

04 – Spatial Demography Concepts and Databases II

J.S. Onésimo Sandoval

Sociology

Saint Louis University

Outline

- Education Attainment Index
- Theil Index of Income Inequality
- GINI Coefficient of Income Inequality
- Normalization of variables and Creating an Index

EDUCATION ATTAINMENT INDEX

Education Attainment

Let's say that 85% of the U.S adult population had at least high school diploma, 27.7% had a bachelor's degree, and 10.2% had a graduate degree.

The education attainment score is
 $.85 + .277 + .102 = 1.228$.

Note:

Maximum score is 3

Minimum score is 0

You should scale the score for ease of interpretation

Saint Louis City (2006-2010 ACS)

80.6% had at least a high school diploma

26.9% had at least a bachelor's degree

04.0% had at least a graduate degree.

The education score

is $.806 + .269 + .040 = 1.115$.

STATA Code

1. Educational Attainment for Population 25 Years and Over

Universe: Population 25 Years and Over

Name: T25

Variables:

T025_001: Population 25 Years and Over:

T025_002: Less than High School

T025_003: High School Graduate

(Includes Equivalency)

T025_004: Some College

T025_005: Bachelor's Degree

T025_006: Master's Degree

T025_007: Professional School Degree

T025_008: Doctorate Degree

*EDUCATION ATTAINMENT.

```
gen Eo3=T025_003/T025_001
```

```
gen Eo4=T025_004/T025_001
```

```
gen Eo5=T025_005/T025_001
```

```
gen Eo6=T025_006/T025_001
```

```
gen Eo7=T025_007/T025_001
```

```
gen Eo8=T025_008/T025_001
```

```
gen ED_HS=(Eo3+Eo4+Eo5+Eo6+Eo7+Eo8)
```

```
gen ED_BS=(Eo5+Eo6+Eo7+Eo8)
```

```
gen ED_GD=(Eo6+Eo7+Eo8)
```

```
gen ED_TOT=ED_HS+ED_BS+ED_GD
```

