



CWRA ACRH
Canadian
Water
Resources
Association Association
Canadienne
des Ressources
Hydriques



CSHS SCSH
Canadian
Society for
Hydrological
Sciences Société
Canadienne
des Sciences
Hydrologiques
Canadian Water
Resources Association Association Canadienne
des Ressources Hydriques



**ecosystem
recovery** inc.
PROFESSIONAL ENGINEERS

Canadian Water Resources Association

CSHS-hydRology: Intro to R Webinar

May 15, 2019

Presenters:
Robert Chlumsky, MAsc., EIT.

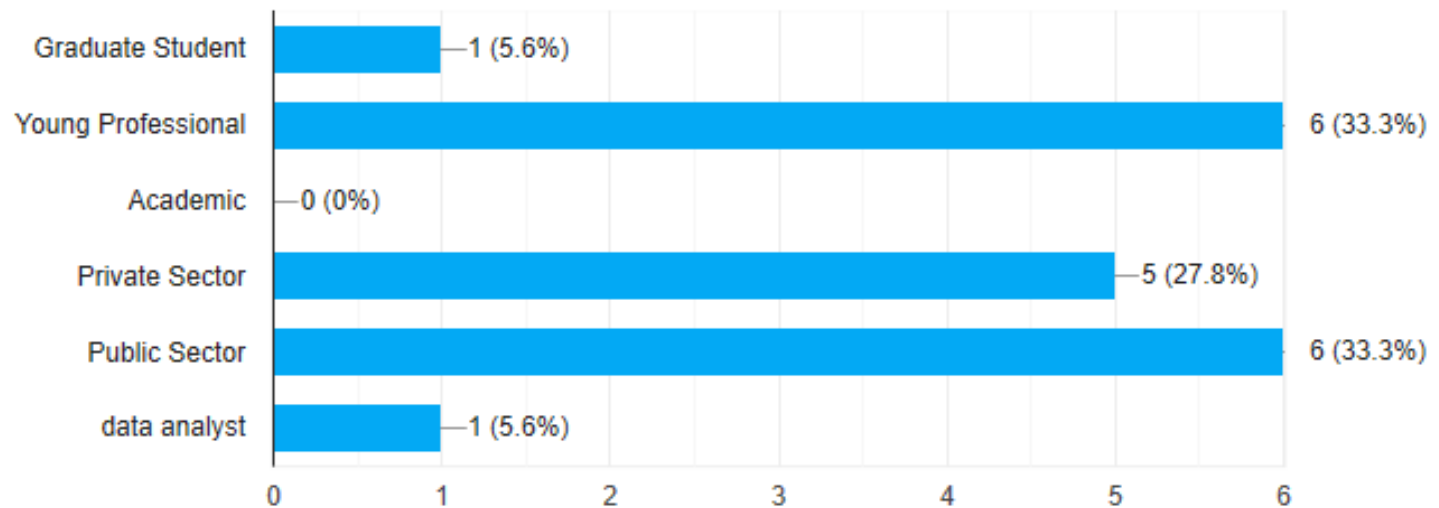
Webinar Outline

- Short presentation on R itself, its advantages, and its applications
- An overview of a popular user interface to R, RStudio
- An interactive introduction to R basics, with opportunities for questions and running R code simultaneously

Survey Responses

What best describes your position?

