R packages

Paul E. Johnson[1] [2]

[1] Department of Political Science

[2] Center for Research Methods and Data Analysis, University of Kansas

2018

KU

Outline

KU

Outline

An Engine with a lot of Packages

Ross Ihaka and Robert Gentleman created the original R (R Core Team, 2017) program in the mid 1990s

- a computational engine that could tolerate the addition of features in the form of "packages"
  - New Zealand junk car story
- User commands followed the style of the S language (Bell Labs) but internal logic different
- R's core functionality, the part that users *think* is R itself, is drawn from the packages "graphics" and "stats".

KU

Outline

Start R, what packages are loaded?

Run "`sessionInfo()`".

```
sessionInfo()
```

```
R version 3.4.4 (2018-03-15)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 18.04 LTS

Matrix products: default
BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.7.1
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.7.1

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
     LC_TIME=en_US.UTF-8
 [4] LC_COLLATE=en_US.UTF-8     LC_MONETARY=en_US.UTF-8
     LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8       LC_NAME=C                  LC_ADDRESS=C
[10] LC_TELEPHONE=C             LC_MEASUREMENT=en_US.UTF-8
    LC_IDENTIFICATION=C

attached base packages:
[1] stats     graphics  grDevices utils     datasets  base
```

Start R, what packages are loaded? ...

```
loaded via a namespace (and not attached):
[1] compiler_3.4.4 tools_3.4.4
```

On a new install of R in MS Windows, I have . . .

```
> library()
Packages in library 'c:/Program Files/R/library':

base                            The R Base Package
boot                            Bootstrap Functions (Originally by Angelo Canty for S)
class                           Functions for Classification
cluster                         "Finding Groups in Data": Cluster Analysis Extended Rousseeuw et al.
codetools                       Code Analysis Tools for R
compiler                        The R Compiler Package
datasets                        The R Datasets Package
foreign                         Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, Weka, dBase,
                                ...
graphics                        The R Graphics Package
grDevices                       The R Graphics Devices and Support for Colours and Fonts
grid                            The Grid Graphics Package
KernSmooth                      Functions for Kernel Smoothing Supporting Wand & Jones (1995)
lattice                         Trellis Graphics for R
MASS                            Support Functions and Datasets for Venables and Ripley's MASS
Matrix                          Sparse and Dense Matrix Classes and Methods
methods                         Formal Methods and Classes
mgcv                            Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation
nlme                            Linear and Nonlinear Mixed Effects Models
nnet                            Feed-Forward Neural Networks and Multinomial Log-Linear Models
parallel                        Support for Parallel computation in R
rpart                           Recursive Partitioning and Regression Trees
spatial                         Functions for Kriging and Point Pattern Analysis
splines                         Regression Spline Functions and Classes
stats                           The R Stats Package
stats4                          Statistical Functions using S4 Classes
survival                        Survival Analysis
tcltk                           Tcl/Tk Interface
tools                           Tools for Package Development
utils                           The R Utils Package
>
```

Listing combines the packages *within* R's distribution, plus the *recommended* packages that were provided with R

Lets check my Laptop

On 2017-05-13:

```
library()
```

```
Packages in library '/home/pauljohn/R/x86_64-pc-linux-gnu-library/3.3':

abind                   Combine Multidimensional Arrays
acepack                 ACE and AVAS for Selecting Multiple Regression
                        Transformations
AER                     Applied Econometrics with R
afex                    Analysis of Factorial Experiments
akima                   Interpolation of Irregularly and Regularly
                        Spaced Data
alr4                    Data to accompany Applied Linear Regression 4rd
                        edition
Amelia                  A Program for Missing Data
animation               A Gallery of Animations in Statistics and
                        Utilities to Create Animations
apsrtable               apsrtable model-output formatter for social
                        science
arm                     Data Analysis Using Regression and
                        Multilevel/Hierarchical Models
assertthat              Easy pre and post assertions.
backports               Reimplementations of Functions Introduced Since
                        R-3.0.0
base64                  Base64 Encoder and Decoder
base64enc               Tools for base64 encoding
BB                      Solving and Optimizing Large-Scale Nonlinear
                        Systems
bdsmatrix               Routines for Block Diagonal Symmetric matrices
BH                      Boost C++ Header Files
```

