university of groningen

R Markdown for fun and for profit

Chris Davis
c.b.davis@rug.nl
http://cbdavis.org
March 31, 2016

Motivations

- You made a plot three months ago, need to revise it tonight and can't remember how you made it
- You need to transparently share code and results with other people
- You need to create a data analytics tutorial for students
- You want students to better document their analysis

university of groningen

Reproducible Research

"Reproducible research is the idea that data analyses, and more generally, scientific claims, are published with their data and software code so that others may verify the findings and build upon them."

university of groningen

Markdown Philosophy

> *"A Markdown-formatted document should be publishable as-is, as plain text, without looking like it's been marked up with tags or formatting instructions."*
>
> -John Gruber

university of groningen

What is R Markdown?

Dynamic Documents for R

R Markdown is an authoring format that enables easy creation of dynamic documents, presentations, and reports from R. It combines the core syntax of markdown (an easy to write plain text format) with embedded R code chunks that are run so their output can be included in the final document. R Markdown documents are fully reproducible (they can be automatically regenerated whenever underlying R code or data changes).

R Markdown has many available output formats including HTML, PDF, MS Word, Beamer, HTML5 presentations, Tufte handouts, R package vignettes, and even entire websites.

http://rmarkdown.rstudio.com/

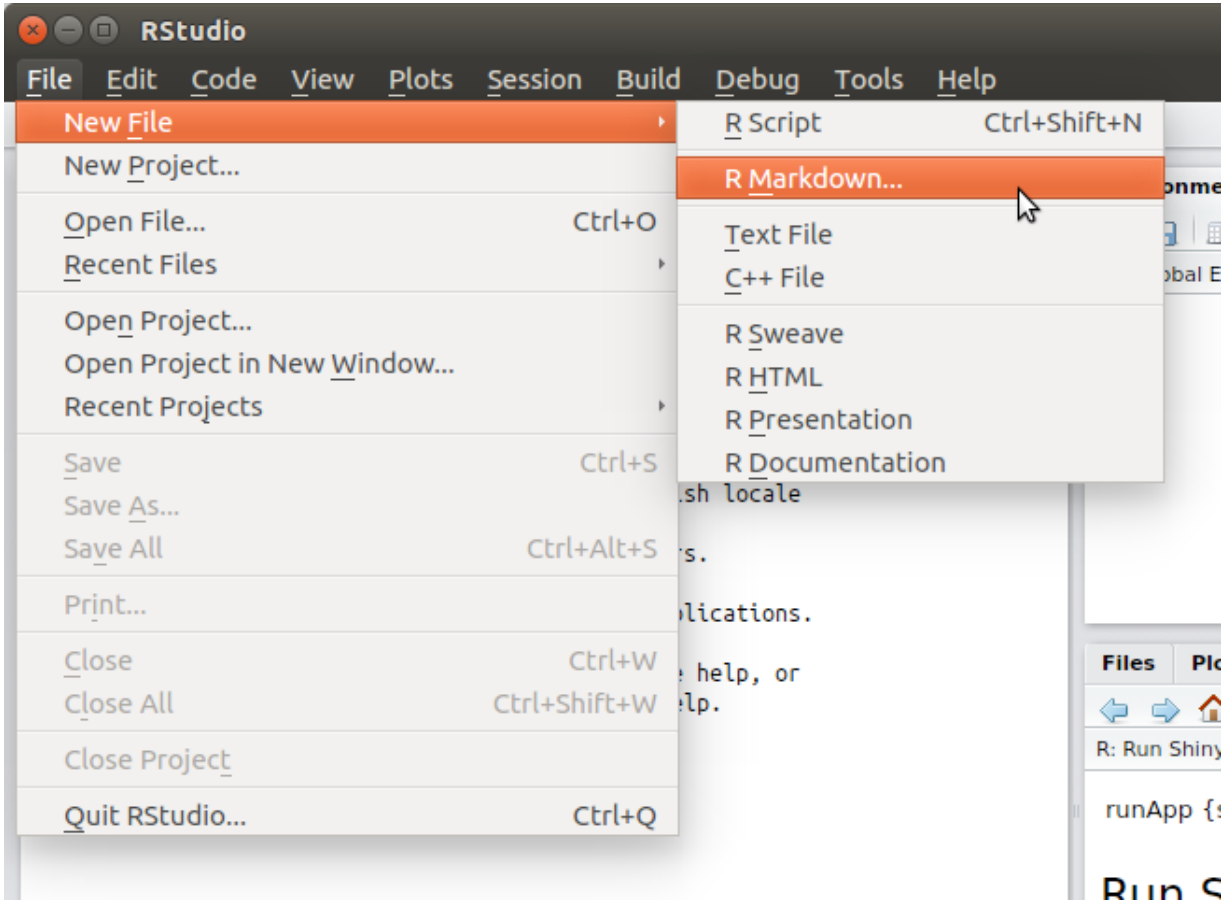
university of groningen

Cool features

- Output to HTML, PDF, Word
- Images encoded within HTML doc
- Caching of code block results
- Control size/resolution of images
- Embed formulas, animations, etc.
- Works well with GitHub

