Visualisation of Complex Social Science Data

Kevin Pulo   and   Rhys Hawkins

kevin.pulo@anu.edu.au   rhys.hawkins@anu.edu.au

Australian National University Supercomputing Facility (ANUSF),
National Computational Infrastructure National Facility (NCI NF),
Canberra, Australia

2009-11-09

ANU Supercomputer Facility

ASSDA

Social Science and Visualisation

- Traditionally: analytically focussed
  - eg. summary statistics, regressions

ANU
Supercomputer
Facility

ASSDA

Social Science and Visualisation

- Traditionally: analytically focussed
  - eg. summary statistics, regressions

- Visualisations: relatively simple
  - eg. pie charts, x-y plots, timeline plots, etc of summary statistics, regression results, etc

Social Science and Visualisation

- Traditionally: analytically focussed
  - eg. summary statistics, regressions

- Visualisations: relatively simple
  - eg. pie charts, x-y plots, timeline plots, etc of summary statistics, regression results, etc

- Visualisation has been under-utilised

Social Science and Visualisation

- Traditionally: analytically focussed
  - eg. summary statistics, regressions

- Visualisations: relatively simple
  - eg. pie charts, x-y plots, timeline plots, etc of summary statistics, regression results, etc

- Visualisation has been under-utilised

- Excellent scope for:
  - detailed, sophisticated advanced visualisations of
  - very large and rich social science datasets
  - eg. 3D, interactive, animated

Why Visualise? Ask Anscombe

Why Visualise? Ask Anscombe

$$\bar{x} = 9.0$$
$$\mathrm{Var}(x) = 10.0$$
$$\bar{y} = 7.5$$
$$\mathrm{Var}(y) = 3.75$$
$$\mathrm{Corr}(x, y) = 0.816$$
$$\mathrm{Regression}: \ y = 3 + 0.5x$$

ANU Supercomputer Facility

ASSDA

Why Visualise? Ask Anscombe

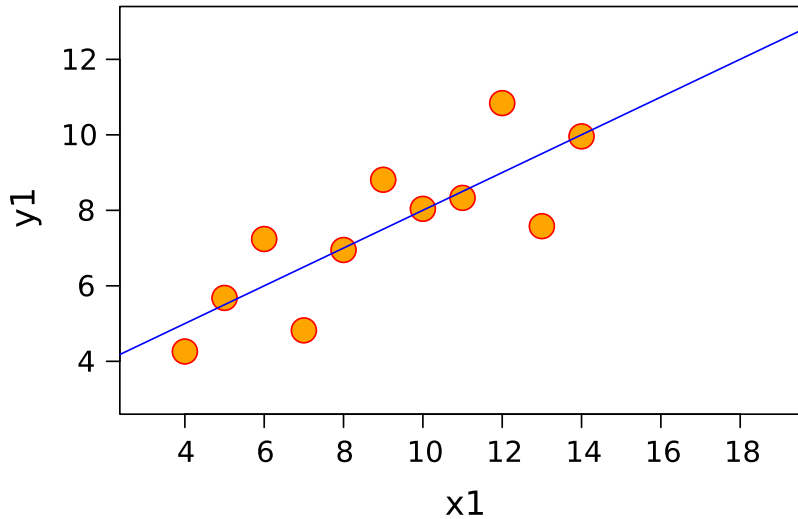|	
$\bar{x} = 9.0$ $\mathrm{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\mathrm{Var}(y) = 3.75$ $\mathrm{Corr}(x, y) = 0.816$ $\text{Regression}: \ y = 3 + 0.5x$	$\bar{x} = 9.0$ $\mathrm{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\mathrm{Var}(y) = 3.75$ $\mathrm{Corr}(x, y) = 0.816$ $\text{Regression}: \ y = 3 + 0.5x$

Why Visualise? Ask Anscombe

$\bar{x} = 9.0$ $\mathrm{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\mathrm{Var}(y) = 3.75$ $\mathrm{Corr}(x, y) = 0.816$ Regression : $y = 3 + 0.5x$	$\bar{x} = 9.0$ $\mathrm{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\mathrm{Var}(y) = 3.75$ $\mathrm{Corr}(x, y) = 0.816$ Regression : $y = 3 + 0.5x$
$\bar{x} = 9.0$ $\mathrm{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\mathrm{Var}(y) = 3.75$ $\mathrm{Corr}(x, y) = 0.816$ Regression : $y = 3 + 0.5x$	