Lets check my Laptop ...

```
biclust                 BiCluster Algorithms
BiocGenerics            S4 generic functions for Bioconductor
bit                     A class for vectors of 1-bit booleans
bit64                   A S3 Class for Vectors of 64bit Integers
bitops                  Bitwise Operations
BMS                     Bayesian Model Averaging Library
bookdown                Authoring Books and Technical Documents with R
                        Markdown
boot                    Bootstrap Functions (Originally by Angelo Canty
                        for S)
brew                    Templating Framework for Report Generation
broom                   Convert Statistical Analysis Objects into Tidy
                        Data Frames
car                     Companion to Applied Regression
caTools                 Tools: moving window statistics, GIF, Base64,
                        ROC AUC, etc.
checkmate               Fast and Versatile Argument Checks
chron                   Chronological Objects which can Handle Dates
                        and Times
cluster                 "Finding Groups in Data": Cluster Analysis
                        Extended Rousseeuw et al.
coda                    Output Analysis and Diagnostics for MCMC
coin                    Conditional Inference Procedures in a
                        Permutation Test Framework
colorspace              Color Space Manipulation
commonmark              High Performance CommonMark and Github Markdown
                        Rendering in R
compare                 Comparing Objects for Differences
coxme                   Mixed Effects Cox Models
crayon                  Colored Terminal Output
crmda                   The KU CRMDA Private Package
cubature                Adaptive Multivariate Integration over
                        Hypercubes
curl                    A Modern and Flexible Web Client for R
```

Lets check my Laptop ...

```
data.table                  Extension of 'data.frame'
DBI                         R Database Interface
ddst                        Data Driven Smooth Tests
deldir                      Delaunay Triangulation and Dirichlet (Voronoi)
                            Tessellation
DEoptimR                    Differential Evolution Optimization in Pure R
desc                        Manipulate DESCRIPTION Files
dfoptim                     Derivative-Free Optimization
dichromat                   Color Schemes for Dichromats
digest                      Create Compact Hash Digests of R Objects
diptest                     Hartigan's Dip Test Statistic for Unimodality -
                            Corrected
doParallel                  Foreach Parallel Adaptor for the 'parallel'
                            Package
dplyr                       A Grammar of Data Manipulation
effects                     Effect Displays for Linear, Generalized Linear,
                            and Other Models
eigenmodel                  Semiparametric factor and regression models for
                            symmetric relational data
emplik                      Empirical Likelihood Ratio for
                            Censored/Truncated Data
ergm                        Fit, Simulate and Diagnose Exponential-Family
                            Models for Networks
ergm.count                  Fit, Simulate and Diagnose Exponential-Family
                            Models for Networks with Count Edges
estimability                Tools for Assessing Estimability of Linear
                            Predictions
evaluate                    Parsing and Evaluation Tools that Provide More
                            Details than the Default
evd                         Functions for Extreme Value Distributions
faraway                     Functions and Datasets for Books by Julian
                            Faraway
flexclust                   Flexible Cluster Algorithms
flexmix                     Flexible Mixture Modeling
```

Lets check my Laptop ...

```
foreach      Provides Foreach Looping Construct for R
formatR      Format R Code Automatically
Formula      Extended Model Formulas
fpc          Flexible Procedures for Clustering
gamm4        Generalized Additive Mixed Models using 'mgcv'
             and 'lme4'
gdata        Various R Programming Tools for Data
             Manipulation
geepack      Generalized Estimating Equation Package
ggnetwork    Geometries to Plot Networks with 'ggplot2'
ggplot2      Create Elegant Data Visualisations Using the
             Grammar of Graphics
ggrepel      Repulsive Text and Label Geoms for 'ggplot2'
GLDEX        Fitting Single and Mixture of Generalised
             Lambda Distributions (RS and FMKL) using
             Various Methods
GLDreg       Fit GLD Regression Model and GLD Quantile
             Regression Model to Empirical Data
gmp          Multiple Precision Arithmetic
goftest      Classical Goodness-of-Fit Tests for Univariate
             Distributions
gplots       Various R Programming Tools for Plotting Data
graph        graph: A package to handle graph data
             structures
gridBase     Integration of base and grid graphics
gridExtra    Miscellaneous Functions for "Grid" Graphics
gsheet       Download Google Sheets Using Just the URL
gtable       Arrange 'Grobs' in Tables
gtools       Various R Programming Tools
haven        Import and Export 'SPSS', 'Stata' and 'SAS'
             Files
HH           Statistical Analysis and Data Display:
             Heiberger and Holland
highr        Syntax Highlighting for R Source Code
```

