



Demystifying R: A Guided Tour

David Keyes
R for the Rest of Us



Before We Start

Please take the survey at <https://rfortherestofus.com/aea/>

American Evaluation Association

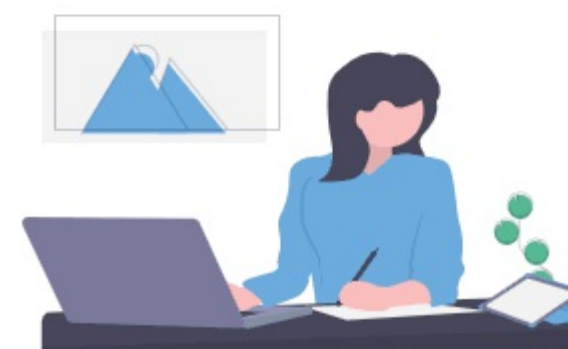
June 11, 2019

Materials for the June 11, 2019 webinar [Demystifying R: A Guided Tour](#) are found on this page. Click any of the items below to see more information.

► R Familiarity Survey

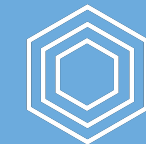
► Follow-Up Q+A

► Materials



Free Course

Getting Started with R is designed to take you through the first, often confusing, steps on your R journey.



What is R?



Download and Install R

The first thing you need to do is download the R software. Go to the [Comprehensive R Archive Network \(aka “CRAN”\) website](#) and download the software for your operating system (Windows, Mac, or Linux).

RStudio



R: Engine



RStudio: Dashboard



Courtesy [Modern Dive](#)



Download and Install RStudio

Download RStudio at the [RStudio website](#). Ignore the various versions listed there. All you need is the latest version of RStudio Desktop.



Packages



Packages

R: A new phone



R Packages: Apps you can download



Courtesy [Modern Dive](#)



Examples of Packages

Tidyverse

[Packages](#)[Articles](#)[Learn](#)[Help](#)[Contribute](#)

R packages for data science

The tidyverse is an opinionated **collection of R packages** designed for data science. All packages share an underlying design philosophy, grammar, and data structures.

Install the complete tidyverse with:

```
install.packages("tidyverse")
```



Examples of Packages

gendercodeR

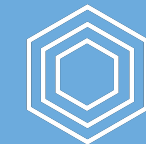
The goal of gendercodeR is to allow simple recoding of freetext gender responses.

Why would we do this?

Researchers who collect self-reported demographic data from respondents occasionally collect gender using a free-text response option. This has the advantage of respecting the gender diversity of respondents without prompting users and potentially including misleading responses. However, this presents a challenge to researchers in that some inconsistencies in typography and spelling create a larger set of responses than would be required to fully capture the demographic characteristics of the sample.

For example, male participants may provide freetext responses as "male", "man", "mail", "mael". Non-binary participants may provide responses as "nonbinary", "enby", "non-binary", "non binary"

This package uses dictionaries of common misspellings to recode these freetext responses into a consistent set of responses.



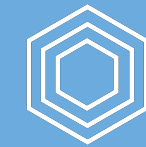
Why Use R?



Data Analysis in a Snap

```
nhanes %>%
  skim(weight)
```

```
## Skim summary statistics
##   n obs: 10000
##   n variables: 22
##
## — Variable type:numeric —————
## variable missing complete      n mean   sd  p0  p25  p50  p75  p100
##   weight      78      9922 10000 70.98 29.13 2.8 56.1 72.7 88.9 230.7
##   hist
##   ┌───┴───┐
```