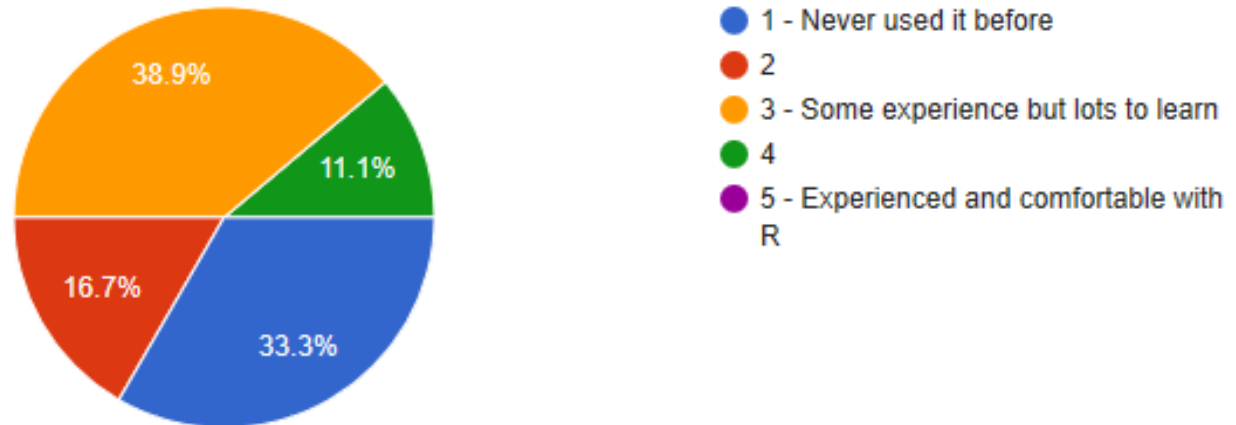
18 responses



Survey Responses

What is your level of experience with using R?

18 responses



What this webinar will do

- Introduce you to what is R, why is R, and who is R
- Get you comfortable writing commands and looking for help in R and RStudio
- Provide you with some resources and links for future learning and lookup
- Give you some useful code snippets and ideas of what R is capable of, particularly with respect to water resources
- Likely provide a few pirate jokes

What this webinar will not do

- Make you an R expert immediately
- Provide you with perfect code for every situation
- Provide a comprehensive list of resources
- Provide a comprehensive overview of every water-related package which exists

Webinar Materials

- Webinar materials will be made available for use following the webinar
- These webinar materials are adapted from Dr. Kevin Shook at the University of Saskatchewan
- Kevin's R training materials can be found on Github as well:
 - https://github.com/CentreForHydrology/Introduction_to_R

Webinar Materials

- Webinar materials for this webinar include:
 - This presentation in PDF format
 - The R package installer file (.R) with commands to install packages used in the webinar
 - R markdown file (.Rmd), which you can use for interactivity with R (used to generate the PDF webinar tutorial file)
 - The PDF of the R markdown file to follow along with

```
CWRA R Webinar May 15.Rmd x
1 |---
2 |title: "CSHS-hydRology: Intro to R webinar"
3 |author: "R. Chlumsky"
4 |date: "May 15, 2019"
5 |output: pdf_document
6 |urlcolor: blue
7 |editor_options:
8 |  chunk_output_type: console
9 |---
10
11 |{r setup, include=FALSE}
12 |knitr::opts_chunk$set(echo = TRUE)
13 |
14 |
15 |# Intro to R Tutorial
16 |
17 |This beginner R tutorial covers the following topics:
18 |
```

CSHS-hydRology: Intro to R Webinar

R. Chlumsky

May 15, 2019

Intro to R Tutorial

This beginner R tutorial covers the following topics:

- Navigating RStudio
- Basic R Syntax
- Data types in R
- Help in R
- Installing and Exploring Packages
- Data checking and sample data sets
- Creating plots
- Basic programming structures
- Building functions

What is R?

- A command-line program
- A GUI-based program
- A programming language
- A general-purpose scientific program
- **NOT JUST STATISTICS**

History of R

- Developed by Ross Ihaka and Robert Gentleman at the University of Auckland, NZ
- Development of R began in 1992, initially released in 1995 and first stable version released in 2000
- R is an implementation of the S programming language (created 1976), which was developed in Bell Labs
 - R bears some resemblance to Unix commands, due in part to this history

Why use R?

- Powerful
- Excellent for:
 - Statistics
 - Data processing
 - Graphing
 - Analysis (GIS, mathematical solvers, etc.)
- Free Open Source Software (FOSS)
 - You can see, test, and trust the code
 - Works with standard file formats – no proprietary or licensing issues
 - Rapid development – huge number of resources at your disposal

Why use R?

- Works well with other programs
 - Interfaces with other languages, such as C and Python
 - Can read/write Excel files directly with packages like `xlsx`
 - Can read in shapefiles, databases, netcdf, many different file formats
- Platform independent
 - Works the same across Windows, Mac OS, and Linux distributions
 - Can develop code in Windows and use on a MAC with minimal issues
 - Can you trust your Excel file once it has been opened on a MAC/Ubuntu computer?

