

# R Graphics

SCS Short Course

March 14, 2008

# Archeology

- Archeological expedition
  - Basic graphics – easy and flexible
  - Lattice (trellis) graphics – powerful but less flexible
  - Rgl – nice 3d but challenging
  - Tons of new stuff: e.g.
    - > install.packages("heplots")
    - > library(heplots)
    - > library(help=heplots)
  - Special purpose: library(car): avp
  - In development: Plot3d.R

# Goals for graphics

Two purposes for graphics:

Exploratory

need something quick that can produce graphs as fast as your mind can grasp them

Presentation

need flexibility to meet publication requirements

# Why 'Lattice=Trellis' graphics

Lattice (trellis) graphics:

developed at Bell Labs in early 80s

Marginality vs Conditional association

Motivation was problem of seeing relevant structures in higher dimensional data, the nub of the problem captured by Simpson's Paradox: two dimensional views of data show *marginal association*

But *conditional association* can be entirely different

Interaction: conditional association can be different for different values of moderator variables

# Smoking and Life Expectancy

Source of demographic data (from the CIA):

<http://www.nationmaster.com/>

Two variables:

- Life Expectancy
- Cigarette Consumption Per Capita

```
> dl = read.csv("http://www.math.yorku.ca/~georges/Data/  
  CigLE.csv")  
> head(dl)
```

# Data

```
> head(dl)
```

	X	Country	Continent	LE	CigCon
1	1	Afghanistan	Asia	35.5	98
2	2	Albania	Europe	61.4	NA
3	3	Algeria	Africa	60.6	1021
4	4	Andorra	Europe	72.2	NA
5	5	Angola	Africa	33.4	571
6	6	Antigua and Barbuda	South America	61.9	NA

```
> tail(dl)
```

	X	Country	Continent	LE	CigCon
184	184	Vietnam	Asia	61.3	NA
185	185	Yemen	Asia	49.3	NA
186	186	Zambia	Africa	34.9	408
187	187	Ethiopia	Africa	NA	87
188	188	Venezuela	South America	NA	1079
189	189	Zimbabwe	Africa	NA	399

# Regression

```
> table( LE = !is.na( dl$LE) , CigCon = !is.na(dl$CigCon))
```

```
      CigCon
LE     FALSE TRUE
FALSE     0    3
TRUE     83   103
```

```
> fit = lm( LE ~ CigCon , dl , na.action = na.omit)
> summary(fit)
```

Call:

```
lm(formula = LE ~ CigCon, data = dl, na.action = na.omit)
```

Residuals:

```
      Min       1Q   Median       3Q      Max
-19.4576  -5.8225   0.8188   5.3636  17.8291
```

Coefficients:

```
              Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.799e+01  1.371e+00  34.995 < 2e-16 ***
CigCon      8.528e-03  9.007e-04   9.468 1.33e-15 ***
```

---

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 8.158 on 101 degrees of freedom
(86 observations deleted due to missingness)
```