Lets check my Laptop ...

```
Hmisc            Harrell Miscellaneous
hms              Pretty Time of Day
HSAUR2           A Handbook of Statistical Analyses Using R (2nd
                 Edition)
htmlTable        Advanced Tables for Markdown/HTML
htmltools        Tools for HTML
htmlwidgets      HTML Widgets for R
httpuv           HTTP and WebSocket Server Library
httr             Tools for Working with URLs and HTTP
ICC              Facilitating Estimation of the Intraclass
                 Correlation Coefficient
igraph           Network Analysis and Visualization
igraphdata       A Collection of Network Data Sets for the
                 'igraph' Package
igraphtosonia    Convert iGraph graps to SoNIA .son files
inline           Functions to Inline C, C++, Fortran Function
                 Calls from R
invgamma         The Inverse Gamma Distribution
irlba            Fast Truncated SVD, PCA and Symmetric
                 Eigendecomposition for Large Dense and Sparse
                 Matrices
iterators        Provides Iterator Construct for R
jsonlite         A Robust, High Performance JSON Parser and
                 Generator for R
kernlab          Kernel-Based Machine Learning Lab
knitr            A General-Purpose Package for Dynamic Report
                 Generation in R
kutils           Project Management Tools
labeling         Axis Labeling
languageR        Data sets and functions with "Analyzing
                 Linguistic Data: A practical introduction to
                 statistics".
latentnet        Latent Position and Cluster Models for
                 Statistical Networks
```

Lets check my Laptop ...

```
latticeExtra          Extra Graphical Utilities Based on Lattice
lavaan                Latent Variable Analysis
lazyeval              Lazy (Non-Standard) Evaluation
leaps                 Regression Subset Selection
lme4                  Linear Mixed-Effects Models using 'Eigen' and
                      S4
lmtest                Testing Linear Regression Models
lpSolve               Interface to 'Lp_solve' v. 5.5 to Solve
                      Linear/Integer Programs
lsmeans               Least-Squares Means
lubridate             Make Dealing with Dates a Little Easier
magrittr              A Forward-Pipe Operator for R
markdown              'Markdown' Rendering for R
MatchIt               MatchIt: Nonparametric Preprocessing for
                      Parametric Casual Inference
MatrixModels          Modelling with Sparse And Dense Matrices
matrixStats           Functions that Apply to Rows and Columns of
                      Matrices (and to Vectors)
mclust                Gaussian Mixture Modelling for Model-Based
                      Clustering, Classification, and Density
                      Estimation
mcmc                  Markov Chain Monte Carlo
MCMCpack              Markov Chain Monte Carlo (MCMC) Package
memisc                Tools for Management of Survey Data and the
                      Presentation of Analysis Results
memoise               Memoisation of Functions
MEMSS                 Data sets from Mixed-effects Models in S
microbenchmark        Accurate Timing Functions
mime                  Map Filenames to MIME Types
miniUI                Shiny UI Widgets for Small Screens
minqa                 Derivative-free optimization algorithms by
                      quadratic approximation
mitools               Tools for multiple imputation of missing data
mixer                 Random graph clustering
```

Lets check my Laptop ...

```
mixtools              Tools for Analyzing Finite Mixture Models
mlmRev                Examples from Multilevel Modelling Software
                      Review
mnormt                The Multivariate Normal and t Distributions
modeltools            Tools and Classes for Statistical Models
multcomp              Simultaneous Inference in General Parametric
                      Models
munsell               Utilities for Using Munsell Colours
mvtnorm               Multivariate Normal and t Distributions
NetCluster            Clustering for networks
NetData               Network Data for McFarland's SNA R labs
network               Classes for Relational Data
networkDynamic        Dynamic Extensions for Network Objects
networkDynamicData    Dynamic (Longitudinal) Network Datasets
nFactors              Parallel Analysis and Non Graphical Solutions
                      to the Cattell Scree Test
nloptr                R interface to NLopt
NLP                   Natural Language Processing Infrastructure
NMF                   Algorithms and Framework for Nonnegative Matrix
                      Factorization (NMF)
np                    Nonparametric kernel smoothing methods for
                      mixed data types
numDeriv              Accurate Numerical Derivatives
openssl               Toolkit for Encryption, Signatures and
                      Certificates Based on OpenSSL
openxlsx              Read, Write and Edit XLSX Files
optextras             Tools to Support Optimization Possibly with
                      Bounds and Masks
optimx                A Replacement and Extension of the optim()
                      Function
ordinal               Regression Models for Ordinal Data
orthopolynom          Collection of functions for orthogonal and
                      orthonormal polynomials
packrat               A Dependency Management System for Projects and
```