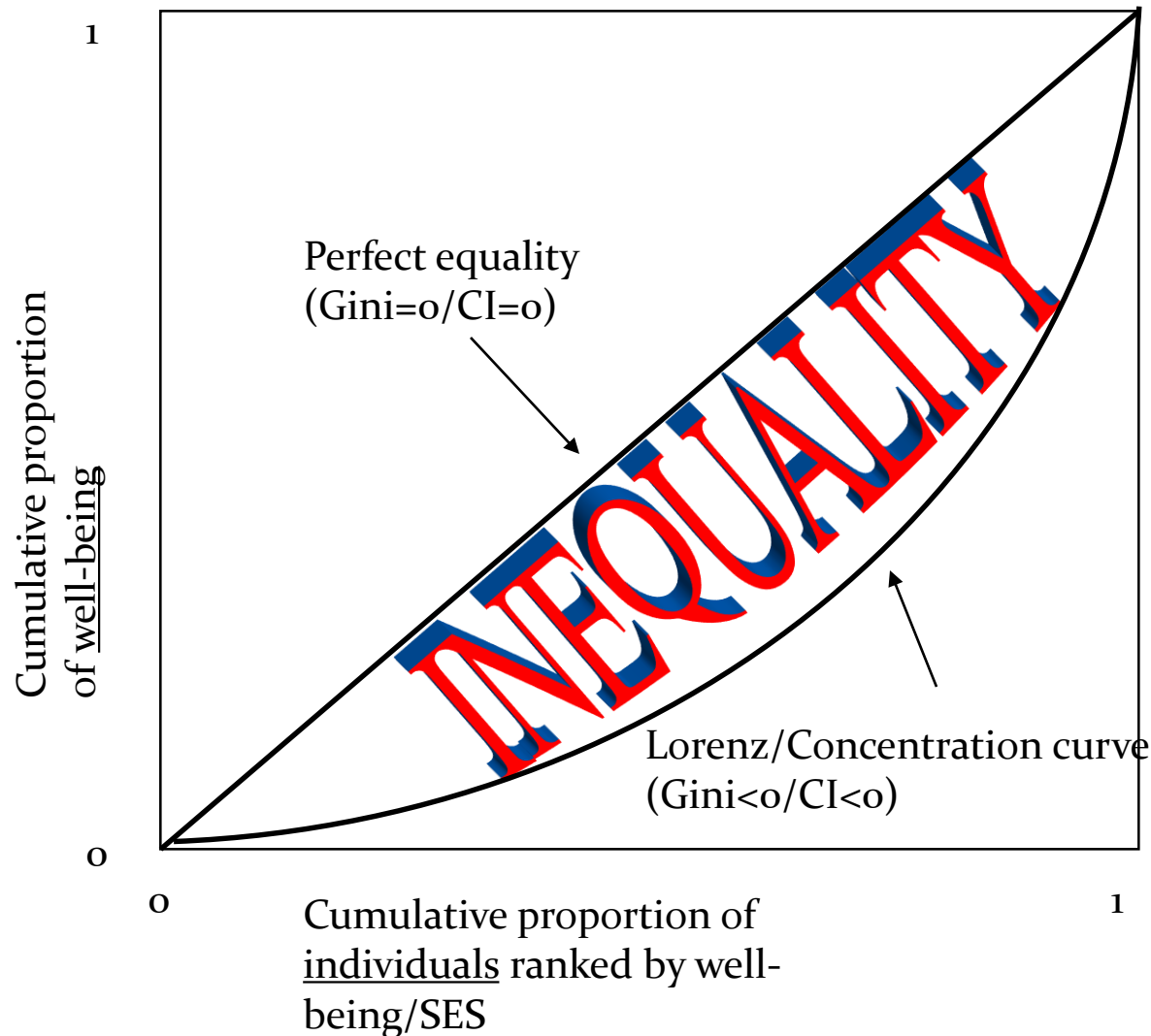
INCOME INEQUALITY

Lorenz Curve

- Lorenz Curve
 - Max O. Lorenz
 - 1905
 - Represent inequality of the wealth distribution
 - Measures inequalities in the distribution of wealth or income
 - Depict the state of concentration of population and of other demographic aggregates'

