Statistics One

Lecture 6

Correlations in R

Two segments

- Scatterplots and correlations in R
- Test/re-test reliability analysis in R

Lecture 6
Segment 1

Scatterplots and correlations in R

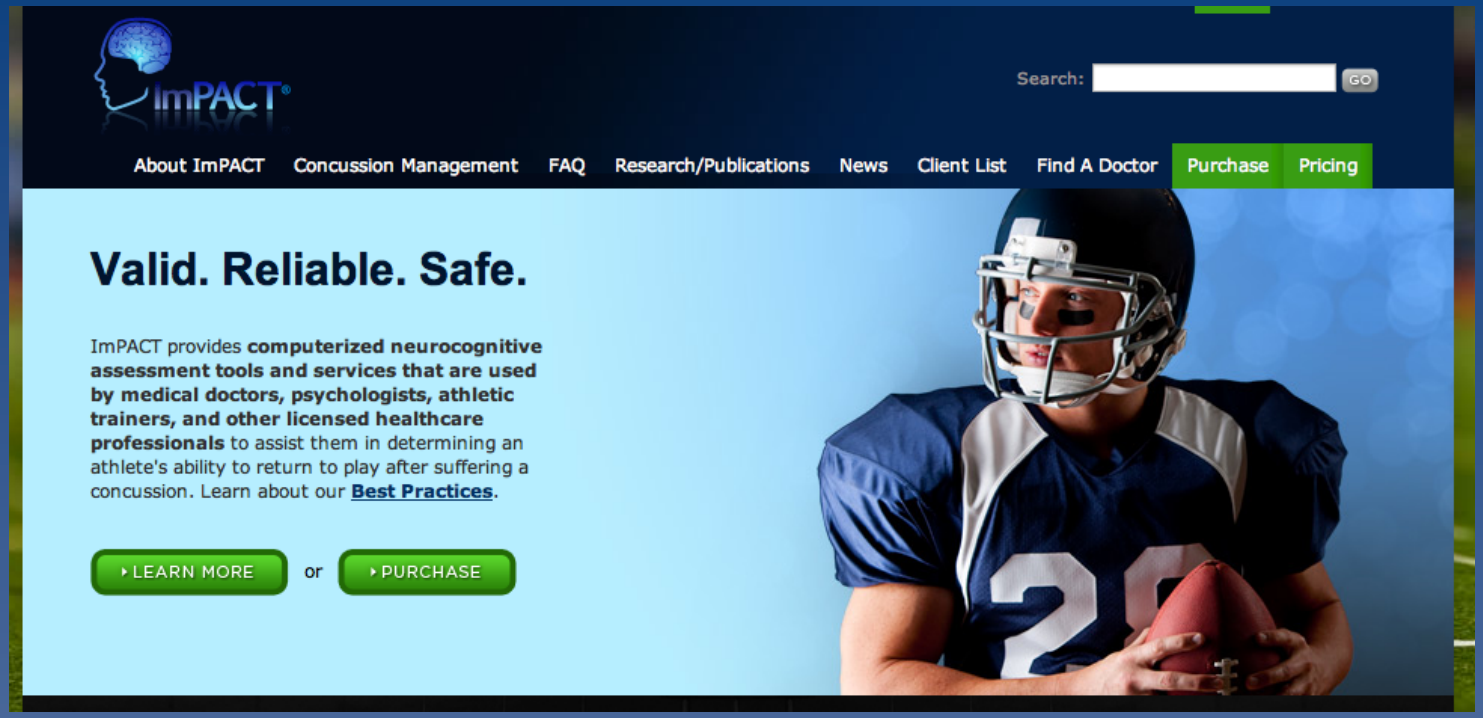
Goal

- Write a script in R
  - Histograms
  - Descriptive statistics
  - Scatterplots
  - Correlations

Example

- Data from ImPACT
  - A computerized neuropsychological assessment of memory and attention
  - Used to assess the cognitive effects of head trauma, for example, sports-related concussion

ImPACT website



ImPACT main measures

- Verbal memory
- Visual memory
- Visual motor speed
- Reaction time
- Impulse control

ImPACT data

- Data are available in the following file:
  - STATS1.EX.02.TXT

Write a script

- First line(s) of code should be comments

  # Statistics One, Lecture 6, example script

  # Read data, plot histograms, get descriptives, examine scatterplots, run correlations

Write a script

- Read data into a dataframe called “impact”

  impact <- read.table(“STATS1.EX.02.TXT”, header = T)

Write a script

- Explore the contents of the dataframe

  class(impact)

    R will return “data.frame”

Write a script

- Explore the contents of the dataframe

  names(impact)