Who uses R?

- “R is a program used by ecologists” ~friend of mine

Certainly popular in ecology, but not unique to it.

R is used by:

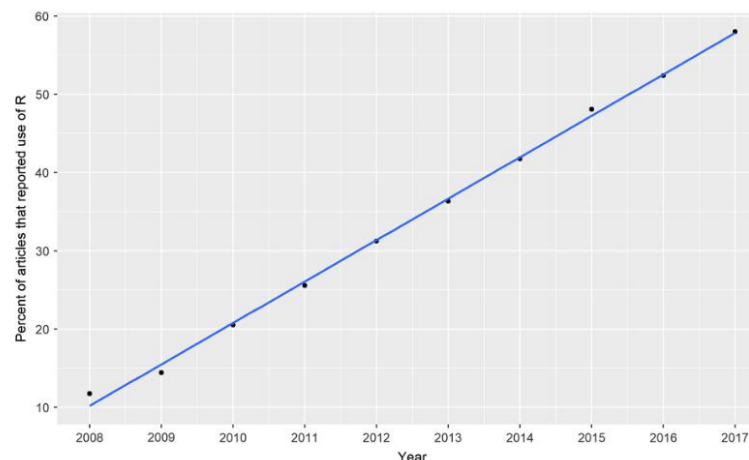
esa

ECOSPHERE

- Statisticians
- Engineers
- Biologists
- Researchers
- Practitioners
- ...lots of people!

Evaluating the popularity of R in ecology

JIANGSHAN LAI^{1,2,†} CHRISTOPHER J. LORTIE^{1,3,4} ROBERT A. MUENCHEN^{1,5} JIAN YANG⁶ AND KEPING MA¹



“Spreadsheets must die!”

~ Dr. Kevin Shook

- Spreadsheets are:
 - Slow
 - Dangerous
 - Unreliable – can’t be trusted to have not been changed once opened
 - No documentation of changes to formulas, organization, data sources, pasting values, etc.
- Proprietary programs and formats (such as Excel) are not trustworthy
 - You can’t know what is happening under the hood
 - You cannot rely on your data being accessible in the future
 - Projects can be lost when stored in proprietary formats
- Caveat – spreadsheets are hard to avoid in many professional settings, but it is important to understand their limitations and their ability to easily hide errors!

What is R used for?

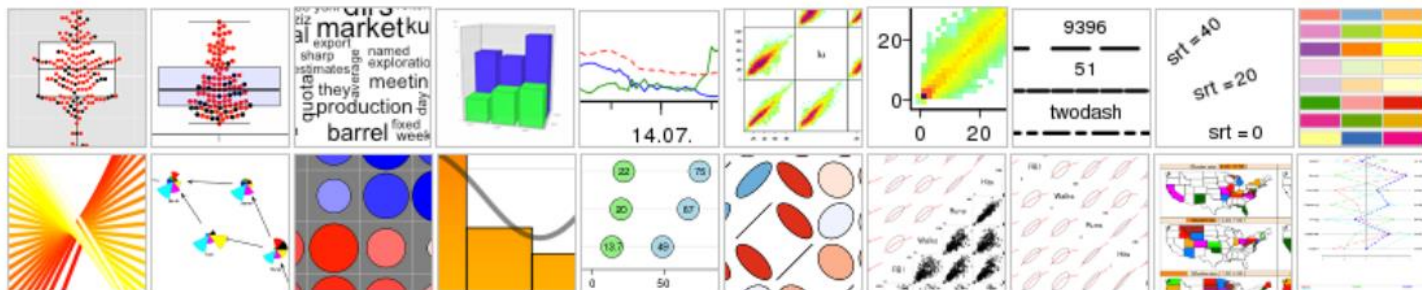
- Statistical analysis
- Data processing
- Publication quality graphing
- GIS analysis
- Data visualization
- “Big Data” processing
- Machine learning models
- ...