Lets check my Laptop ...

```
                          their R Package Dependencies
PAFit                     Joint Inference of Preferential Attachment and
                          Node Fitness in Temporal Complex Networks
pander                    An R Pandoc Writer
pbivnorm                  Vectorized Bivariate Normal CDF
pbkrtest                  Parametric Bootstrap and Kenward Roger Based
                          Methods for Mixed Model Comparison
pcse                      Panel-Corrected Standard Error Estimation in R
pkgmaker                  Package development utilities
PKI                       Public Key Infrastucture for R Based on the
                          X.509 Standard
PKPDmodels                Pharmacokinetic/pharmacodynamic models
plm                       Linear Models for Panel Data
plogr                     The 'plog' C++ Logging Library
plotrix                   Various Plotting Functions
plyr                      Tools for Splitting, Applying and Combining
                          Data
polyclip                  Polygon Clipping
polynom                   A Collection of Functions to Implement a Class
                          for Univariate Polynomial Manipulations
portableParallelSeeds     Allow Replication of Simulations on Parallel
                          and Serial Computers
prabclus                  Functions for Clustering of Presence-Absence,
                          Abundance and Multilocus Genetic Data
praise                    Praise Users
pscl                      Political Science Computational Laboratory,
                          Stanford University
psidR                     Build Panel Data Sets from PSID Raw Data
psych                     Procedures for Psychological, Psychometric, and
                          Personality Research
qdapTools                 Tools for the 'qdap' Package
quadprog                  Functions to solve Quadratic Programming
                          Problems.
quantmod                  Quantitative Financial Modelling Framework
```

Lets check my Laptop ...

```
quantreg          Quantile Regression
R.cache           Fast and Light-Weight Caching (Memoization) of
                  Objects and Results to Speed Up Computations
R.methodsS3       S3 Methods Simplified
R.oo              R Object-Oriented Programming with or without
                  References
R.rsp             Dynamic Generation of Scientific Reports
R.utils           Various Programming Utilities
R2OpenBUGS        Running OpenBUGS from R
R6                Classes with Reference Semantics
RBGL              An interface to the BOOST graph library
Rcgmin            Conjugate Gradient Minimization of Nonlinear
                  Functions
RColorBrewer      ColorBrewer Palettes
Rcpp              Seamless R and C++ Integration
RcppArmadillo     'Rcpp' Integration for the 'Armadillo'
                  Templated Linear Algebra Library
RcppEigen         'Rcpp' Integration for the 'Eigen' Templated
                  Linear Algebra Library
RCurl             General Network (HTTP/FTP/...) Client Interface
                  for R
readr             Read Rectangular Text Data
registry          Infrastructure for R Package Registries
relevent          Relational Event Models
reshape2          Flexibly Reshape Data: A Reboot of the Reshape
                  Package
rgl               3D Visualization Using OpenGL
Rgraphviz         Provides plotting capabilities for R graph
                  objects
rJava             Low-Level R to Java Interface
rjson             JSON for R
RJSONIO           Serialize R objects to JSON, JavaScript Object
                  Notation
rmarkdown         Dynamic Documents for R
```

Lets check my Laptop ...

Rmpfr	R MPFR - Multiple Precision Floating-Point Reliable
rngtools	Utility functions for working with Random Number Generators
robustbase	Basic Robust Statistics
rockchalk	Regression Estimation and Presentation
ROCR	Visualizing the Performance of Scoring Classifiers
roxygen2	In-Line Documentation for R
RPostgreSQL	R interface to the PostgreSQL database system
rprojroot	Finding Files in Project Subdirectories
rsconnect	Deployment Interface for R Markdown Documents and Shiny Applications
RSiena	Siena - Simulation Investigation for Empirical Network Analysis
RSQLite	'SQLite' Interface for R
rstan	R Interface to Stan
rstudioapi	Safely Access the RStudio API
Rvmmin	Variable Metric Nonlinear Function Minimization
sand	Statistical Analysis of Network Data with R
sandwich	Robust Covariance Matrix Estimators
SAScii	Import ASCII files directly into R using only a SAS input script
scales	Scale Functions for Visualization
scatterplot3d	3D Scatter Plot
segmented	Regression Models with Breakpoints/Changepoints Estimation
setRNG	Set (Normal) Random Number Generator and Seed
shiny	Web Application Framework for R
slam	Sparse Lightweight Arrays and Matrices
sna	Tools for Social Network Analysis
snow	Simple Network of Workstations
SnowballC	Snowball stemmers based on the C libstemmer UTF-8 library

Lets check my Laptop ...

```
sourcetools          Tools for Reading, Tokenizing and Parsing R
                     Code
sp                   Classes and Methods for Spatial Data
SparseM              Sparse Linear Algebra
spatstat             Spatial Point Pattern Analysis, Model-Fitting,
                     Simulation, Tests
StanHeaders          C++ Header Files for Stan
stargazer            Well-Formatted Regression and Summary
                     Statistics Tables
statnet              Software Tools for the Statistical Analysis of
                     Network Data
statnet.common       Common R Scripts and Utilities Used by the
                     Statnet Project Software
stringi              Character String Processing Facilities
stringr              Simple, Consistent Wrappers for Common String
                     Operations
survey               Analysis of Complex Survey Samples
survival             Survival Analysis
svDialogs            SciViews GUI API - Dialog boxes
svGUI                SciViews GUI API - Functions to manage GUIs
svUnit               SciViews GUI API - Unit testing
tables               Formula-Driven Table Generation
tensor               Tensor product of arrays
tergm                Fit, Simulate and Diagnose Models for Network
                     Evolution Based on Exponential-Family Random
                     Graph Models
testthat             Unit Testing for R
texreg               Conversion of R Regression Output to LaTeX or
                     HTML Tables
TH.data              TH's Data Archive
tibble               Simple Data Frames
tidyr                Easily Tidy Data with 'spread()' and 'gather()'
                     Functions
tm                   Text Mining Package
```

