popular'=as.character(year[which.max(prop)]),
              'Proportion that year'=max(prop)) %>%
    arrange(-Total) %>%
    rename(Name=name) %>%
    kable(caption='Top 10 names, 1880-2015\\label{tbl:top10}',
          digits=4,
          booktabs=TRUE,
          vline='',
          linesep='',
          format='latex',
          format.args=list(big.mark=','),
          align='lccc'
          )
kable_styling(tbl, "latex")
```

A.2 Figure code

Here is the code used to produce Figure 1:

```
library(tidyverse)
library(babynames)
```

```r
library(knitr)
library(kableExtra)
opts_chunk$set(message=FALSE,
               comment=NA)
names <- c('Mary', 'James')
babynames %>%
    filter(name %in% c('James', 'Mary')) %>%
    ggplot(aes(x=year, y=prop, color=sex)) +
    geom_point() +
    facet_wrap( ~ name, scales='fixed') + theme_bw()
```

References

Schmuller, Joseph. 2017. *Statistical Analysis with R for Dummies*. Wiley.

Wickham, Hadley, and Garrett Grolemund. 2017. *R for Data Science*. O'Reilly.

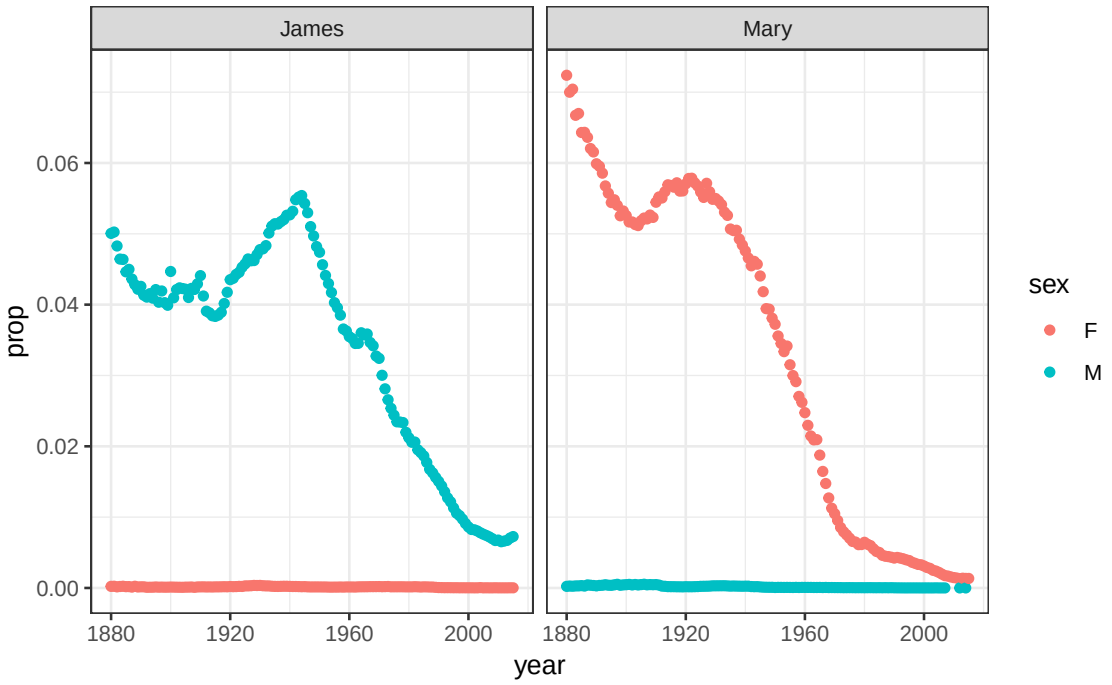

Figure 1: Popularity of James and Mary from 1880 to 2015