Statistical Analysis

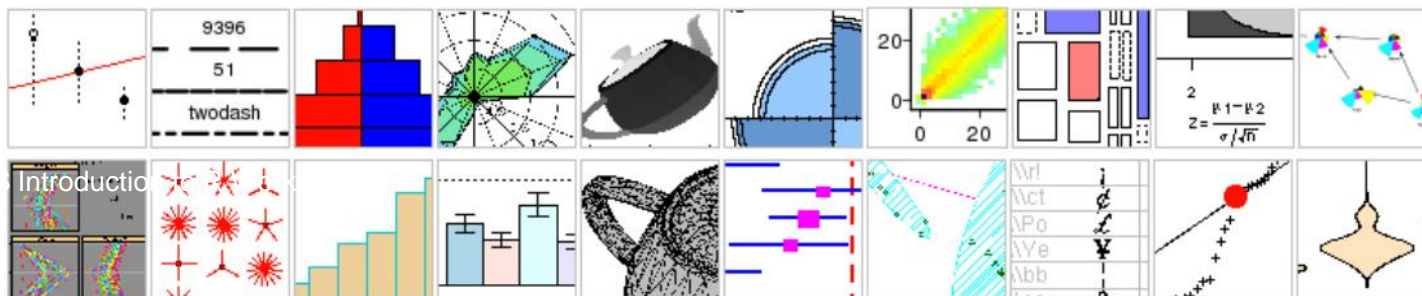
- R supports a vast number of statistical analyses:
 - T-tests
 - Random sampling (rnorm, runif, etc.)
 - Tests for normality, qqplot
 - Mann-Kendall non-parametric test
 - ANOVA analysis
 - ... many more

Graphing

- R is one of the best programs for scientific/technical graphing



» Random entries



GIS

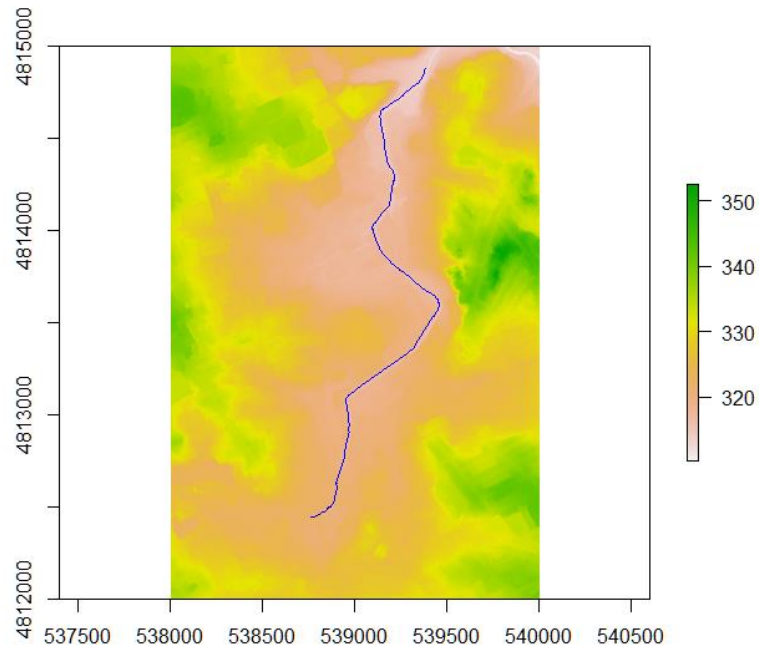
R is capable of performing GIS analysis directly