```
Multiple R-squared:  0.4702,    Adjusted R-squared:  0.465
```

```
F-statistic: 89.64 on 1 and 101 DF,  p-value: 1.333e-15
```

# Interpreting coefficient

Regression coefficient for CigCon is

0.008529

- An increase in 1 cigarette per year increases Life Expectancy by 0.008529 years
- An increase in 1 cigarette per day increases Life Expectancy by  
 $365 \times 0.008529 = 3.11$  years



All it takes is 3 cigarettes a  
day to add 10 years to your  
life.

# Why graphics!

Always look at your data

The **plot** function:

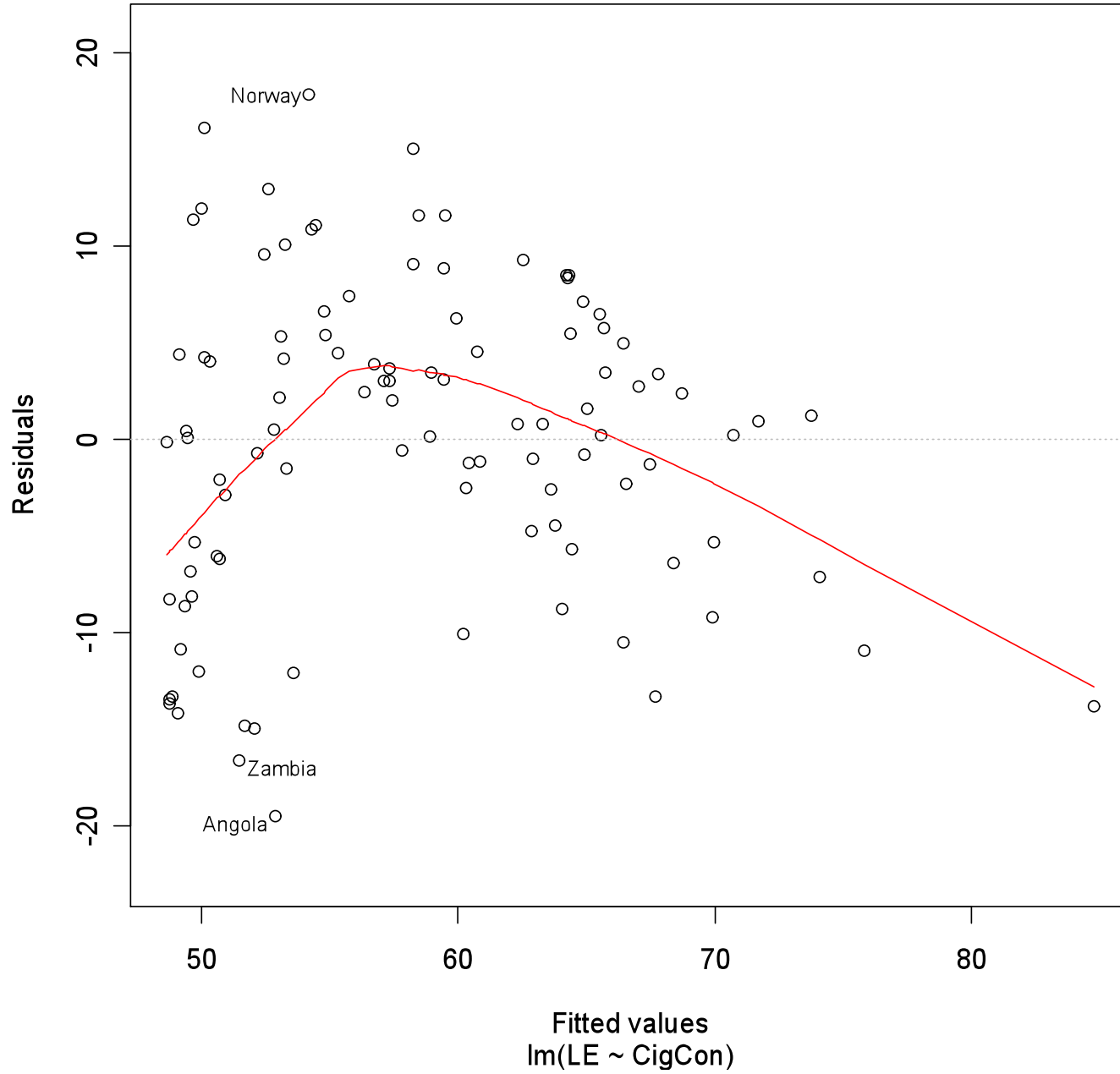
- The Joys of Object-Oriented Programming: plot is a **generic** function so it does different things to different kinds of objects