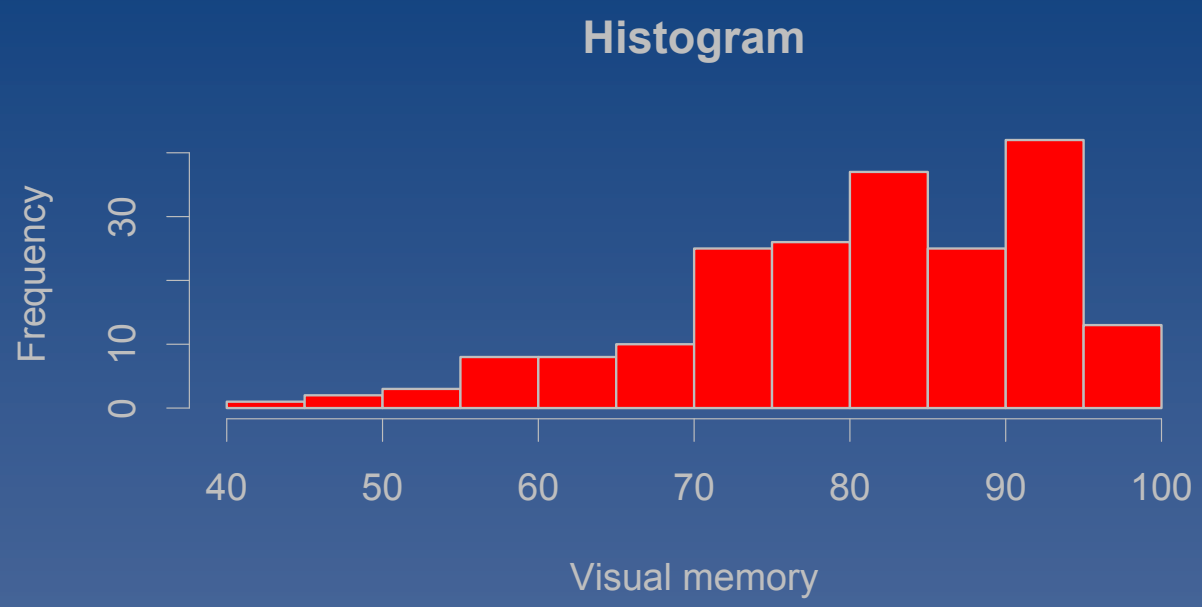
    R will return names of variables

Write a script

- Plot histograms

```
hist(impact$memory.visual, xlab = "Visual memory", main =
"Histogram", col = "red”)
```

Visual memory distribution



Write a script

- Get descriptive statistics

  describe(impact)

Descriptive statistics

```
> describe(impact)
                 var   n  mean    sd median trimmed   mad   min    max range  skew kurtosis   se
memory.verbal      1 200 91.27  9.20  94.00   92.71  7.41 58.00 100.00 42.00 -1.26     1.19 0.65
memory.visual      2 200 81.20 11.75  83.00   82.34 11.86 41.00 100.00 59.00 -0.80     0.29 0.83
speed.vismotor     3 200 45.56  6.25  47.04   46.25  5.74 19.10  54.50 35.40 -1.03     1.08 0.44
speed.general      4 200  0.54  0.08   0.53    0.54  0.06  0.41   0.98  0.57  1.77     6.33 0.01
impulse.control    5 200  4.64  3.56   4.00    4.17  2.97  0.00  24.00 24.00  1.70     4.93 0.25
```

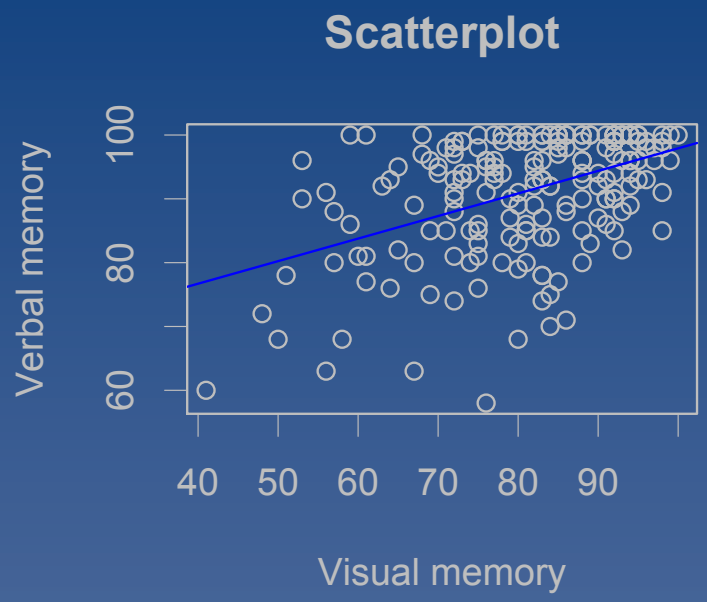
Write a script

- Scatterplots

```
plot(impact$memory.verbal ~ impact$memory.visual, main =
"Scatterplot", ylab = "Verbal memory", xlab = "Visual memory”)


abline(lm(impact$memory.verbal ~ impact$memory.visual),
col="blue”)
```

Verbal by visual memory



Write a script

- Correlations
  - One pair at a time

```
cor(impact$memory.verbal, impact$memory.visual)
```

Write a script

- Correlations
  - One pair at a time, testing for significance

```
cor.test(impact$memory.verbal, impact$memory.visual)
```

Verbal and visual memory *r*

```
   Pearson's product-moment correlation

data:  impact$memory.verbal and impact$memory.visual
t = 7.1196, df = 198, p-value = 1.951e-11
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.3336243 0.5554192
sample estimates:
      cor
0.4514681
```