university of groningen

Simple Example



Tufte Handouts

Overview

Tufte Handouts are documents formatted in the style that Edward Tufte uses in his books and handouts. Tufte's style is known for its extensive use of sidenotes, tight integration of graphics with text, and well-set typography:

Figures

Margin Figures

Images and graphics play an integral role in Tufte's work. To place figures or tables in the margin you can use the `fig.margin` knitr chunk option. For example:

```
library(ggplot2)
qplot(Sepal.Length, Petal.Length, data = iris,
    color = Species)
```

Note the use of the `fig.cap` chunk option to provide a figure caption. You can adjust the proportions of figures using the `fig.width` and `fig.height` chunk options. These are specified in inches, and will be automatically scaled down to fit within the handout margin.

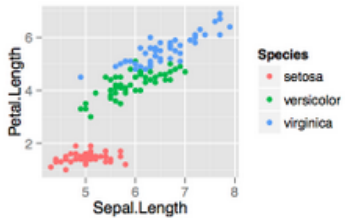


Figure 1: Sepal length vs. petal length, colored by species

Equations

You can also include LaTeX equations in the margin by explicitly invoking the `marginfigure` environment.

Note the use of the `\caption` command to add additional text below the equation.

$$\frac{d}{dx}\left(\int_0^x f(u)\,du\right) = f(x).$$

Figure 2: An equation

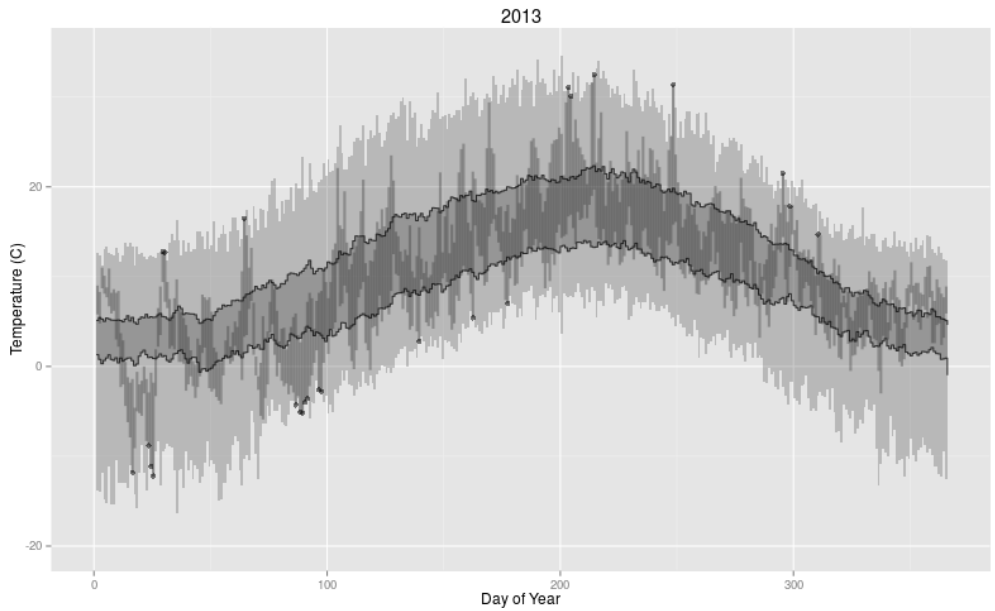
More Advanced Example...

A Visual Exploration of Rotterdam Temperature Data

Chris Davis

November 15, 2015

TL;DR Explore temperature data and learn how to make plots like these using R:



https://github.com/cbdavis/Visual-Exploration-of-Rotterdam-Temperature-Data

Resources

- R Markdown
  - Main page
  - Cheatsheet
  - Reference guide
- Jupyter
  - Similar project covering more languages
- Pandoc
  - Swiss army knife of file markup format converters