Introduction to R

Demo session

Arni Magnusson

Hafro, 8 Nov 2010

Outline

1. Import - get data into R
2. Explore - plot and summarize
3. Analyze - fit a regression model
4. Export - embed results into a report
5. Philosophy - repeatable and scalable

Outline

Save spreadsheet as text file

Download `cars.xls` and open in Excel/OpenOffice

Save as `cars.txt` (tab-separated) and view in editor

Read text file into R

Open R and type:

```
read.table("c:/shop/cars.txt", header=T)
```

```
cars <- read.table("c:/shop/cars.txt", header=T)
```

Check if data look OK

```
cars
```

```
head(cars)
```

Outline

Scatterplot

```
plot(cars$speed, cars$dist)
```

Range, median, mean, etc.

```
summary(cars)
```

Outline

Fit linear regression model

```
cars.lm <- lm(dist ~ speed, data=cars)
```

```
abline(cars.lm)
```

Outline

Paste model summary

```
summary(cars.lm)
```

Copy from R and paste into Word/OpenOffice document

Write plot to PNG file

```
png("c:/shop/cars.png")

plot(dist ~ speed, data=cars)

abline(cars.lm)

dev.off()
```

Insert `cars.png` into document

Save command history

```
history()
```

Save the command history as `cars.R`

To make this a proper script, you could:

- delete all unnecessary lines
- add comments
- make sure the whole script runs without errors

Outline

Repeatable analysis

If you send the input files and script to colleagues, they can repeat the analysis

- foundation of the scientific method

Easy to make changes and redo the entire analysis

- add one year of data
- explore different modelling choices
- respond to a reviewer's comment

Easy to repeat the analysis for many datasets