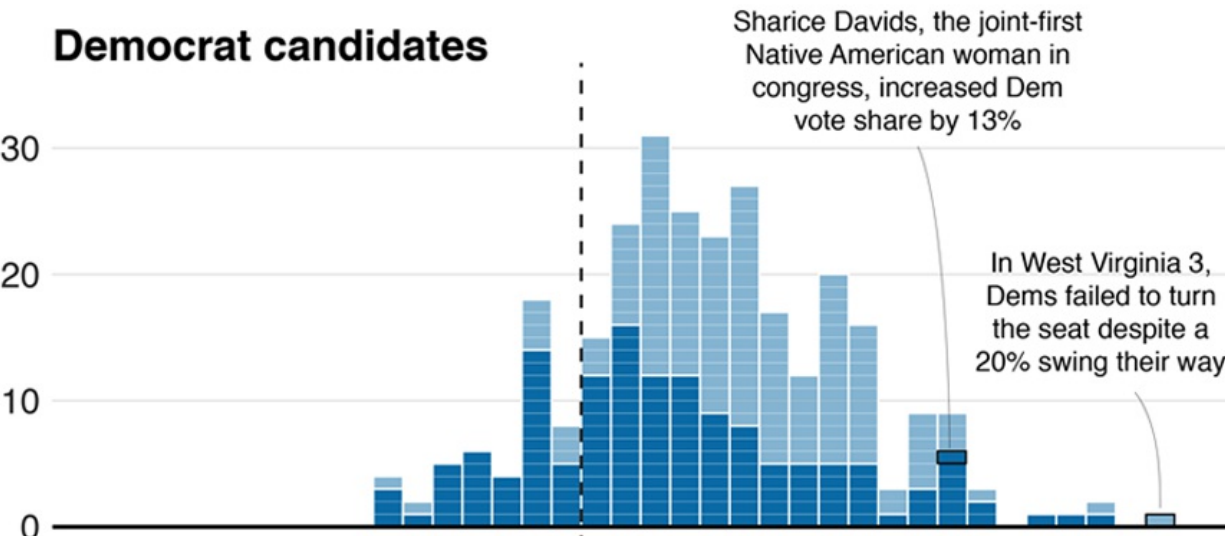
High-Quality Data Visualization



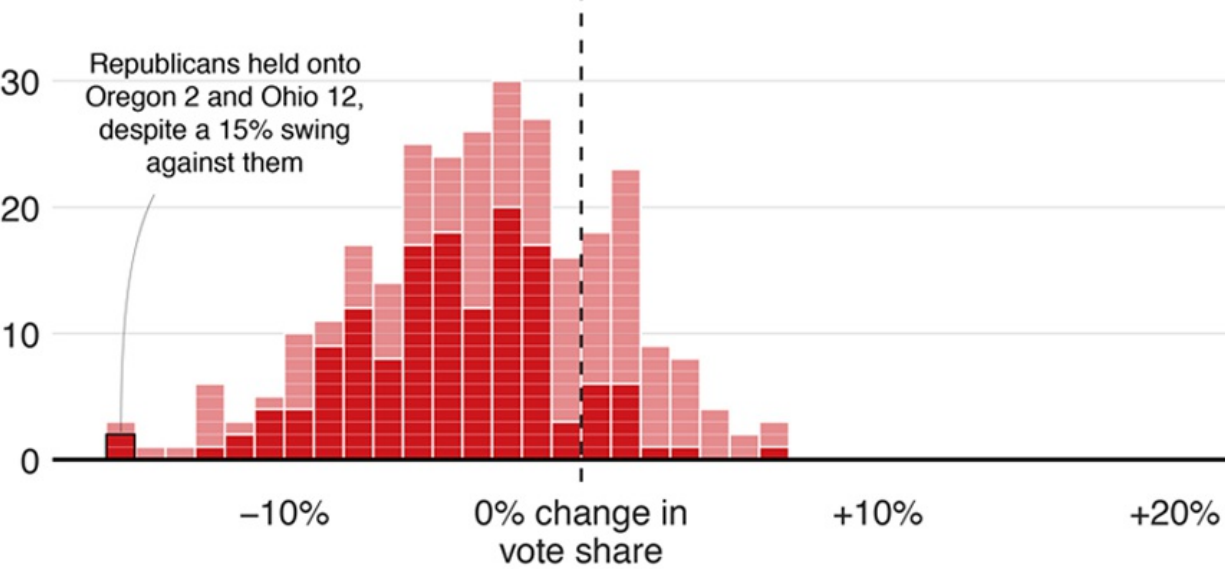
Blue wave

■ Won seat ■ Didn't win

Democrat candidates



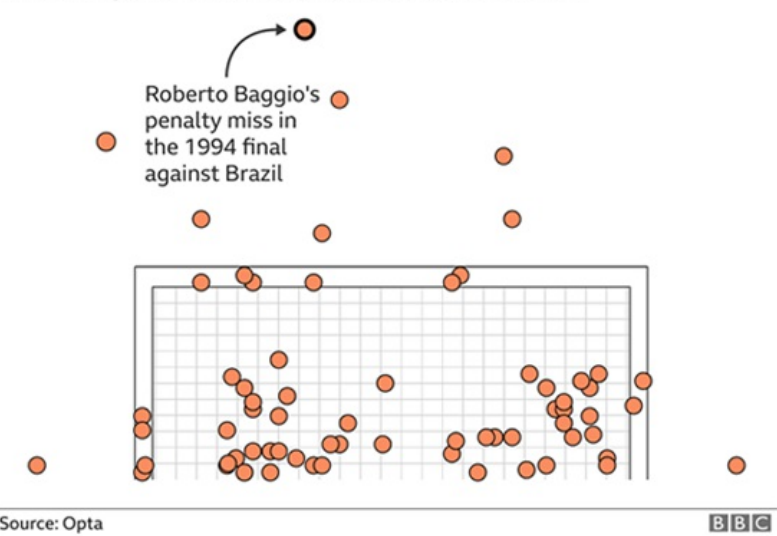
Republican candidates



Source: AP, 19:01 ET

Where penalties are saved