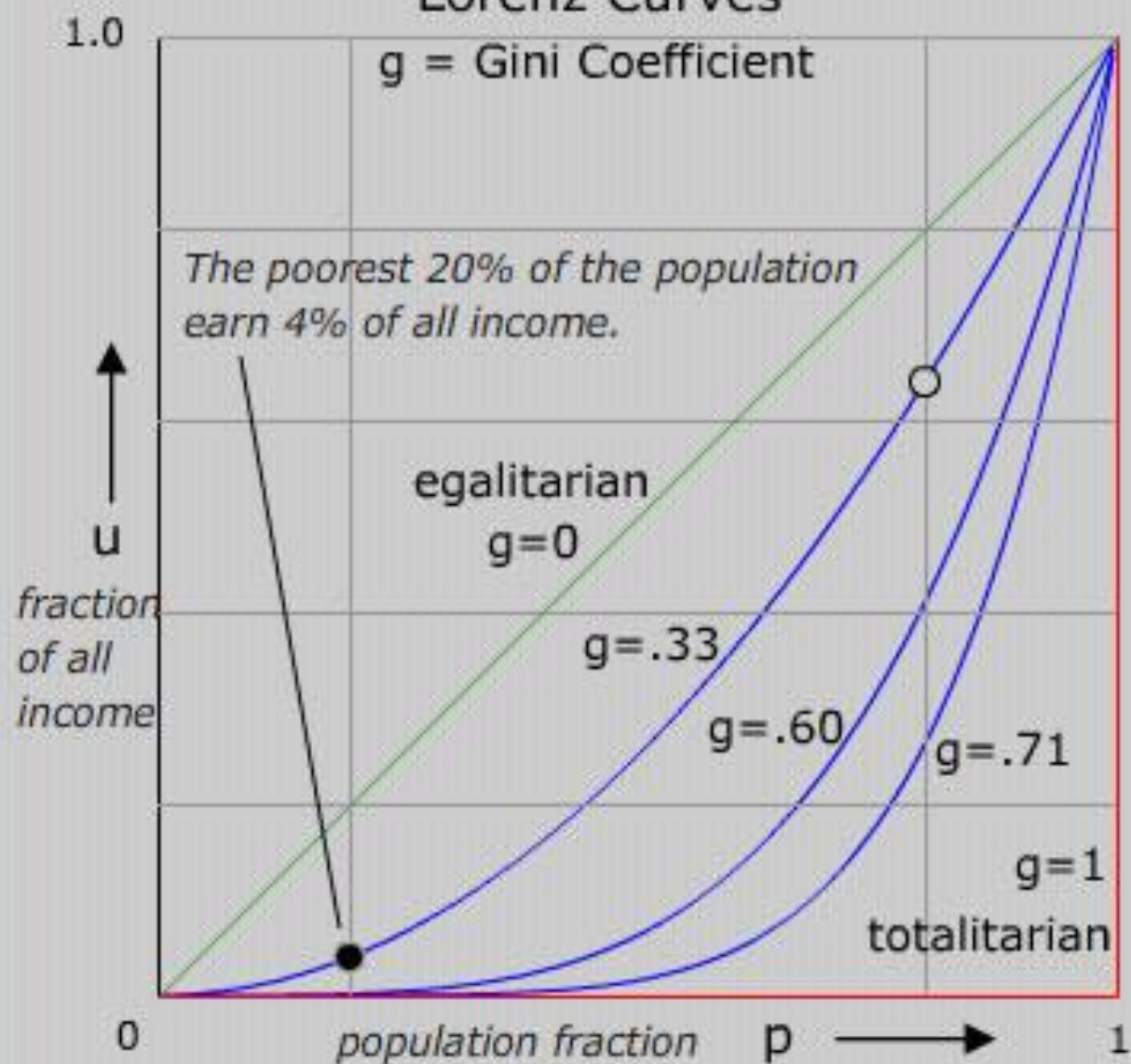


Lorenz/Concentration Curve for Well-Being



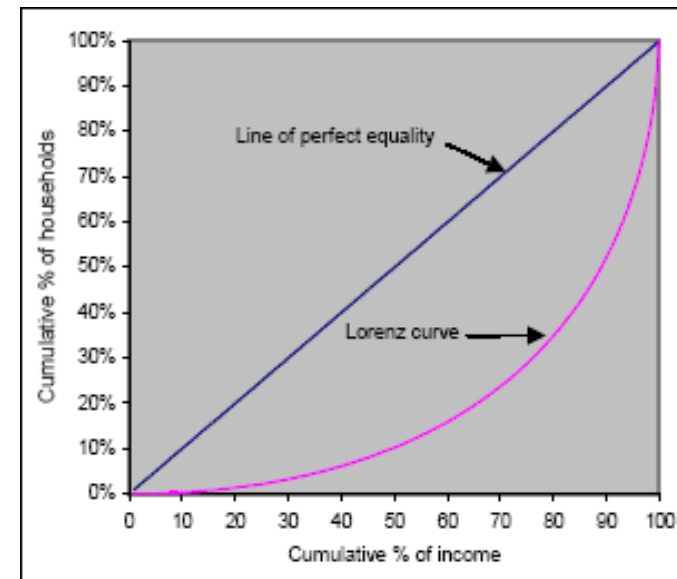
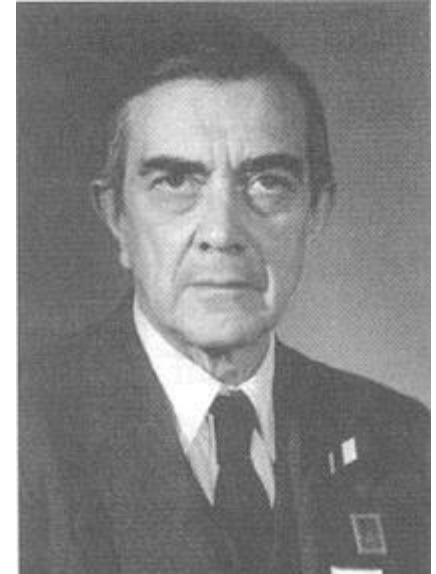
Lorenz Curves

g = Gini Coefficient



Gini Concentration Ratio

- Corrado Gini – 1912
- Measures the proportion of the total area under the diagonal that lies in the area between the diagonal and the Lorenz Curve.



Gini Concentration Ratio

- The coefficient varies between
- 0 which reflects complete equality
- 1 which indicates complete inequality
- (one person has all the income or consumption, all others have none).
- The Gini coefficient can be used to indicate how the distribution of income has changed within a country over a period of time
- It can be used to compare income distributions across different population sectors as well as countries

GINI Coefficient

$$G_i = 1 - \sum_{i=1}^N (X_i - X_{i-1})(Y_i + Y_{i+1})$$

X=Cum. Per. of Income Distribution
(i.e, households or family)

Y=Cum. Per. of Aggregate Income Distribtuion
(need to use midpoint)

Theil-index

- While less commonly used than the Gini coefficient, the Theil-index of inequality has the advantage of being additive across different subgroups or regions in the country.
- The Theil index, however, does not have a straightforward representation and lacks the appealing interpretation of the Gini coefficient.
- The Theil index is part of a larger family of measures referred to as the General Entropy class.



