PSY 503: Foundations of Statistical Methods in Psychological Science

Lab 02

R-Markdown, ggplot

Suyog Chandramouli

Zoom & 311 PSH (Princeton University)

17th September, 2025

R markdown = R + Markdown

Markdown: a simple text format for writing documents.

R markdown: a tool for mixing R code within markdown

R markdown is a “chimera”

Input combines:

- Text in markdown
- R code
- “metadata” with YAML
- Equations with LaTeX

R markdown is a “chimera”

Output produced can be:

- HTML
- PDF
- .docx

.. *to name a few*

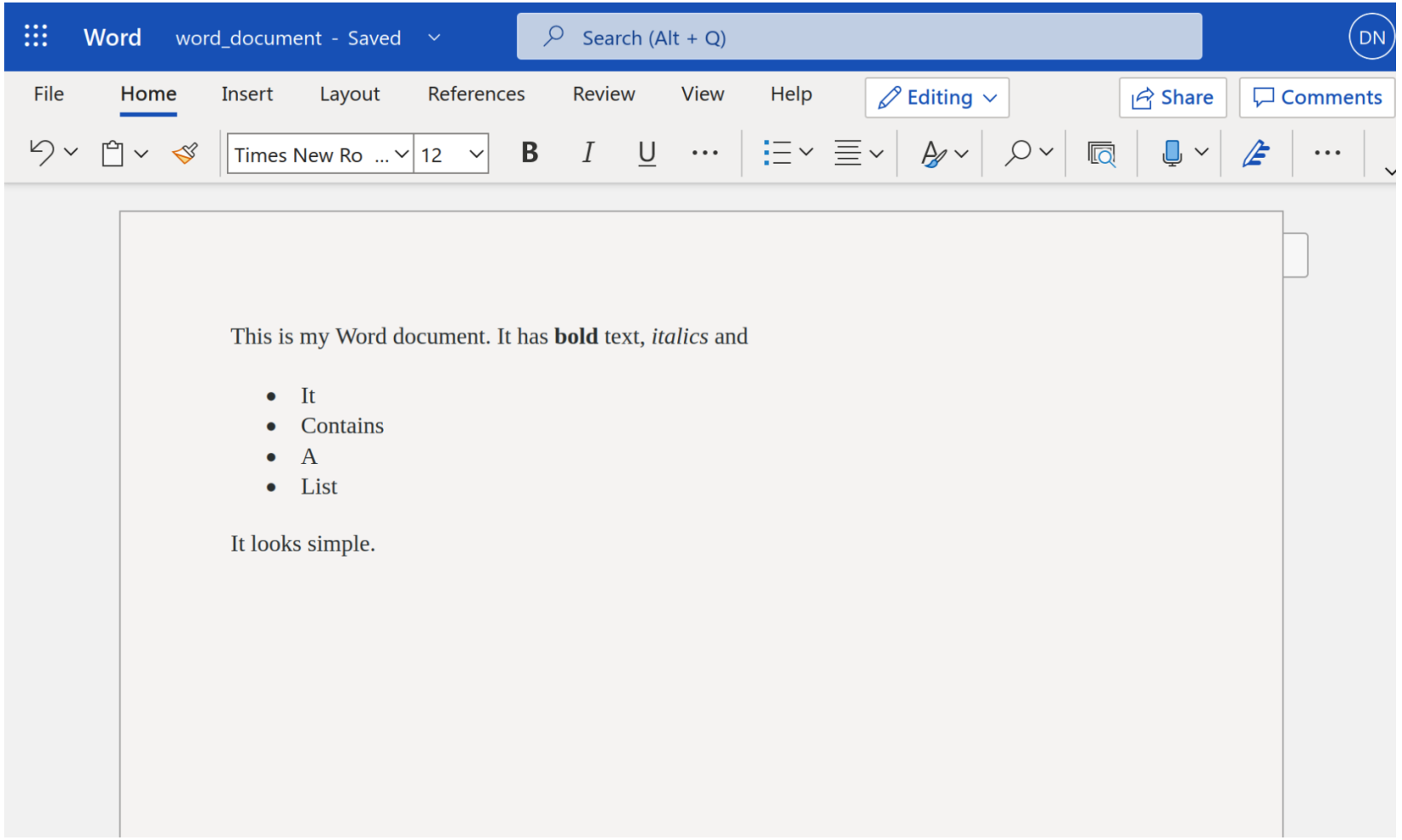
Markdown 

- What is the problem it is trying to solve?