What does it do to a regression object?

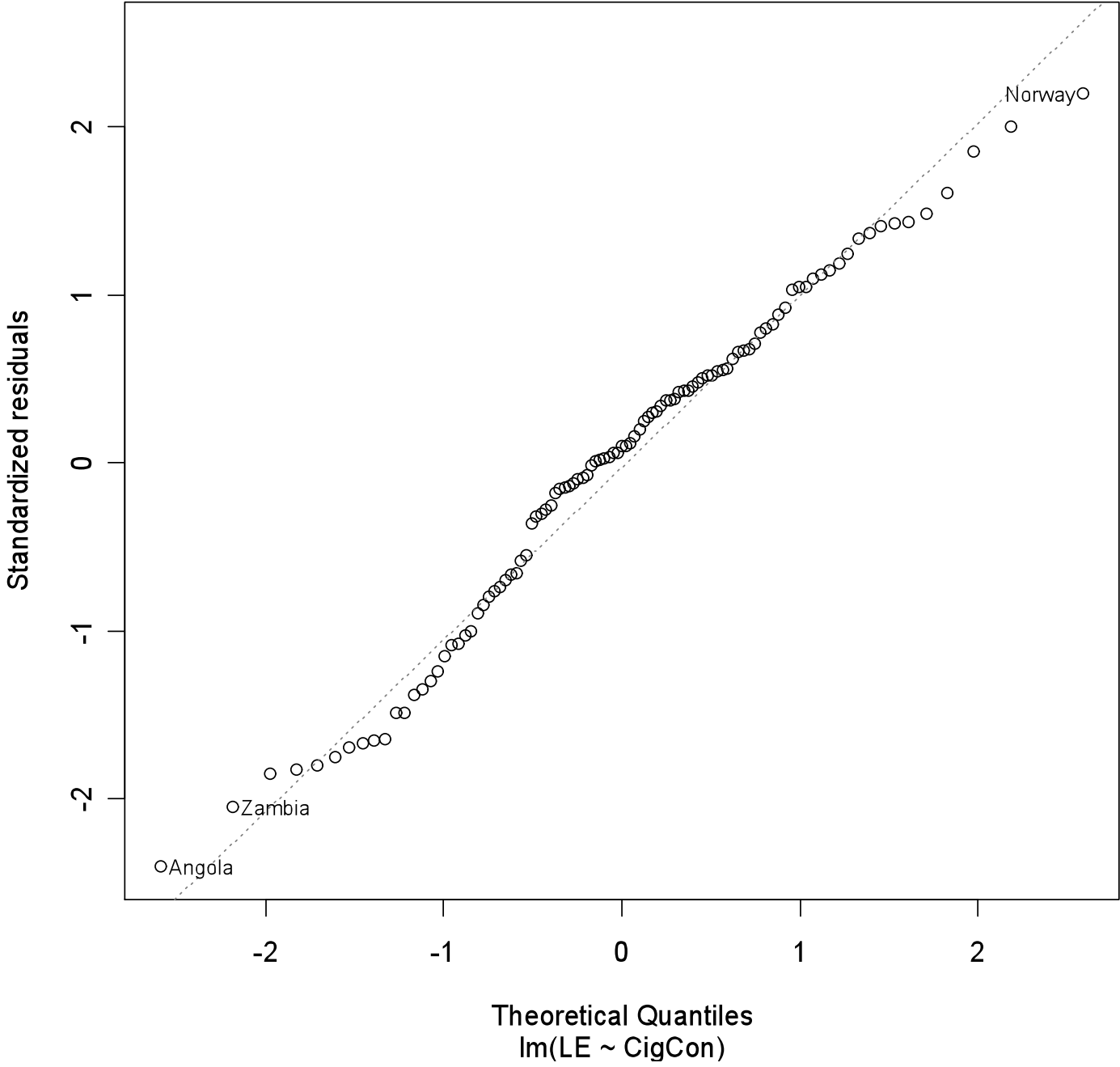
```
> plot ( fit )
```

# Basic regression diagnostics

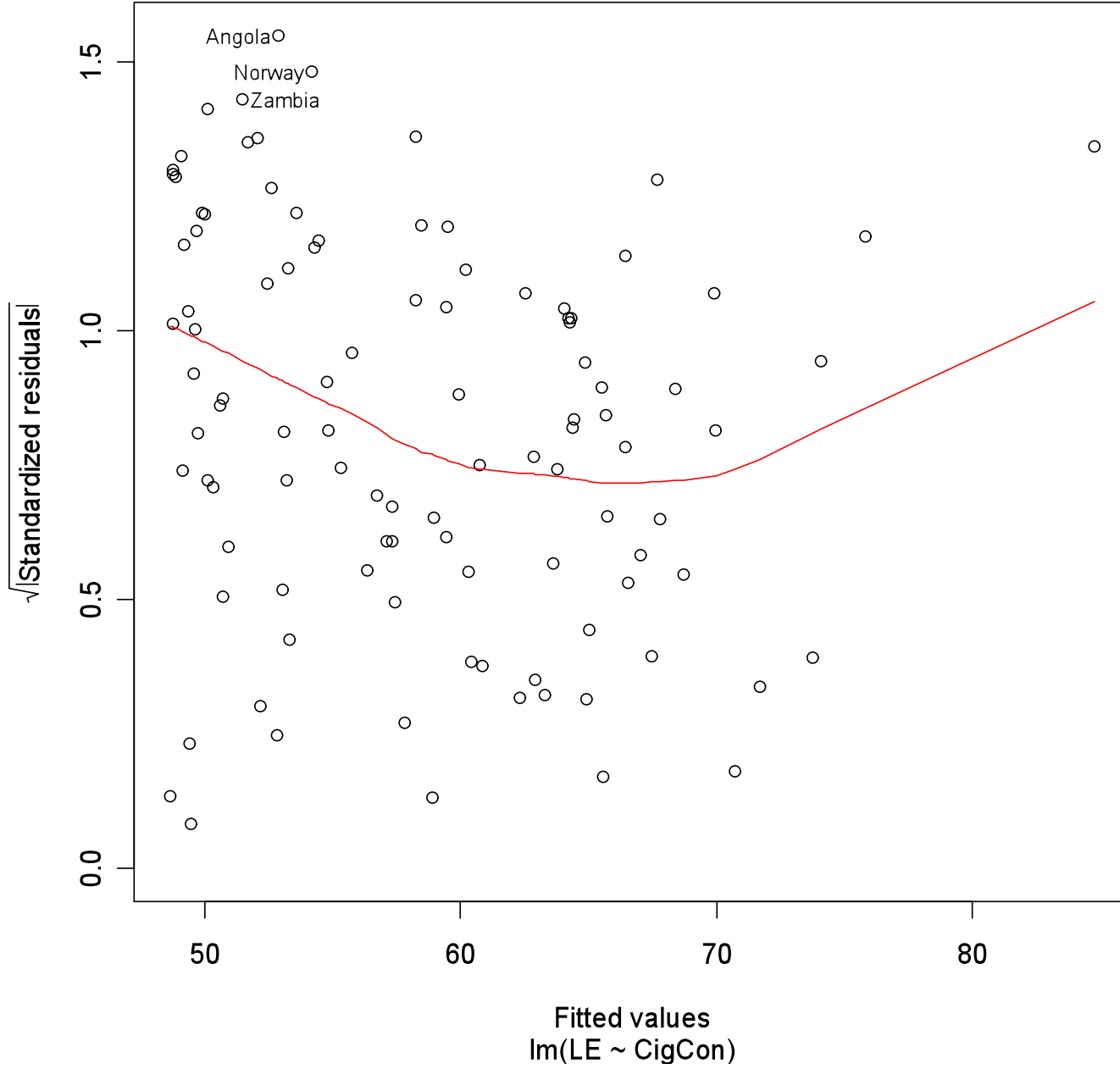
Residuals vs Fitted



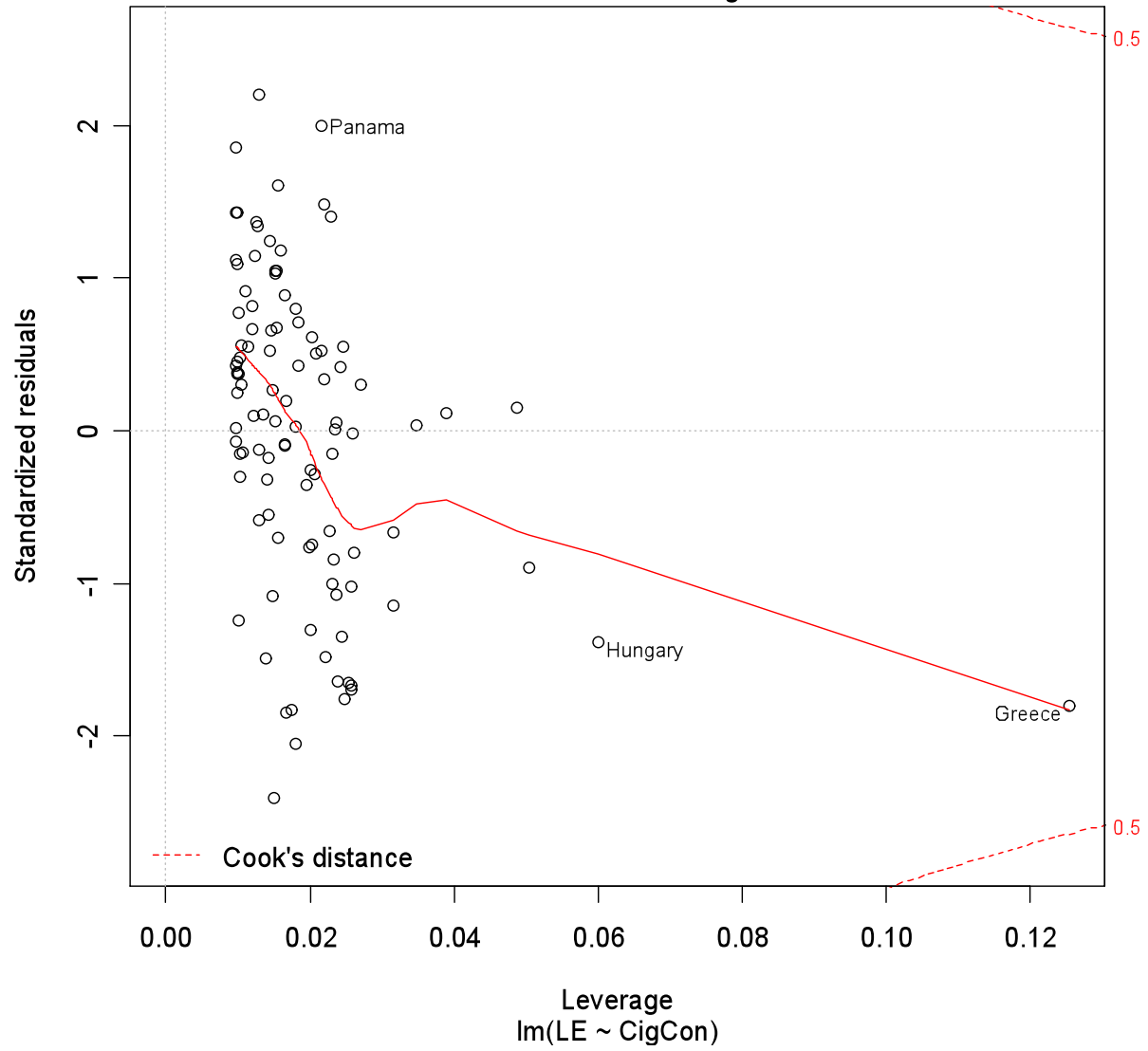
Normal Q-Q