Theil Income Inequality T statistic

$$H_{(y)} = \sum_{i=1}^m \left\{ \left(\frac{p_i}{P} \right) * \left(\frac{y_i}{\mu} \right) * \ln \left(\frac{y_i}{\mu} \right) \right\}$$

where :

p_i is the population of the group i ,

P is the total population,

y_i is the average income in group i ,

μ is the average income across the entire population

Decomposition of the Global City Index

- The sum of the neighborhoods will equal the index for the city
- Between Income Groups (within tracts)
- Within Income Groups (Between tracts)
- For example
- .146 is the index for a city, we may ask how much of the inequality is within the tracts or between the tracts

(1) Calculate the weights for each tract P_{I_i}

(2) Calculate the between group score $A = H_y * P_{I_i} = .062$ (43%)

(3) Calculate the within group score $B = P_{I_i} * \ln \left(P_{I_i} * \left(\sum_1^5 IG / IG_i \right) \right) = .084$ (57%)

Example of Theil Income Inequality

Mid Point Average Income	\$25,000	\$45,000	\$60,000	\$75,000	\$95,000	Total Population	Average Income	Inequality Within Tract
Neighborhood 1	130	1	1	1	47	180	\$43,861	0.213
Neighborhood 2	20	40	60	40	20	180	\$60,000	0.054
Neighborhood 3	0	0	0	0	180	180	\$95,000	0.000
Neighborhood 4	180	0	0	0	0	180	\$25,000	0.000
Neighborhood 5	36	36	36	36	36	180	\$60,000	0.085

Example NH 2 - Individuals in the top two income groups contribute positive elements. Individuals in the middle income group contributes nothing to the Theil's T Statistic because the group average salary is equal to the population average. Individuals in the bottom two groups contribute negative elements

STATA Code – Part 1 – Calculate Midpoint for the Income Group

2. Household Income (In <DollarYear> Inflation Adjusted Dollars)

Universe: Households

Name: T56

Variables:

To56_001: Households:

To56_002: Less than \$10,000

To56_003: \$10,000 to \$14,999

To56_004: \$15,000 to \$19,999

To56_005: \$20,000 to \$24,999

To56_006: \$25,000 to \$29,999

To56_007: \$30,000 to \$34,999

To56_008: \$35,000 to \$39,999

To56_009: \$40,000 to \$44,999

To56_010: \$45,000 to \$49,999

To56_011: \$50,000 to \$59,999

To56_012: \$60,000 to \$74,999

To56_013: \$75,000 to \$99,999

To56_014: \$100,000 to \$124,999

To56_015: \$125,000 to \$149,999

To56_016: \$150,000 to \$199,999

To56_017: \$200,000 or More

*compute midpoint values for each income category aka as Yi.

gen mp01=5000

gen mp02=12499

gen mp03=17499

gen mp04=22499

gen mp05=27499

gen mp06=32499

gen mp07=37499

gen mp08=42499

gen mp09=47499

gen mp10=54499

gen mp11=67499

gen mp12=87499

gen mp13=112499

gen mp14=137499

gen mp15=174999

gen mp16=325000

Data and Methods

- Theil Income Inequality T statistic

$$H_{(y)} = \sum_{i=1}^m \left\{ \left(\frac{p_i}{P} \right) * \left(\frac{y_i}{\mu} \right) * \ln \left(\frac{y_i}{\mu} \right) \right\}$$

where :

p_i is the population of the group i ,

P is the total population,

y_i is the average income in group i ,

μ is the average income across the entire population

STATA Code – Part 2 – Calculate Average Income for census tracts

2. Household Income (In <DollarYear> Inflation Adjusted Dollars)

Universe: Households

Name: T56

Variables:

T056_001: Households:

T056_002: Less than \$10,000

T056_003: \$10,000 to \$14,999

T056_004: \$15,000 to \$19,999

T056_005: \$20,000 to \$24,999

T056_006: \$25,000 to \$29,999

T056_007: \$30,000 to \$34,999

T056_008: \$35,000 to \$39,999

T056_009: \$40,000 to \$44,999

T056_010: \$45,000 to \$49,999

T056_011: \$50,000 to \$59,999

T056_012: \$60,000 to \$74,999

T056_013: \$75,000 to \$99,999

T056_014: \$100,000 to \$124,999

T056_015: \$125,000 to \$149,999

T056_016: \$150,000 to \$199,999

T056_017: \$200,000 or More

*Compute total income in each category. This will allow us to calculate average income.

```
gen lo1=T056_002*mp01
```

```
gen lo2=T056_003*mp02
```

```
gen lo3=T056_004*mp03
```

```
gen lo4=T056_005*mp04
```

```
gen lo5=T056_006*mp05
```

```
gen lo6=T056_007*mp06
```

```
gen lo7=T056_008*mp07
```

```
gen lo8=T056_009*mp08
```

```
gen lo9=T056_010*mp09
```

```
gen lo10=T056_011*mp10
```

```
gen lo11=T056_012*mp11
```

```
gen lo12=T056_013*mp12
```

```
gen lo13=T056_014*mp13
```

```
gen lo14=T056_015*mp14
```

```
gen lo15=T056_016*mp15
```

```
gen lo16=T056_017*mp16
```

```
gen
```

```
INC=(lo1+lo2+lo3+lo4+lo5+lo6+lo7+lo8+lo9+lo10+lo11+lo12+lo13+lo14+lo15+lo16)
```

```
gen AVEINC=INC/T056_001
```

Data and Methods

- Theil Income Inequality T statistic

$$H_{(y)} = \sum_{i=1}^m \left\{ \left(\frac{p_i}{P} \right) * \left(\frac{y_i}{\mu} \right) * \ln \left(\frac{y_i}{\mu} \right) \right\}$$

where :

p_i is the population of the group i ,

P is the total population,

y_i is the average income in group i ,

μ is the average income across the entire population

STATA Code – Part 3 – Calculate Part 1 of the formula

*Compute Part 1 of the formula - Proportion Breakdown within Neighborhood for Income Groups.

```
gen P01=T056_002/T056_001
```

```
gen P02=T056_003/T056_001
```

```
gen P03=T056_004/T056_001
```

```
gen P04=T056_005/T056_001
```

```
gen P05=T056_006/T056_001
```

```
gen P06=T056_007/T056_001
```

```
gen P07=T056_008/T056_001
```

```
gen P08=T056_009/T056_001
```

```
gen P09=T056_010/T056_001
```

```
gen P10=T056_011/T056_001
```

```
gen P11=T056_012/T056_001
```

```
gen P12=T056_013/T056_001
```

```
gen P13=T056_014/T056_001
```

```
gen P14=T056_015/T056_001
```

```
gen P15=T056_016/T056_001
```

```
gen P16=T056_017/T056_001
```

```
gen PTOT=(P01+P02+P03+P04+P05+P06+P07+P08+P09+P10+P11+P12+P13+P14+P15+P16)
```

Data and Methods

- Theil Income Inequality T statistic

$$H_{(y)} = \sum_{i=1}^m \left\{ \left(\frac{p_i}{P} \right) * \left(\frac{y_i}{\mu} \right) * \ln \left(\frac{y_i}{\mu} \right) \right\}$$

where :

p_i is the population of the group i ,

P is the total population,

y_i is the average income in group i ,

μ is the average income across the entire population

STATA Code – Part 4 – Calculate Part 2 of the formula

*Compute Part 2 of the formula Y_i/m m =average income.

```
gen T101=mp01/AVEINC
```

```
gen T102=mp02/AVEINC
```

```
gen T103=mp03/AVEINC
```

```
gen T104=mp04/AVEINC
```

```
gen T105=mp05/AVEINC
```

```
gen T106=mp06/AVEINC
```

```
gen T107=mp07/AVEINC
```

```
gen T108=mp08/AVEINC
```

```
gen T109=mp09/AVEINC
```

```
gen T110=mp10/AVEINC
```

```
gen T111=mp11/AVEINC
```

```
gen T112=mp12/AVEINC
```

```
gen T113=mp13/AVEINC
```

```
gen T114=mp14/AVEINC
```

```
gen T115=mp15/AVEINC
```

```
gen T116=mp16/AVEINC
```


