

36-315: Statistical Graphics and Visualization

Handout 9

Date: February 12, 2003

Predictable - y is predictable from x if the mean of y changes with x , when only x is known. Thus x reduces (but doesn't necessarily eliminate) the uncertainty in y . Note that x may or may not be predictable from y .

Trend line/prediction line/regression line - Depicts the mean of y as a function of x . If this line is flat, y is not predictable from x .

Aspect ratio - Height divided by width of the data rectangle, as seen on the page.

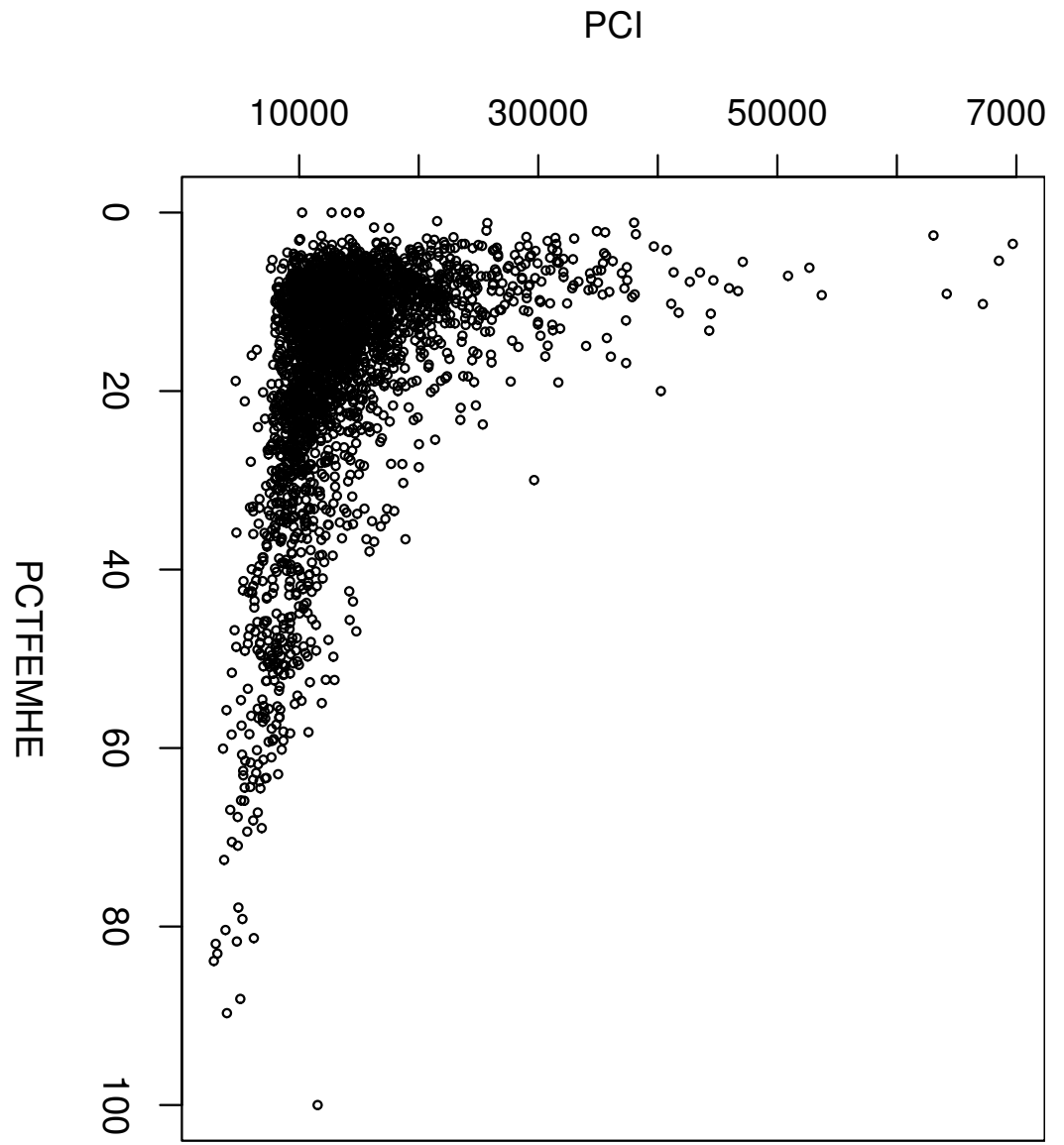
45° rule for aspect ratio - To maximize the perception of curvature, center the slopes on 45 degrees. (Both positive and negative slopes.)