Scale-Location



Residuals vs Leverage



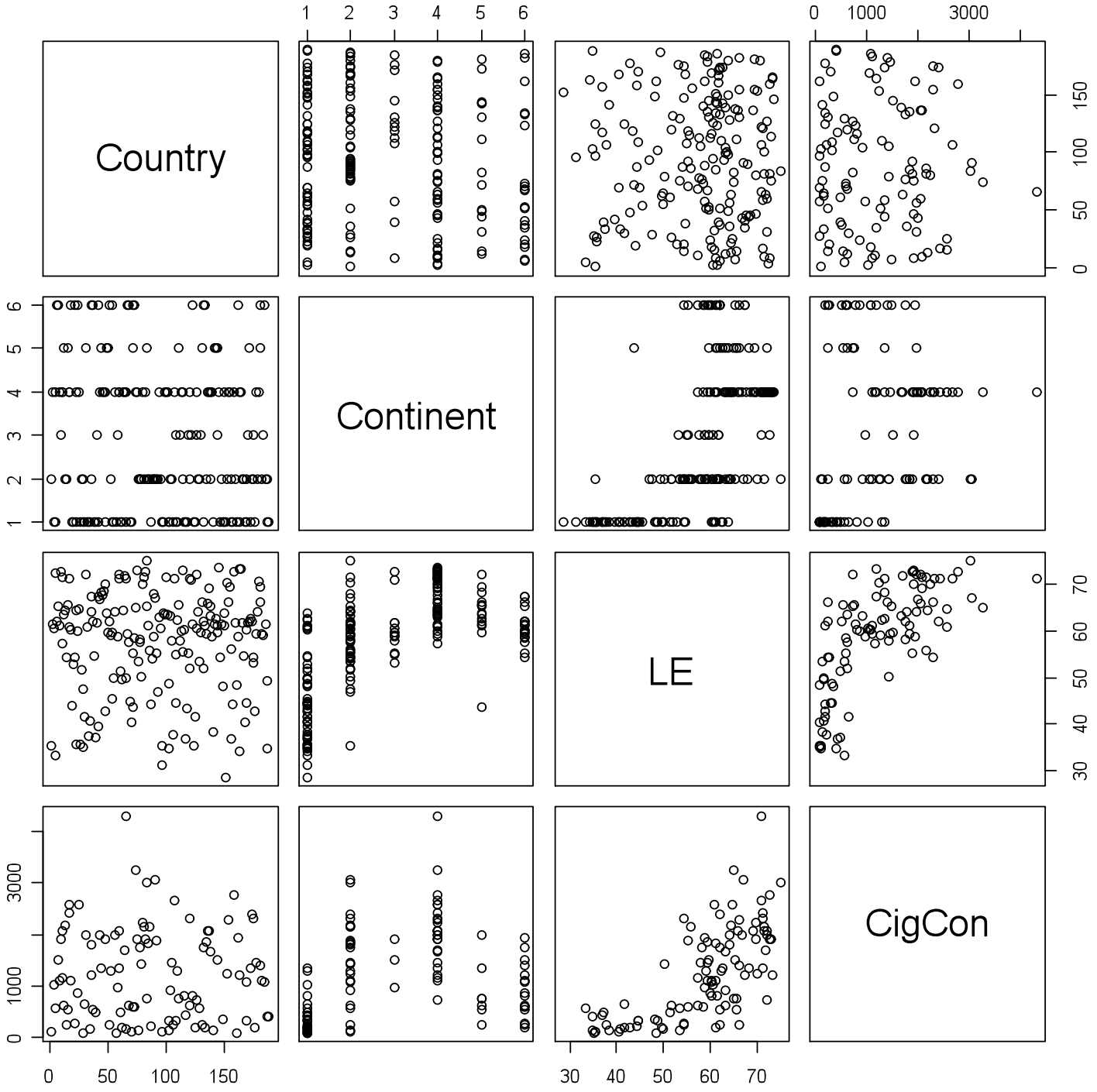
So 'plot' a fitted regression to get some diagnostic plots

But diagnostics not enough to inspire understanding

What happens if we plot a data frame?