Data and Methods

- Theil Income Inequality T statistic

$$H_{(y)} = \sum_{i=1}^m \left\{ \left(\frac{p_i}{P} \right) * \left(\frac{y_i}{\mu} \right) * \ln \left(\frac{y_i}{\mu} \right) \right\}$$

where :

p_i is the population of the group i ,

P is the total population,

y_i is the average income in group i ,

μ is the average income across the entire population

STATA Code – Part 5 – Calculate Part 3 of the formula

*Compute Part 3 of the fomula $\ln(Y_i/m)$ m =average income.

```
gen T201=ln(T101)
```

```
gen T202=ln(T102)
```

```
gen T203=ln(T103)
```

```
gen T204=ln(T104)
```

```
gen T205=ln(T105)
```

```
gen T206=ln(T106)
```

```
gen T207=ln(T107)
```

```
gen T208=ln(T108)
```

```
gen T209=ln(T109)
```

```
gen T210=ln(T110)
```

```
gen T211=ln(T111)
```

```
gen T212=ln(T112)
```

```
gen T213=ln(T113)
```

```
gen T214=ln(T114)
```

```
gen T215=ln(T115)
```

```
gen T216=ln(T116)
```

Data and Methods

- Theil Income Inequality T statistic

$$H_{(y)} = \sum_{i=1}^m \left\{ \left(\frac{p_i}{P} \right) * \left(\frac{y_i}{\mu} \right) * \ln \left(\frac{y_i}{\mu} \right) \right\}$$

where :

p_i is the population of the group i ,

P is the total population,

y_i is the average income in group i ,

μ is the average income across the entire population

STATA Code – Part 6 – Calculate Part 4 of the formula

*Compute Part 4 of the formula $Parto1 * Parto2 * Parto3$.

```
gen T301=P01*T101*T201
gen T302=P02*T102*T202
gen T303=P03*T103*T203
gen T304=P04*T104*T204
gen T305=P05*T105*T205
gen T306=P06*T106*T206
gen T307=P07*T107*T207
gen T308=P08*T108*T208
gen T309=P09*T109*T209
gen T310=P10*T110*T210
gen T311=P11*T111*T211
gen T312=P12*T112*T212
gen T313=P13*T113*T213
gen T314=P14*T114*T214
gen T315=P15*T115*T215
gen T316=P16*T116*T216
```

*Create your income inequality score

```
gen theil5=(T301+T302+T303+T304+T305+T306+T307+T308+T309+T310+T311+T312+T313+T314+T315+T316)
```

OTHER STATA CODE

5. Ratio of Income in to Poverty Level

Universe: Population for whom poverty status is determined

Name: A13004

Variables:

A13004_001: Population for Whom Poverty Status is Determined:

A13004_002: Under .50

A13004_003: .50 to .74

A13004_004: .75 to .99

A13004_005: 1.00 to 1.49

A13004_006: 1.50 to 1.99

A13004_007: 2.00 and Over

Anything under 1 is in poverty

$$(A13004_002 + A13004_003 + A13004_004) / A13004_001$$

6. Health Insurance

Universe: Civilian Noninstitutionalized Population

Name: A20001

Variables:

A20001_001: Total:

A20001_002: No Health Insurance Coverage

A20001_003: With Health Insurance Coverage:

A20001_004: Public Health Coverage

A20001_005: Private Health Insurance

*percent no health insurance= $\frac{\text{Health Insurance Coverage}}{\text{Total}}$
gen nhi= A20001_002/ A20001_001

NORMALIZATION

Index

Conceptual Formula for the Index

$$(1) V_i = \left(\frac{X_j - Y_i}{Z_i - Y_i} \right)$$

V_i is the standardized index score for each observation

X_j = actual value for each observation

Y_i = minimum value in the universe of observations

Z_i = maximum value in the universe of observations

$$(2) S_i = \frac{\sum_i^N V_i}{N}$$

S_i is the index score for each grid cell

N is the number of dimensions

*gen ed_sc15=(ED_TOT-min)/(max-min)

*gen mhi_sc15=(log(mhi15)-log(min))/(log(max)-log(min))

*Income Inequality Scale - High values are bad and low values are good

*gen inc_sc15=(theil15-min)/(max-min)