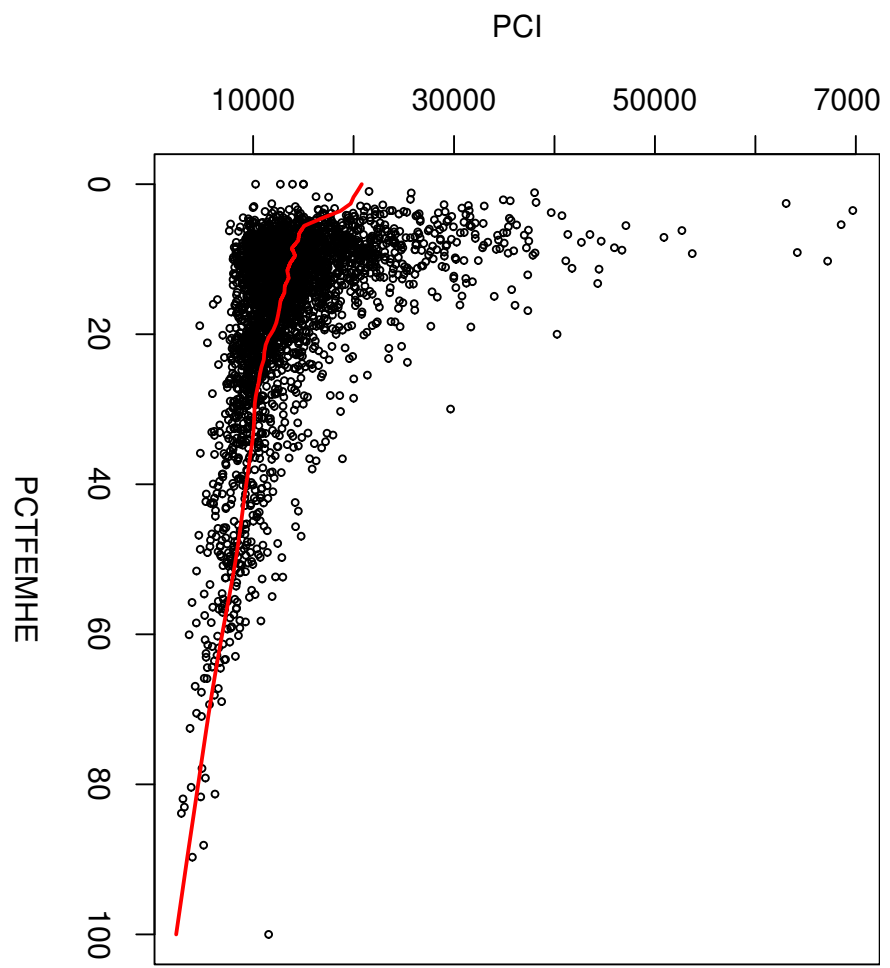
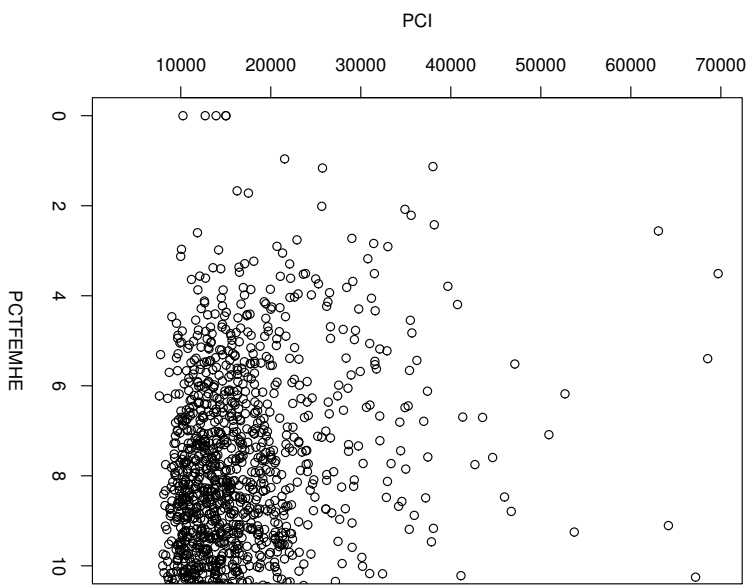
Timeline:

- 1686 Edmond Halley (England) makes first scatterplot of empirical data, to study weather patterns.
- 1760 J.H. Lambert (Germany) draws "best curves" through scatterplots, to study laws of light and reflectance.
- 1832 Scatterplots and fitted curves become widely used in science.
- 1912 Henry Russell's (US) scatterplot of stars revolutionizes astrophysics.
- 1977 Automatic drawing of trend lines.
- 1994 William Cleveland develops 45° rule.