Lets check my Laptop ...

```
triangle                Provides the Standard Distribution Functions
                        for the Triangle Distribution
trimcluster             Cluster analysis with trimming
tripack                 Triangulation of Irregularly Spaced Data
truncnorm               Truncated normal distribution
trust                   Trust Region Optimization
TTR                     Technical Trading Rules
tweedie                 Tweedie Exponential Family Models
ucminf                  General-Purpose Unconstrained Non-Linear
                        Optimization
vcd                     Visualizing Categorical Data
VGAM                    Vector Generalized Linear and Additive Models
viridis                 Default Color Maps from 'matplotlib'
visreg                  Visualization of Regression Models
wordcloud               Word Clouds
xlsx                    Read, write, format Excel 2007 and Excel
                        97/2000/XP/2003 files
xlsxjars                Package required POI jars for the xlsx package
XML                     Tools for Parsing and Generating XML Within R
                        and S-Plus
xml2                    Parse XML
xtable                  Export Tables to LaTeX or HTML
xts                     eXtensible Time Series
yaml                    Methods to Convert R Data to YAML and Back
zoo                     S3 Infrastructure for Regular and Irregular
                        Time Series (Z's Ordered Observations)

Packages in library '/usr/lib/R/library':

base                    The R Base Package
boot                    Bootstrap Functions (Originally by Angelo Canty
                        for S)
class                   Functions for Classification
cluster                 "Finding Groups in Data": Cluster Analysis
```

Lets check my Laptop ...

```
                        Extended Rousseeuw et al.
codetools               Code Analysis Tools for R
compiler                The R Compiler Package
datasets                The R Datasets Package
foreign                 Read Data Stored by Minitab, S, SAS, SPSS,
                        Stata, Systat, Weka, dBase, ...
graphics                The R Graphics Package
grDevices               The R Graphics Devices and Support for Colours
                        and Fonts
grid                    The Grid Graphics Package
KernSmooth              Functions for Kernel Smoothing Supporting Wand
                        & Jones (1995)
lattice                 Trellis Graphics for R
MASS                    Support Functions and Datasets for Venables and
                        Ripley's MASS
Matrix                  Sparse and Dense Matrix Classes and Methods
methods                 Formal Methods and Classes
mgcv                    Mixed GAM Computation Vehicle with GCV/AIC/REML
                        Smoothness Estimation
nlme                    Linear and Nonlinear Mixed Effects Models
nnet                    Feed-Forward Neural Networks and Multinomial
                        Log-Linear Models
parallel                Support for Parallel computation in R
rpart                   Recursive Partitioning and Regression Trees
spatial                 Functions for Kriging and Point Pattern
                        Analysis
splines                 Regression Spline Functions and Classes
stats                   The R Stats Package
stats4                  Statistical Functions using S4 Classes
survival                Survival Analysis
tcltk                   Tcl/Tk Interface
tools                   Tools for Package Development
utils                   The R Utils Package
```

Lets check my Laptop ...

```
Warning message:
In library() :
  libraries '/usr/local/lib/R/site-library', '/usr/lib/R/site-library'
contain no packages
```

The package collections are in "libraries"

- Did you notice sections in previous output? Those are library folders
- R packages are sorted among folders (AKA "libraries")
  1. System-wide folders
     1. The main R distribution (e.g., stats, graphics)
     2. recommended packages distributed with R (e.g., MASS, parallel)
     3. Others added by system administrator
  2. Personal user folders, often within user HOME folder
- To see your current R PATH folders, run the function " `.libPaths()` "
  - The first letter in that functions name is a period! (UNIX tradition)

```
.libPaths()
```

```
[1] "/home/pauljohn/R/x86_64-pc-linux-gnu-library/3.4"
[2] "/usr/local/lib/R/site-library"
[3] "/usr/lib/R/site-library"
[4] "/usr/lib/R/library"
```

To load a package, run library(package-name-here)

- Almost every time I run

```
library(package-name-here)
```

- Whenever I run `library`, I almost always also run

```
help(package = "package-name-here")
```

- `MASS` is a famous recommended package distributed with R.

```
library(MASS)
help(package = "MASS")
```

Outline

Repository: a package server

- When you start R, an object named "options" is created in the workspace. It supplies settings, which you can change.
- Do you have any repositories configured?

```
options("repos")
```

  - On some computers I use, I'll see a list of servers, which will be searched in order
  - But sometimes I see

```
$repos
CRAN
"@CRAN@"
```

  - This means my default repository is undefined.