*Recode income Inequality Scale - High values are good and low values are bad

*gen rinc_sc15=1-inc_sc15

*gen index01=(ed_sc15+rinc_sc15+mhi_sc15)/3

STATA Code

```
summarize theil15 ED_TOT mhi15
```

```
* Variable |      Obs      Mean   Std. Dev.    Min     Max
*-----+-----
* theil15 |      614   .2980129   .0840118   .1302825   .744298
* ED_TOT |      614   1.315936   .3371135   .5626781   2.433559
* mhi15 |      614  57267.08  26618.61    9782   196286
```

```
*Education Scale - High values are good and low values are bad
gen ed_sc15=(ED_TOT-.5626781)/(2.433559-.5626781)
```

```
*Income Inequality Scale - High values are bad and low values are good
gen inc_sc15=(theil15-.1302825)/(.744298-.1302825)
```

```
*Recode income Inequality Scale - High values are good and low values are bad
gen rinc_sc15=1-inc_sc15
```

```
*Median Income - High is good and low is bad
gen mhi_sc15=(log(mhi15)-log(9782))/(log(196286)-log(9782))
```

```
*High is good and low is bad - all variables are equal
gen indexo1=(ed_sc15+rinc_sc15+mhi_sc15)/3
```

```
*High is good and low is bad - education is 50%, inequality is 25% and income is 25%
gen indexo2=(ed_sc15*.5)+(rinc_sc15*.25)+(mhi_sc15*.25)
```

PROCEDURE

Methodology

1. Run STATA code to produce parto1.xlsx
 - We did last week
2. Work to modify shapefiles to the Saint Louis MSA
 - We did this last week
3. Run STATA code to produce parto2.xlsx
4. Join new data to shapefile for the Saint Louis MSA
 - Create a variable called keep
 - Input a value of 1
 - Make sure you clean and delete suspect data
5. Export this shapefile to a folder called “gis”
6. Run STATA code to read shapefile
 - Install the command if needed
7. Run STATA code to merge, produce final.xlsx
8. Join files in ArcMap
9. Project shapefile to UTM 15

1. Add a new variable

2. Input “1”

Add Field

Name:

keep

Type:

Short Integer

Field Properties

Alias

Allow NULL Values

Default Value

Yes

OK

Cancel

Field Calculator

Parser

VB Script

Python

Fields:

FIPS_1
theil
edtot
pov
nhi
mhi
Shape_Length
Shape_Area
keep

Type:

Number

String

Date

Functions:

Abs()
Atn()
Cos()
Exp()
Fix()
Int()
Log()
Sin()
Sqr()
Tan()

Show Codebook

keep =
1

About calculating fields

Clear

Load...

Save...