Why Visualise? Ask Anscombe

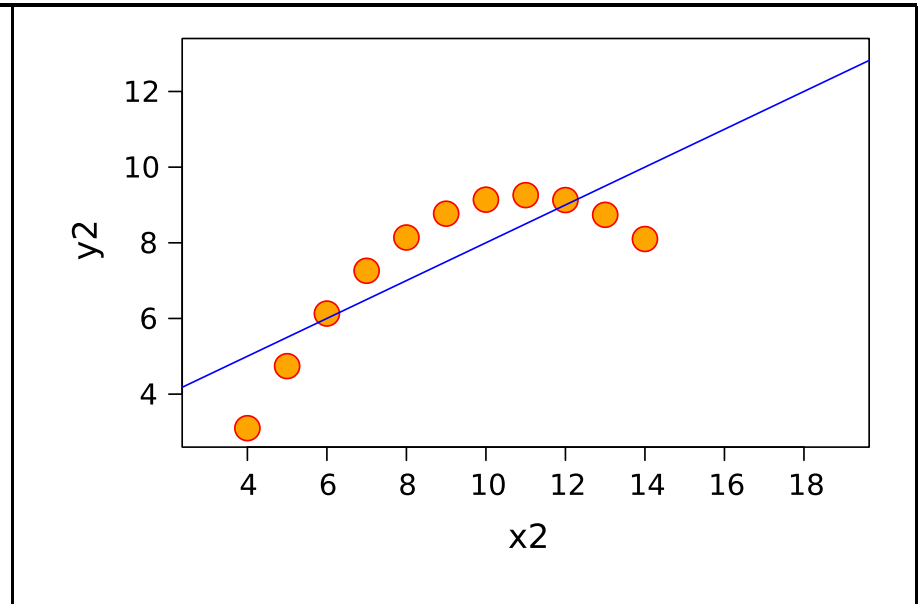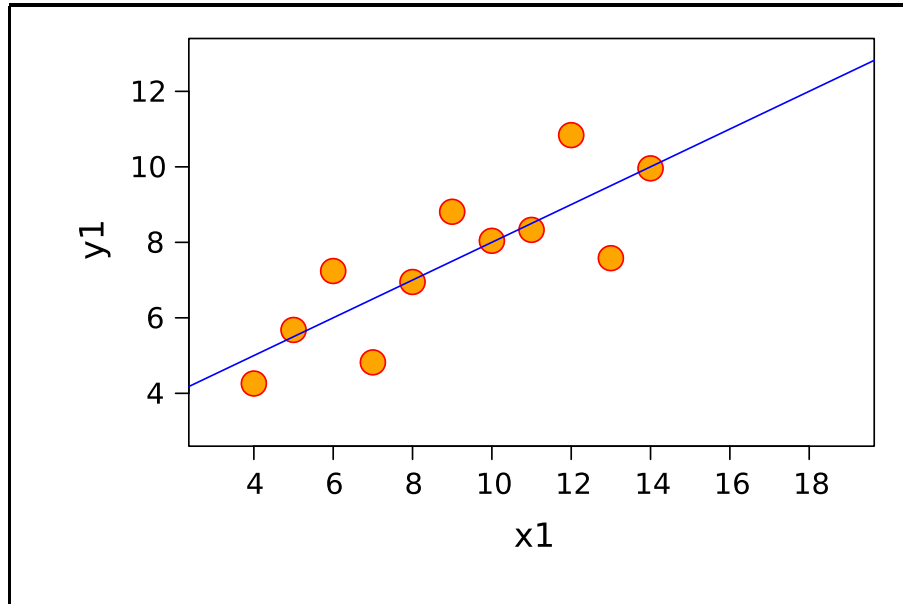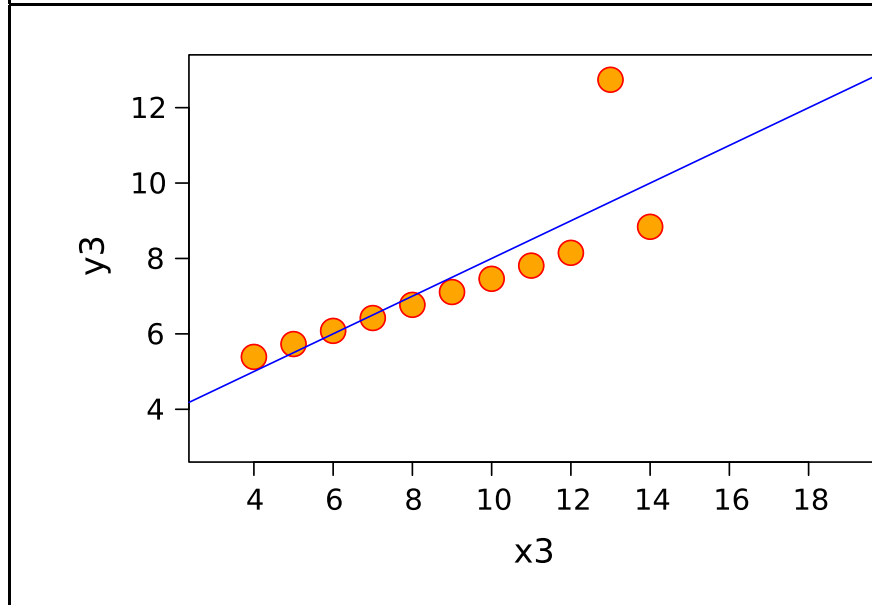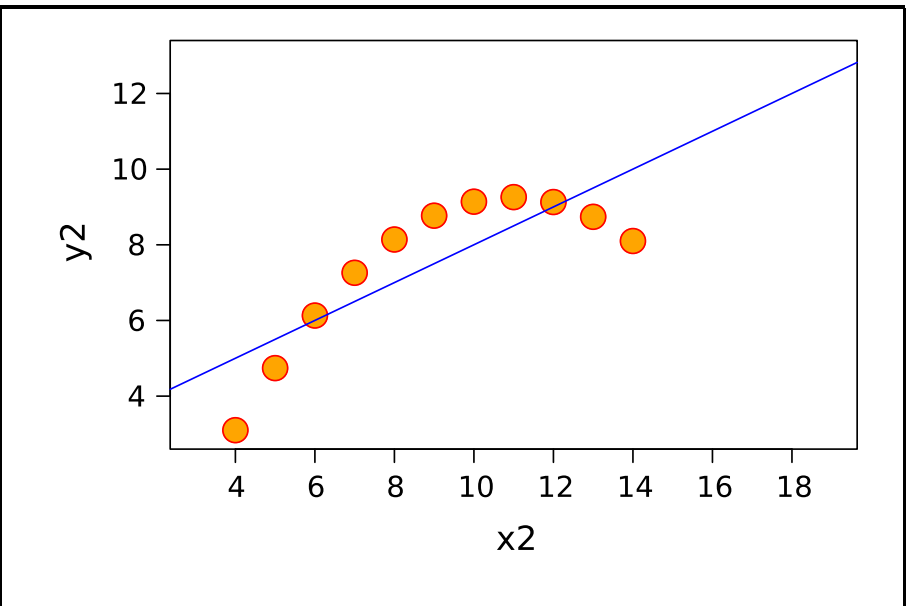|	
$\bar{x} = 9.0$ $\text{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\text{Var}(y) = 3.75$ $\text{Corr}(x, y) = 0.816$ $\text{Regression}: \ y = 3 + 0.5x$	$\bar{x} = 9.0$ $\text{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\text{Var}(y) = 3.75$ $\text{Corr}(x, y) = 0.816$ $\text{Regression}: \ y = 3 + 0.5x$
$\bar{x} = 9.0$ $\text{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\text{Var}(y) = 3.75$ $\text{Corr}(x, y) = 0.816$ $\text{Regression}: \ y = 3 + 0.5x$	$\bar{x} = 9.0$ $\text{Var}(x) = 10.0$ $\bar{y} = 7.5$ $\text{Var}(y) = 3.75$ $\text{Corr}(x, y) = 0.816$ $\text{Regression}: \ y = 3 + 0.5x$

