Exercise 6

The file *stockton96.gdt* contains 940 observations on home sales in Stockton, CA in 1996.

a) Use least squares to estimate a linear equation that relates house price *PRICE* to the size of the house in square feet *SQFT* and the age of the house in years *AGE*. Interpret all the estimates.
ols price const age sqft

```
Model 1: OLS, using observations 1-940
Dependent variable: price

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const       5193.15      3586.64        1.448    0.1480
  age         -217.843       35.0976     -6.207    8.11e-010 ***
  sqft          68.3907       2.16868    31.54     2.39e-149 ***

Mean dependent var   97937.83   S.D. dependent var   34179.37
Sum squared resid    4.76e+11   S.E. of regression   22539.63
R-squared            0.566050   Adjusted R-squared   0.565124
F(2, 937)            611.1178   P-value(F)           1.4e-170
Log-likelihood      -10753.95   Akaike criterion     21513.90
Schwarz criterion    21528.43   Hannan-Quinn         21519.44
```

b) Suppose that you own two houses. One has 1400 square feet; the other has 1800 square feet. Both are 20 years old. What price do you estimate you will get for each house?

$$p_1 = 5193 + 20 * (-217) + 68.39 * 1400$$
$$p_2 = 5193 + 20 * (-217) + 68.39 * 1800$$

c) Test the hypothesis that the size and the age of the house are important determinants of its price (separately as well as jointly). **Both have three stars. Also jointly significant according to above output**

d) Using the Breusch-Pagan test for heteroscedasticity, test whether the model satisfies the homoscedasticity assumption by using the command for the BP test in Gretl.
series yhat=$yhat
genr resid=price-yhat
modtest --breusch-pagan

e) Use the White test to test for heteroskedasticity.
modtest --white

f) What do you conclude regarding the heteroskedasticity? Does your conclusion depend on the choosing a specific test? Discuss also drawbacks of the BP and White tests.
There is heteroskedasticity

A weakness of the BP test is that it assumes the heteroskedasticity is a linear function of the independent variables. Failing to find evidence of heteroskedasticity with the BP doesn't rule out a nonlinear relationship between the independent variable(s) and the error variance.
The weakness of white test is that if you have many variables, the number of possible interactions plus the squared variables plus the original variables can be quite high.

g) Test the hypothesis that the size and the age of the house are important determinants of its price (separately as well as jointly). Hint: choose appropriate standard errors. Does your conclusion differ from part (c)?

ols price const age sqft –robust

compare the robust and non-robust standard errors and parameters. You can see that the parameters did not change, while standard errors increased

```
? ols price const sqft age --robust

Model 10: OLS, using observations 1-940
Dependent variable: price
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const       5193.15      3648.56       1.423    0.1550
  sqft          68.3907       2.46807   27.71     6.35e-124 ***
  age         -217.843       36.3142    -5.999    2.84e-09  ***

Mean dependent var   97937.83   S.D. dependent var   34179.37
Sum squared resid    4.76e+11   S.E. of regression   22539.63
R-squared            0.566050   Adjusted R-squared   0.565124
F(2, 937)            476.5571   P-value(F)           1.7e-143
Log-likelihood      -10753.95   Akaike criterion     21513.90
Schwarz criterion    21528.43   Hannan-Quinn         21519.44

? ols price const sqft age

Model 11: OLS, using observations 1-940
Dependent variable: price

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const       5193.15      3586.64       1.448    0.1480
  sqft          68.3907       2.16868   31.54     2.39e-149 ***
  age         -217.843       35.0976    -6.207    8.11e-010 ***

Mean dependent var   97937.83   S.D. dependent var   34179.37
Sum squared resid    4.76e+11   S.E. of regression   22539.63
R-squared            0.566050   Adjusted R-squared   0.565124
F(2, 937)            611.1178   P-value(F)           1.4e-170
Log-likelihood      -10753.95   Akaike criterion     21513.90
Schwarz criterion    21528.43   Hannan-Quinn         21519.44
```