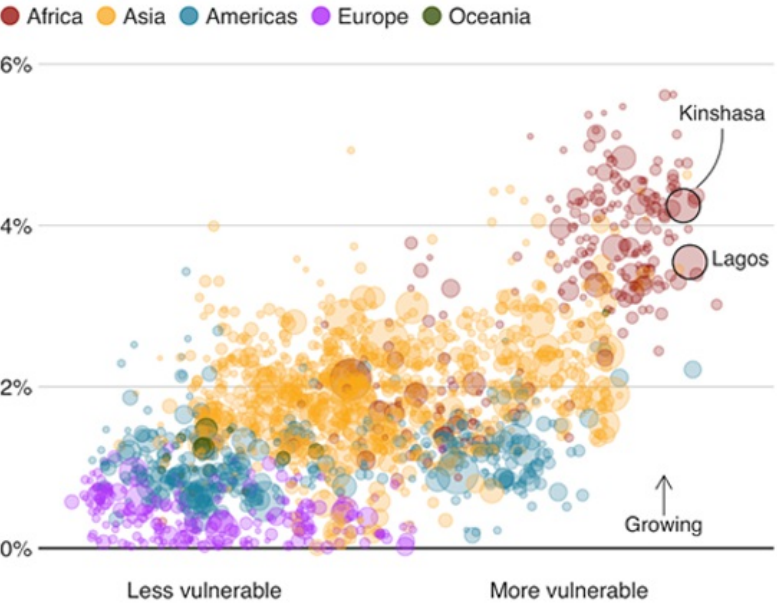
World Cup shootout misses and saves, 1982-2014



Source: Opta

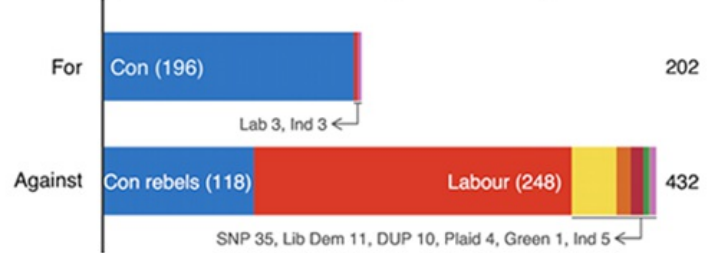
Fast-growing cities face worse climate risks

Population growth 2018-2035 over climate change vulnerability



Source: Verisk Maplecroft. Circle size represents current population.

