

Library

Library

plyr: Tools for Splitting, Applying and Combining Data

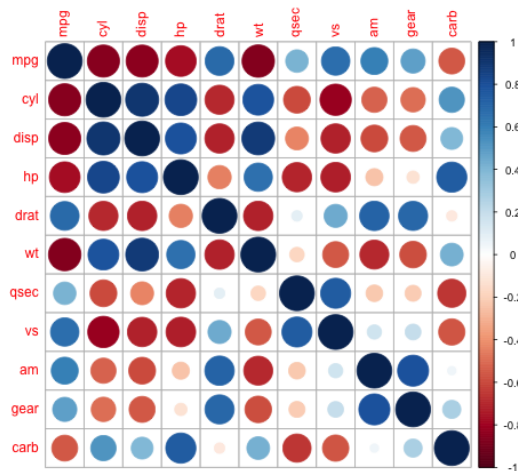
The **corrplot** package is a graphical display of a correlation matrix, confidence interval.

It also contains some algorithms to do matrix reordering.

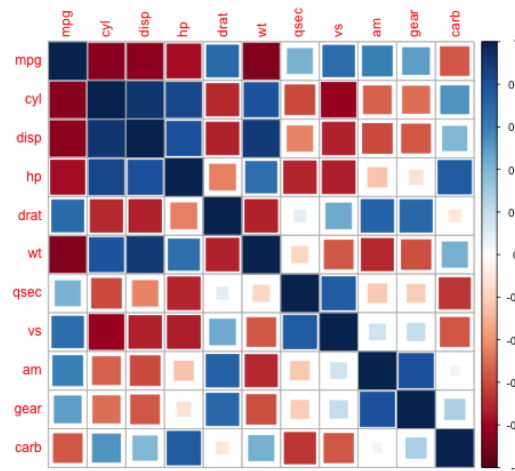
In addition, corrplot is good at details, including choosing color, text labels, color labels, layout, etc.

There are seven visualization methods (parameter method) in **corrplot** package, named "circle", "square", "ellipse", "number", "shade", "color", "pie".

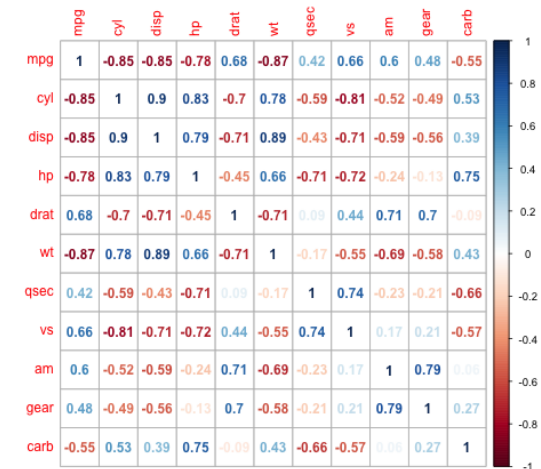
- library(corrplot)
-
- M <- cor(mtcars)
- corrplot(M, method = "circle")



```
M <- cor(mtcars)
corrplot(M, method =
"circle")
```