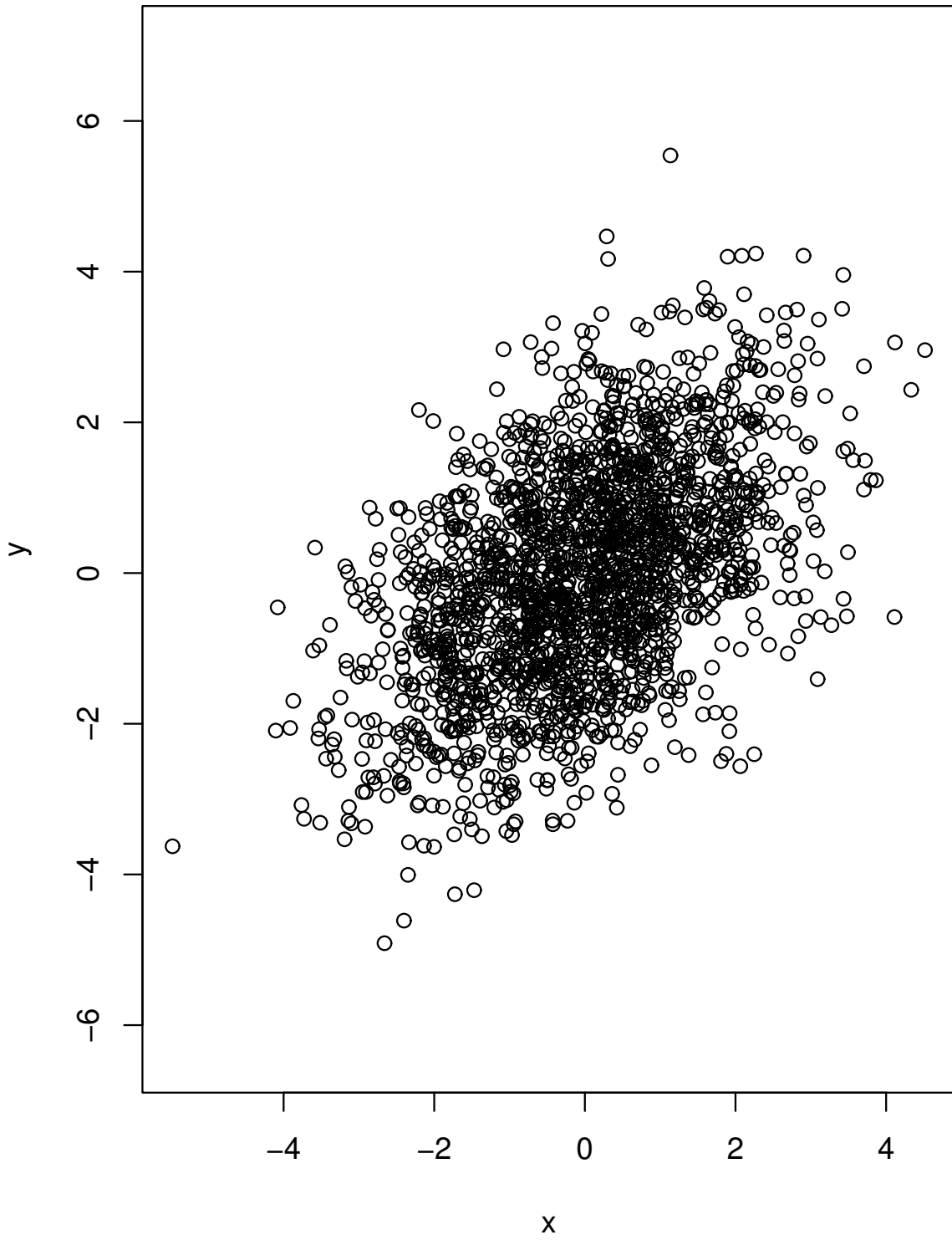
References

- [1] William S. Cleveland. *The Elements of Graphing Data*. Second Edition. Hobart Press, NJ, 1994.
- [2] Michael Friendly and Daniel J. Denis. *Milestones in the History of Thematic Cartography, Statistical Graphics, and Data Visualization*.
<http://www.math.yorku.ca/SCS/Gallery/milestone/>
- [3] I. Spence and R.F. Garrison. "A remarkable scatterplot" *The American Statistician* 47: 12-19, 1993. http://www.psych.utoronto.ca/~spence/Spence_Garrison_1993.pdf