GIS

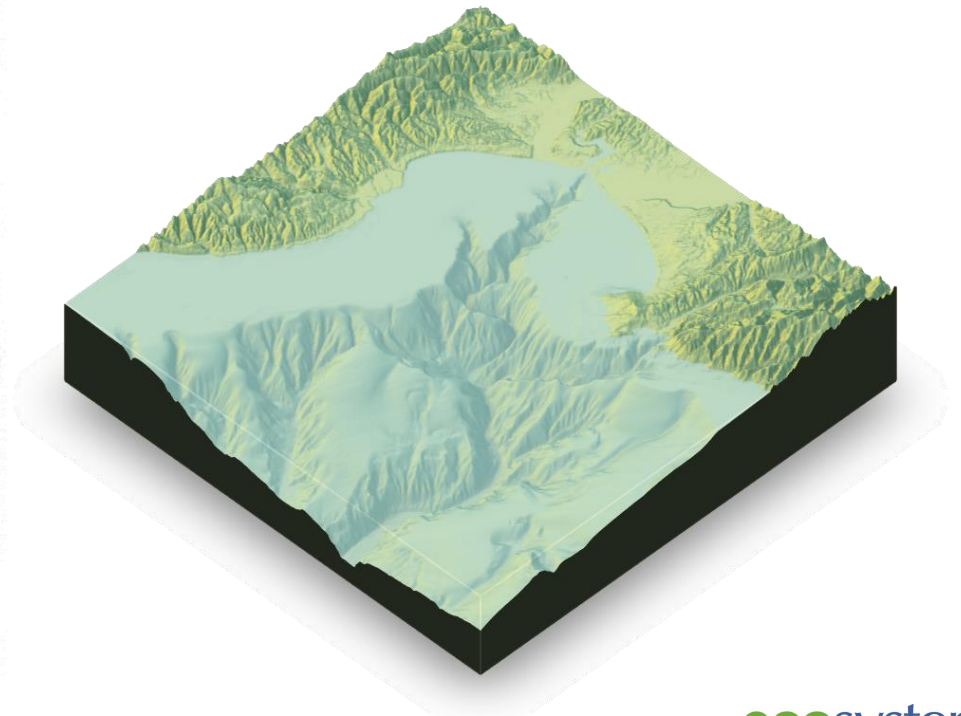
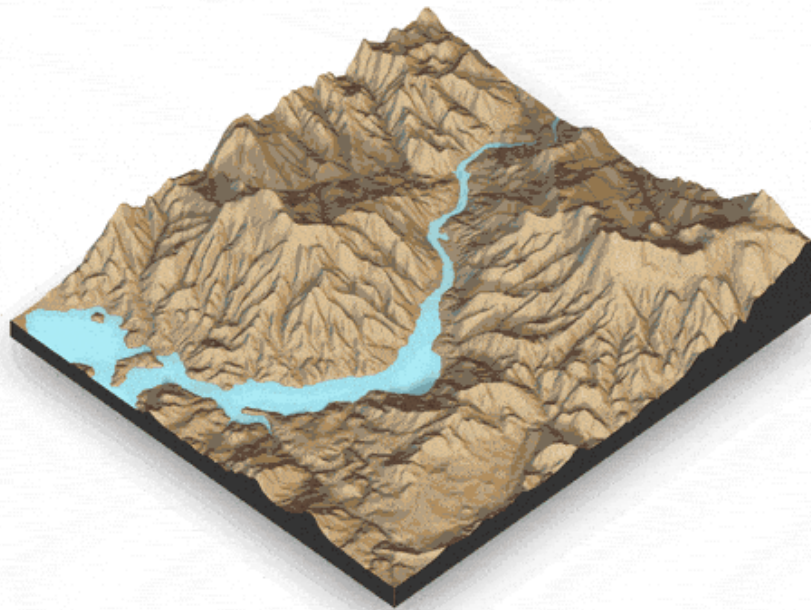
Not as difficult to plot GIS files as one might think...

```
plot(raster("LaurelCreekMerged.tif"))  
plot(read_sf("LaurelCreek_watercourse.shp")$geometry,col='blue',add=T)
```



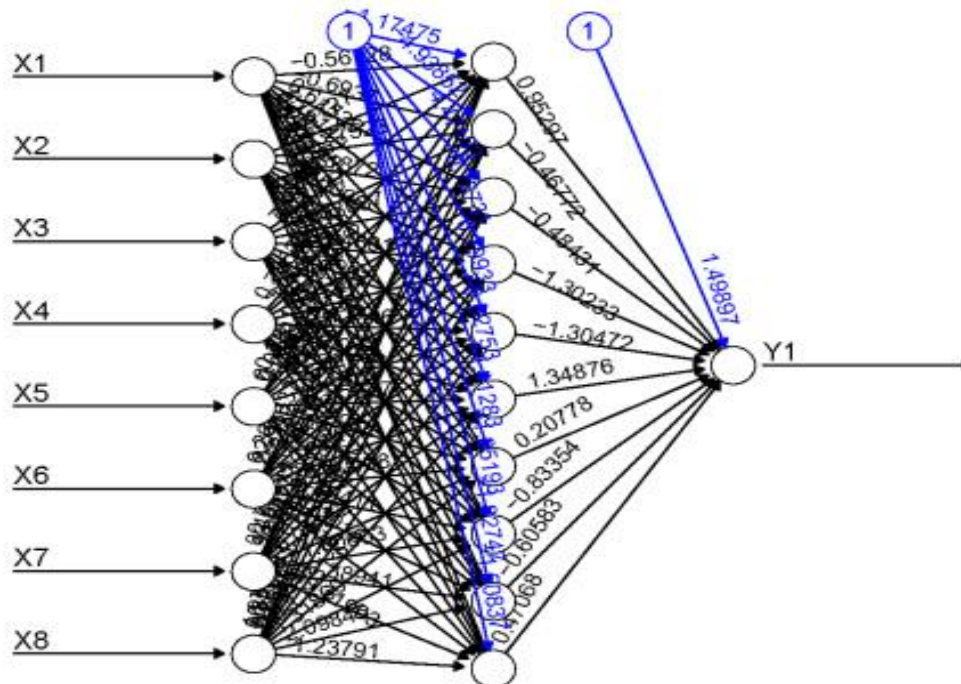
Visualizations

Incredible visualizations with packages such as rayshader



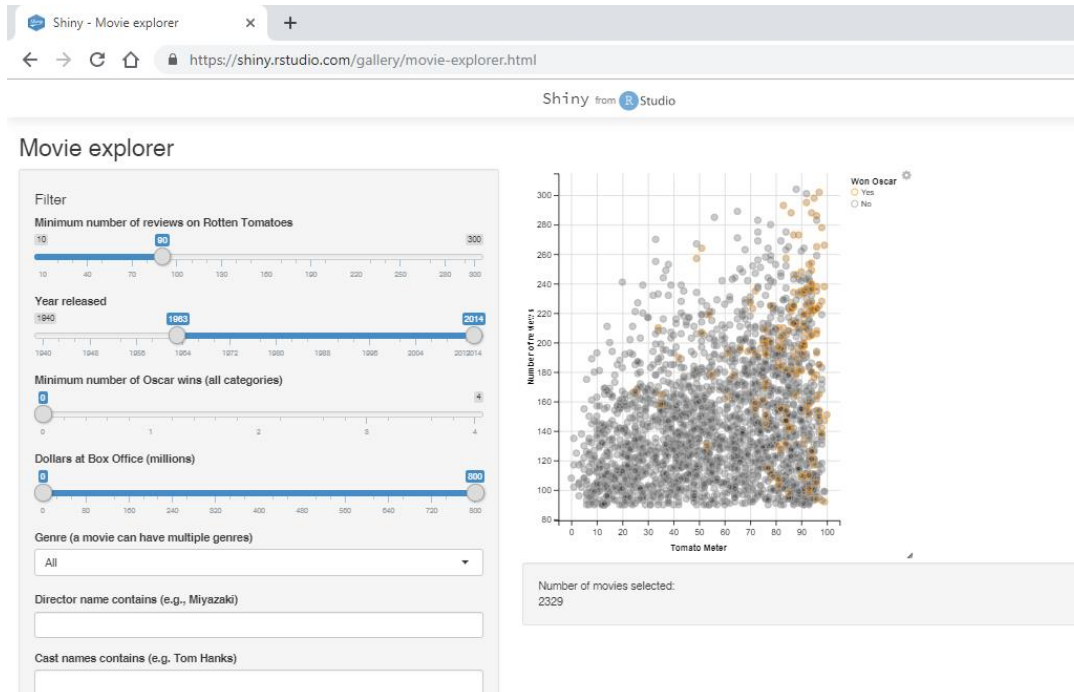
Machine Learning

- Many packages exist to support machine learning and neural networks, hot topic in literature and practice these days



Web Applications

- Interactive web applications can be developed directly in R with Shiny
 - <https://shiny.rstudio.com/gallery/movie-explorer.html>



R Typical Workflow

- Read in data (textfiles, database, xlsx, csv, shapefile, etc.)
- Massage and quality control data
- Explore and visualize data
- Perform calculations
- Save results
- Export data to other programs
- Create high quality graphics for report/presentations/papers

