# Week of March 7 Tutorial

Textbook notes

Also see the textbook comment slide in the March 4 lecture notes. The textbook goes into much more mathematical depth regarding correlation in section 11.12. We'll use correlation mostly as an informal tool. I suppose 11.44(a), 11.45(a,c), 11.47(a,c), 11.49(a,c), are suitable, if not somewhat pointless exercises. Correlation can (and should) be calculated for any bivariate dataset.

We have covered sections 12.1, 12.2, 12.3, and part of 12.4. Under no circumstances do I expect you do practice manually estimating multiple regression parameters. This is always a task for a computer.

Questions 12.1 to 12.16 are are suitable exercises, provided you have a computer handy to perform the calculations. See this week's Lab notes for help with fitting multiple regression models using R, which is very easy. Note that the sample sizes for these questions are incredibly small for fitting regression models with so many variables, which is probably not such a good idea in practice.

Tutorial questions

1. (2014 midterm) Here are scatterplots of four bivariate datasets called A, B, C, and D.
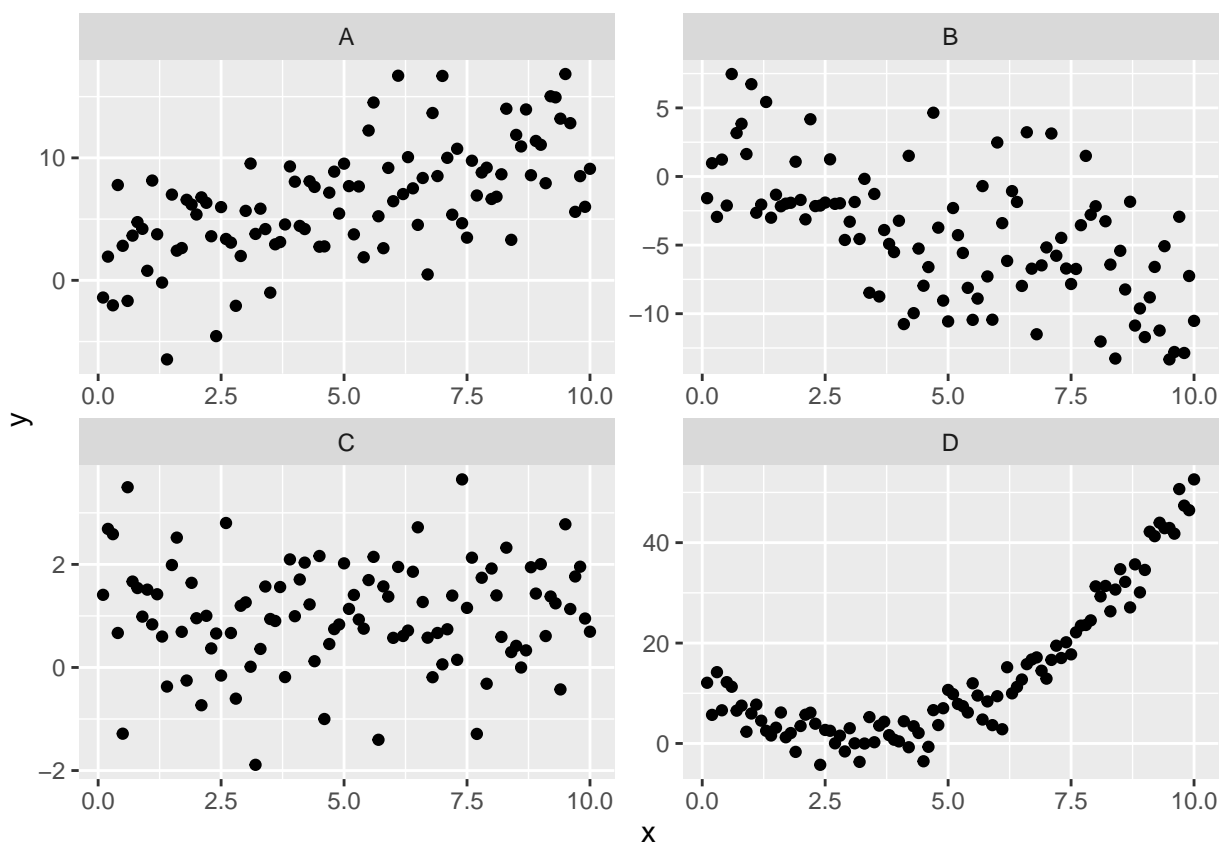
   a. The sample correlation coeffients for the four datasets are listed among the following seven numbers. Match dataset to number.

```
## [1]  0.811  0.359 -0.591  0.621  0.021 -0.827
```

b. Which of the four datasets has the strongest linear relationship between its two variables?

c. (added... not on 2014 midterm) The sample size for each dataset is 100. The total sum of squares for dataset A is 2102.1148565. Construct the ANOVA table for the regression of y on x for Dataset A.

2. Data from textbook question 12.16. Fit the model, estimate σ^2, and interpret the p-values for the indivudual $H_0 : \beta_i = 0$ tests using the following output:

```
##
## Call:
## lm(formula = Y ~ X1 + X2 + X3, data = gain)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -6.9517 -2.3992  0.3168  1.8602  7.8955
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -21.4696    51.6756  -0.415    0.684
## X1           -3.3243     3.5888  -0.926    0.369
## X2            0.2465     0.2747   0.897    0.384
## X3           20.3448     1.2576  16.178 6.65e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.151 on 15 degrees of freedom
## Multiple R-squared:  0.9916, Adjusted R-squared:  0.9899
## F-statistic: 588.1 on 3 and 15 DF,  p-value: 8.976e-16
```