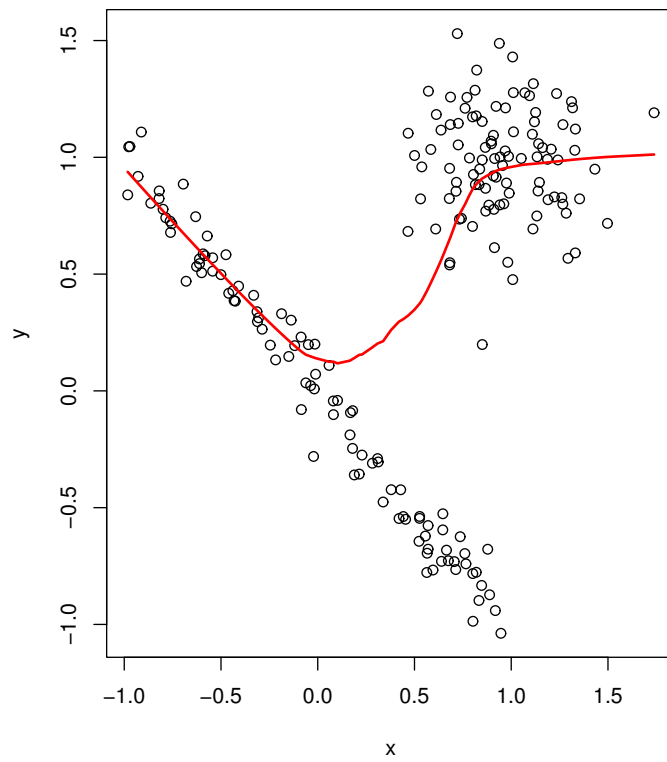
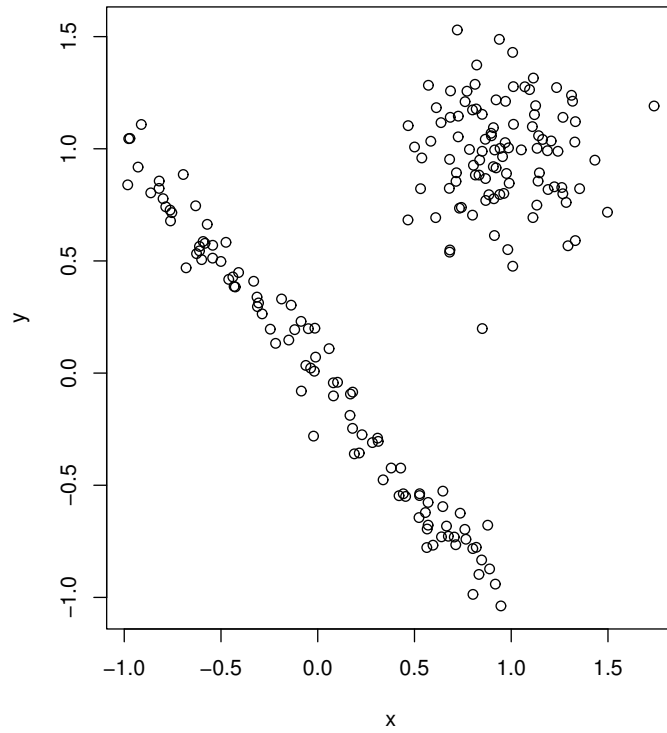