R Workflow Advantages

- All tasks can be manual (one-off script) or automated
- Logging of calculations/results
- All data/ results stored in standard, non-proprietary files
- Data/ results easily exported
- No data silos, format issues
 - R scripts stored as text files, can be read with Notepad if needed

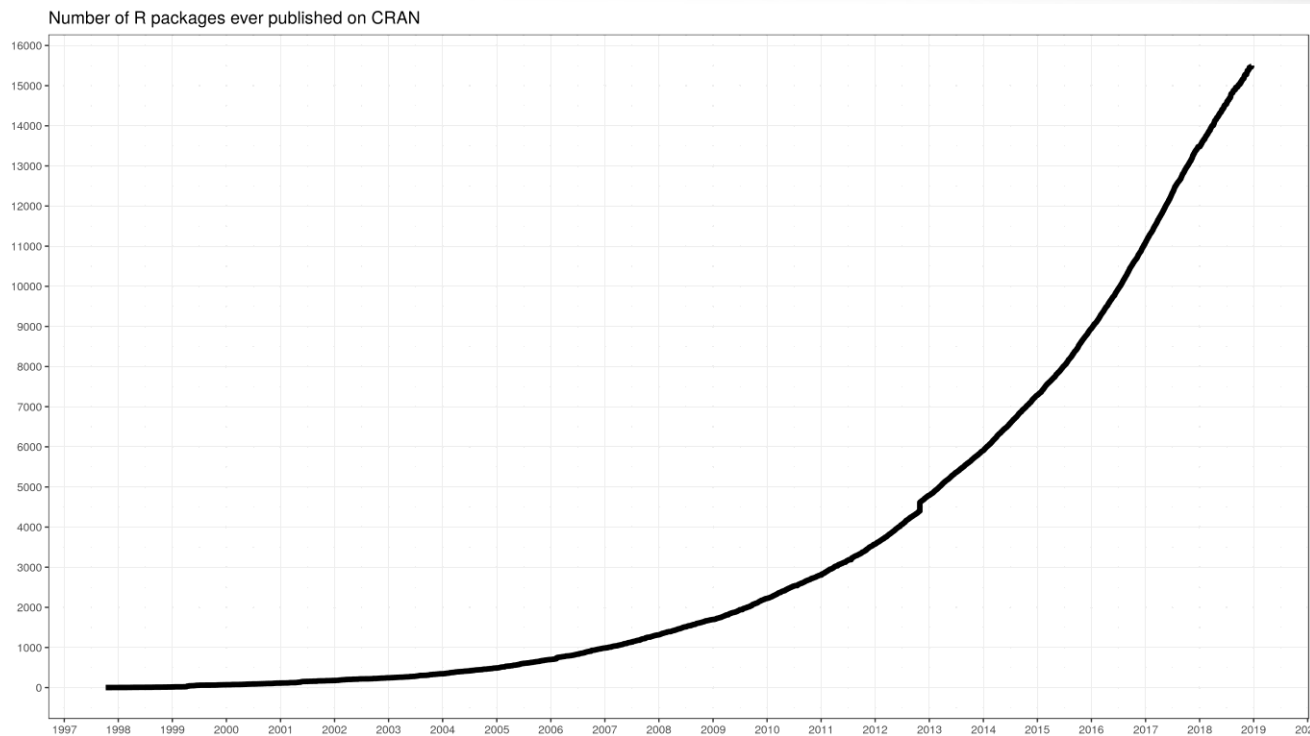
Challenges with R

- Steep learning curve
 - Different from other programs, even other programming languages
 - Need to learn commands
 - Different way of thinking from seeing spreadsheets
- However,
 - Lots of help available
 - Huge number of users
 - GUIs make it easier to manage
 - Reference cards and cheat sheets make it easy to learn commands
 - <https://cran.r-project.org/doc/contrib/Short-refcard.pdf>

Packages

- R had thousands of built-in functions, readily available with the base R install
- Thousands more functions available in downloadable packages
- Each package contains:
 - Functions
 - Sample data
 - Documentation
 - Help files and examples
- Downloading directly through R, no struggling with dependencies (cough, Python)
- Number of R packages growing exponentially

