

Download extended slides:

<http://bit.ly/rro-crug>

Reproducibility with Revolution R Open

David Smith
@revodavid

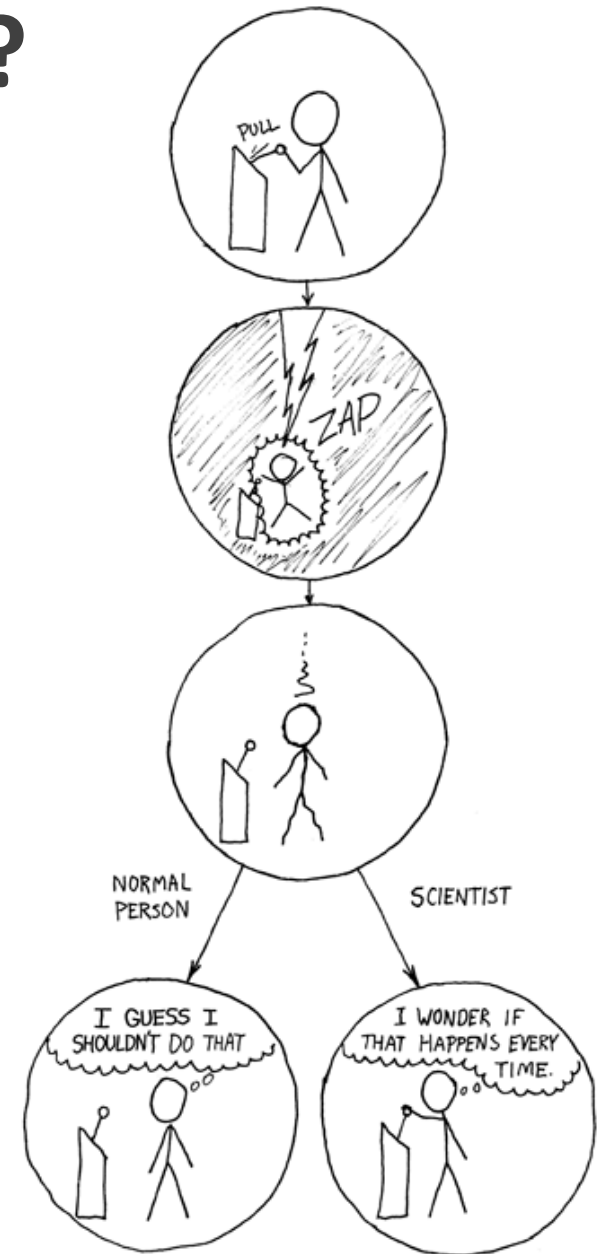
Bay Area R User Group
February 2015

What is Reproducibility?

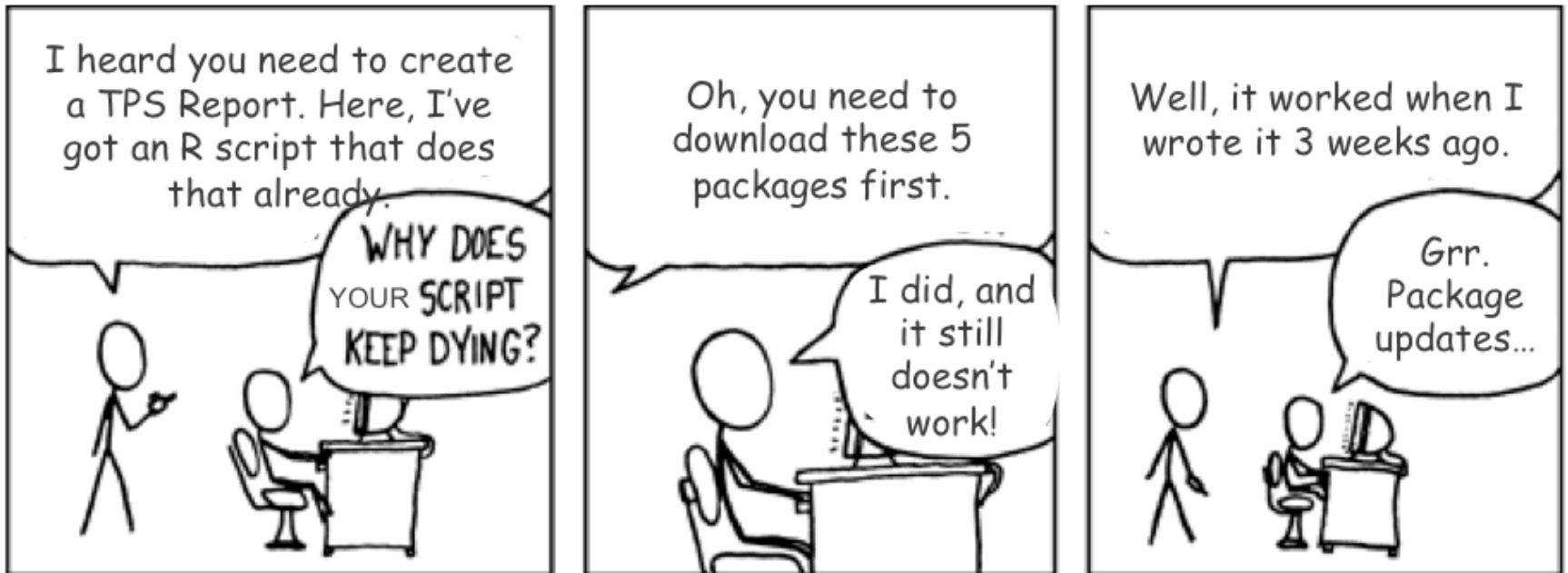
“The goal of reproducible research is to tie specific instructions to data analysis and experimental data so that scholarship can be recreated, better understood and verified.”

