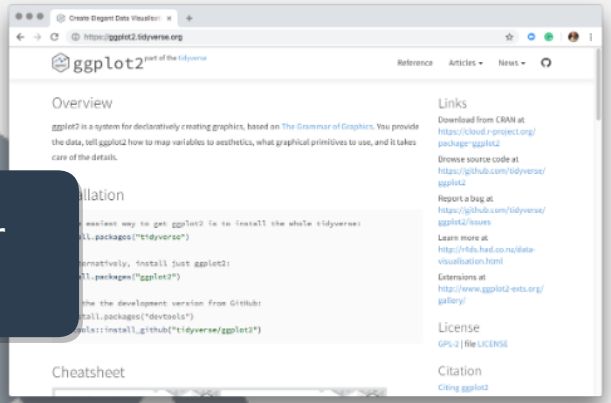
Data Science with R Workflow

If you want to learn R and this workflow ***for business analysis***, take the **R For Business Analysis (DS4B 101-R) course through Business Science University.**



Click the links for Documentation

Import

readr (CS)
readxl / writexl
odbc / DBI
rvest

Tidy

tibble (CS)
tidyr (CS)

Transform

dplyr (CS)
stringr (CS)
lubridate (CS)
forcats (CS)
Base R (CS)

Visualize

ggplot2 (CS)

Model

purrr (CS)
(iteration)

recipes
rsample
parsnip
broom
yardstick
dials

Communicate

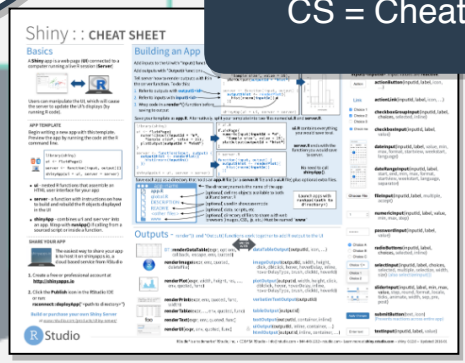
RMarkdown (CS)
Shiny (CS)

RStudio IDE (CS)

fs (file system)

reticulate (CS)

CS = Cheat Sheet



Important Resources

- **R For Data Science Book:** http://r4ds.had.co.nz/
- **Rmarkdown Book:** https://bookdown.org/yihui/rmarkdown/
- **Data Visualization Book**: https://r-graphics.org/
- **More Cheatsheets:** https://www.rstudio.com/resources/cheatsheets/
- **tidyverse packages:** https://www.tidyverse.org/
- **Connecting to databases:** https://db.rstudio.com/
- **RMarkdown website:** https://rmarkdown.rstudio.com/
- **Shiny web applications website:** http://shiny.rstudio.com/
- **Jenny Bryan's purrr tutorial:** https://jennybryan.org/

"Data Science Courses for Business"

Business Science University
university.business-science.io

version: 2.0

Data Science with R

Web Applications & the "Shinyverse"

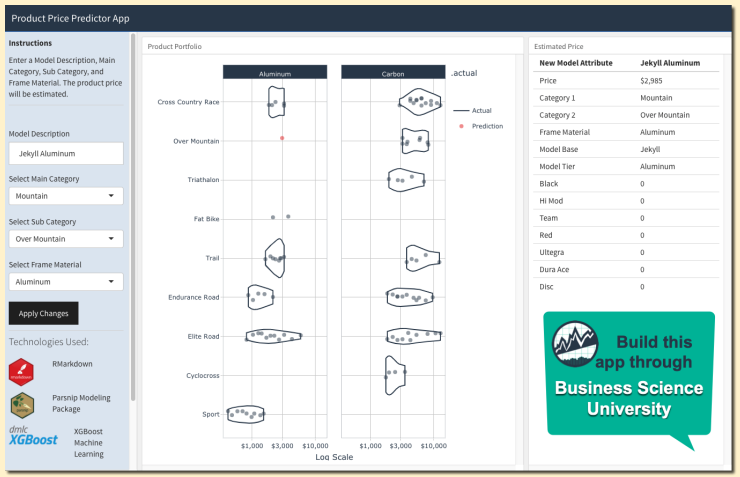
Start	Components	Advanced Concepts (Optional)	Testing	Publish
Shiny (CS) Flexdashboard	Widgets shinyWidgets htmlWidgets shinyjs	crosstalk shinyrouter Async (promises)	profvis shinytest shinyloadtest	shinyapps.io Shiny Server Connect (Enterprise)

Flexdashboard Apps

Flexdashboard is an RMarkdown-based dashboard tool that can be used to integrate shiny components at runtime. Development is fast and efficient, but layouts are not as open to modification as building an app using Shiny.