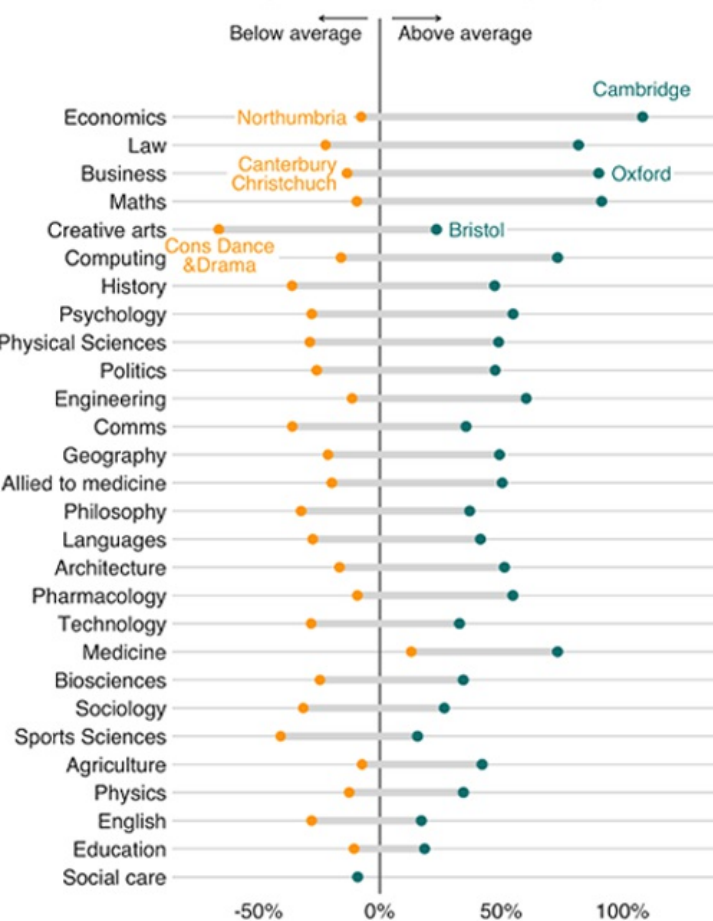
MPs rejected Theresa May's deal by 230 votes



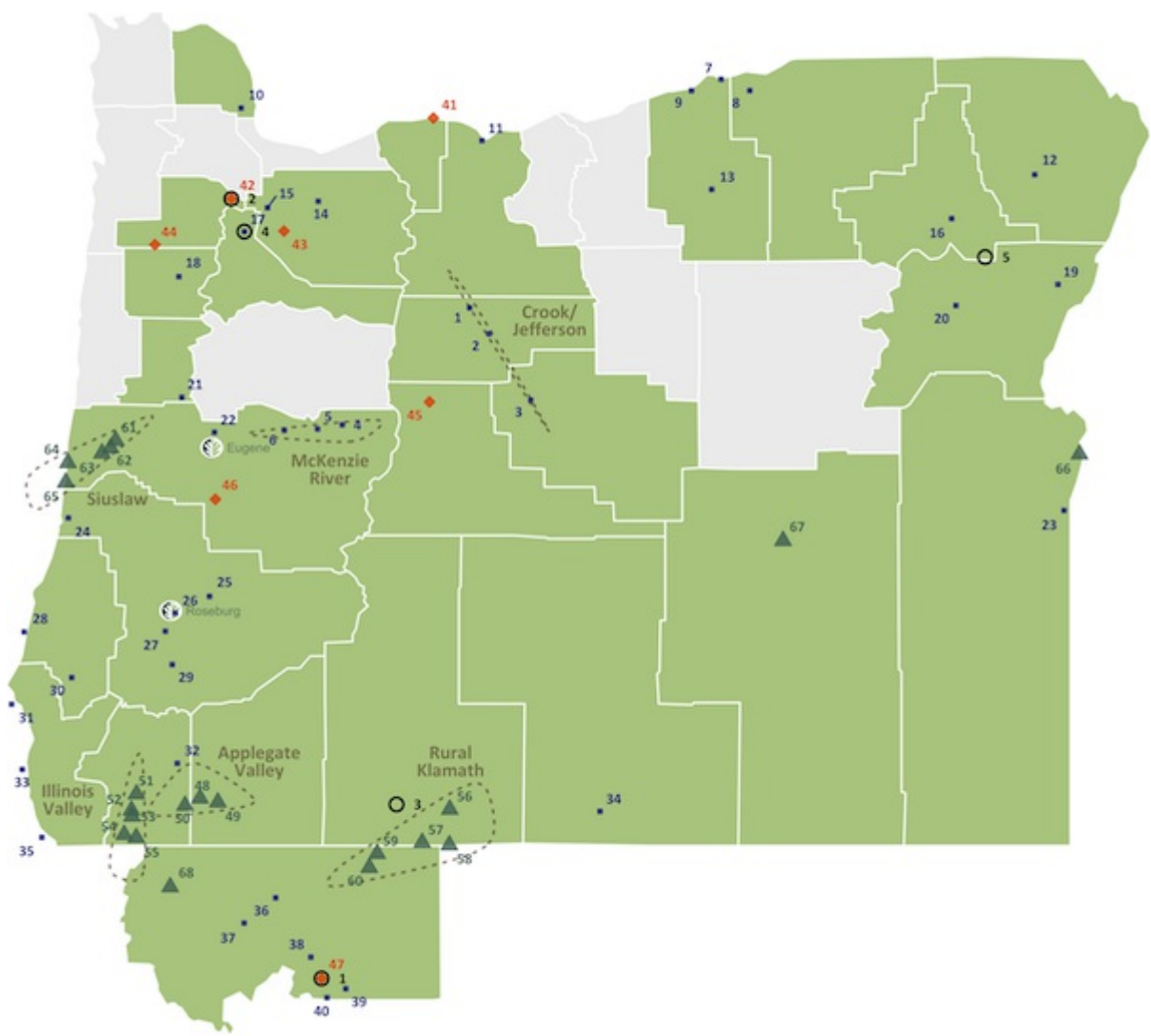
Source: Commons Votes Services. Excludes 'tellers', the Speaker and deputies