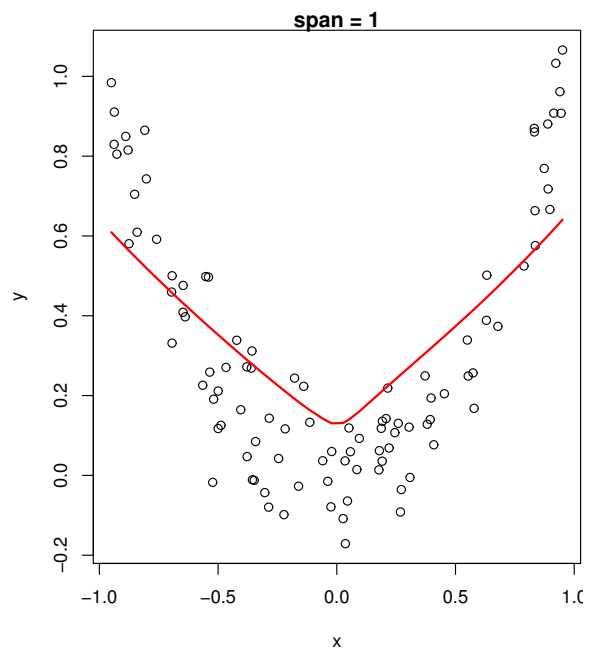
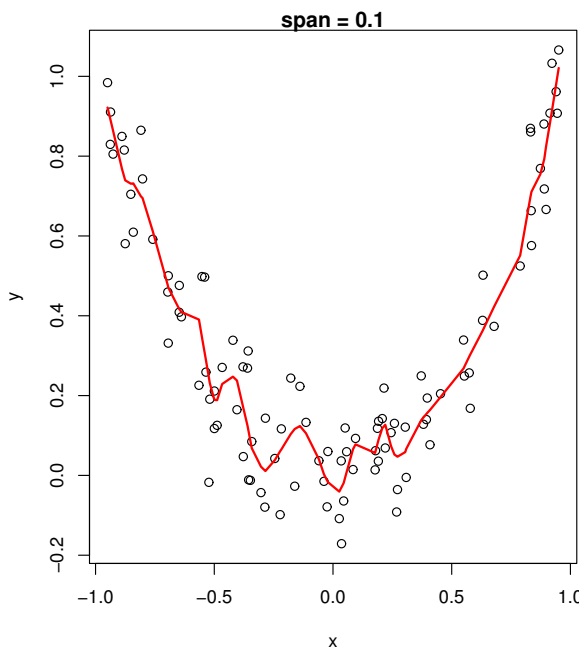
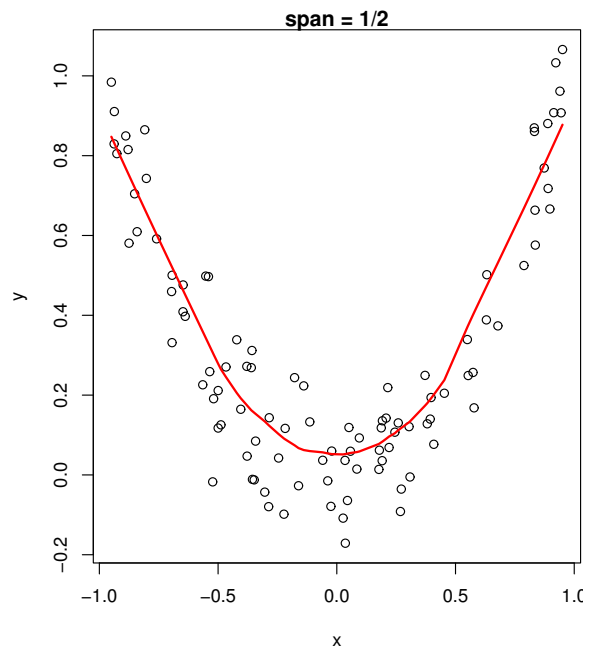
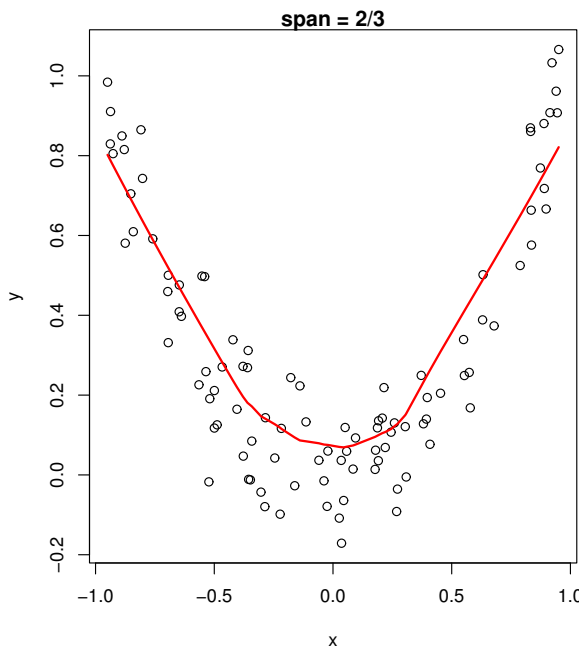
Can you draw the correct prediction line?