Shiny Apps

Shiny is an R-Package that enables web app development from R. Contains R functions for common HTML structures, UI Controls (Components/Widgets), and web framework tools. The framework if highly flexible, but users require more knowledge of HTML & CSS.



Flexdashboard App
Built in DS4B 102-R (Demo Here)

Themes, Dashboards, & Examples

Flexdashboard Gallery

Themes
Layouts

Shiny Gallery

shinythemes
shinydashboard
shinydashboardPlus

Semantic

shiny.semantic
semantic.dashboard

Bootstrap 4

bs4Dash

Argon

argonR
argonDash

"Data Science Courses for Business"

Business Science University
university.business-science.io

Data Science with R

Special Topics

Time Series Analysis

- Time-aware tibbles: tibbletime & tsibble
- Convert between classes: timetk & tsbox
- Time Series Index Summary: timetk
- Generating Future Series: timetk

Forecasting

- ARIMA, ETS, etc: forecast & fable
- Tidy, glance, augment for forecast models: sweep
- Converting forecast prediction to tibble: sweep

Anomaly Detection

- Identify anomalies: anomalize

Financial Analysis

- Getting financial data: tidyquant & quantmod
- Quantitative Analysis: tidyquant & xts/TTR
- Portfolio Analysis: tidyquant & PerformanceAnalytics

Financial & Time Viz

- Static:
  - tidyquant - Financial ggplot2 geoms
- Interactive:
  - highcharter - highchart.js in R
  - dygraphs - xts plotting
  - plotly (CS) - plotly.js (financial) in R

Text Analysis & NLP

- Text Mining with R (Book): tidytext
- NLP:
  - H2O word2vec: Word embeddings
  - text2vec: fast vectorization, topic modeling
  - udpipe: UDPipe C++ lib in R

Network Analysis

- Network Data Transformations (Tidy): tidygraph
- Network Data Transformations: igraph

Network Viz

- Static:
  - ggraph - Graph plotting utilities for ggplot2
- Interactive (JavaScript):
  - networkD3 - D3 Networks in R
  - plotly (CS) - plotly.js (network graphs) in R

Geospatial Analysis

- Geocoding (getting lat/long, bboxes, & sf's):
  - ggmap - Google API (requires key)
  - osmdata - OpenStreet Overpass API
  - tmaptools - OpenStreet Nominatum API
- Simple Features (sf objects): sf (CS) (tidy)
- Spatial Objects (sp objects): sp (non-tidy)

Geospatial Viz

- Static:
  - ggmap - Google API (requires key)
  - osmplotr - Impressive Maps via OSM
  - tmap - Thematic Maps
  - cartography (CS) - Thematic Maps
- Interactive (JavaScript):
  - leaflet (CS) - leaflet.js in R
  - plotly (CS) - plotly.js (maps) in R

Machine Learning

- Multi-Threaded/Scalable/Production ML:
  - H2O (CS)
  - Extreme Gradient Boosting: xgboost
  - R + Spark: sparklyr (CS)
  - Sparkling Water (Spark + H2O): rsparkling
- ML (Tidy): parsnip
- ML: caret (CS)

Deep Learning

- R Interface to TensorFlow Homepage:
  - Keras (CS)
  - TF Estimators
  - TensorFlow (Core)

Speed & Scale

- Faster than dplyr & pandas: data.table (CS)
- Distributed Cluster (Spark): sparklyr (CS)
- Parallel Processing: furrr

Interoperability

- Python: reticulate (**CS**)
- Java: rJava
- C++: Rcpp

Miscellaneous Tools

- **Interactive Plotting:** htmlwidgets for R
- **Building R Packages:** R packages Book
  - Pkg Development Tools: devtools (CS)
  - R Templates: usethis
  - Build Web Doc's: pkgdown
- **Advanced Concepts** (Advanced R Book)
  - rlang & Tidy Evaluation (CS)
- **Making Blogs & Books:**
  - blogdown, bookdown
- **Posting Code (GitHub, Stack Overflow):** reprex