Markdown 

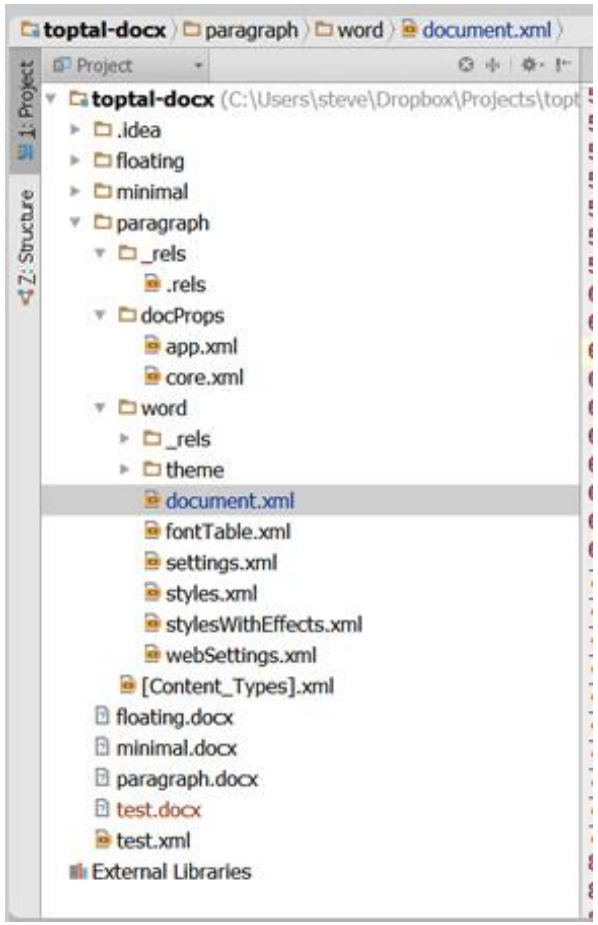
- What is the problem it is trying to solve?

A simple word document..



Is in fact, quite complicated

```
<?xml version="1.0" encoding="utf-8" standalone="yes"?><w:document xmlns:o="
urn:schemas-microsoft-com:office:office" xmlns:r="http://schemas.openxmlformats.org/
officeDocument/2006/relationships" xmlns:v="urn:schemas-microsoft-com:vml" xmlns:w="
http://schemas.openxmlformats.org/wordprocessingml/2006/main" xmlns:w10="
urn:schemas-microsoft-com:office:word" xmlns:wp="http://schemas.openxmlformats.org/
drawingml/2006/wordprocessingDrawing" xmlns:wps="http://schemas.microsoft.com/office/
word/2010/wordprocessingShape" xmlns:wpg="http://schemas.microsoft.com/office/word/2010/
wordprocessingGroup" xmlns:mc="http://schemas.openxmlformats.org/
markup-compatibility/2006" xmlns:wp14="http://schemas.microsoft.com/office/word/2010/
wordprocessingDrawing" xmlns:w14="http://schemas.microsoft.com/office/word/2010/wordml"
mc:Ignorable="w14 wp14"><w:body><w:p w:rsidP="3DB3F1AF" w14:paraId="53FDABE3" xmlns:wp14
="http://schemas.microsoft.com/office/word/2010/wordml" wp14:textId="6B19D258"><w:pPr><w
:pStyle w:val="Normal" /><w:bidi w:val="0" /><w:jc w:val="left" /><w:rPr><w:rFonts w:
ascii="Times New Roman" w:hAnsi="Times New Roman" w:eastAsia="Times New Roman" w:cs="
Times New Roman" /></w:rPr></w:pPr><w:r w:rsidRPr="3DB3F1AF" w:rsidR="3DB3F1AF"><w:rPr><
w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman" w:eastAsia="Times New Roman
" w:cs="Times New Roman" /></w:rPr><w:t xml:space="preserve">This is my Word document.
</w:t></w:r><w:r w:rsidRPr="3DB3F1AF" w:rsidR="3DB3F1AF"><w:rPr><w:rFonts w:ascii="
Times New Roman" w:hAnsi="Times New Roman" w:eastAsia="Times New Roman" w:cs="Times New
Roman" /></w:rPr><w:t xml:space="preserve">It has </w:t></w:r><w:r w:rsidRPr="3DB3F1AF"
w:rsidR="3DB3F1AF"><w:rPr><w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman"
w:eastAsia="Times New Roman" w:cs="Times New Roman" /><w:b w:val="1" /><w:bCs w:val="1"
/><w:i w:val="0" /><w:iCs w:val="0" /></w:rPr><w:t xml:space="preserve">bold </w:t></w:r
><w:r w:rsidRPr="3DB3F1AF" w:rsidR="3DB3F1AF"><w:rPr><w:rFonts w:ascii="Times New Roman"
 w:hAnsi="Times New Roman" w:eastAsia="Times New Roman" w:cs="Times New Roman" /></w:rPr
><w:t xml:space="preserve">text, </w:t></w:r><w:r w:rsidRPr="3DB3F1AF" w:rsidR="3DB3F1AF
```



Markdown 

- What is the problem it is trying to solve?
  - In most document formats, text is buried in with a lot of formatting information.

A simple html document..

This is my HTML document. It has **bold** text, *italics* and

- It
- Contains
- A
- List

It looks simple.

