Exercise 7C.1

Linear regression was used to investigate the relationship between age (years) and systolic blood pressure (mmHg) in a sample of adults ranging from 29 to 69 years old. Stata output is below.

```
. regress sbp age

      Source |       SS       df       MS              Number of obs =      29
-------------+------------------------------           F(  1,    27) =   66.81
       Model |  6110.10173     1  6110.10173            Prob > F      =  0.0000
    Residual |  2469.34654    27  91.4572794            R-squared     =  0.7122
-------------+------------------------------           Adj R-squared =  0.7015
       Total |  8579.44828    28  306.408867            Root MSE      =  9.5633

------------------------------------------------------------------------------
         sbp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         age |   .9493225   .1161445     8.17   0.000     .7110137    1.187631
       _cons |   97.07708   5.527552    17.56   0.000     85.73549    108.4187
------------------------------------------------------------------------------
```

(a) Calculate the sample correlation between age and SBP.

$r = sign(\hat{\beta}_1) \times \sqrt{R^2} = +\sqrt{0.7122} = 0.844$

(b) If we convert age from years to months, will the sample correlation change? Will the slope estimate change? Will the R^2 change?

Correlation will not change, slope will change, R^2 will not change.

Exercise 7C.2

A summary of some of the variables in the New Hampshire student survey data is on the next page. Use this to answer the questions.

(a) If I perform a linear regression using the miles away from school a student lives (`miles`) to predict his/her GPA (`gpa`), will the slope estimate be positive or negative?

positive, because $r = 0.1558 > 0$

(b) If I perform a linear regression using the miles away from school a student lives (`miles`) to predict his/her GPA (`gpa`), what will the coefficient of determination for the regression be?

$R^2 = r^2 = 0.1558^2 = 0.0243$

(c) If I perform a linear regression using the miles away from school a student lives (`miles`) to predict his/her drinking score (`drink`), will the slope estimate be positive or negative?

negative, because $r = -0.2702 < 0$

(d) If I perform a linear regression using the miles away from school a student lives (`miles`) to predict his/her drinking score (`drink`), what will the coefficient of determination for the regression be?

$R^2 = r^2 = (-0.2702)^2 = 0.073$

(e) If I perform a linear regression using drinking score (`drink` to predict the miles away from school a student lives (`miles`), what will the coefficient of determination for the regression be?

Same as in (d), $R^2 = 0.073$ *– switching X and Y doesn't change* r *or* R^2

Output for problem on previous page:

```
. corr gpa drink miles
(obs=206)

             |      gpa    drink    miles
-------------+---------------------------
         gpa |   1.0000
       drink |  -0.2591   1.0000
       miles |   0.1558  -0.2702   1.0000
```