## Section 1: Introduction

|  |
| --- |
| 1.1 Welcome to ggplot2  1.2 Other resources  1.3 What is the “Grammar of graphics”  1.4  1.5 About this book  1.6 Installation  1.7 acknowledgments |

## Quickplot

#useful site: Exploratory Data Analysis SI 618 (university of Michigan)

Quickqplot is a convenient wrapper function

Use colour attribute to discern patterns

Contrived data set (called simple “DataSet”)

* Xvar : the predictor variable
* Yvar : the response variable
* DataSet : the data set name
* Grp : grouping (Control and Experimental)

|  |
| --- |
| qplot(Xvar, Yvar,data=DataSet,color=Grp) |

## Chapter 8 Polishing your plots for Publication

## Chapter 9 Manipulating Data

9.1 An Introduction to plyr

9.2 Converting data from “wide” to “long”

9.3 ggplot() methods

Linear Models

Writing your own

The plyr package provides tools for breaking and combining lists, arrays and data frames.

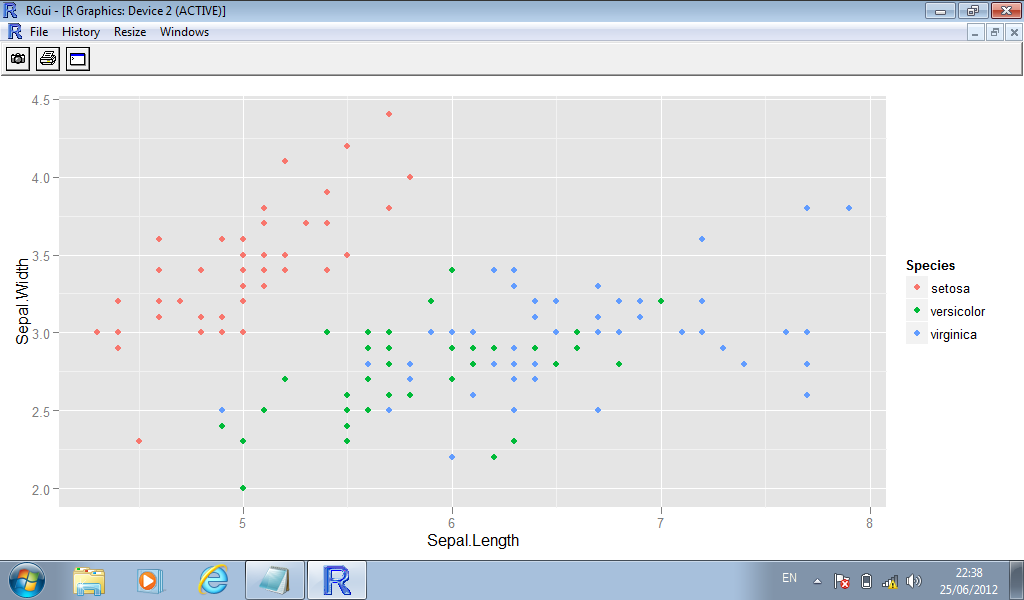
* ddply() breaks up a data-frame based on row values.
* subset()
* transform()
* colwise()
* numcolwise()

## Chapter 10 Reducing Duplication

|  |
| --- |
| 10.1 Introduction  10.2 Iteration  10.3 Plot templates  10.4 Plot functions |

# Scatterplot (ggplot)

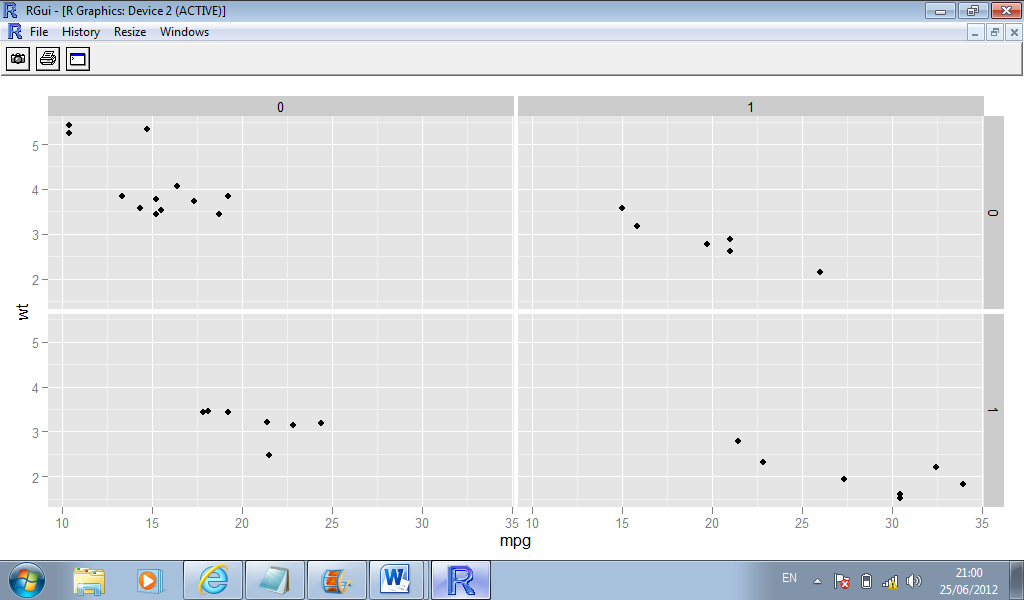
|  |
| --- |
| ggplot(iris,aes(Sepal.Length,Sepal.Width))+geom\_point(aes(color=Species)) |



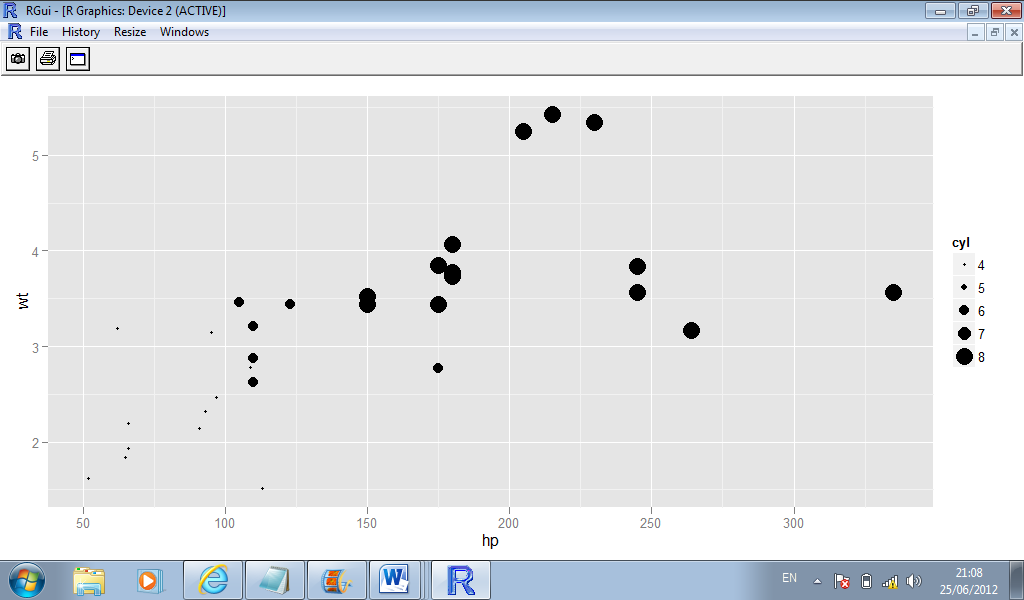
|  |
| --- |
| ggplot(iris,aes(Sepal.Length,Sepal.Width))+  geom\_point(aes(color=Species))+  scale\_x\_continuous(limits=c(5,7),breaks=c(5:7))  #Warning message:  #Removed 34 rows containing missing values (geom\_point). |

# Faceting (qplot)

|  |
| --- |
| >qplot(mpg, wt, data=mtcars, facets=vs ~ am) |



|  |
| --- |
| qplot(hp, wt, data=mtcars,vsize=cyl) |



# The *ggplot()* Command

* The command ggplot() initializes a ggplot object.
* It can be used to declare the **input data frame** for a graphic.
* It can also be used to **specify** the set of plot **aesthetics** intended to be **common** throughout all subsequent layers (unless specifically overridden).
* The actual plots are built with subsequent commands.
* ggplot() is typically used to construct a plot incrementally, using the + operator to add layers to the existing *ggplot* object.
* This is advantageous in that the code is explicit about which layers are added and the order in which they are added.
* We can define common aesthetics using the *aes* argument for this command.

|  |
| --- |
| df <- data.frame(gp = factor(rep(letters[1:3], each = 10)),  y = rnorm(30))  # Compute sample mean and standard deviation in each group  library(plyr)  ds <- ddply(df, .(gp), summarise, mean = mean(y), sd = sd(y)) |

* Declare the data frame and common aesthetics.
* The summary data frame ds is used to plot larger red points in a second geom\_point() layer.
* If the data = argument is not specified, it uses the declared data frame from ggplot(); ditto for the aesthetics.

|  |
| --- |
| ggplot(df, aes(x = gp, y = y)) +  geom\_point() +  geom\_point(data = ds, aes(y = mean),  colour = 'red', size = 3) |