ANU
Supercomputer
Facility

ASSDA

Why Visualise? Ask Anscombe

$$\bar{x} = 9.0$$
$$\mathrm{Var}(x) = 10.0$$
$$\bar{y} = 7.5$$
$$\mathrm{Var}(y) = 3.75$$
$$\mathrm{Corr}(x, y) = 0.816$$
$$\mathrm{Regression}: \ y = 3 + 0.5x$$

$$\bar{x} = 9.0$$
$$\mathrm{Var}(x) = 10.0$$
$$\bar{y} = 7.5$$
$$\mathrm{Var}(y) = 3.75$$
$$\mathrm{Corr}(x, y) = 0.816$$
$$\mathrm{Regression}: \ y = 3 + 0.5x$$

$$\bar{x} = 9.0$$
$$\mathrm{Var}(x) = 10.0$$
$$\bar{y} = 7.5$$
$$\mathrm{Var}(y) = 3.75$$
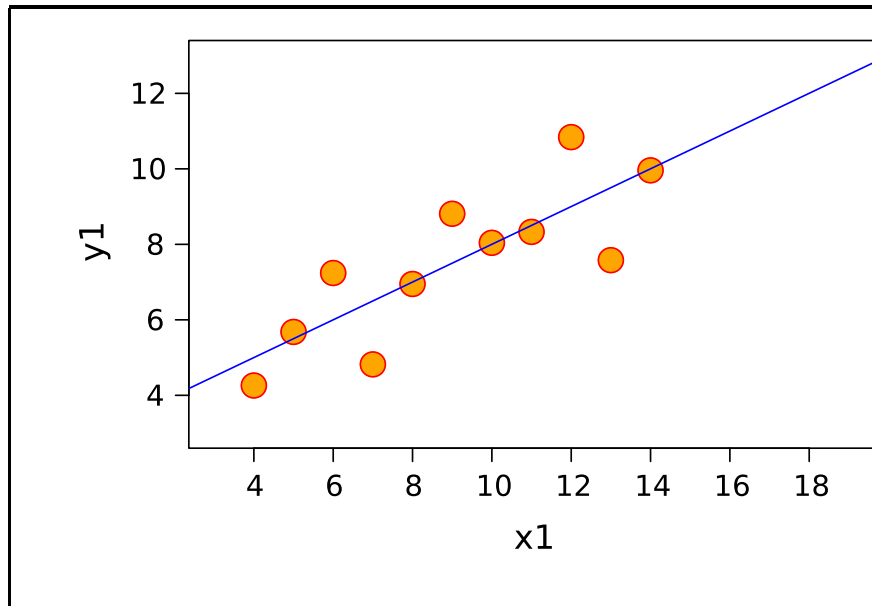$$\mathrm{Corr}(x, y) = 0.816$$
$$\mathrm{Regression}: \ y = 3 + 0.5x$$