- If you do not have specified repos, then when you run any functions like "update.packages()", "install.packages()", and so forth, R will ask you to choose a repository

Repository: a package server ...

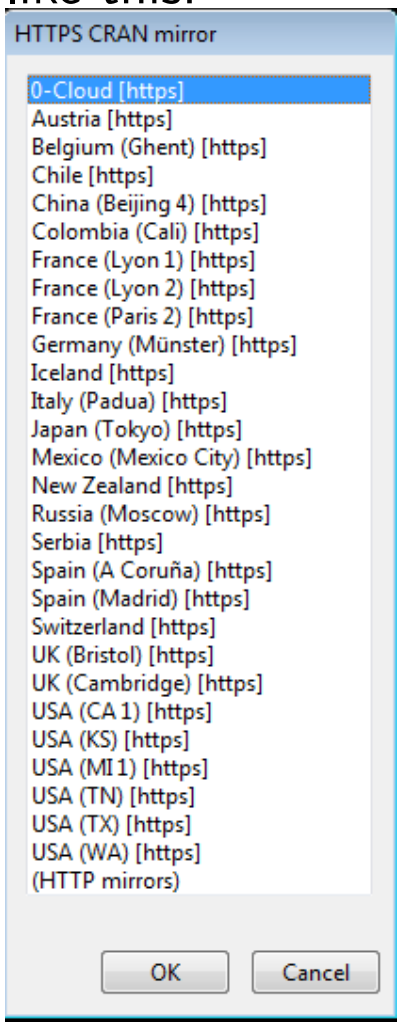
```
--- Please select a CRAN mirror for use in this session ---
HTTPS CRAN mirror

 1: 0-Cloud [https]                  2: Algeria [https]
 3: Australia (Canberra) [https]     4: Australia (Melbourne) [https]
 5: Australia (Perth) [https]        6: Austria [https]
 7: Belgium (Ghent) [https]          8: Brazil (PR) [https]
 9: Brazil (RJ) [https]             10: Brazil (SP 1) [https]
11: Bulgaria [https]                12: Canada (MB) [https]
13: Chile 1 [https]                 14: Chile 2 [https]
15: China (Beijing) [https]         16: China (Hefei) [https]
17: China (Lanzhou) [https]         18: Colombia (Cali) [https]
19: Czech Republic [https]          20: Denmark [https]
21: Estonia [https]                 22: France (Lyon 1) [https]
23: France (Lyon 2) [https]         24: France (Marseille) [https]
25: France (Montpellier) [https]    26: France (Paris 2) [https]
27: Germany (G
29: Iceland [https]                 30: India [https]
31: Indonesia (Jakarta) [https]     32: Ireland [https]
33: Italy (Padua) [https]           34: Japan (Tokyo) [https]
35: Malaysia [https]                36: Mexico (Mexico City) [https]
37: New Zealand [https]             38: Norway [https]
39: Philippines [https]             40: Russia (Moscow) [https]
41: Serbia [https]                  42: Spain (A Coru
43: Spain (Madrid) [https]          44: Sweden [https]
45: Switzerland [https]             46: Taiwan (Chungli) [https]
47: Turkey (Denizli) [https]        48: Turkey (Mersin) [https]
49: UK (Bristol) [https]            50: UK (Cambridge) [https]
51: UK (London 1) [https]           52: USA (CA 1) [https]
53: USA (IA) [https]                54: USA (IN) [https]
55: USA (KS) [https]                56: USA (MI 1) [https]
57: USA (OR) [https]                58: USA (TN) [https]
59: USA (TX 1) [https]              60: USA (TX 2) [https]
61: Vietnam [https]                 62: (HTTP mirrors)
```

Repository: a package server ...

- CRMDA hosts server 55.

Repository: a package server

In Windows, looks like this:



Ways to avoid that interactive repos choice

1. Put this at the top of your R script:

```
CRAN <-
    "https://rweb.crmda.ku.edu/cran"
options(repos = CRAN)
```

To check result, look at ALL of the options with

```
options()
```

or just check that one:

```
options("repos")
```

```
$repos
[1] "https://rweb.crmda.ku.edu/cran"
```

2. Specify the `repos` argument in "install.packages" or "update.packages" (example coming up)

Perhaps the Company that provided your GUI configured "repos"

- I recently installed RStudio in Windows and found that they decided for me

```
> options("repos")
$repos
                                    CRAN
          "https://cran.rstudio.com/"
                                CRANextra
"http://www.stats.ox.ac.uk/pub/RWin"
attr(,"RStudio")
[1] TRUE
```

- Whereas R for Windows GUI has

```
> options("repos")
$repos
                                 CRAN
                            "@CRAN@"
                            CRANextra
"http://www.stats.ox.ac.uk/pub/RWin"
```

Perhaps the Company that provided your GUI configured "repos" ...