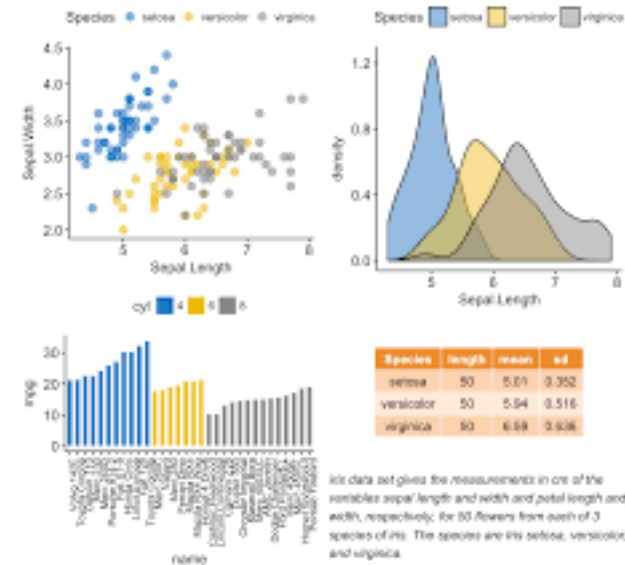
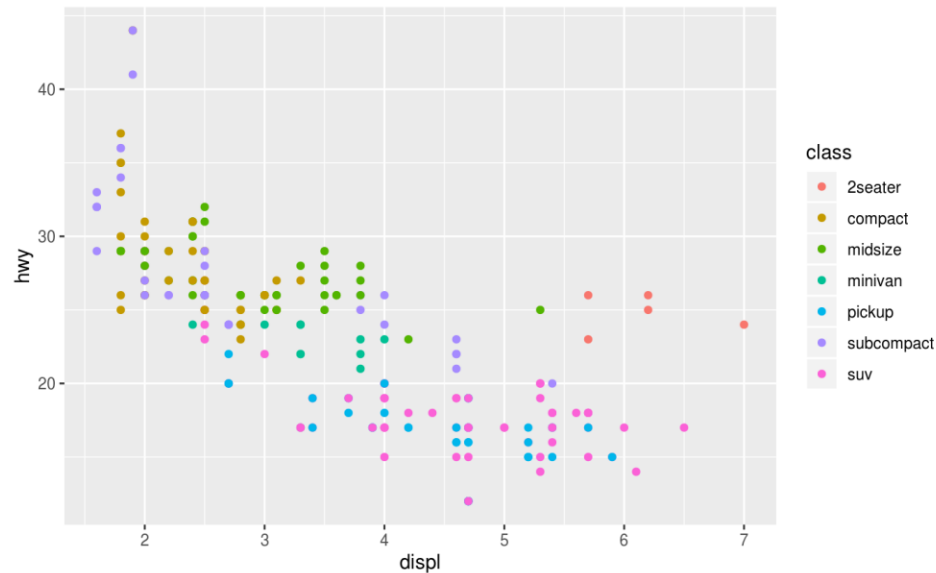
```
corrplot(M, method =
"square")
```



```
corrplot(M, method =
"number") # Display the
correlation coefficient
```