A simple html document..

This is my HTML document. It has **bold** text, *italics* and

- It
- Contains
- A
- List

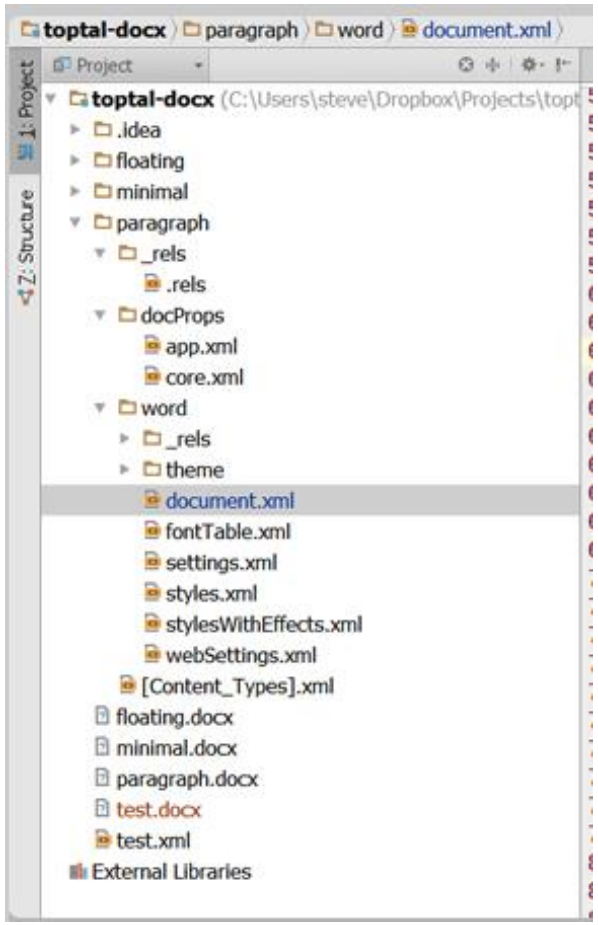
It looks simple.

With most formatting removed, is still unpleasant to type, edit, etc.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<html>
<head>
    <meta http-equiv="content-type" content="text/html; charset=utf-8"/>
</head>
<body>
    <p>This is my HTML document. It has <b>bold</b> text, <i>italics</i> and</p>
    <br>
    <ul>
        <li>It</li>
        <li>Contains</li>
        <li>A</li>
        <li>List</li>
    </ul>
    <br>
    <p>It looks simple.</p>
</body>
</html>
```

Is in fact, quite complicated

```
<?xml version="1.0" encoding="utf-8" standalone="yes"?><w:document xmlns:o="
urn:schemas-microsoft-com:office:office" xmlns:r="http://schemas.openxmlformats.org/
officeDocument/2006/relationships" xmlns:v="urn:schemas-microsoft-com:vml" xmlns:w="
http://schemas.openxmlformats.org/wordprocessingml/2006/main" xmlns:w10="
urn:schemas-microsoft-com:office:word" xmlns:wp="http://schemas.openxmlformats.org/
drawingml/2006/wordprocessingDrawing" xmlns:wps="http://schemas.microsoft.com/office/
word/2010/wordprocessingShape" xmlns:wpg="http://schemas.microsoft.com/office/word/2010/
wordprocessingGroup" xmlns:mc="http://schemas.openxmlformats.org/
markup-compatibility/2006" xmlns:wp14="http://schemas.microsoft.com/office/word/2010/
wordprocessingDrawing" xmlns:w14="http://schemas.microsoft.com/office/word/2010/wordml"
mc:Ignorable="w14 wp14"><w:body><w:p w:rsidP="3DB3F1AF" w14:paraId="53FDABE3" xmlns:wp14
="http://schemas.microsoft.com/office/word/2010/wordml" wp14:textId="6B19D258"><w:pPr><w
:pStyle w:val="Normal" /><w:bidi w:val="0" /><w:jc w:val="left" /><w:rPr><w:rFonts w:
ascii="Times New Roman" w:hAnsi="Times New Roman" w:eastAsia="Times New Roman" w:cs="
Times New Roman" /></w:rPr></w:pPr><w:r w:rsidRPr="3DB3F1AF" w:rsidR="3DB3F1AF"><w:rPr><
w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman" w:eastAsia="Times New Roman
" w:cs="Times New Roman" /></w:rPr><w:t xml:space="preserve">This is my Word document.
</w:t></w:r><w:r w:rsidRPr="3DB3F1AF" w:rsidR="3DB3F1AF"><w:rPr><w:rFonts w:ascii="
Times New Roman" w:hAnsi="Times New Roman" w:eastAsia="Times New Roman" w:cs="Times New
Roman" /></w:rPr><w:t xml:space="preserve">It has </w:t></w:r><w:r w:rsidRPr="3DB3F1AF"
w:rsidR="3DB3F1AF"><w:rPr><w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman"
w:eastAsia="Times New Roman" w:cs="Times New Roman" /><w:b w:val="1" /><w:bCs w:val="1"
/><w:i w:val="0" /><w:iCs w:val="0" /></w:rPr><w:t xml:space="preserve">bold </w:t></w:r
><w:r w:rsidRPr="3DB3F1AF" w:rsidR="3DB3F1AF"><w:rPr><w:rFonts w:ascii="Times New Roman"
 w:hAnsi="Times New Roman" w:eastAsia="Times New Roman" w:cs="Times New Roman" /></w:rPr
><w:t xml:space="preserve">text, </w:t></w:r><w:r w:rsidRPr="3DB3F1AF" w:rsidR="3DB3F1AF
```



Markdown 

- What is the problem it is trying to solve?
  - In most document formats, text is buried in with a lot of formatting information.

- We want something that..
  - can be written in **plain text**
  - is **human readable**
  - allows **formatting**
  - is as **capable** as the formats it is trying to replace.

Example of a markdown document