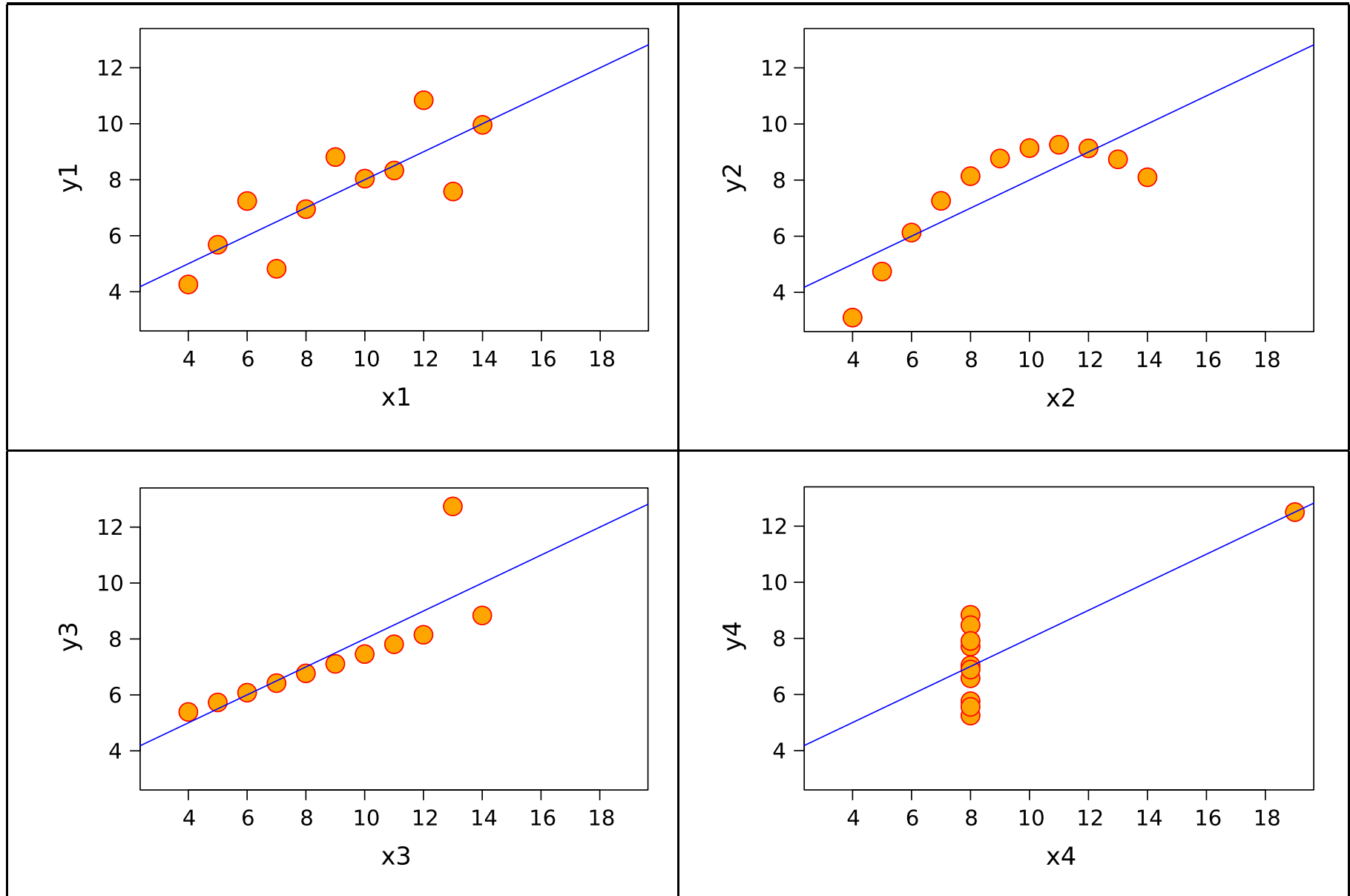# Why Visualise? Ask Anscombe

$$\bar{x} = 9.0$$
$$\text{Var}(x) = 10.0$$
$$\bar{y} = 7.5$$
$$\text{Var}(y) = 3.75$$
$$\text{Corr}(x, y) = 0.816$$
$$\text{Regression}: \ y = 3 + 0.5x$$

$$\bar{x} = 9.0$$
$$\text{Var}(x) = 10.0$$
$$\bar{y} = 7.5$$
$$\text{Var}(y) = 3.75$$
$$\text{Corr}(x, y) = 0.816$$
$$\text{Regression}: \ y = 3 + 0.5x$$

ANU
Supercomputer
Facility

ASSDA

Why Visualise? Ask Anscombe

$$\bar{x} = 9.0$$
$$\mathrm{Var}(x) = 10.0$$
$$\bar{y} = 7.5$$
$$\mathrm{Var}(y) = 3.75$$
$$\mathrm{Corr}(x, y) = 0.816$$
$$\mathrm{Regression}: \ y = 3 + 0.5x$$

ANU
Supercomputer
Facility

ASSDA

Why Visualise? Ask Anscombe

Visualisation supports

- Exploration

ANU
Supercomputer
Facility

ASSDA

Visualisation supports

- Exploration
- Information discovery

ANU
Supercomputer
Facility

ASSDA

Visualisation supports

- Exploration

- Information discovery

- Validation

ANU
Supercomputer
Facility

ASSDA

Visualisation supports

- Exploration
- Information discovery
- Validation
- Cross-disciplinary relationships (ie. mashups)

ANU Supercomputer Facility

ASSDA

Visualisation supports

- Exploration
- Information discovery
- Validation
- Cross-disciplinary relationships (ie. mashups)
- Complementary to traditional analysis

ANU
Supercomputer
Facility

ASSDA