Earnings vary across unis even within subjects

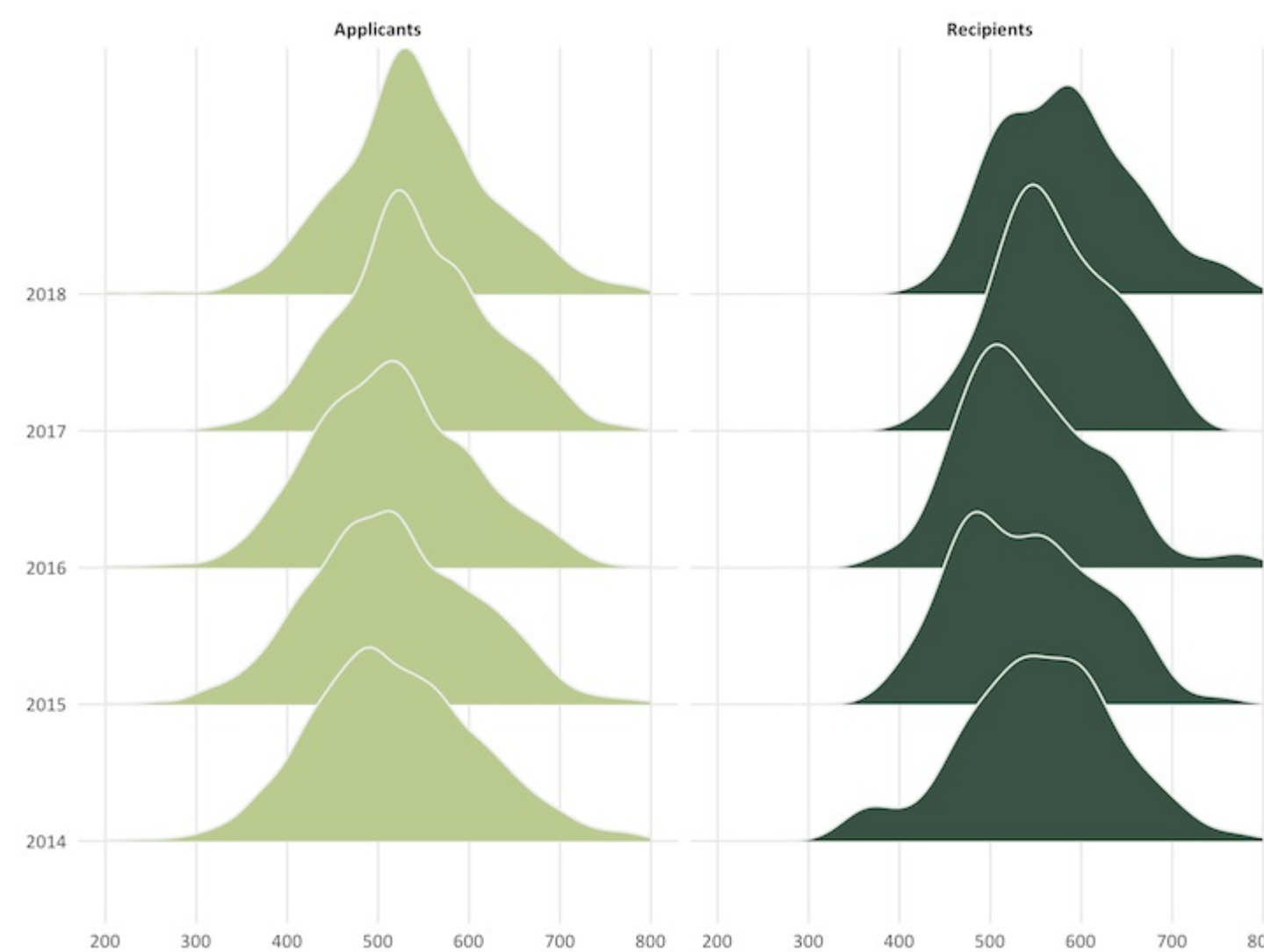
Impact on men's earnings relative to the average degree