```
# Introduction

Welcome to my **awesome** class. You will learn
all kinds of useful things about R markdown.

## Why should you care?

- Markdown is simple and reproducible
- You can make it pretty if you want to
- The R Markdown variant lets you add R code
```

Output

Introduction

Welcome to my awesome class. You will learn all kinds of useful things about R markdown.

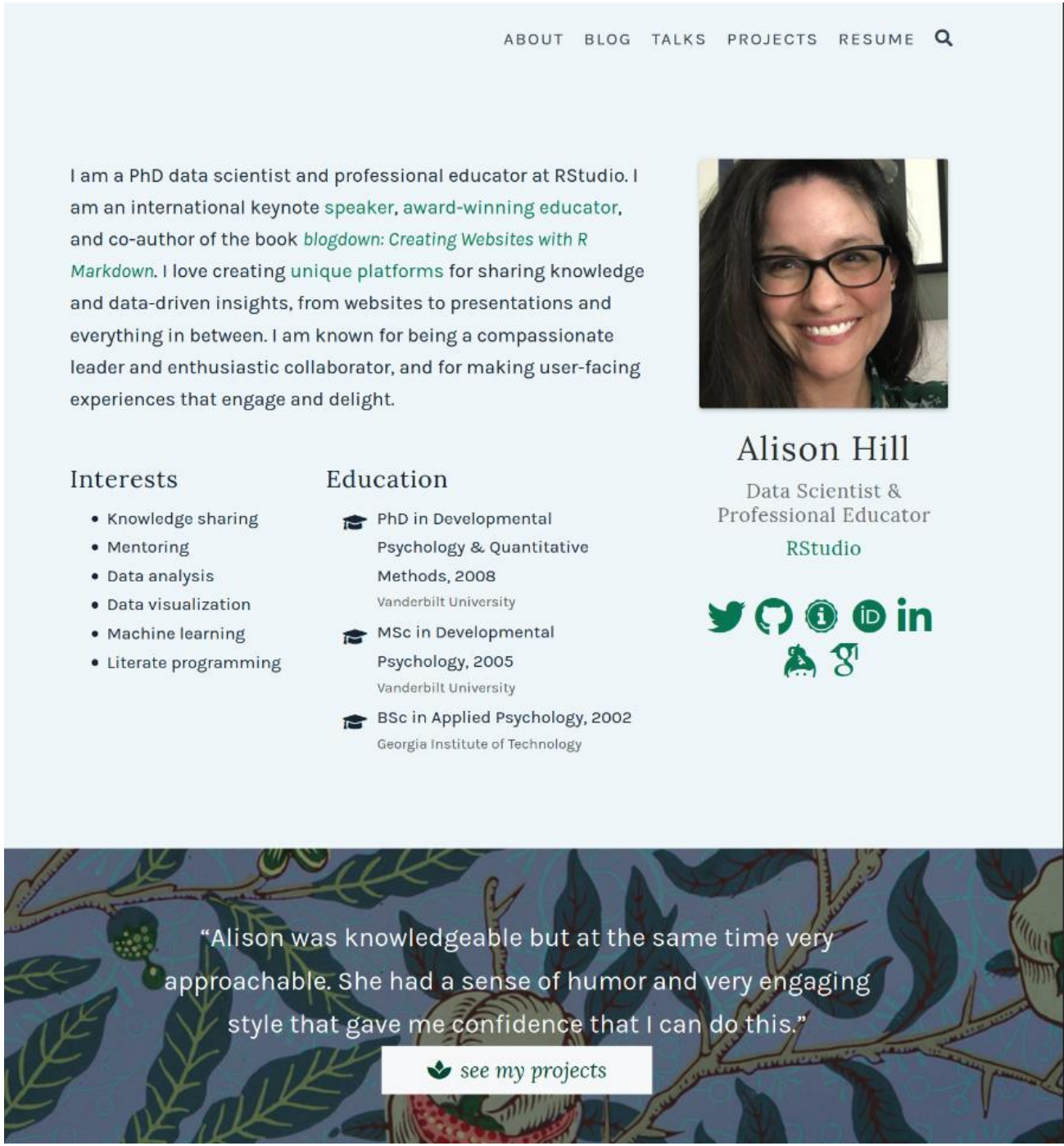
Why should you care?

- Markdown is simple and reproducible
- You can make it pretty if you want to
- The R Markdown variant lets you add R code

Boring?

- Turns out it is both capable and generally usable

Websites


ABOUT BLOG TALKS PROJECTS RESUME
I am a PhD data scientist and professional educator at RStudio. I am an international keynote speaker, award-winning educator, and co-author of the book blogdown: Creating Websites with R Markdown. I love creating unique platforms for sharing knowledge and data-driven insights, from websites to presentations and everything in between. I am known for being a compassionate leader and enthusiastic collaborator, and for making user-facing experiences that engage and delight.
Alison Hill
Data Scientist & Professional Educator
RStudio
Interests
Knowledge sharing
Mentoring
Data analysis
Data visualization
Machine learning
Literate programming
Education
PhD in Developmental Psychology & Quantitative Methods, 2008
Vanderbilt University
MSc in Developmental Psychology, 2005
Vanderbilt University
BSc in Applied Psychology, 2002
Georgia Institute of Technology
"Alison was knowledgeable but at the same time very approachable. She had a sense of humor and very engaging style that gave me confidence that I can do this."
see my projects

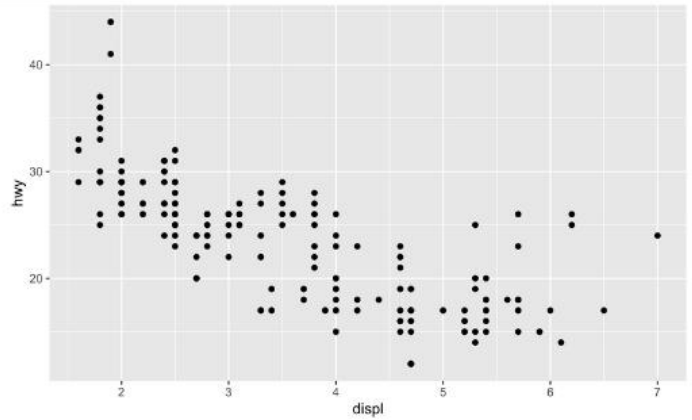
Books

3.2.2 Creating a ggplot

To plot `mpg`, run this code to put `displ` on the x-axis and `hwy` on the y-axis:

```
ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy))
```

Copy



The plot shows a negative relationship between engine size (`displ`) and fuel efficiency (`hwy`). In other words, cars with big engines use more fuel. Does this confirm or refute your hypothesis about fuel efficiency and engine size?

With ggplot2, you begin a plot with the function `ggplot()`. `ggplot()` creates a coordinate system that you can add layers to. The first argument of `ggplot()` is the dataset to use in the graph. So `ggplot(data = mpg)` creates an empty graph, but it's not very interesting so I'm not going to show it here.

You complete your graph by adding one or more layers to `ggplot()`. The function `geom_point()` adds a layer of points to your plot, which creates a scatterplot. ggplot2 comes with many geom functions that each add a different type of layer to a plot. You'll learn a whole bunch of them throughout this chapter.

Papers

Between the devil and the deep blue sea: Tensions between scientific judgement and statistical model selection

Danielle J. Navarro[1]

[1] University of New South Wales

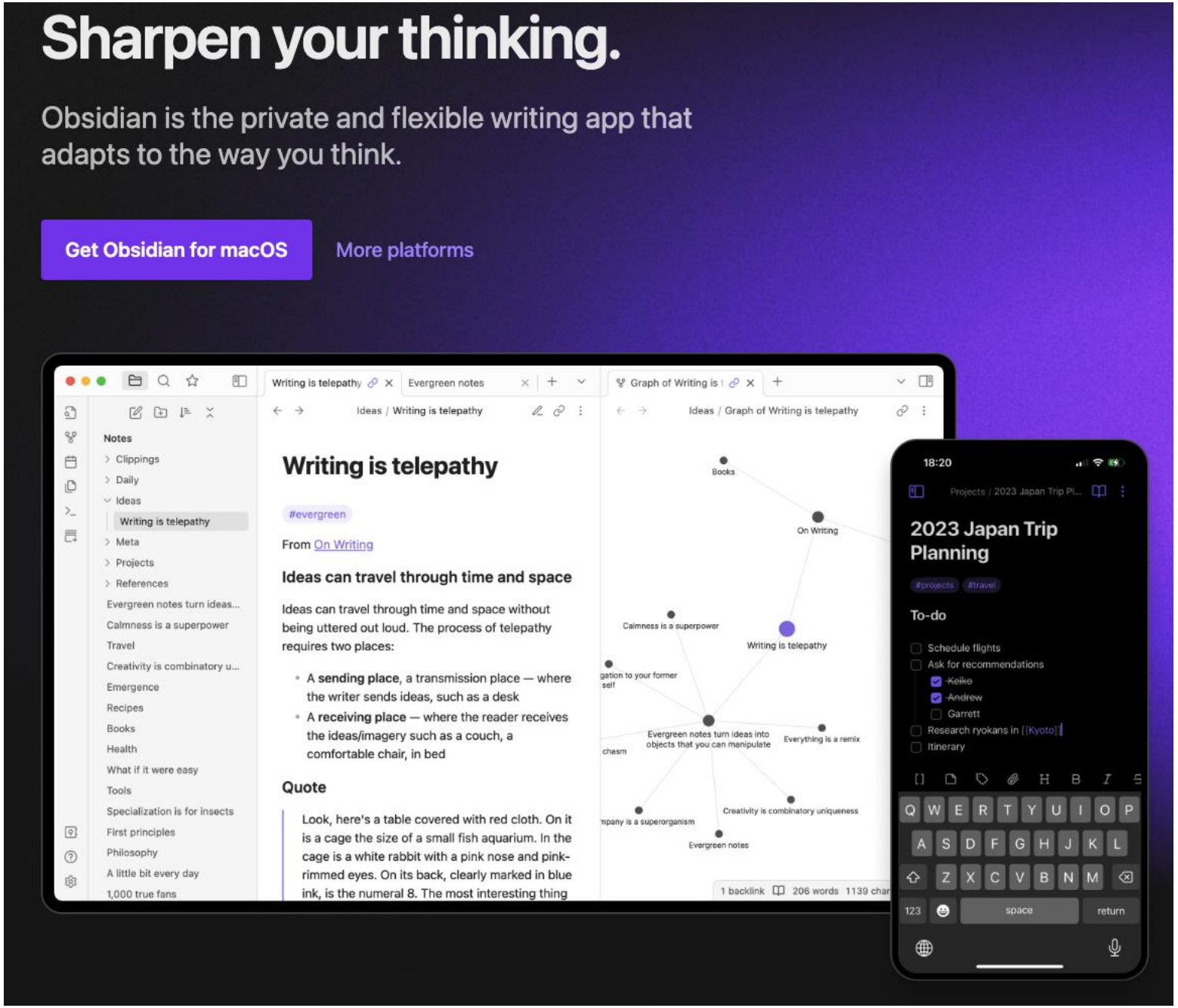
Abstract