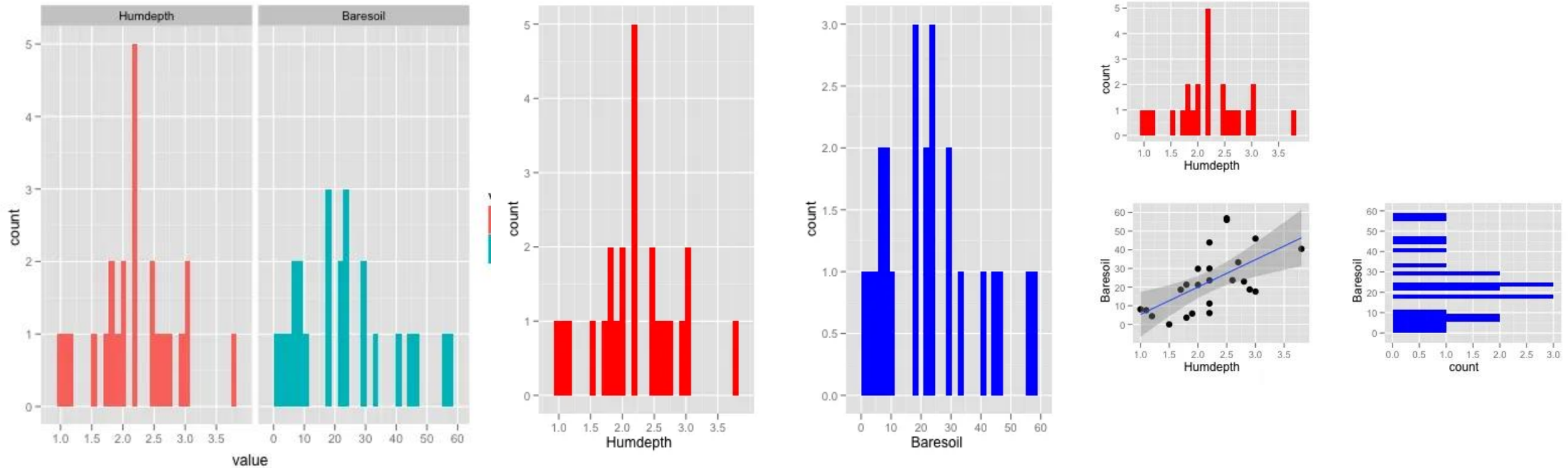
Data Visualization

- ggplot2: is a data visualization package for the statistical programming language R.
- [library\(ggplot2\)](#)
- [ggplot\(mpg, aes\(displ, hwy, colour = class\)\)](#) + [geom_point\(\)](#)



- gridExtra: Miscellaneous Functions for "Grid" Graphics
- Provides a number of user-level functions to work with "grid" graphics, notably to arrange multiple grid-based plots on a page, and draw tables.

Extension of QQ Plot--Extra! Extra! Get Your gridExtra!



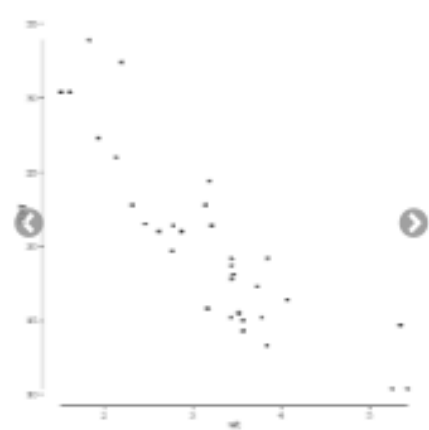
Through its function `grid.arrange`, you can make a multi-paneled graph using `ggplot2` plots, lattice plots, and more

- ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'
- Some extra themes, geoms, and scales for 'ggplot2'.
- Provides 'ggplot2' themes and scales that replicate the look of plots by Edward Tufte, Stephen Few, 'Fivethirtyeight', 'The Economist', 'Stata', 'Excel', and 'The Wall Street Journal', among others. Provides 'geoms' for Tufte's box plot and range frame.