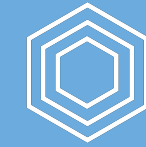


Source: Institute for Fiscal Studies

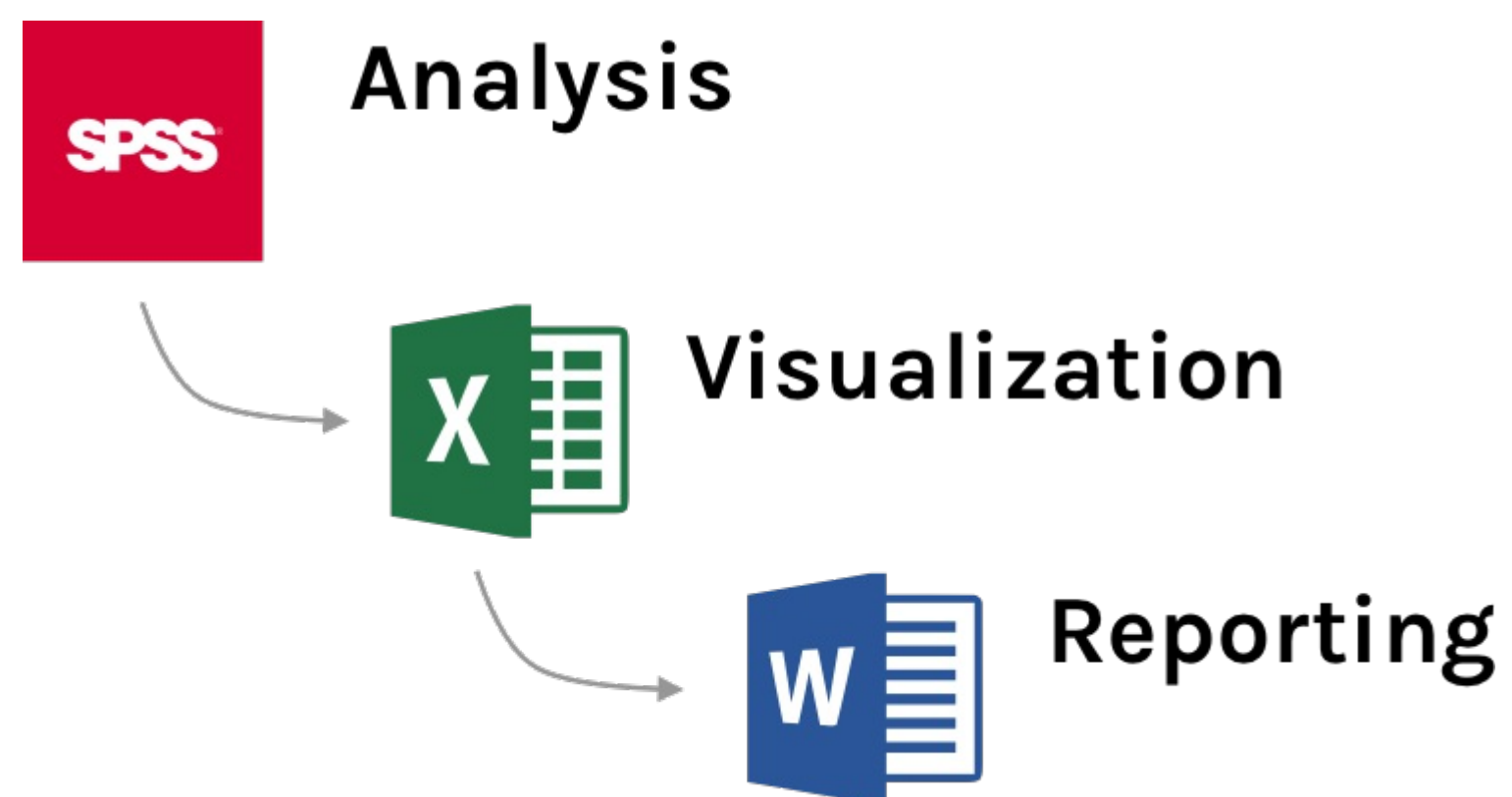


R for the Rest of Us





Use a Single Tool From Data
Import to Final Report





RMarkdown

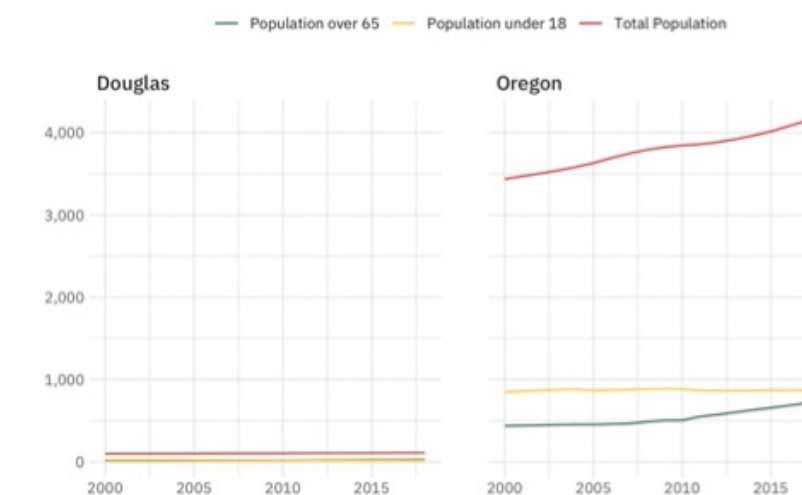
```
77 # Population
78
79 The populations of Douglas County and Oregon are shown below.
80
81 ```{r}
82 population <- read_excel("data/dc-data.xlsx",
83   sheet = "Population") %>%
84   clean_names() %>%
85   gather("geography", "number", -c(indicator, year)) %>%
86   mutate(number = number / 1000) %>%
87   dk_replace_dc() %>%
88   mutate(group = paste(indicator, geography))
89
90 ggplot(population, aes(year, number,
91   group = indicator,
92   color = indicator)) +
93   geom_line() +
94   facet_wrap(~geography) +
95   scale_y_continuous(labels = comma_format()) +
96   dk_remove_color_title +
97   dk_set_colors
98
99 ```
100
```



RMarkdown

Population

The populations of Douglas County and Oregon are shown below.