Discussions of model selection in the psychological literature typically frame the issues as a question of statistical inference, with the goal being to determine which model makes the best predictions about data. Within this setting, advocates of leave-one-out cross-validation and Bayes factors disagree on precisely which prediction problem model selection questions should aim to answer. In this comment, I discuss some of these issues from a scientific perspective. What goal does model selection serve when all models are known to be systematically wrong? How might "toy problems" tell a misleading story? How does the scientific goal of explanation align with (or differ from) traditional statistical concerns? I do not offer answers to these questions, but hope to highlight the reasons why psychological researchers cannot avoid asking them.

Keywords: model selection, science, statistics

Model selection seems to be an evergreen topic in mathematical psychology. Given two or more competing theories about the world, each instantiated as parameterised computational models that provide different accounts of a data set, how should we decide which model is better supported by the data? Typically we formulate this as a statistical inference problem, with various authors arguing for Bayes factors (e.g., Wagenmakers 2007), minimum description length (e.g., Grünwald 2007), cross-validation (e.g., Browne 2000) and a variety of other possibilities besides. To highlight the behaviour of different model selec-

Note-taking software



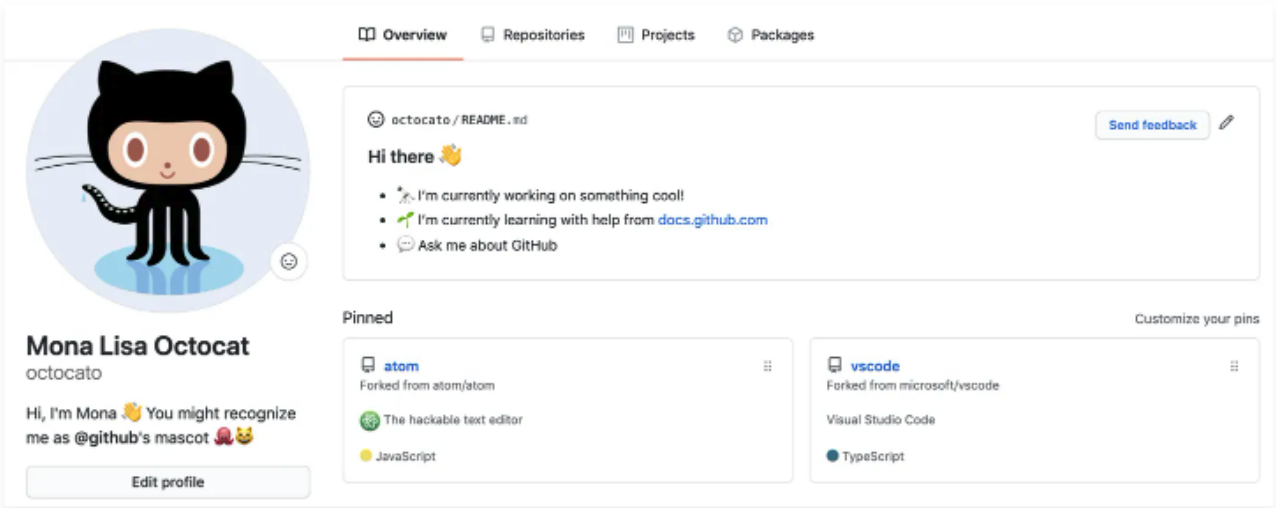
Readme.md

About your profile README

You can share information about yourself with the community on GitHub by creating a profile README. GitHub shows your profile README at the top of your profile page.

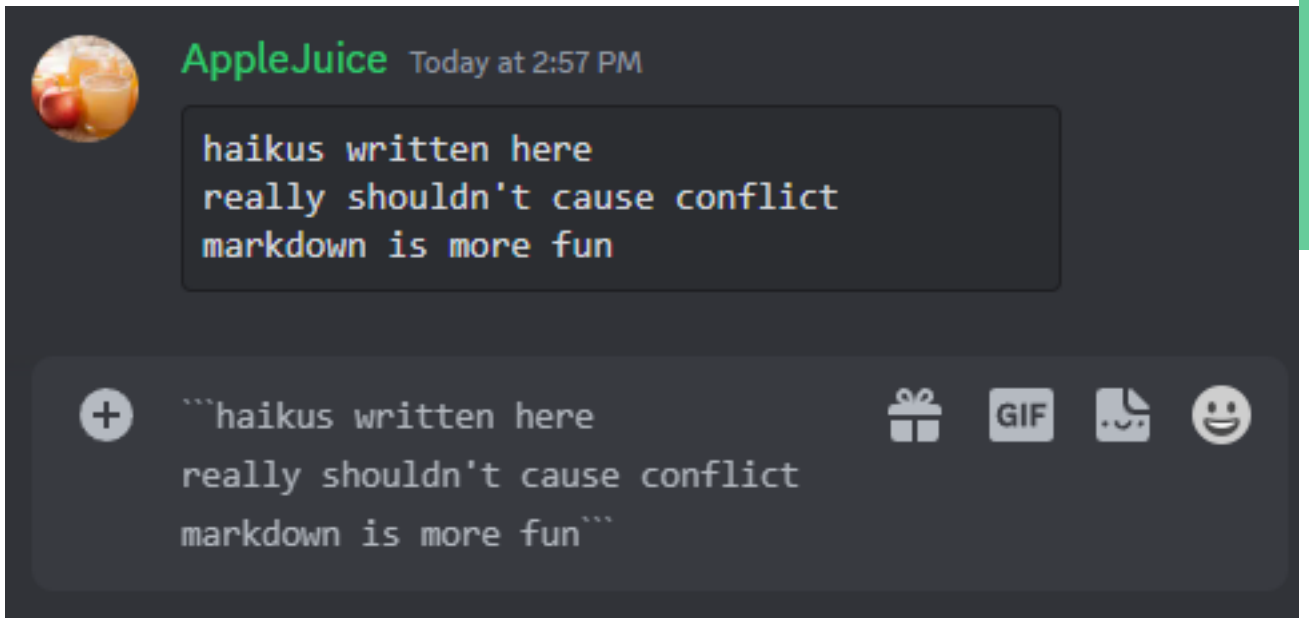
You decide what information to include in your profile README, so you have full control over how you present yourself on GitHub. Here are some examples of information that visitors may find interesting, fun, or useful in your profile README.

- An "About me" section that describes your work and interests
- Contributions you're proud of, and context about those contributions
- Guidance for getting help in communities where you're involved



You can format text and include emoji, images, and GIFs in your profile README by using GitHub Flavored Markdown. For more information, see "Getting started with writing and formatting on GitHub." For a hands-on guide to customizing your profile README, see "Quickstart for writing on GitHub."

Whatsapp, Discord..



Common markdown formatting details..

Text emphasis

What you type...

