36-315: Statistical Graphics and Visualization

Handout 13
Date: March 3, 2003

Three-dimension plots

3D data does not imply the use of a 3D visualization—are the variables are inherently *spatial* or they are merely *statistical*?

3D scatterplot—Direct analog of 2D scatterplot. Requires special tricks to maintain the illusion of depth: lines to the floor, motion. Does not allow quantitative judgements.

Bubble plot—Scatterplot where symbol size encodes the third dimension (smaller = farther away). Susceptible to overplotting.

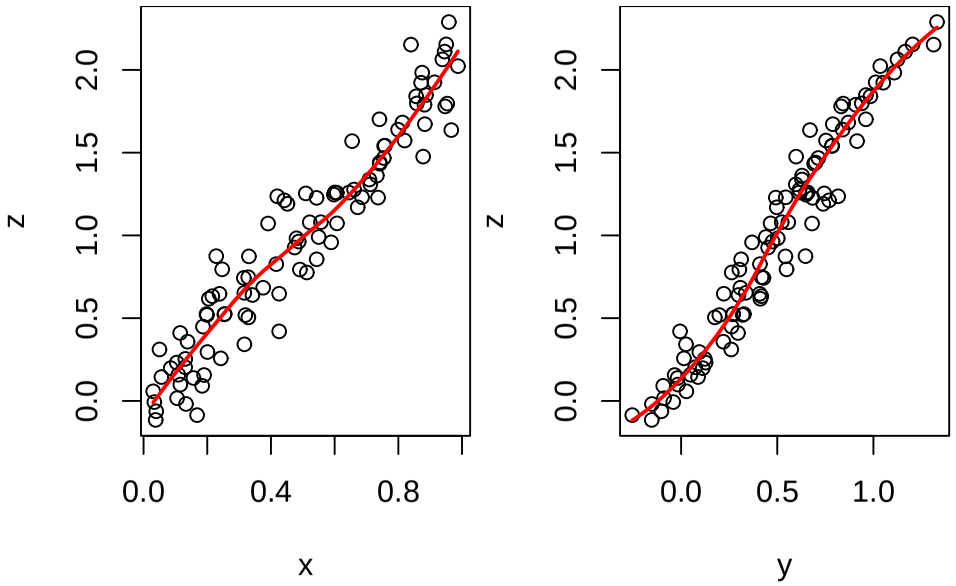
Vane plot—Scatterplot where symbol orientation encodes the third dimension (taller = larger).

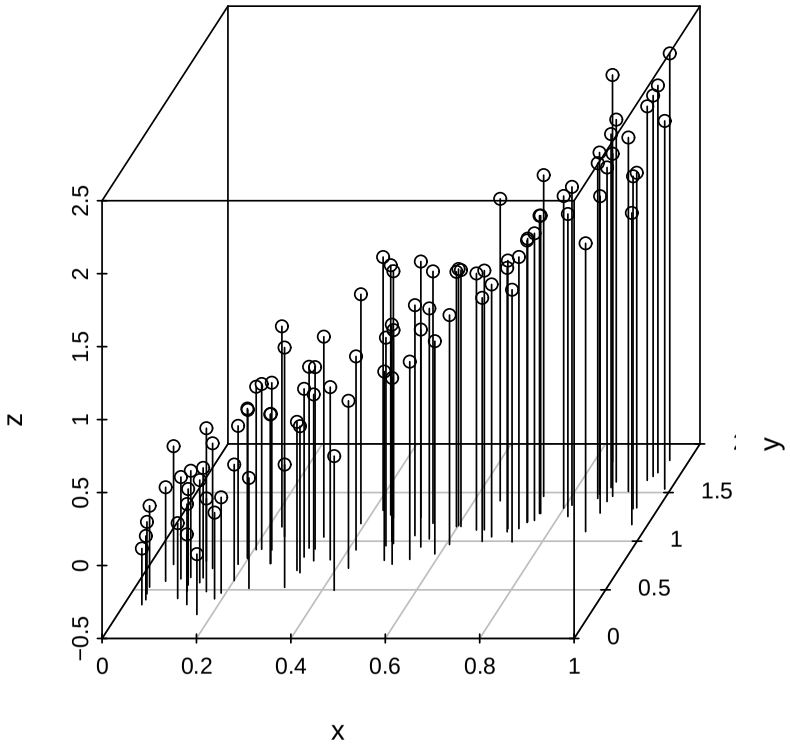
Color plot—Scatterplot where symbol color encodes the third dimension (darker = larger).