Write a script

- Correlations
  - All in a matrix

```
cor(impact)
```

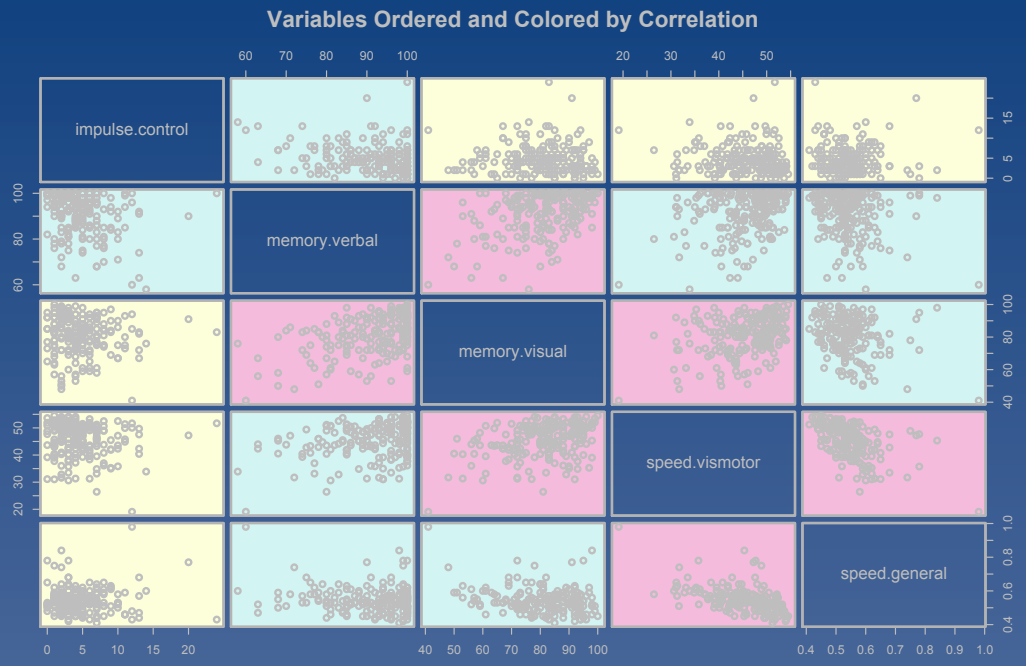
Correlation matrix

```
> # Correlations (all in a matrix)
> cor(impact)
                 memory.verbal memory.visual speed.vismotor speed.general impulse.control
memory.verbal        1.0000000    0.45146807     0.31402025   -0.18140863     -0.28531086
memory.visual        0.4514681    1.00000000     0.37008353   -0.25851977     -0.07425865
speed.vismotor       0.3140202    0.37008353     1.00000000   -0.54662062     -0.09116694
speed.general       -0.1814086   -0.25851977    -0.54662062    1.00000000      0.00739429
impulse.control     -0.2853109   -0.07425865    -0.09116694    0.00739429      1.00000000
```

Write a script

- Fancy scatterplot matrix
  - For code, see final script

Scatterplot matrix



Final script (part 1)

```
# Statistics One, Lecture 6, example script
# Read data, plot historgrams, get descriptives, examine scatterplots, run correlations
library(psych)

# Read the data into a dataframe called impact
impact <- read.table("STATS1.EX.02.TXT", header = T)

# What type of object is impact?
class(impact)

# List the names of the variables in the dataframe called impact
names(impact)
```

Final script (part 2)

```
# Change default settings for graphics
par(cex = 2, lwd = 2, col.axis = 200, col.lab = 200, col.main = 200, col.sub = 200, fg = 200)

# Plot histograms
hist(impact$memory.verbal, xlab = "Verbal memory", main = "Histogram", col = "red")
hist(impact$memory.visual, xlab = "Visual memory", main = "Histogram", col = "red")
hist(impact$speed.vismotor, xlab = "Visual-motor speed", main = "Histogram", col = "red")
hist(impact$speed.general, xlab = "General speed", main = "Histogram", col = "red")
hist(impact$impulse.control, xlab = "Impulse control", main = "Histogram", col = "red")

# Descriptive statistics for the variables in the dataframe called impact
describe(impact)
```

Final script (part 3)

```
# Scatterplots (one pair at a time)
plot(impact$memory.verbal~impact$memory.visual, main = "Scatterplot", ylab = "Verbal memory", xlab = "Visual memory")
abline(lm(impact$memory.verbal~impact$memory.visual), col="blue")

# Correlations (one pair at a time)
cor.test(impact$memory.verbal, impact$memory.visual)

# Correlations (all in a matrix)
cor(impact)
```

Final script (part 4)

```
# Scatterplot matrix
library(gclus)
impact.r = abs(cor(impact))
impact.col = dmat.color(impact.r)
impact.o <- order.single(impact.r)
cpairs(impact, impact.o, panel.colors=impact.col, gap=.5,
       main="Variables Ordered and Colored by Correlation")
```

Image in slide 6 was retrieved from
http://www.impacttest.com/