

# BUILDING YOUR PERSONAL PORTFOLIO

Steven Mortimer | March 1, 2018

# About Me

I am a Wahoo!



BA Statistics - Dec '09

MS Statistics - May '11

# About Me



**2011-2015**  
Advisory Board Company  
Washington, DC  
*Data Scientist*

---



**2015-2017**  
Homes.com, Inc.  
Norfolk, VA  
*Senior Manager of Data Science*

Getting back to building portfolios...

# Why Building a Digital Profile is Good

- Practice coding skills
- Practice communication skills
- Practice collaboration skills
- Practice!

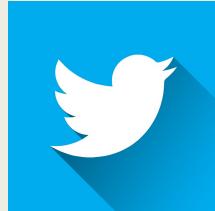


# Other Reasons to Build a Digital Profile



- To prove you are a skilled job candidate.
- You like it.
- You use it to bookmark learnings.

# Modes – GitHub & Everything Else



# Put Your GitHub URL on Your Resume

**YOUR NAME**

Your Email Address • Your Phone Number • [Your GitHub URL HERE](#) • City, State

---

A resume with a strong GitHub profile will open more doors than anything else.

# Repositories – Quality over Quantity

Hiring managers and future teammates are looking for:

1. Folder Structure
2. Code Structure
  1. *Functions?*
  2. *Comments?*
  3. *Readable?*
3. Aesthetics
4. User in mind (collaboration)?



# Upload Projects that Interest You

Love Shopping?



Love Fantasy Football?



Passion is  
interesting

# Make a Good README

Badges

README.md

dplyr

build failing build failing CRAN 0.7.4 codecov 79%

Overview

dplyr is a grammar of data manipulation, providing a consistent set of verbs that help you solve the most common data manipulation challenges:

- `mutate()` adds new variables that are functions of existing variables
- `select()` picks variables based on their names.

Code Examples

Images



# Add README Table of Contents

## Table of Contents

- [Getting Started](#)
- [Repository Folder Structure](#)
- [Data Details](#)

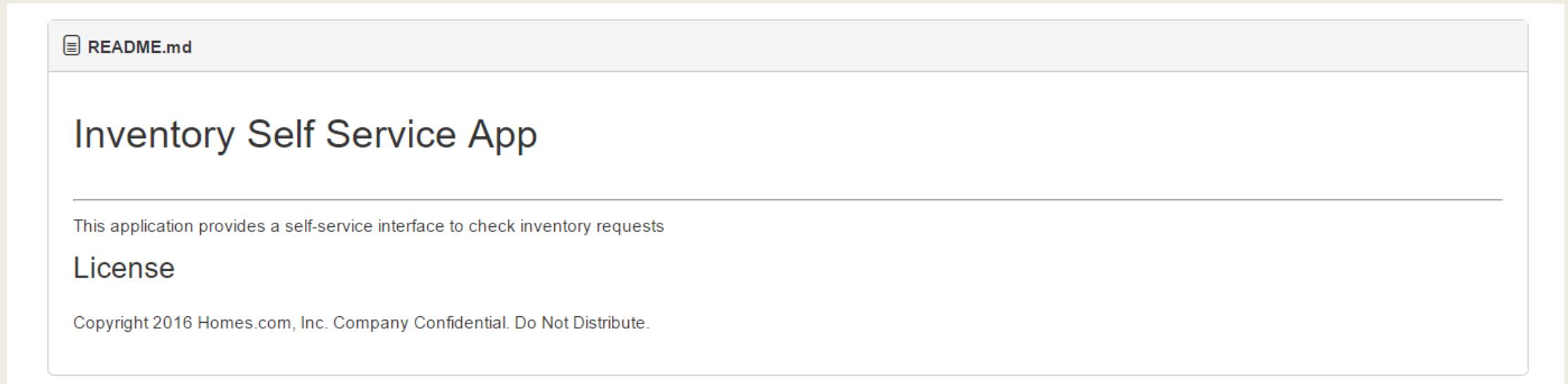
## Getting Started

```
## Table of Contents
- [Getting Started](#getting-started)
- [Repository Folder Structure](#repository-folder-structure)
- [Data Details](#data-details)

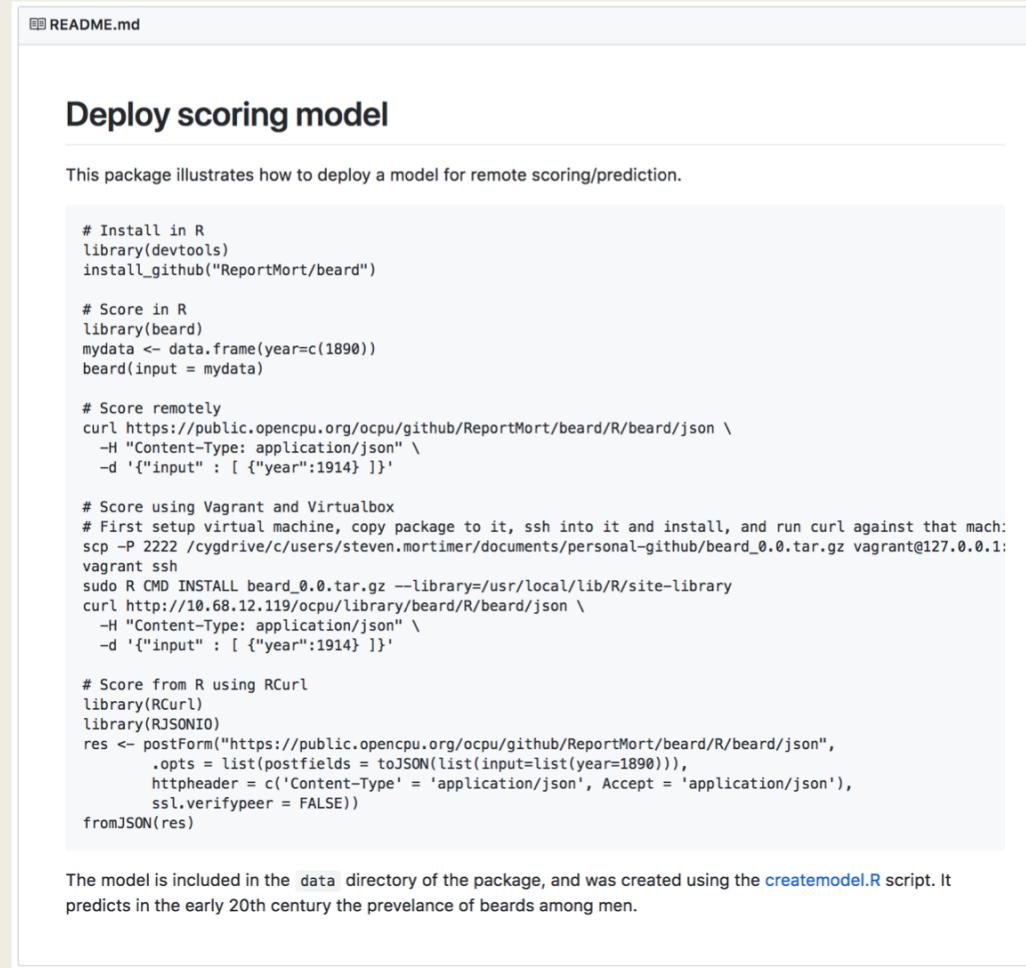
### Getting Started
```

**Note:** Adding a “Quick Start/Getting Started” section shows you’re proactive about collaborating with others and getting people up and running.

# Avoid the “Bare Bones” README



# Avoid the “Wall of Text” README



The screenshot shows a window titled "README.md" containing a section titled "Deploy scoring model". The text within the section reads:

This package illustrates how to deploy a model for remote scoring/prediction.

```
# Install in R
library(devtools)
install_github("ReportMort/beard")

# Score in R
library(beard)
mydata <- data.frame(year=c(1890))
beard(input = mydata)

# Score remotely
curl https://public.opencpu.org/opcpu/github/ReportMort/beard/R/beard/json \
-H "Content-Type: application/json" \
-d '{"input" : [ {"year":1914} ]}' 

# Score using Vagrant and Virtualbox
# First setup virtual machine, copy package to it, ssh into it and install, and run curl against that machine
scp -P 2222 /cygdrive/c/users/steven.mortimer/documents/personal-github/beard_0.0.tar.gz vagrant@127.0.0.1:
vagrant ssh
sudo R CMD INSTALL beard_0.0.tar.gz --library=/usr/local/lib/R/site-library
curl http://10.68.12.119/opcpu/library/beard/R/beard/json \
-H "Content-Type: application/json" \
-d '{"input" : [ {"year":1914} ]}' 

# Score from R using RCurl
library(RCurl)
library(RJSONIO)
res <- postForm("https://public.opencpu.org/opcpu/github/ReportMort/beard/R/beard/json",
.opts = list(postfields = toJSON(list(input=list(year=1890))),
httpheader = c('Content-Type' = 'application/json', Accept = 'application/json'),
ssl.verifypeer = FALSE))
fromJSON(res)
```

The model is included in the `data` directory of the package, and was created using the [createmodel.R](#) script. It predicts in the early 20th century the prevalence of beards among men.

# Summary of GitHub Recommendations

1. Resume + GitHub Link =  🔥
2. Assume your future employer is reading your code
3. Upload things you can speak passionately about
4. Make a great README

# One More Thing Concerning GitHub...

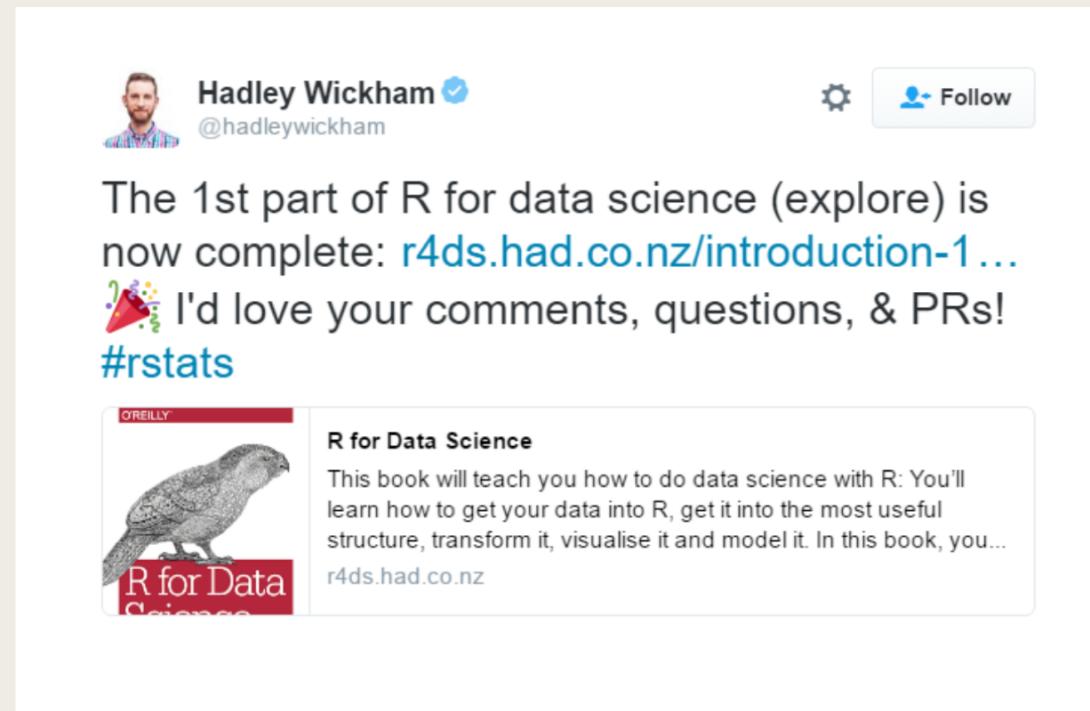
Think of an open source project you use.

Could the documentation be better?

Commit to it!

Add to your resume

# Taking the Open Source Invitation



Hadley Wickham  @hadleywickham

The 1st part of R for data science (explore) is now complete: [r4ds.had.co.nz/introduction-1...](http://r4ds.had.co.nz/introduction-1.html)

 I'd love your comments, questions, & PRs!

#rstats



**R for Data Science**

This book will teach you how to do data science with R: You'll learn how to get your data into R, get it into the most useful structure, transform it, visualise it and model it. In this book, you...

[r4ds.had.co.nz](http://r4ds.had.co.nz)

# Even Tiny Contributions are Valid!

Change `atually` to `actually` in intro.Rmd #175

Merged hadley merged 1 commit into hadley:master from unknown repository on Jul 23

Conversation 1 Commits 1 Files changed 1

ReportMort commented on Jul 22  
Change atually to actually

Fix Typo ...  
Change 'atually' to 'actually'

hadley merged commit 3800ebe into hadley:master on Jul 23  
View details  
1 check passed

hadley commented on Jul 23  
Owner +  
Thanks!

+1 -1 0 0 0

Projects  
None yet

Labels  
None yet

Milestone  
No milestone

Assignees  
No one assigned

2 participants

This screenshot shows a GitHub pull request titled "Change `atually` to `actually` in intro.Rmd #175". The pull request has been merged by "hadley" into the "hadley:master" branch from an "unknown repository" on July 23. The conversation section shows a comment from "ReportMort" on July 22 suggesting a typo correction. This led to a commit from "hadley" merging the changes. The pull request summary indicates 1 check passed. The right sidebar displays project, label, milestone, and assignee information, all of which are currently empty.

# Personal Website/Blog

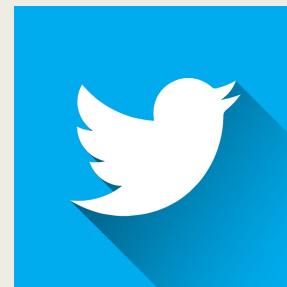
1. No website is better than a bad website. Hiring managers would rather look at your GitHub.

Data Scientist != Web Developer

2. It's hard to keep up! Serious time investment.

What is your vision and purpose?

# The Social Sites



# Uses for Twitter



- Promote your own blog posts
- Bookmark links or articles with “like”/“heart”
- Track influencers to stay current/knowledgeable

Likes

 **Jerry Gamblin**  @JGamblin · Feb 23  
Google gives anyone with Gmail account a free Debian 7 VM with 5Gb of storage. [console.cloud.google.com/cloudshell](http://console.cloud.google.com/cloudshell)

 36  1.1K  2.3K 

 **Patrick Boueri** @P\_Boueri · Feb 18  
Replying to @Maude\_LL @hadleywickham  
options(error=recover) allows you to jump into any part of the call stack right before the error occurs, giving you an exact state of your program. Highly recommend

 1  20  91 

# Leveraging Twitter for Job Interviews

## Interview Tip #1:

*Bond over recent community news.*

- “The Anaconda Distribution 5.0 came out and introduced JupyterLab, the next-gen version of Jupyter Notebooks. What do you think?!”
- “PyCon 2018 is coming up in May, is anyone from the office going?”

## Interview Tip #2:

*Have a preference on common tools.*

- “I like using PyCharm over Jupyter notebooks because of the better git functionality”
- “I like using Seaborn over Bokeh for plotting because it’s visually cleaner to my eye”

# Stack Overflow

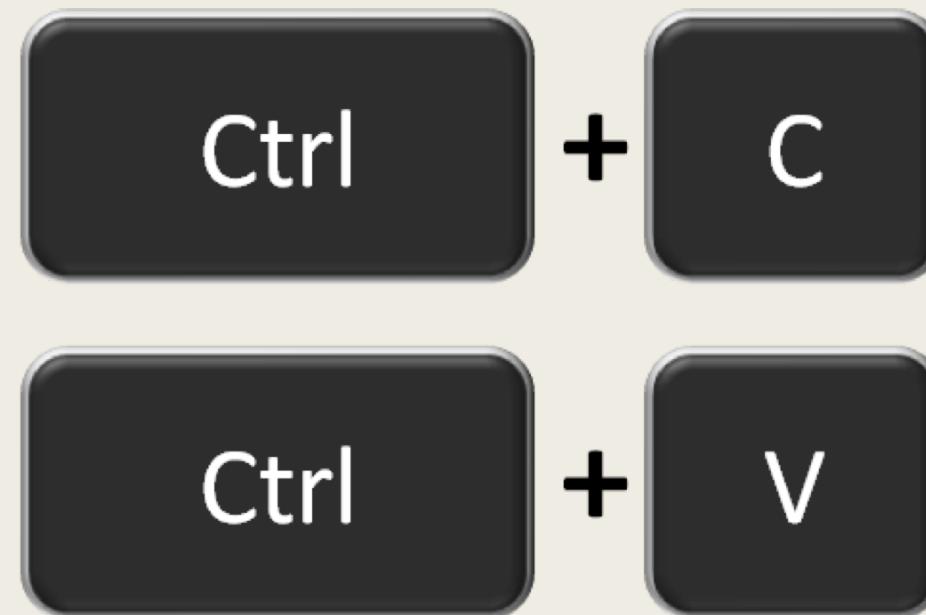


- A **code kata** is an exercise in programming which helps programmers hone their skills through practice and repetition.\*
- Helping others = expanding network



# LinkedIn

- Don't neglect. Any non-tech folks will be looking here for your information (Recruiters, Managers, etc.)
- It doesn't have to be hard, just copy/paste from resume



# Final Thoughts

- Building a profile builds skills and others' awareness of your skills.
- If you could devote time to anything, go for GitHub.
- Don't be scared to contribute to open source. It's not just for "experts".
- Other mediums for promoting your "brand" and profile are good.
  - *Personal Website: Strike balance that works for you*
  - *Twitter: Use it to stay current and promote yourself*
  - *Stack Overflow: Code Kata + Showing off skills to others with questions*
  - *LinkedIn: Staying in front of your less technical audiences*