Packages



<https://gist.github.com/daroczgi/3cf06d6db4be2bbe3368>

R and hydRology

- Large number of research groups and institutes using R in the field of water resources and hydrology
- Many packages exist specifically for water resources
 - CSHS-hydRology (package for Canadian hydrologists)
 - EcohydRology (ecohydrology functions, e.g. baseflow partitioning)
 - HydroTSM (hydrologic time series management)
 - hydroGOF (goodness of fit statistics)
 - Tidyhydat (scraping Water Survey Canada station data)
- + Many (many) more not listed

Webinar Agenda

- Introduction to R Studio
- Basic R syntax
- Getting help in R
- Exploring and installing packages
- Reading in data
- Creating plots
- If statements, functions, loops

Time to Switch Gears to Actual R Content...



Switch to Rstudio...

Hopefully your computers already have:

- **Base R**, installed from <https://www.r-project.org/>
- **RStudio**, R GUI, installed from <https://www.rstudio.com/>
- Run the installer file to install packages needed for the webinar
- Downloaded the R Markdown PDF Webinar document

RStudio...

Webinar Wrap-up

Other things to touch on as we wrap up:

- Working with other R packages
- Building packages
- Working on repositories
- hydRology project
- CWRA National Conference Sessions

CRAN Packages

- Super easy to install directly from R, as you will have seen
- Help information is required on each function
- The submission process for CRAN packages is rigorous, resulting in high quality packages with lots of resources

Other R Packages

- Many more packages exist outside of CRAN, which people simply publish themselves
- Many of these exist on Github, which has excellent integration with R and RStudio
- R has a function to install packages directly from GitHub with ease, e.g.
 - `devtools::install_github("rchlumsk/RavenR")`

Building Your Own R Packages

- RStudio makes building packages straightforward
- Packages exist to help make this process easier (packages to build packages??)
- Not as hard as it sounds!
- Packages are helpful for:
 - Forcing you to document your own work
 - Creating online repositories of your work to back up projects
 - Sharing your work easily with other colleagues and R users
- **Come by the post-conference CSHS-hydRology session for more information on how to build packages**

Help Resources

- Plenty of resources for water resources practitioners and hydrologist users of R
 - Kevin Shook (USask) [R Training](#) resources on Github
 - Dan Moore (UBC) [R routines](#) for hydrologists
 - ROpenSci discussion on hydrology [R packages](#)
 - [CRAN Task View](#) on hydrological data and modelling packages
 - Many useful blog posts and other webpages in results for [‘hydrology+R’](#)

Other Resources

- Google your R problem – there is likely a solution already discovered!
- Stack Overflow is a great resource
- The Rproject.org has a lot of helpful resources and guides
- Free online R courses
- Talk to friends and colleagues for help!

CSHS-hydRology Project

- Collaboration by Canadian hydrologists to build an R package which addresses issues in Canadian climates
- Community built, tested, and coded
- Goal is to submit the package to CRAN
- No R experience necessary to join!
- Check out [commentary](#) in CWRA journal, “R-functions for Canadian hydrologists: a Canada wide collaboration”

CWRA Conference 2019

CSHS-hydRology Sessions

- Wednesday, May 29th are multiple R-related sessions at the upcoming National Conference
 - Two part workshop to discuss the CSHS-hydRology package, meet in person, learn how to use Github to work on the package repository, and learn how to use RStudio to build packages in R
 - Afternoon session `Exploring R in Water Science` with short presentations on specific R functions being developed

CWRA Pre- and Post-Conference Workshops

- Spaces still available for some the pre-conference workshops
 - FLOW FORECASTING USING WATFLOOD(R) AND GREENKENUE(TM)
 - UNDERSTANDING INDIGENOUS CULTURE
 - ENHANCING ENGAGEMENT IN CANADIAN WATER MANAGEMENT
 - INDIGENOUS APPROACHES & ENGAGEMENT IN WATERSHED PLANNING
 - NASH FLOW REGATTA

Thanks!

Follow up contact with any questions/ comments/
requests:

Robert Chlumsky

robert.chlumsky@ecosystemrecovery.ca

rchlumsk@uwaterloo.ca

Thank you.