```
this is *italics*

this is **bold**

this is ***bold italics***
```

What you get...

this is *italics*

this is **bold**

this is ***bold italics***

Lists

What you type...

```
- unnumbered lists
- look like this

1. numbered lists
2. look like this
```

What you get...

- unnumbered lists
- look like this

1. numbered lists
2. look like this

Headings

What you type...

```
# Level 1 heading
## Level 2 heading
### Level 3 heading
```

What you get...

Level 1 heading
Level 2 heading
Level 3 heading

And more..

- Hyperlinks
- Image insertion
- Block quotes

...

Comments in .md

<!-- -->

Comments in .md

<!– ***text to comment out*** **-->**

R markdown = R + Markdown

Markdown: a simple text format for writing documents.

R markdown: a tool for mixing R code within markdown

Demo: MD

R markdown = R + Markdown

Markdown: a simple text format for writing documents.

R markdown: a tool for mixing R code within markdown

What are the benefits of mixing code with markdown?

What are the benefits of mixing code with markdown? (Discuss)

- Why do you use it?
- Have you found it useful?

Some use-cases

- Report generation
  - Across formats
  - Removes manual work for each iteration
  - Submitting homeworks
- Interactive Tutorials
- Reproducibility
  - Code connected to outputs
  - Code explained (***literate programming***)
  - Not only at publication, but even for e.g. lab-presentations

R markdown

Three main elements

- R code
- Text of the report (markdown)
- Metadata

Demo: RMD

````
---
title: "Lab 02 - RMD Demo"
author: "Suyog Chandramouli"
date: "2024-09-18"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```{r cars}
summary(cars)
```

## Including Plots

You can also embed plots, for example:

```{r pressure, echo=FALSE}
plot(pressure)
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
````

```
---
title: "Lab 02 - RMD Demo"
author: "Suyog Chandramouli"
date: "2024-09-18"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```{r cars}
summary(cars)
```

Including Plots

You can also embed plots, for example:

```{r pressure, echo=FALSE}
plot(pressure)
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
```

The "YAML" header / Metadata

YAML = "Yet another mark-up language"

YAML = "YAML ain't mark-up language"

.. doesn't really matter

```
---
title: "Lab 02 - RMD Demo"
author: "Suyog Chandramouli"
date: "2024-09-18"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```{r cars}
summary(cars)
```

Including Plots

You can also embed plots, for example:

```{r pressure, echo=FALSE}
plot(pressure)
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
```

The “YAML” header / Metadata

```{r}
# holds R code
```

```
---
title: "Lab 02 - RMD Demo"
author: "Suyog Chandramouli"
date: "2024-09-18"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```{r cars}
summary(cars)
```

Including Plots

You can also embed plots, for example:

```{r pressure, echo=FALSE}
plot(pressure)
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
```

The “YAML” header / Metadata

```{r}
# holds R code
```

R chunk output suppression

The following table summarises which types of output each option supressess:

Option	Run code	Show code	Output	Plots	Messages	Warnings
`eval = FALSE`	-		-	-	-	-
`include = FALSE`		-	-	-	-	-
`echo = FALSE`		-				
`results = "hide"`			-			
`fig.show = "hide"`				-		
`message = FALSE`					-	
`warning = FALSE`						-

```
---
title: "Lab 02 - RMD Demo"
author: "Suyog Chandramouli"
date: "2024-09-18"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```{r cars}
summary(cars)
```

Including Plots

You can also embed plots, for example:

```{r pressure, echo=FALSE}
plot(pressure)
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.
```

The "YAML" header / Metadata

The rest is markdown text

```{r}
# holds R code
```

R markdown

Three main elements

- R code
- Text of the report (markdown) – the narrative
- Metadata

- These elements are “knit” together into the output
  - Behind the hood:
    - rmarkdown::render() is called

      (a) compiled to Markdown (*.md) through knitr.

      (b) *.md is compiled to other output formats (such as LaTeX or HTML) through Pandoc

rmarkdown::render()

To render an R Markdown file, you can use the following basic command

```
rmarkdown::render("path/to/your/file.Rmd")
```

Particularly relevant if you’re not using RStudio

Equations

- Anything inside dollar signs $ is treated as "inline" maths
- Anything inside two dollar $$ is a standalone equation
- Whitespace matters: x is an equation, $x $ is not
- Equations follow "LaTeX" rules

Equations are special

```
This $x^2$ is inline

This equation is standalone
$$
a^2 + b^2 = c^2
$$
```

This x^2 is inline

This equation is standalone

$$a^2 + b^2 = c^2$$

Inline Math

The solution to $\sqrt{x} = 26$ is $x = 5$.

This formula $f(y) = x^2$ is another example.

Display Math

```
$$
P(E) = {n \choose k} p^k (2-p)^{n-k}
$$
```

$$P(E) = \binom{n}{k} p^k (2-p)^{n-k}$$

Some disadvantages

What have you disliked about rmd?

Sometimes we just want r files.

- See:
  - ***knitr::purl()*** extracts R code chunks from a **knitr** document and save the code to an R script.

RMD vs. QMD (Quarto)

You can use QMD exactly like RMD for your R projects, but it opens doors for more.

QMD is like RMD++

- Developed by Posit (formerly RStudio)
- You can also use Python & Julia
- More features
  - More layouts
  - Cross-referencing & citations