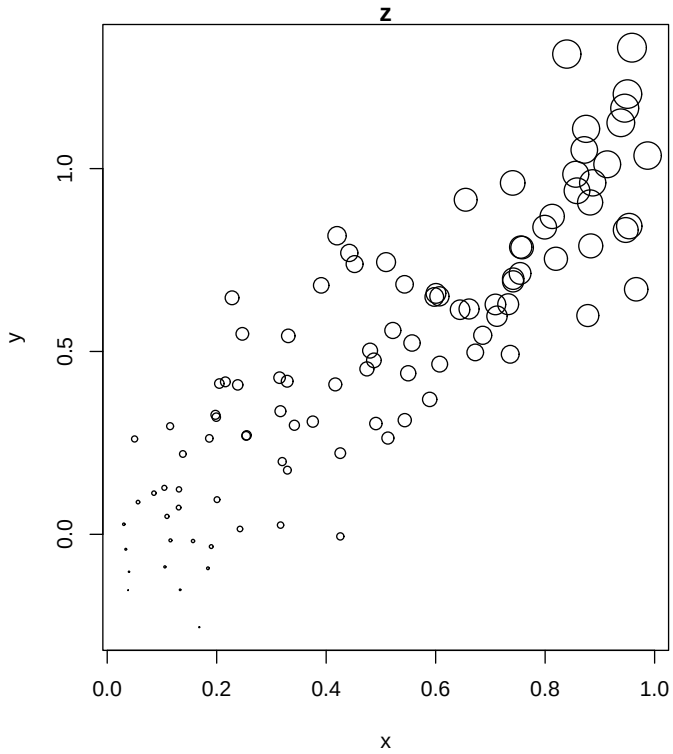
Uses of three-dimension plots:

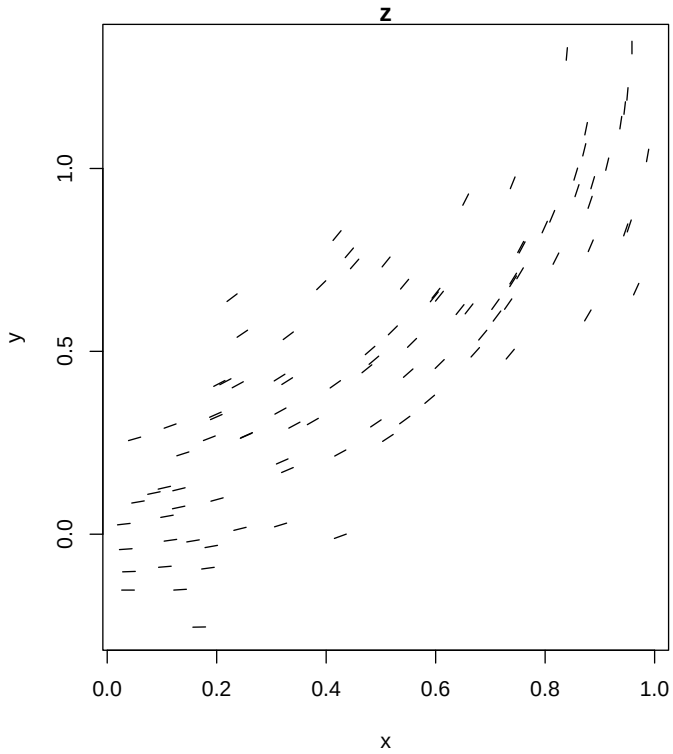
- Explaining the outliers and clusters in a 2D scatterplot. Response is the y axis.
- Determining which of two variables is a stronger predictor of the response. Response is the z axis.

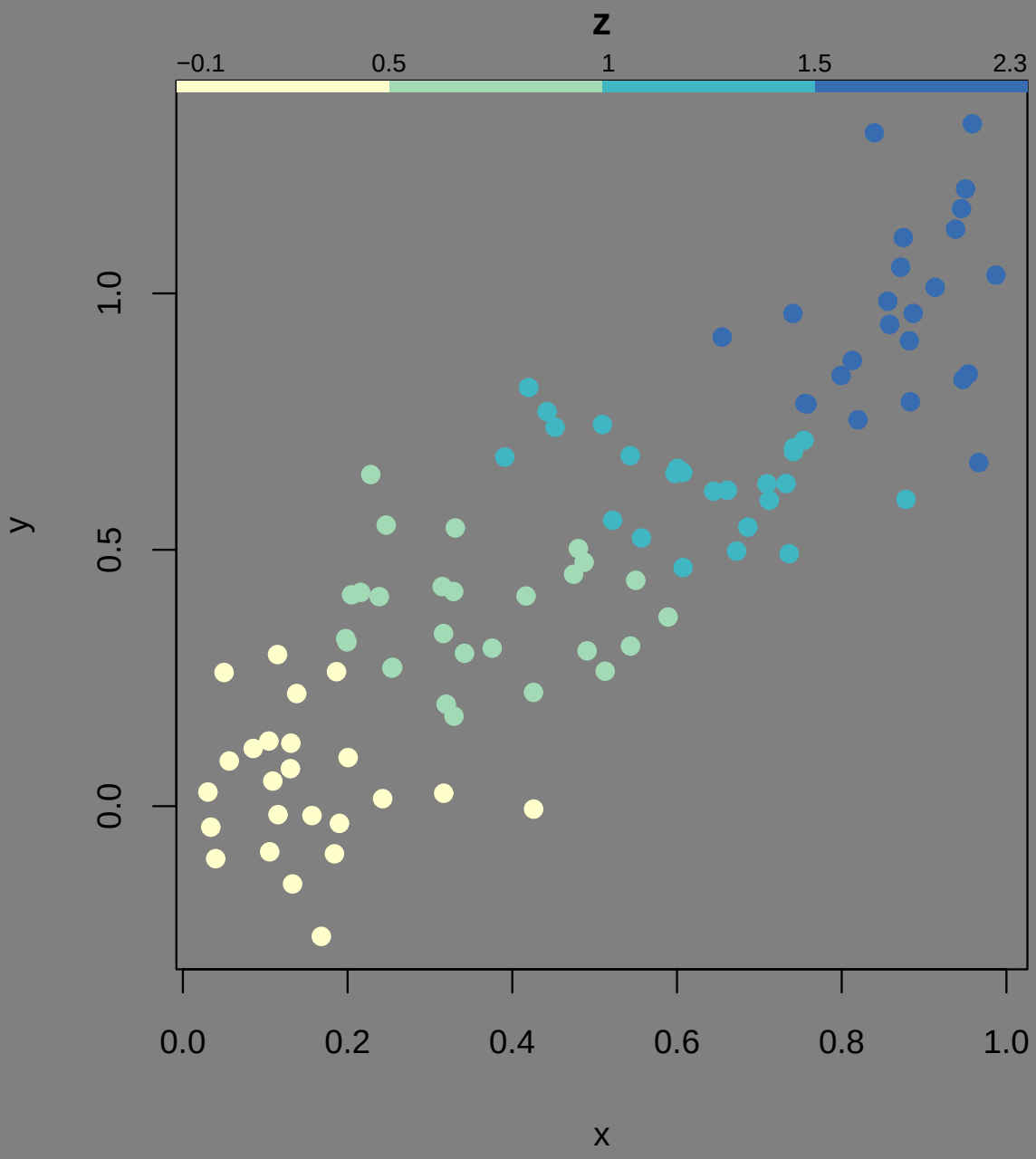
Data set 1:


z
0.0
0.5
1.0
1.5
2.0
0.0
0.4
0.8
x
z
0.0
0.5
1.0
1.5
2.0
0.0
0.5
1.0
y

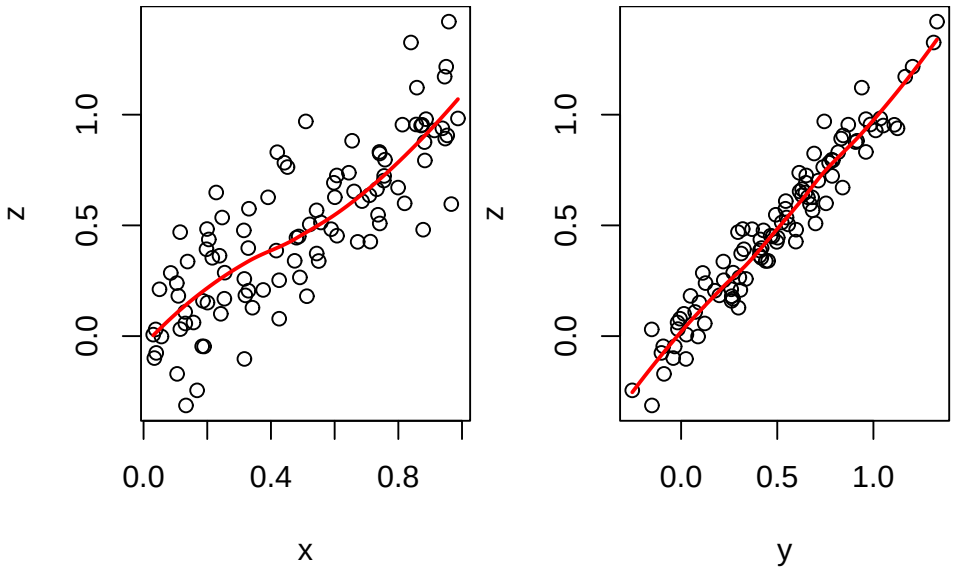

z
−0.5
0
0.5
1
1.5
2
2.5
0
0.2
0.4
0.6
0.8
1
x
y
0
0.5
1
1.5

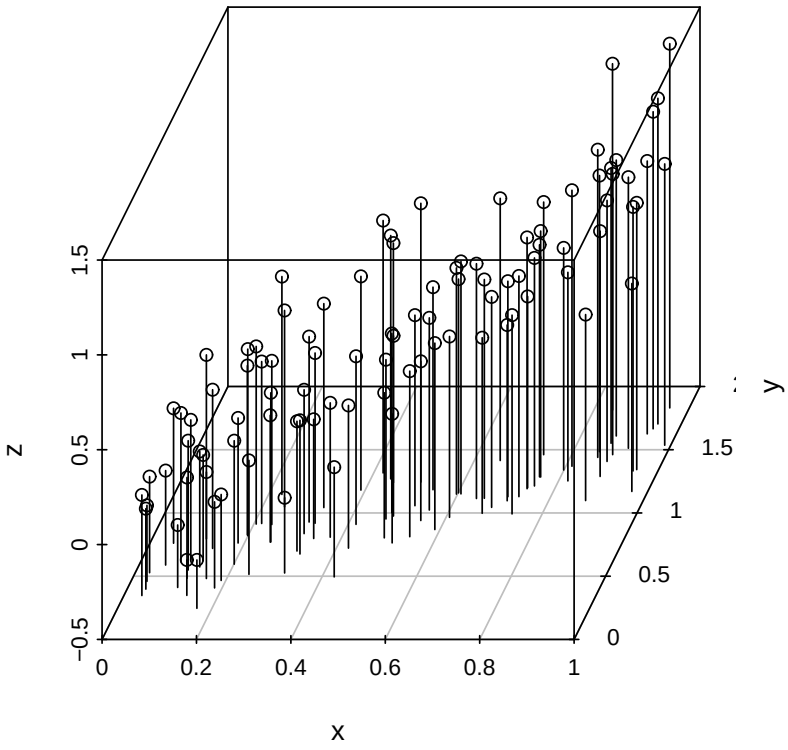

z
y
x
0.0
0.5
1.0
0.0
0.2
0.4
0.6
0.8
1.0

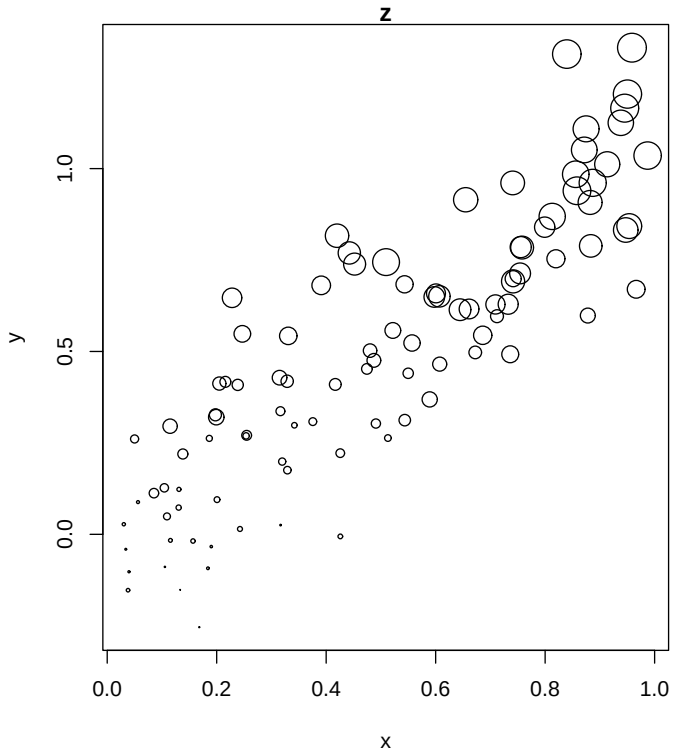

z
y
x
0.0
0.5
1.0
0.0
0.2
0.4
0.6
0.8
1.0

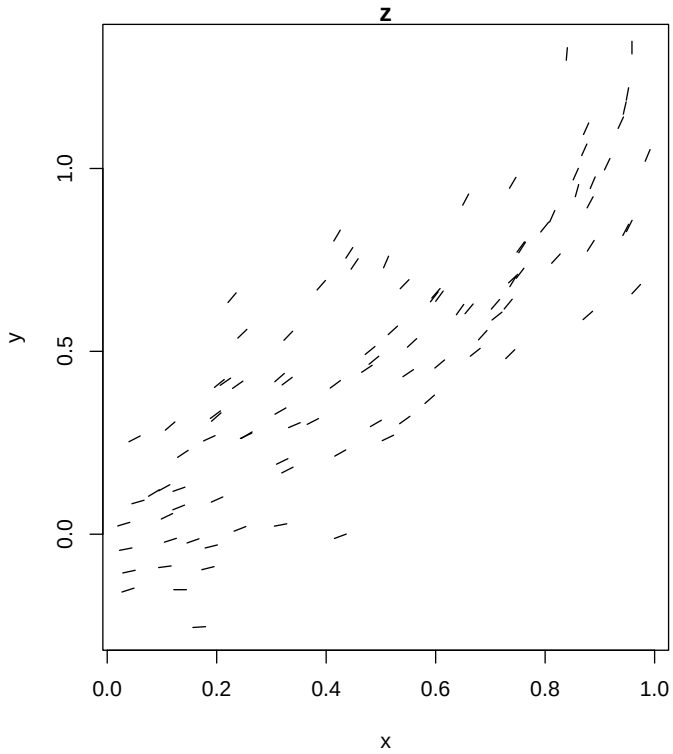

z
−0.1
0.5
1
1.5
2.3
y
0.0
0.5
1.0
0.0
0.2
0.4
0.6
0.8
1.0
x

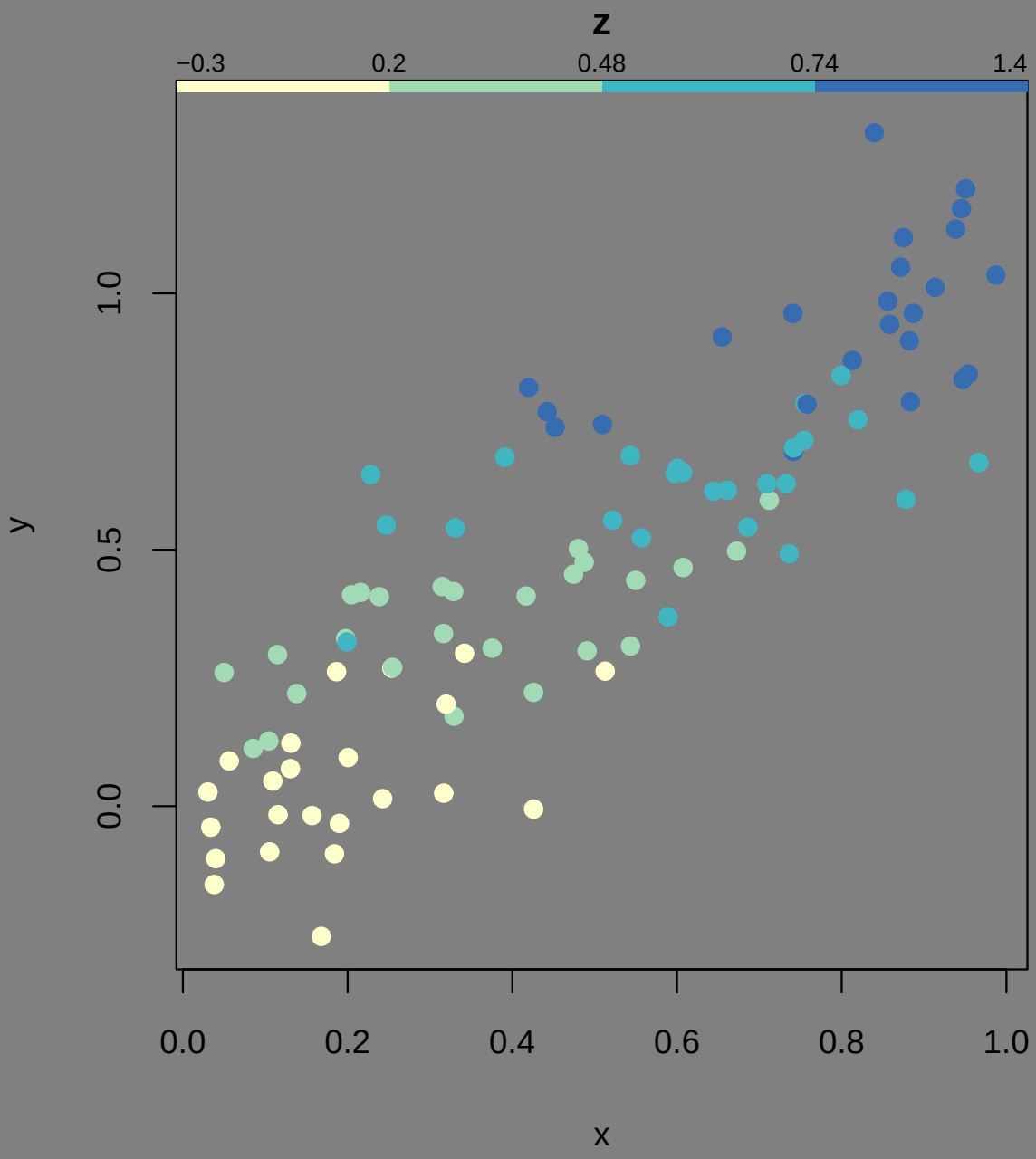
Data set 2:


z
0.0
0.5
1.0
0.0
0.4
0.8
x
z
0.0
0.5
1.0
0.0
0.5
1.0
y

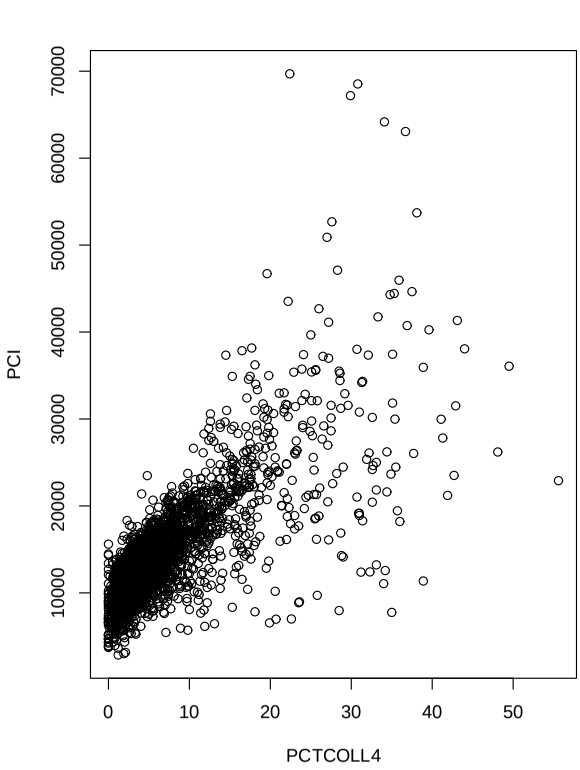

z
−0.5
0
0.5
1
1.5
0
0.2
0.4
0.6
0.8
1
x
y
0
0.5
1
1.5

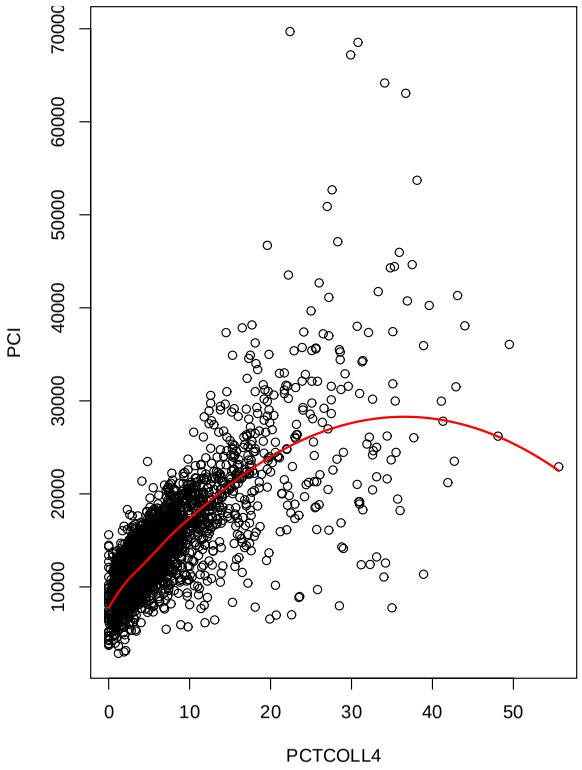

z
y
x
0.0
0.5
1.0
0.0
0.2
0.4
0.6
0.8
1.0

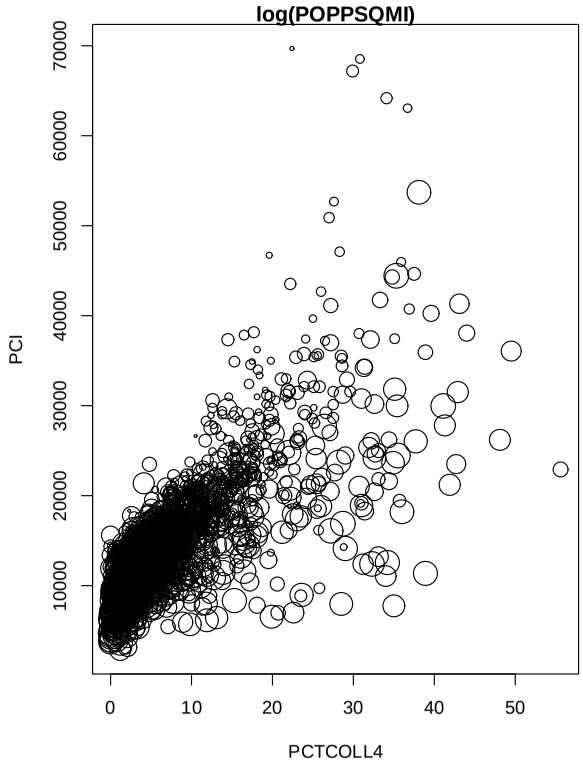

z
y
x
0.0
0.5
1.0
0.0
0.2
0.4
0.6
0.8
1.0

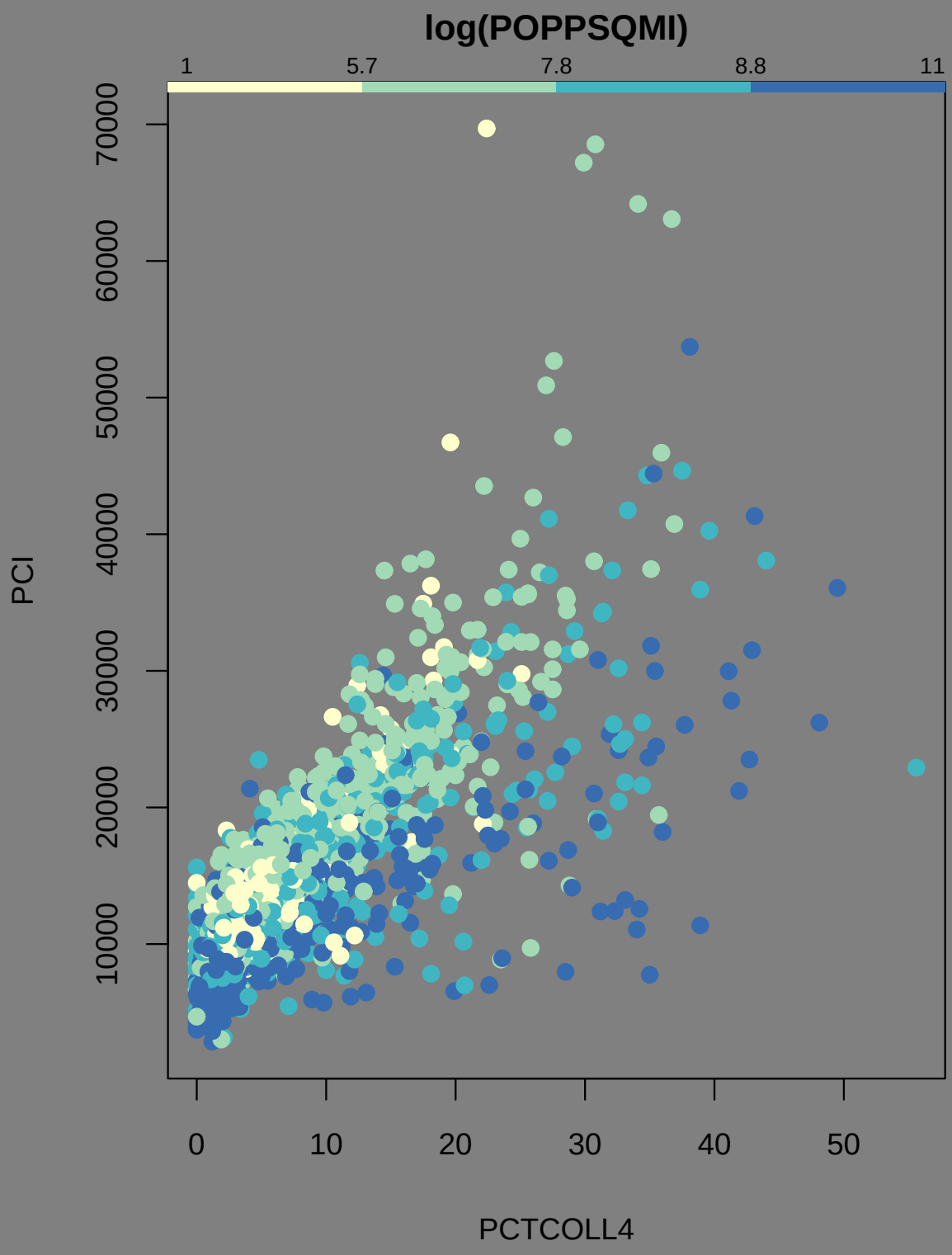

z
−0.3
0.2
0.48
0.74
1.4
y
0.0
0.5
1.0
0.0
0.2
0.4
0.6
0.8
1.0
x

Education vs. Income:

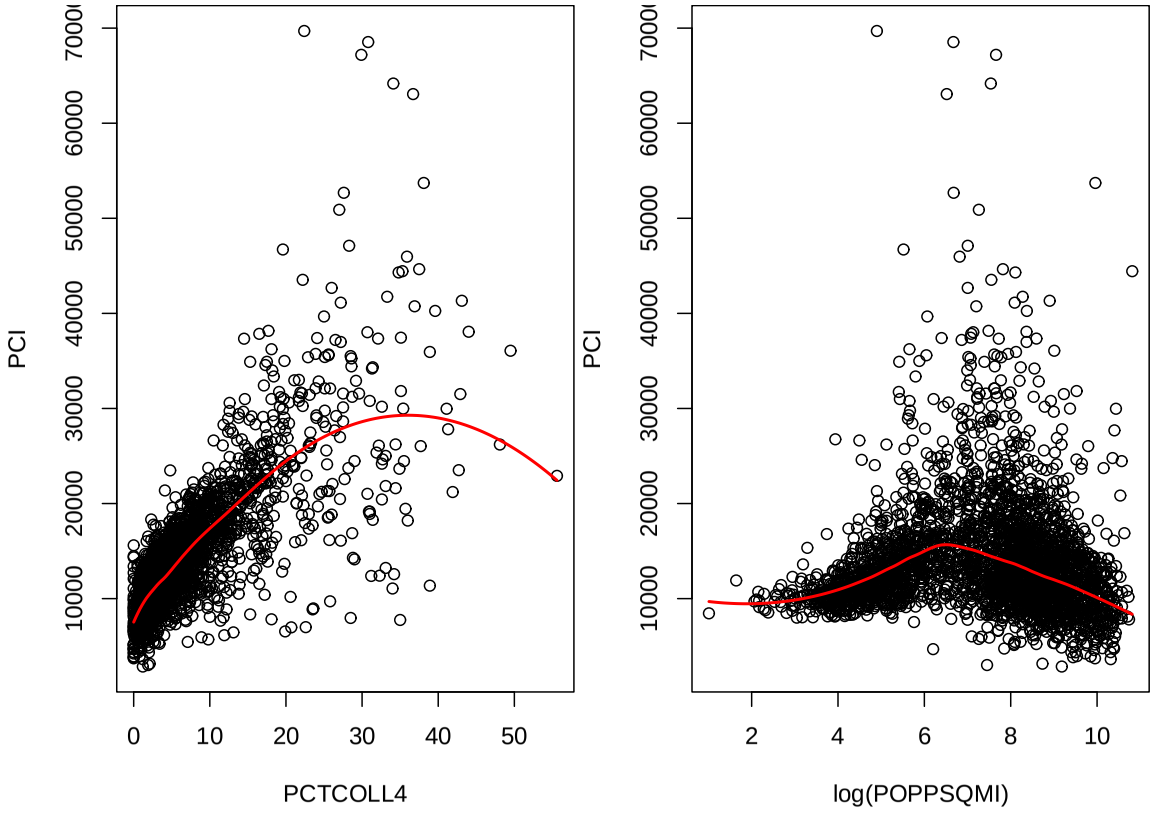


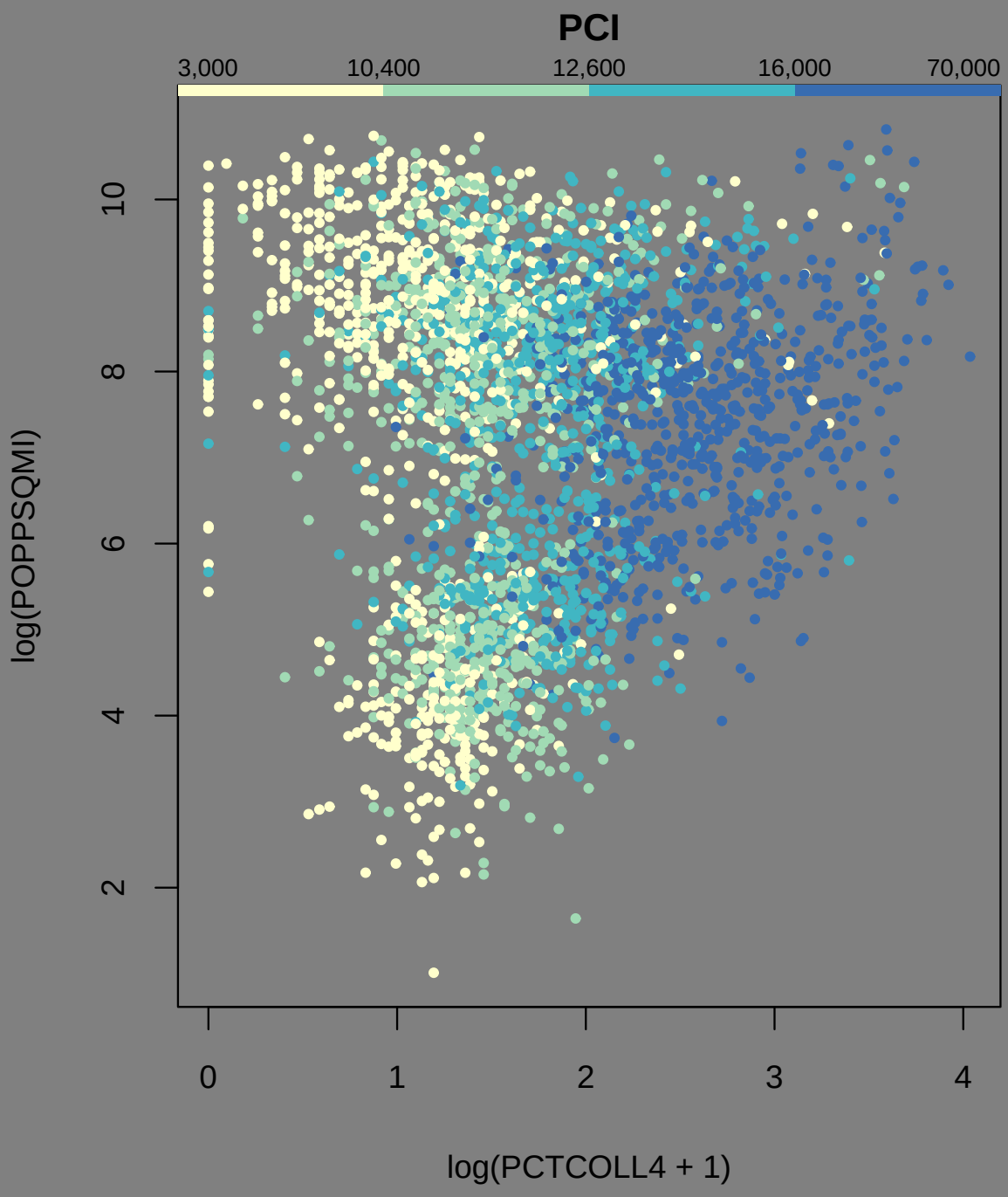





log(POPPSQMI)
1
5.7
7.8
8.8
11
PCI
10000
20000
30000
40000
50000
60000
70000
0
10
20
30
40
50
PCTCOLL4

Which is the better predictor of income?




PCI
3,000
10,400
12,600
16,000
70,000
log(POPPSQMI)
2
4
6
8
10
0
1
2
3
4
log(PCTCOLL4 + 1)