2.3 (Weighted least squares) Suppose that the observations y_i have different precisions and so we would like to weight them using different weights $w_i > 0$. For example, each y_i may be a sample mean of n_i i.i.d. observations so that their variances are inversely proportional to the n_i and hence we use the n_i as the weights. Show that the **weighted least squares (WLS)** estimator of the slope β for regression through the origin obtained by minimizing the LS criterion $Q = \sum_{i=1}^{n} w_i(y_i - \beta x_i)^2$ equals

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} w_i x_i y_i}{\sum_{i=1}^{n} w_i x_i^2}.$$

2.9 (Beta coefficients of stocks) The β of a stock is a coefficient that describes how the return on that stock is related to the return on a diversified stock portfolio. It is the slope coefficient in the simple linear regression model, $y = \alpha + \beta x$, where y is the return on that stock and x is the return on a benchmark stock market index representing a diversified portfolio. In this exercise we want to compare the β's of IBM and Apple versus S&P 500.

The file `IBM-Apple-SP500 RR Data.csv` contains data on percentage monthly rates of return (adjusted for dividends and stock splits) from February 2005 until September 2013 for IBM, Apple and S&P 500. These rates were calculated by downloading historical monthly prices from the Yahoo Finance website (`http://finance.yahoo.com/`). (If you prefer, you can download the more current data and use that to do the problem.)

a) Make scatter plots of rates of return of IBM versus S&P 500 and Apple versus S&P 500 and comment on them.

b) Calculate the β's for IBM and Apple versus S&P 500. Comment on the relative magnitudes of the β's. Which stock had a higher expected return relative to S&P 500?

c) Calculate the sample standard deviations (SD's) of rates of return for S&P 500, IBM and Apple. Also calculate the correlation matrix. Check that $\widehat{\beta} = rs_y/s_x$

for each stock where r is the correlation coefficient between S&P 500 and the given stock, s_x is the sample SD of S&P 500 and s_y is the sample SD of the given stock.

d) Explain based on the statistics calculated how a higher expected return is accompanied by higher volatility of the stock relative to S&P 500.

2.10 (Price elasticities of steaks) Data file `steakprices.csv` gives time series data on the prices and quantities sold of three types of beef steaks, chuck, porterhouse and rib eye,
(`http://www.aabri.com/manuscripts/08118.pdf`).

a) Estimate the price elasticities of all three steaks. Given that chuck is the least expensive cut and rib eye is the most expensive cut of beef, are the price elasticities of the three cuts in the expected order?

b) Estimate how much the demand will change if the price is increased by 10% for each cut.

2.11 (Smoking versus cancer) Data file `smoking-cancer.csv` contains data from 43 states and Washington, D.C. on the average number of cigarettes smoked (hundreds/capita) and number of deaths per 100,000 population due to four types of cancer.

a) Make scatter plots of the number of deaths due to each type of cancer versus cigarettes smoked to see what types of relationships (linear, nonlinear) exist and if there are any outliers.

b) Perform tests on the correlations to see which type of cancer deaths are most significantly correlated with cigarette smoking.