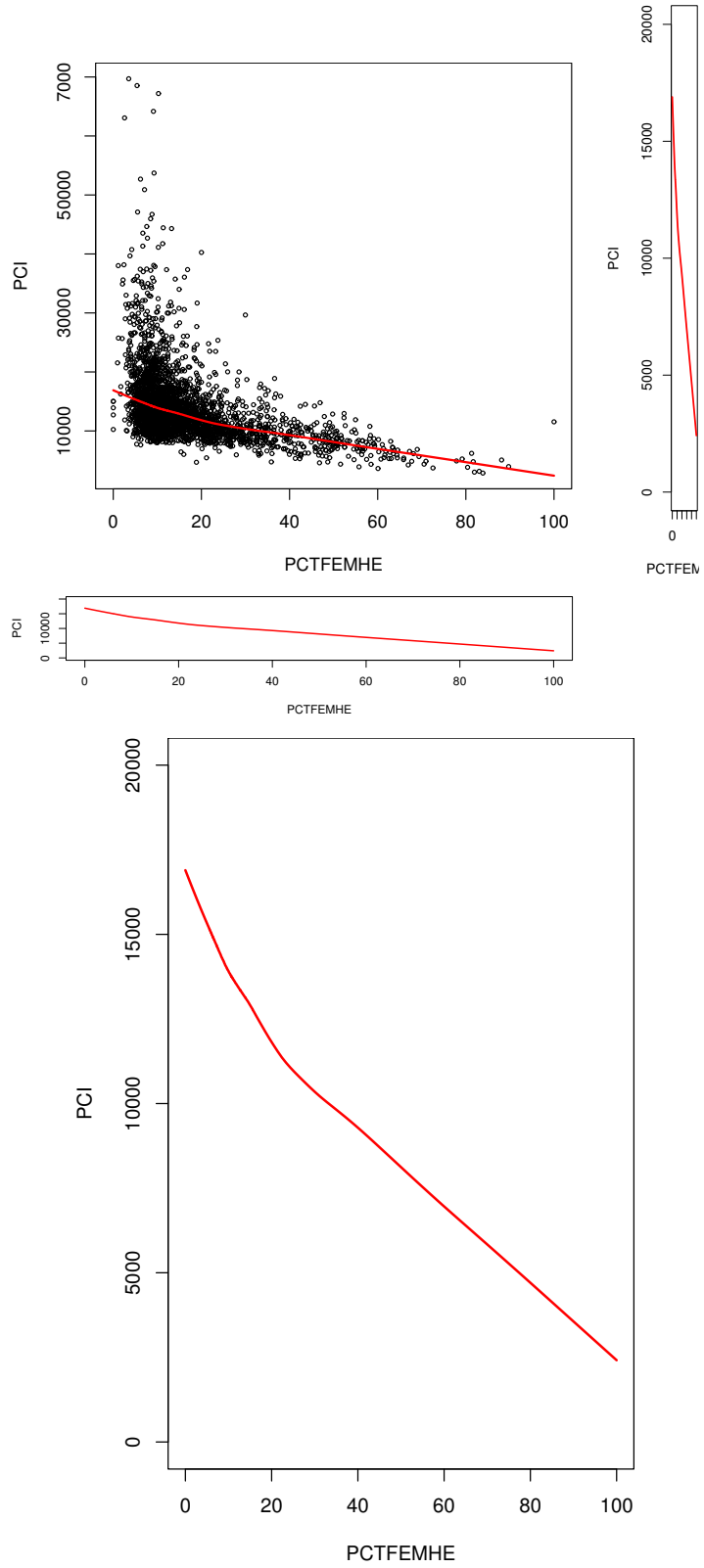
Data where a prediction line is meaningless:



Smoothing span:



Aspect ratio and the 45° rule:



Fixing a bad aspect ratio:

