**Lab Problems**

1. Write a do file named lab1problemset.do that does the following:

lab1problemset.do file:

/\* A. Shawn Bandy

Lab #1

January 31, 2013

\*/

/\* close previously run do-files \*/

cap log close

//executes without pausing to paginate

set more 1

//clears variables

clear

//set semi-colon as command line delimiter

#delimit ;

//change directory

cd "C:\Users\cla-spa206.CAMPUS-DOMAIN\Downloads\econ485-master\econ485-master\lab1";

//a. Create a log file named lab1problemset.log on your flash drive that records all output.

log using lab1problemset.log , replace;

//b.Imports the BEA data from your flash drive

insheet using "BEA per capita income 1969 to 2008.csv", names;

foreach v of varlist fips year\* {;

gen `v'\_new=real(`v');

drop `v';

rename `v'\_new `v';

};

drop if missing(fips);

//e. Renames the variables year1970 and year2008 as pci70 and pci08

rename year1970 pci70;

rename year2008 pci08;

//f. Creates a new variable called perch\_pci7008 that looks at the % change in per capita income from 1970 to 2008.

generate perch\_pci7008 = (pci08 - pci70) / pci70 \* 100;

//g. Saves the FIPS code, the name of each county, pci70, pci08, and perch\_pci7008 in a new dataset called pci70and08.dta in the folder on your flash drive.

keep fips pci70 pci08 perch\_pci7008;

save pci70and08, replace;

//h. Reorders the variables so that the first variable is the FIPS code, followed by name of the county, then the % change in per capita income, and then per capita income in 1970 and 2008.

order fips;

//i. Sorts by the change in per capita income from 1970 to 2008

sort perch\_pci7008;

//j. Uses the simple summarize function to look at all the variables left in the sample.

summarize;

//k. Uses the detailed summarize function for only pci70, pci08, and perch\_pci7008

summarize pci70 pci08 perch\_pci7008, detail;

//l. Examines the correlation between pci70 and pci08.

correlate pci70 pci08;

//m. Examines the correlation between pci70 and the change in per capita income from 1970 to 2008.

correlate pci70 perch\_pci7008;

//n. Creates a histogram of pci70 based on fractions

histogram pci70, fraction;

//p. Exports the histogram in n. as a Windows Enhanced Metafile (emf) with the name pcihistogram. Print a copy of the histogram and include with your assignment.

graph export pcihistogram.emf, replace;

//o. Creates a histogram of pci70 based on fractions and with the normal distribution added.

histogram pci70, fraction normal;

log close;

lab1problemset.log:

--------------------------------------------------------------------------------

name: <unnamed>

log: C:\Users\cla-spa206.CAMPUS-DOMAIN\Downloads\econ485-master\econ485-

> master\lab1\lab1problemset.log

log type: text

opened on: 31 Jan 2013, 11:00:39

. //b.Imports the BEA data from your flash drive

> insheet using "BEA per capita income 1969 to 2008.csv", names;

(43 vars, 3140 obs)

. foreach v of varlist fips year\* {;

2. gen `v'\_new=real(`v');

3. drop `v';

4. rename `v'\_new `v';

5. };

(5 missing values generated)

(31 missing values generated)

(31 missing values generated)

(31 missing values generated)

(31 missing values generated)

(31 missing values generated)

(31 missing values generated)

(31 missing values generated)

(31 missing values generated)

(31 missing values generated)

(31 missing values generated)

(37 missing values generated)

(37 missing values generated)

(37 missing values generated)

(36 missing values generated)

(35 missing values generated)

(35 missing values generated)

(35 missing values generated)

(35 missing values generated)

(35 missing values generated)

(34 missing values generated)

(33 missing values generated)

(33 missing values generated)

(31 missing values generated)

(31 missing values generated)

(30 missing values generated)

(30 missing values generated)

(30 missing values generated)

(30 missing values generated)

(30 missing values generated)

(30 missing values generated)

(30 missing values generated)

(30 missing values generated)

(30 missing values generated)

(29 missing values generated)

(29 missing values generated)

(29 missing values generated)

(29 missing values generated)

(29 missing values generated)

(29 missing values generated)

(28 missing values generated)

. drop if missing(fips);

(5 observations deleted)

. //e. Renames the variables year1970 and year2008 as pci70 and pci08

> rename year1970 pci70;

. rename year2008 pci08;

. //f. Creates a new variable called perch\_pci7008 that looks at the % change

> in per capita income from 1970 to 2008.

> generate perch\_pci7008 =(pci08 - pci70) / pci70 \* 100;

(46 missing values generated)

. //g. Saves the FIPS code, the name of each county, pci70, pci08, and perch\_

> pci7008 in a new dataset called pci70and08.dta in the folder on your flash dri

> ve.

> keep fips pci70 pci08 perch\_pci7008;

. save pci70and08, replace;

file pci70and08.dta saved

. //h. Reorders the variables so that the first variable is the FIPS code, fo

> llowed by name of the county, then the % change in per capita income, and then

> per capita income in 1970 and 2008.

> order fips;

. //i. Sorts by the change in per capita income from 1970 to 2008

> sort perch\_pci7008;

. //j. Uses the simple summarize function to look at all the variables left i

> n the sample.

> summarize;

Variable | Obs Mean Std. Dev. Min Max

-------------+--------------------------------------------------------

fips | 3135 30028.73 15222.11 1001 56045

pci70 | 3109 3252.609 775.2473 1330 8537

pci08 | 3112 33093.17 8530.69 12558 140275

perch\_p~7008 | 3089 935.7032 203.3258 339.2445 4607.215

. //k. Uses the detailed summarize function for only pci70, pci08, and perch\_

> pci7008

> summarize pci70 pci08 perch\_pci7008, detail;

pci70

-------------------------------------------------------------

Percentiles Smallest

1% 1817 1330

5% 2131 1400

10% 2309 1407 Obs 3109

25% 2700 1411 Sum of Wgt. 3109

50% 3214 Mean 3252.609

Largest Std. Dev. 775.2473

75% 3700 6936

90% 4154 7041 Variance 601008.3

95% 4589 7441 Skewness .8335436

99% 5550 8537 Kurtosis 5.181562

pci08

-------------------------------------------------------------

Percentiles Smallest

1% 20026 12558

5% 23347 14553

10% 24950 15408 Obs 3112

25% 27807.5 15452 Sum of Wgt. 3112

50% 31551.5 Mean 33093.17

Largest Std. Dev. 8530.69

75% 36601.5 95764

90% 42148 120766 Variance 7.28e+07

95% 47045 129956 Skewness 2.862153

99% 62559 140275 Kurtosis 24.05697

perch\_pci7008

-------------------------------------------------------------

Percentiles Smallest

1% 565.1243 339.2445

5% 663.0004 348.9991

10% 723.7597 367.0411 Obs 3089

25% 810.7478 371.6119 Sum of Wgt. 3089

50% 914.9754 Mean 935.7032

Largest Std. Dev. 203.3258

75% 1034.645 2056.145

90% 1172.608 2179.003 Variance 41341.39

95% 1262.627 2289.321 Skewness 2.790385

99% 1553.517 4607.215 Kurtosis 39.70566

. //l. Examines the correlation between pci70 and pci08.

> correlate pci70 pci08;

(obs=3089)

| pci70 pci08

-------------+------------------

pci70 | 1.0000

pci08 | 0.6583 1.0000

. //m. Examines the correlation between pci70 and the change in per capita in

> come from 1970 to 2008.

> correlate pci70 perch\_pci7008;

(obs=3089)

| pci70 per~7008

-------------+------------------

pci70 | 1.0000

perch\_p~7008 | -0.3787 1.0000

. //n. Creates a histogram of pci70 based on fractions

> histogram pci70, fraction;

(bin=34, start=1330, width=211.97059)

. //p. Exports the histogram in n. as a Windows Enhanced Metafile (emf) with

> the name pcihistogram. Print a copy of the histogram and include with your as

> signment.

> graph export pcihistogram.emf, replace;

(note: file pcihistogram.emf not found)

(file C:\Users\cla-spa206.CAMPUS-DOMAIN\Downloads\econ485-master\econ485-master\

> lab1\pcihistogram.emf written in Enhanced Metafile format)

. //o. Creates a histogram of pci70 based on fractions and with the normal di

> stribution added.

> histogram pci70, fraction normal;

(bin=34, start=1330, width=211.97059)

. log close;

name: <unnamed>

log: C:\Users\cla-spa206.CAMPUS-DOMAIN\Downloads\econ485-master\econ485-

> master\lab1\lab1problemset.log

log type: text

closed on: 31 Jan 2013, 11:00:51

--------------------------------------------------------------------------------

2. *What are the minimum, maximum, range, mean, standard deviation and variance of per capita income of U.S. counties in 2008?*

For U.S. counties in 2008, the minimum per capita income is $12,558; the maximum per capita income is $140,275; the mean is $33093.17; the standard deviation is $8530.69.

1. *What are the mean and variance of the percent change in per capita income of U.S. counties from 1970 to 2008?*

For U.S. counties, the percent change in per capita income from 1970 to 2008 was 935.70 percent with a standard deviation of 203.33 percent.

1. *What county experienced had the lowest percent change in per capita income from 1970 to 2008? Do an internet search for this county. What factor about this county do you think contributed to it being last in per capita income growth rate?*

The county that experienced the lowest percent change in per capital income from 1970 to 2008 was Buffalo, South Dakota. According to Wikipedia (data from the United Census Bureau), ten counties in South Dakota are in the bottom one hundred for per capita income. Buffalo County is at the very bottom at $5,213; 40.8% of the population falls below the poverty line.[[1]](#footnote-1) [[2]](#footnote-2) The percent of people 25 years or older without a high-school diploma is about twice as high as the average for South Dakota.[[3]](#footnote-3) I suspect that the county is locked in persistent cycles of poverty. Those who seek out an education, career or entrepreneurship almost certainly leave the county, relocating to a more hospital economic environment. That drain leaves the area with relatively little capital and increasingly less reason for outside private investment.

*What is the n or the number of observations for the % change in per capita income from 1970 to 2008?*

There were 3089 observations for the percent change in per capita income from 1970 to 2008. There are few observations for this series than pci70 and pci08 because it requires a numeric value for both 1970 and 2008 per capita income.

1. *Is the distribution of per capita income in 1970 symmetric? Explain.*

Skewness for the per capita income series for 1970 is 0.83. When skewness is non-zero the distribution is not symmetric. When skewness is greater than zero the observations have more smaller values than larger.

1. *Does the distribution of per capita income in 1970 look like a normal distribution?*



Yes, it does appear to be normally distributed.

1. *What is the relationship between per capita income in 1970 and per capita income in 2008? Does this make sense? Explain.*

The correlation coefficient for pci70 and pci08 is 0.66. A positive value means that counties with a higher per capita income in 1970 tended to have a higher per capita income in 2008 and vice versa.

1. *Do counties with high per capita incomes in 1970 appear to also have a higher than average per capita income growth rate from 1970 to 2008? Explain how you know and why this might be the case.*

No, the counties with higher per capita income in 1970 did not tend to have higher growth rate from 1970 to 2008 as the correlation coefficient is negative (-0.37). In general, per capita income growth is subject to decreasing returns.

1. <http://en.wikipedia.org/wiki/List_of_lowest-income_counties_in_the_United_States> [↑](#footnote-ref-1)
2. http://en.wikipedia.org/wiki/Buffalo\_County,\_South\_Dakota [↑](#footnote-ref-2)
3. http://quickfacts.census.gov/qfd/states/46/46017.html [↑](#footnote-ref-3)