- R for Windows GUI has a pull down:
  Packages -> Set CRAN Mirror
  which has the same effect as options() function previously described.

Check Available Packages

- Browse a CRAN server's list of packages.
  - http://rweb.crmda.ku.edu/cran/web/packages/index.html
- Number of packages in CRAN is growing
  - 2017-05-13: 10606 packages
  - 2016-05-17: 8500
  - 2015-04-28: 6400.
- Difficult to know which packages are worthwhile. *Buyer Beware!*
- Inside R, check for the list of packages on CRAN

```
giantList <-
    available.packages(repos="https://rweb.crmda.ku.
row.names(giantList)
```

What Packages Do You Really Need?

- When we started with R, there were 100 packages and our policy was to install everything and keep it up to date. Today, that is impractical.
- CRMDA workstations have a script that runs nightly to make sure that about 500 packages are installed and up-to-date.
- R program will be found in the R folder of this project.
- Beyond that, users can install packages for themselves in their own HOME directories (that's the next topic).

KU

How to install new packages

- Run install.packages().

```
install.packages(c("rockchalk", "kutils"),
    dependencies = TRUE, repos =
    "https://rweb.crmda.ku.edu/cran")
```

- All good KU students install the package "rockchalk" and look at its beautiful vignettes.
- In that example, I set dependencies = TRUE. That means I also get others on which this
  - "Depends"
  - "Imports" (package you want loads features from a package)
  - "LinkingTo" (package you want uses C/Fortran libraries from others)
  - "Suggests" (package you want has examples that require these packages)
  - "Enhances" (other packages that might benefit from the one you are installing)

KU

How to install new packages ...

- "Suggests" or "Enhances" sometimes cause in trouble, pull in too many things you don't want.
- For example, when I install the package "`rms`" with `dep = TRUE`, I see

```
also installing the dependencies 'stringi', 'magrittr',
    'colorspace', 'RColorBrewer', 'chron', 'stringr', 'dichromat',
    'munsell', 'labeling', 'zoo', 'Formula', 'latticeExtra',
    'acepack', 'gtable', 'gridExtra', 'data.table', 'digest',
    'plyr', 'reshape2', 'scales', 'mvtnorm', 'TH.data',
    'sandwich', 'Hmisc', 'ggplot2', 'polspline', 'multcomp'
```

- As a result, these days I'm more likely to run with a leaner mixture:

```
install.packages("rockchalk", dep = c("Depends",
   "Imports", "LinkingTo"))
```

- The default for dependencies in R is now
  `c("Depends", "Imports", "LinkingTo", "Suggests")`

Installing several packages: Example

```
install.packages(c("lme4", "car"), dep =
    c("Depends", "Imports", "LinkingTo"), repos =
    "http://rweb.crmda.ku.edu/cran")
```

:: and :::

- Packages have functions in 2 categories
  - exported, meaning you can use them by typing their names
  - non-exported (private) functions, which are more difficult to access.
- Same-named functions may exist in many packages. I put "`summarize()`" in the rockchalk package, without realizing that functions of same name existed in the packages "Hmisc" and "plyr".
- The package that is loaded last will "steal" control of that name, blocking access to previously loaded functions of same name.
- To access those functions, use the *disambiguator* "::"

  Syntax package::function().
  Namespace The name of the package is also a "namespace" where functions can be found.
- Paranoid people might call a base-R function like plot by naming its namespace as well:

:: and ::: ...

```
graphics::plot(whatever, whatever)
```

- I don't generally write out the full package::function name unless I'm sure there is a conflict

- The ":::" access method
  - Some functions are "private", not exported.
  - package:::function() is syntax to access those functions.