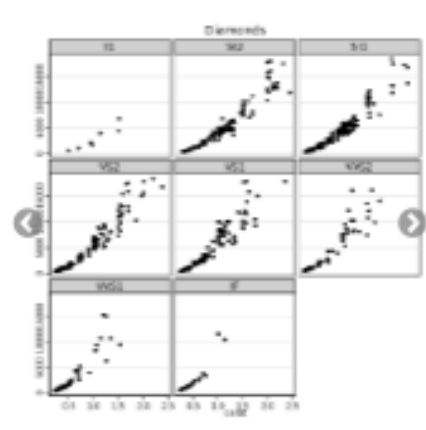
theme_wsj

Wall Street Journal theme



theme_tufte

Tufte Maximal Data, Minimal Ink Theme



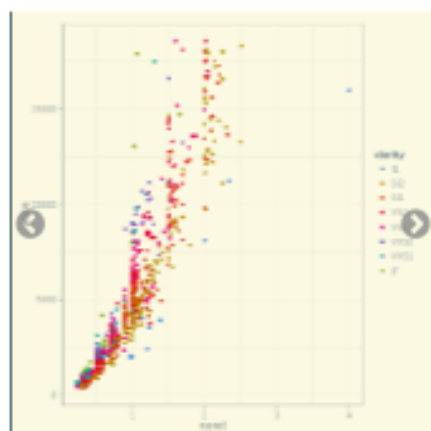
theme_stata

Themes based on Stata graph schemes



theme_solid

Theme with nothing other than a background color



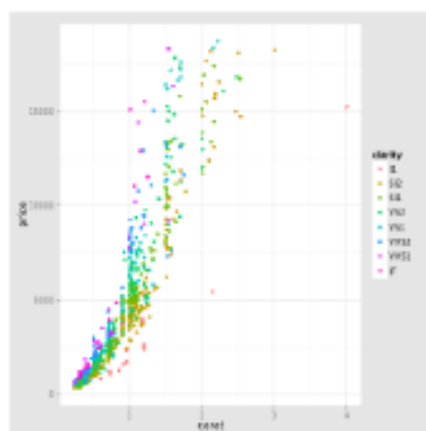
theme_solarized

ggplot color themes based on the Solarized palette



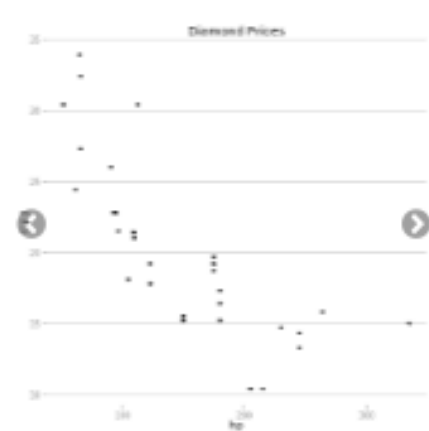
theme_map

Clean theme for maps



theme_igray

Inverse gray theme



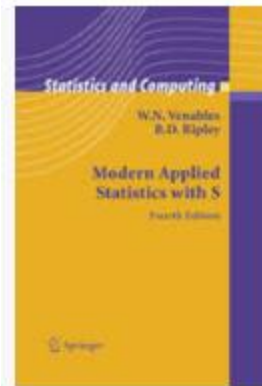
theme_hc

Highcharts JS theme

- caret Package
- The **caret** package (short for Classification And REgression Training) contains functions to streamline the model training process for complex regression and classification problems.
- *Caret Package is a comprehensive framework for building machine learning models in R.*
- <https://www.machinelearningplus.com/machine-learning/caret-package/>

MASS

- Package 'MASS'
- Functions and datasets to support Venables and Ripley, "Modern Applied Statistics with S" (4th edition, 2002).



Modern Applied Statistic.
pringer.com

- party:
- use 'party' R package to train model using decision tree
- **library**(party)

Tree-Based Models

Recursive partitioning is a fundamental tool in data mining. It helps us explore the structure of a set of data, while developing easy to visualize decision rules for predicting a categorical (classification tree) or continuous (regression tree) outcome. This section briefly describes CART modeling, conditional inference trees, and random forests.

CART Modeling via rpart

Classification and regression trees (as described by Brieman, Freidman, Olshen, and Stone) can be generated through the [rpart](#) package. Detailed information on rpart is available in [An Introduction to Recursive Partitioning Using the RPART Routines](#). The general steps are provided below followed by two examples.

- library(RColorBrewer)
- RColorBrewer: ColorBrewer Palettes
- Provides color schemes for maps (and other graphics)
- Using RColorBrewer to colour your figures in R

### Load the package or install if not present	
	if (!require("RColorBrewer")) {
	install.packages("RColorBrewer")
	library(RColorBrewer)
	}

ROCR: Visualizing the Performance of Scoring Classifiers

- `library(ROCR)`
- is helpful at estimating performance measures and plotting these measures over a range of cutoffs.

- class: Functions for Classification
- Various functions for classification, including k-nearest neighbour, Learning Vector Quantization and Self-Organizing Maps.

- Recursive Partitioning and Regression Trees
- Recursive partitioning for classification, regression and survival trees. An implementation of most of the functionality of the 1984 book by Breiman, Friedman, Olshen and Stone.
- `rpart(formula, data, weights, subset, na.action = na.rpart, method, model = FALSE, x = FALSE, y = TRUE, parms, control, cost, ...)`

- `library(rattle)`
- The R Analytic Tool To Learn Easily (Rattle) provides a collection of utilities functions for the data scientist.