```
> plot( dl )
```





# Basic graphics functions in R

Axis	Generic function to add an Axis to a Plot
<b>abline</b>	<b>Add Straight Lines to a Plot</b>
arrows	Add Arrows to a Plot
assocplot	Association Plots
axTicks	Compute Axis Tickmark Locations
axis	Add an Axis to a Plot
axis.POSIXct	Date and Date-time Plotting Functions
<b>barplot</b>	<b>Bar Plots</b>
box	Draw a Box around a Plot
<b>boxplot</b>	<b>Box Plots</b>
bxp	Draw Box Plots from Summaries
cdplot	Conditional Density Plots
contour	Display Contours
coplot	Conditioning Plots
curve	Draw Function Plots
<b>dotchart</b>	<b>Cleveland Dot Plots</b>
filled.contour	Level (Contour) Plots
fourfoldplot	Fourfold Plots
frame	Create / Start a New Plot Frame
graphics-package	The R Graphics Package
grid	Add Grid to a Plot
<b>hist</b>	<b>Histograms</b>
hist.POSIXt	Histogram of a Date or Date-Time Object

# Basic graphics functions in R

**identify**

image

layout

**legend**

**lines**

**locator**

**matplot**

**mosaicplot**

mtext

**pairs**

panel.smooth

**par**

persp

pie

**Identify Points in a Scatter Plot**

Display a Color Image

Specifying Complex Plot Arrangements

**Add Legends to Plots**

**Add Connected Line Segments to a Plot**

**Graphical Input**

**Plot Columns of Matrices**

**Mosaic Plots**

Write Text into the Margins of a Plot

**Scatterplot Matrices**

Simple Panel Plot

**Set or Query Graphical Parameters**

Perspective Plots

Pie Charts

# Basic graphics functions in R

Generic function *plot* and its methods:

<code>plot</code>	<b>Generic X-Y Plotting</b>
<code>plot.data.frame</code>	<b>Plot Method for Data Frames</b>
<code>plot.default</code>	<b>The Default Scatterplot Function</b>
<code>plot.design</code>	Plot Univariate Effects of a 'Design' or Model
<code>plot.factor</code>	<b>Plotting Factor Variables</b>
<code>plot.formula</code>	<b>Formula Notation for Scatterplots</b>
<code>plot.histogram</code>	<b>Plot Histograms</b>
<code>plot.table</code>	<b>Plot Methods for 'table' Objects</b>
<code>plot.window</code>	Set up World Coordinates for Graphics Window
<code>plot.xy</code>	Basic Internal Plot Function

# Basic graphics functions in R

## **points**

polygon

rect

rug

screen

segments

spineplot

stars

stem

## **stripchart**

strwidth

sunflowerplot

symbols

## **text**

title

xinch

## **Add Points to a Plot**

Polygon Drawing

Draw One or More Rectangles

Add a Rug to a Plot

Creating and Controlling Multiple Screens on a Single Device

Add Line Segments to a Plot

Spine Plots and Spinograms

Star (Spider/Radar) Plots and Segment Diagrams

Stem-and-Leaf Plots

## **1-D Scatter Plots**

Plotting Dimensions of Character Strings and Math Expressions

Produce a Sunflower Scatter Plot

Draw Symbols (Circles, Squares, Stars, Thermometers, Boxplots) on a Plot

## **Add Text to a Plot**

Plot Annotation

Graphical Units

Continue with example script file:

From R:

```
> download.file(http://  
www.math.yorku.ca/~georges/R/R-Graphics.R,  
"R-Graphics.R")
```

Then load in R via “File | Open script ...”

# Links

Local wiki: (contact [georges@yorku.ca](mailto:georges@yorku.ca) for account to edit)

<http://wiki.math.yorku.ca>

navigate to R Index

R Graphics: <http://csg.sph.umich.edu/docs/R/graphics-1.pdf>

Gallery of R Graphics:

<http://addictedtor.free.fr/graphiques/>

- very fancy, it would be nice to have an ordinary version

R Graphics Gallery:

- <http://research.stowers-institute.org/efg/R/>