Difference between library load and ::

- The `library()` function to loads all of a package's functions

```
library(rockchalk)
```

- But loading full package not absolutely necessary. Can access functions from a package with the "`::`" notation.
  Example

```
data(swiss)
## Assign shorter names to fit on one page
colnames(swiss) <- c("Agric", "Cath", "Educ",
    "Exam", "Fert", "Inf.Mort")
rockchalk::summarize(swiss)
```

Difference between library load and :: ...

```
Numeric variables
            Agric      Cath      Educ      Exam      Fert     Inf.Mort
min          35         1.20      3         1         2.15     10.80
med          70.40     54.10     16         8        15.14     20
max          92.50     89.70     37        53       100        26.60
mean         70.14     50.66     16.49     10.98     41.14     19.94
sd           12.49     22.71      7.98      9.62     41.70      2.91
skewness     -0.46     -0.32      0.45      2.27      0.48     -0.33
kurtosis      0.26     -0.89     -0.14      6.14     -1.67      0.78
nobs         47        47        47        47        47        47
nmissing      0         0         0         0         0         0
```

- Some experts avoid loading whole package if possible–R session is more "responsive" if list of immediately accessible functions is smaller. Don't worry about this until you see a problem.

Staying up to date

- Please remember that the package world is always changing.

```
update.packages()
```

Triggers 2 interactive steps

1. R asks which CRAN server do you use?
2. R will ask about each and every package to be updated.

- Don't let R ask you one-by-one about packages.

```
update.packages(ask = FALSE, checkBuilt =
    TRUE)
```

- I generally insert checkBuilt as well

```
update.packages(ask = FALSE, checkBuilt =
    TRUE)
```

Packages built with old R version are replaced, even if package's version is not updated.

Staying up to date ...

- Avoid the repo-chooser with

```
update.packages(ask = FALSE, checkBuilt =
    TRUE, repos =
    "http://rweb.crmda.ku.edu/cran")
```

Other repositories exist as well

- The R script to maintain CRMDA computers looks at 3 repositories:

```
CRAN <- "http://rweb.crmda.ku.edu/cran"
KRAN <- "http://rweb.crmda.ku.edu/kran"
BIOC <-
    "http://www.bioconductor.org/packages/3.3/bioc"
options(repos = c(KRAN, CRAN, BIOC))
update.packages(ask = FALSE, checkBuilt =
    TRUE)
```

- KRAN is our testing server, for fine packages like "rockchalk" and "kutils"
- After CRAN, Bioconductor (BIOC) is the most notable package repository. (*Setting that up is bothersome/tricky*).

System administration questions

- When you run this, where does your package "go"?

```
install.packages(c("lme4"), dep = TRUE)
```

- R selects among folders in .libPaths().
- But which one? That's a puzzler!
- If you are an Administrator, R may (usually will) try to put new packages in a System-wide directory so other users can use as well.
- If you are a limited user, R it will ask if you want to install in your own user account (only for you, not for other users).
- The argument "lib" can be used to specify the install directory, as in

```
install.packages(c("lme4"), dep = TRUE, lib =
    "/some/folder/you/like")
```

Don't forget about help after loading packages

- Make a package's functions immediately accessible.

```
library(lme4)
```

- Dont forget to always run

```
help(package = "lme4")
```

- See helps on specifics

```
?lmer
?glmer
```

- I just got some bad news
  - On Mac and Windows, the default help shows in the Web Browser
  - That's less than ideal because running the examples line-by-line is a problem. (I'll show how to avoid that if you use Emacs)

KU

Outline

Points of emphasis

- R is an "extensible" software framework built on a packaging system
- `help.start()` to review docs and review installed packages
- Retrieve a list of installed packages:

```
> library()
```

- Use the functions `install.packages()` and `update.packages()` to keep everything in order
- When you have trouble, ALWAYS include output from "sessionInfo()" in your requests for help.

References

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Session

```
sessionInfo()
```

```
R version 3.4.4 (2018-03-15)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 18.04 LTS

Matrix products: default
BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.7.1
LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.7.1

locale:
 [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
     LC_TIME=en_US.UTF-8
 [4] LC_COLLATE=en_US.UTF-8     LC_MONETARY=en_US.UTF-8
     LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=en_US.UTF-8       LC_NAME=C                  LC_ADDRESS=C
[10] LC_TELEPHONE=C             LC_MEASUREMENT=en_US.UTF-8
    LC_IDENTIFICATION=C

attached base packages:
[1] stats     graphics  grDevices utils     datasets  base

other attached packages:
[1] MASS_7.3-49
```

Session ...

```
loaded via a namespace (and not attached):
 [1] Rcpp_0.12.15         lattice_0.20-35     grid_3.4.4
     MatrixModels_0.4-1   nlme_3.1-137
 [6] rockchalk_1.8.110    SparseM_1.77        minqa_1.2.4
     nloptr_1.0.4         car_2.1-6
[11] Matrix_1.2-14        splines_3.4.4       lme4_1.1-15
     tools_3.4.4          pbkrtest_0.4-7
[16] parallel_3.4.4       compiler_3.4.4      mgcv_1.8-23
     nnet_7.3-12          quantreg_5.35
[21] methods_3.4.4
```