```
*Clear data
clear
*make log
log using "C:\Users\Gtjohnso\Documents\Pset3.smcl", replace
*import data
use "C:\Users\Gtjohnso\Documents\county_health-1.dta"

*Avrg obesity rate
mean adultobesity

*count counties in texas
count if state == "TX"

*estimate Model A
reg healthcarecostsvalue median_ink_10K

*estimate Model B
reg healthcarecostsvalue smoking

*estimate Model C
reg healthcarecostsvalue median_ink_10K smoking adultobesity

*have stata check work
local median_income 5.05  // Median household income in $10,000 units
($50,500)
local smoking_rate 10    // Smoking rate in percentage (10%)
local obesity_rate 40    // Obesity rate in percentage (40%)

*Predict and list the data
predict healthcare_predicted if e(sample), xb
list healthcare_predicted in 1/1

*Predict healthcare spending for elderly in Harris County TX
*values in Harris County
local median_income 5.423  // Median household income in $10,000 units
($54,230)
local smoking_rate 13.9    // Smoking rate in percentage (13.9%)
local obesity_rate 27.4    // Obesity rate in percentage (27.4%)

*Regression Coeff from Model C
local beta0 5773.707
local beta1 -75.70479
local beta2 45.98631
local beta3 98.83699

*using the equation from model c
g hc_pred = `beta0' + `beta1' * `median_income' + `beta2' *
`smoking_rate' + `beta3' * `obesity_rate'
display hc_pred
*actual observed value
local hc_obsv 11488

*calculate residual
g residual = `hc_obsv' - hc_pred
```

```
display residual

*convert to pdf
translate  Pset3.smcl  Pset3.pdf

*Close log
log close
```

```
. do "C:\Users\Gtjohnso\AppData\Local\Temp\18\STD8784_000000.tmp"

. *Clear data
. clear

. *make log
. log using "C:\Users\Gtjohnso\Documents\Pset3.smcl", replace
-------------------------------------------------------------------------------
      name:  <unnamed>
       log:  C:\Users\Gtjohnso\Documents\Pset3.smcl
  log type:  smcl
 opened on:  20 Feb 2024, 20:44:27

. *import data
. use "C:\Users\Gtjohnso\Documents\county_health-1.dta"

.
. *Avrg obesity rate
. mean adultobesity

Mean estimation                   Number of obs    =      3,141

--------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
adultobesity |    30.9482   .0796839      30.79196    31.10444
--------------------------------------------------------------

.
. *count counties in texas
. count if state == "TX"
  254

.
. *estimate Model A
. reg healthcarecostsvalue median_ink_10K

      Source |       SS           df       MS      Number of obs   =
3,133
-------------+----------------------------------   F(1, 3131)      =
254.35
       Model |   521930318         1   521930318   Prob > F        =
0.0000
    Residual |  6.4248e+09     3,131  2051989.59   R-squared       =
0.0751
-------------+----------------------------------   Adj R-squared   =
0.0748
       Total |  6.9467e+09     3,132  2217978.84   Root MSE        =
1432.5

-------------------------------------------------------------------------------
-------
```

```
healthcareco~e |      Coef.   Std. Err.      t    P>|t|     [95% Conf.
Interval]
---------------+----------------------------------------------------------
-------
median_ink_10K |   -337.9777   21.19188   -15.95   0.000    -379.5291   -
296.4263
         _cons |    10914.82   103.0382   105.93   0.000     10712.79
11116.84
------------------------------------------------------------------------------
-------

.
. *estimate Model B
. reg healthcarecostsvalue smoking

      Source |       SS           df       MS      Number of obs   =
3,134
-------------+----------------------------------   F(1, 3132)      =
396.27
       Model |   780716271         1   780716271   Prob > F        =
0.0000
    Residual |  6.1706e+09     3,132  1970173.85   R-squared       =
0.1123
-------------+----------------------------------   Adj R-squared   =
0.1120
       Total |  6.9513e+09     3,133  2218736.28   Root MSE        =
1403.6

------------------------------------------------------------------------------
-----
healthcare~e |      Coef.   Std. Err.      t    P>|t|     [95% Conf.
Interval]
-------------+----------------------------------------------------------
-----
     smoking |    131.742    6.61805    19.91   0.000     118.7659
144.7182
       _cons |    6898.201   124.3301    55.48   0.000     6654.424
7141.978
------------------------------------------------------------------------------
-----

.
. *estimate Model C
. reg healthcarecostsvalue median_ink_10K smoking adultobesity

      Source |       SS           df       MS      Number of obs   =
3,133
-------------+----------------------------------   F(3, 3129)      =
219.57
       Model |  1.2081e+09         3   402693069   Prob > F        =
0.0000
    Residual |  5.7386e+09     3,129  1834014.23   R-squared       =
0.1739
```

```
.
. *Regression Coeff from Model C
. local beta0 5773.707

. local beta1 -75.70479

. local beta2 45.98631

. local beta3 98.83699

.
. *using the equation from model c
. g hc_pred = `beta0' + `beta1' * `median_income' + `beta2' *
`smoking_rate' + `beta3' * `obesity_rate'

. display hc_pred
8710.5029

. *actual observed value
. local hc_obsv 11488

.
. *calculate residual
. g residual = `hc_obsv' - hc_pred

. display residual
2777.4971

.
. *convert to pdf
. translate  Pset3.smcl  Pset3.pdf
```