Word



Provide Immediate Feedback

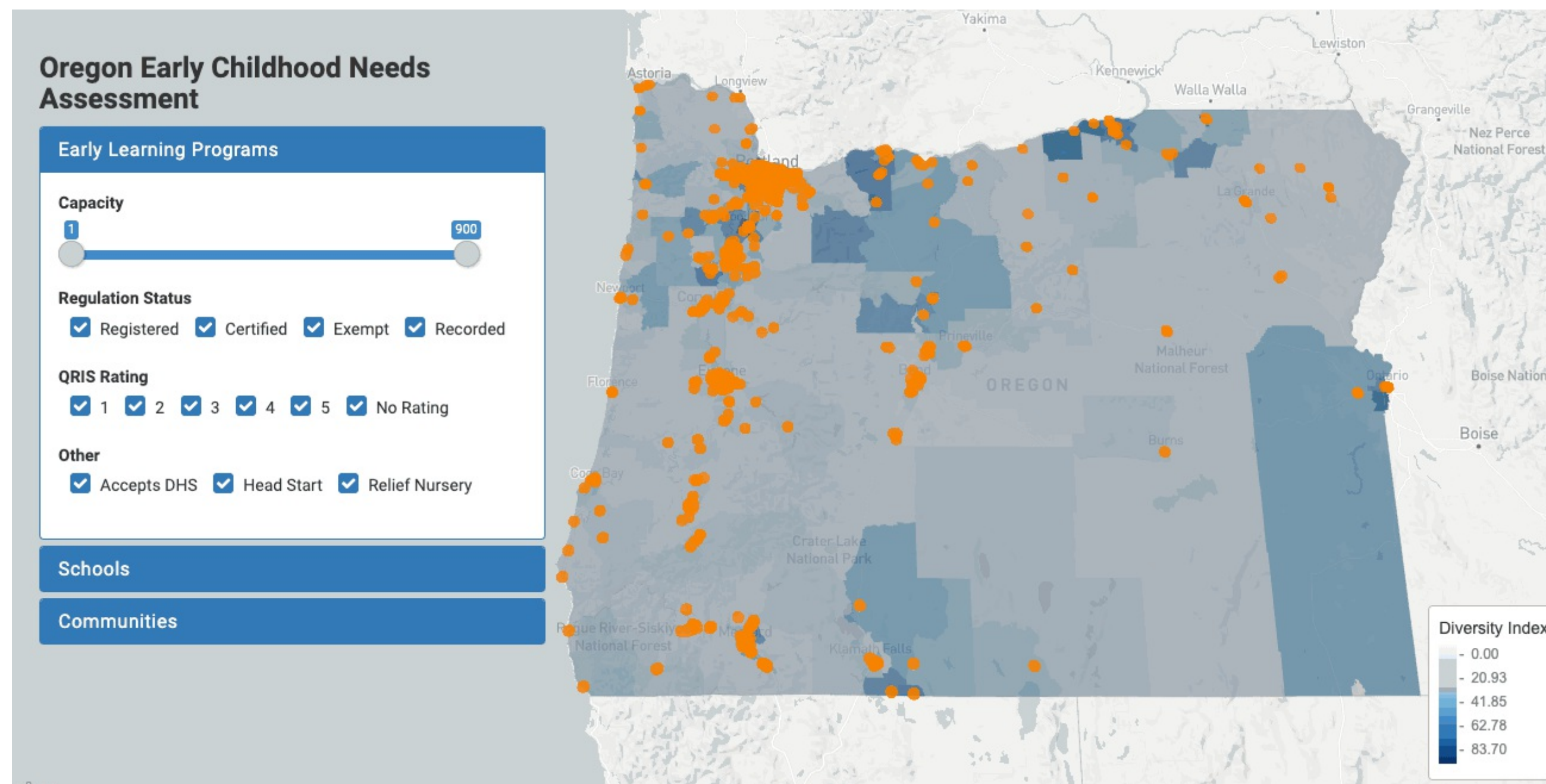
I love this method mainly because all the work is done up front and then for every session the client has I only need to spend 15 minutes generating the report and sending it to them.

They can get their data immediately after data collection!

— [Dana Wanzer](#)



Unique Reporting Possibilities





R Familiarity Survey



The Best Reason to Learn R



Eileen Parkes
@eileen_parkes

Follow

▼

Today I downloaded "R" 🤔
Can a 37 year old mother of two learn to code?
Watch this space...

9:00 AM - 8 May 2019

262 Retweets 5,830 Likes



💬 555

↻ 262

📄

❤️ 5.8K

✉️


🔒



Tweet your reply



Questions?



ONLINE COURSESWORKSHOPSCUSTOM TRAINING1-ON-1 COACHINGABOUTMY ACCOUNTFREE COURSE

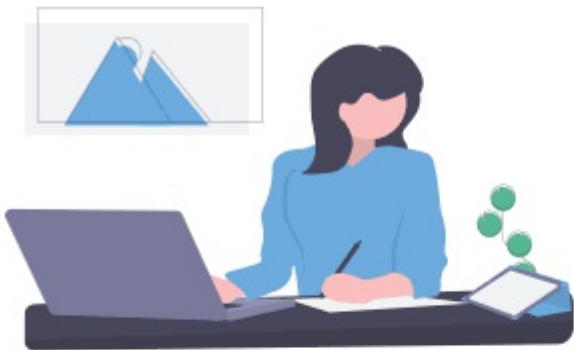
American Evaluation Association

June 11, 2019

Materials for the June 11, 2019 webinar [Demystifying R: A Guided Tour](#) are found on this page. Click any of the items below to see more information.

▶ R Familiarity Survey

▶ Follow-Up Q+A



Free Course

Getting Started with R is designed to take you through the first, often confusing, steps of your R

R for the Rest of Us



Start Your R Journey

To join the Q+A and to get all materials from this webinar, please visit <https://rfortherestofus.com/aea/>

Please email me at any point with questions: david@rfortherestofus.com

I have made a coupon code (AEAJUNE2019) that will give you \$50 off the just released [Fundamentals of R course](#). It expires June 30, 2019.