CRAN Task View on Reproducible Research (Kuhn)

- Method + Environment
→ Results
- A **process** for:
 - Sharing the method
 - Describing the environment
 - Recreating the results



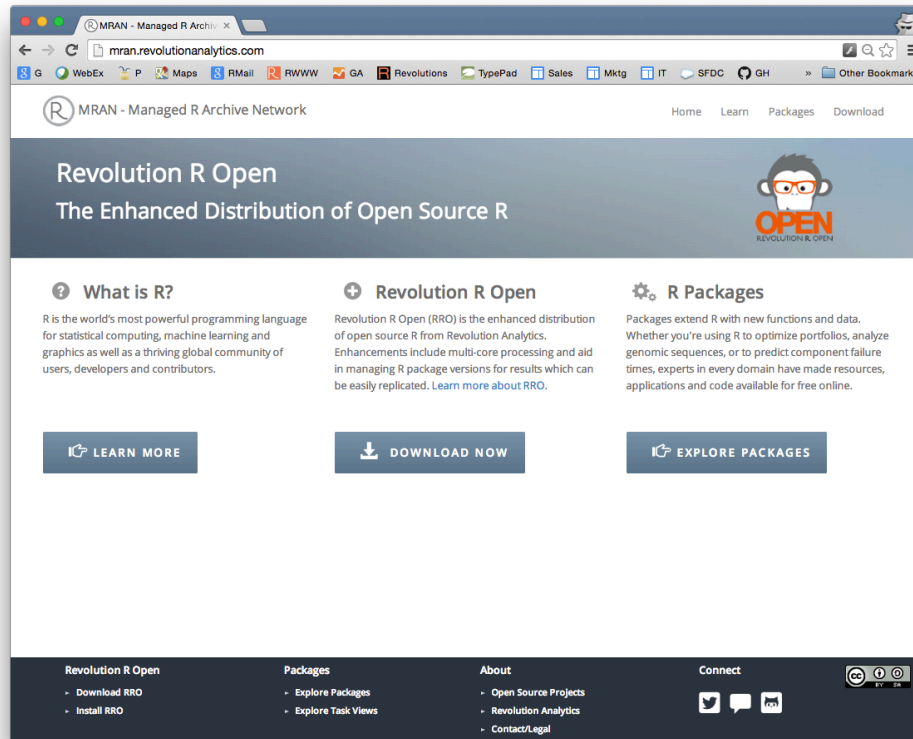
The problem with packages



Reproducibility with Revolution R Open

- Static CRAN mirror in RRO
 - CRAN packages fixed with each Revolution R Open update
- Daily CRAN snapshots
 - Storing every package version since September 2014
 - Hosted at mran.revolutionanalytics.com/snapshot
- `checkpoint` package installed with RRO
 - Write and share scripts synced to a specific snapshot date

MRAN: The Managed R Archive Network



- Download Revolution R Open
- Learn about R and RRO
- Explore Packages
 - and dependencies
- Explore Task Views
- R tips and applications
- Daily CRAN snapshots

`mran.revolutionanalytics.com`

Using checkpoint

- Add 2 lines to the top of your script

```
library(checkpoint)
```

```
checkpoint("2015-01-28")
```

Or, whichever date you want

- Err, that's it.

The checkpoint magic

The `checkpoint()` call does all this:

- Scans project for required packages
- Installs required packages and dependencies
 - Packages installed specific to project
 - Versions specific to checkpoint date
 - Installed in `~/ .checkpoint/DATE`
 - Skips packages if already installed (2nd run through)
- Reconfigures package search path
 - Points only to project-specific library

Sharing projects with checkpoint

- Just share your script or project folder!
- Recipient only needs:
 - compatible R version
 - checkpoint package (installed with RRO)
 - Internet connection to MRAN (at least first time)
- Checkpoint takes care of:
 - (optionally) checking R version
 - Installing CRAN packages
 - Binaries (ease of installation)
 - Correct versions (reproducibility)
 - Dependencies (ease of installation)
 - Eliminating conflicts with other installed packages

New features

- Progress bars
- Checks on R versions

```
checkpoint("2015-02-17",  
           R.version="3.1.2")
```

- Better RStudio experience
- Get latest version from GitHub:

```
library(devtools)  
install_github("RevolutionAnalytics/checkpoint")
```

Download extended slides:

<http://bit.ly/rro-crug>

Thank You!

David Smith

@revodavid

david@revolutionanalytics.com

blog.revolutionanalytics.com