OK

Cancel

NAME	NAMLSAD	MTFCC	FUNCTAT	ALAND	AWATER	INTPTLAT	INTPTLON	FIPS	dp	cdp	sdp	ai	e	wht	blk	lat	oth	tot	pwht	pbk	FIPS_1	theil	edtot	pov	nhi	mhi	Shape_Length	Shape_Area	keep	
2154	Census Tract 2154	05020	S	11066541	0	-38.6496278	-90.3799454	29189215400	0.687401	0.352289	0.335132	0.951351	0.326	4674	126	68	205	5317	0.879097	0.023688	29189215400	0.184617	2.289173	0.017115	0.004138	193239	0.159234	0.001145	1	
2178.42	Census Tract 2178.42	05020	S	7074546	25100	-38.6323217	-90.5313350	29189217842	0.647040	0.342288	0.398133	1.639175	0.599314	5492	403	128	238	7299	0.716526	0.055902	29189217842	0.339101	2.063504	0.038285	0.062605	99810	0.126037	0.000735	1	
2153	Census Tract 2153	05020	S	4623051	0	-38.7330292	-90.2703437	29189215300	0.528499	0.444937	0.121862	0.302027	0.776602	3171	3166	620	434	7872	0.42052	0.400915	29189215300	0.371078	1.047365	0.194022	0.164116	30214	0.10121	0.000479	1	
2182.01	Census Tract 2182.01	05020	S	2917397	0	-38.5743782	-90.4343847	29189218201	0.747392	0.331504	0.427238	1.334477	0.203244	2973	22	46	84	3182	0.934310	0.060914	29189218201	0.273242	1.807205	0.026713	0.018228	85429	0.082415	0.000302	1	
2178.52	Census Tract 2178.52	05020	S	6879264	0	-38.6171991	-90.565560	29189217852	0.733565	0.327632	0.405933	1.238889	0.247675	5939	117	712	647	815523	0.103036	29189217852	0.265834	1.872303	0.019143	0.02025	111524	0.125039	0.000712	1		
2151.41	Census Tract 2151.41	05020	S	6833209	1480737	-38.6994285	-90.4906804	29189215141	0.693936	0.196923	0.484655	2.432	0.368837	3614	175	50	122	4225	0.855335	0.04142	29189215141	0.26627	1.863031	0.011999	0.037877	85994	0.134735	0.00004	1	
2149	Census Tract 2149	05020	S	5373798	0	-38.7096023	-90.3699954	29189214900	0.524393	0.357360	0.167027	0.447218	0.692441	3564	1440	1080	70	6343	0.518719	0.227022	29189214900	0.229499	1.153206	0.123534	0.161391	46507	0.096387	0.000440	1	
2178.41	Census Tract 2178.41	05020	S	4161202	17623	-38.6411430	-90.3468177	29189217841	0.656221	0.281105	0.375116	1.334437	0.393172	4427	116	60	155	5338	0.829337	0.021731	29189217841	0.250757	1.945939	0.007254	0.033533	126250	0.109684	0.000432	1	
2204.31	Census Tract 2204.31	05020	S	4952580	0	-38.4927831	-90.3316771	29189220431	0.461663	0.161285	0.300379	1.062416	0.289644	7225	271	118	8102	0.817555	0.033449	29189220431	0.359236	1.462185	0.063087	0.072349	64186	0.117604	0.000511	1		
2152.01	Census Tract 2152.01	05020	S	11598797	0	-38.6551687	-90.401528	29189215201	0.651258	0.237545	0.413713	1.741622	0.585993	4658	501	576	139	6430	0.727417	0.077916	29189215201	0.276437	2.037102	0.031203	0.058551	115739	0.155084	0.001197	1	
2178.02	Census Tract 2178.02	05020	S	6226159	0	-38.6088425	-90.5223568	29189217802	0.736075	0.327145	0.402811	1.26	0.441739	5544	153	330	131	7387	0.894657	0.020712	29189217802	0.229563	1.941524	0.008935	0.013331	105711	0.104833	0.000944	1	
2179.32	Census Tract 2179.32	05020	S	4479647	47975	-38.5550910	-90.5324396	29189217932	0.670423	0.335837	0.334585	0.996772	0.421104	4502	62	132	203	5337	0.819565	0.011617	29189217932	0.249968	1.915327	0.011446	0.02229	103139	0.100991	0.000468	1	
2158	Census Tract 2158	05020	S	3145603	0	-38.6625792	-90.3479207	29189215800	0.726842	0.302887	0.423955	1.399216	0.501766	6125	1068	182	164	8016	0.764097	0.133234	29189215800	0.402868	2.46588	0.074709	0.063869	72043	0.091643	0.000325	1	
1023	Census Tract 1023	05020	S	1262334	48882	-38.5642270	-90.284473	2901102300	0.384115	0.192700	0.191408	0.993243	0.556426	1589	97	118	73	2126	0.747413	0.045626	2901102300	0.232869	1.295878	0.165999	0.082314	53641	0.050343	0.000138	1	
1024	Census Tract 1024	05020	S	504895	0	-38.5717153	-90.2778545	2901102400	0.437935	0.192575	0.245336	1.274086	0.294901	2184	237	48	2479	0.889334	0.056503	2901102400	0.234933	1.360852	0.160449	0.072417	40105	0.042645	0.00006	1		
2173	Census Tract 2173	05020	S	2236423	0	-38.6202341	-90.341242	29189217300	0.421329	0.275466	0.145803	0.529614	0.448952	2381	103	106	48	2972	0.801144	0.034657	29189217300	0.287732	2.084135	0.069314	0.011104	83958	0.061656	0.000231	1	
2174	Census Tract 2174	05020	S	1978023	0	-38.6225316	-90.364949	29189217400	0.364949	0.211778	0.153171	0.723262	0.29781	4292	88	146	38	4821	0.890272	0.020328	29189217400	0.301963	2.40786	0.047967	0.017349	79507	0.069682	0.000205	1	
2178.51	Census Tract 2178.51	05020	S	2626209	0	-38.6324989	-90.4786158	29189217851	0.749716	0.344411	0.455775	1.778512	0.340378	2688	102	89	57	3076	0.873862	0.03318	29189217851	0.29584	1.926995	0.02006	0.011053	125091	0.091434	0.000272	1	
1945	Census Tract 1945	05020	S	1939714	0	-38.6265570	-90.3088064	2901104500	0.225597	0.096166	0.12744	1.296343	0.572697	1827	109	71	303	2280	0.718912	0.047788	2901104500	0.361037	1.889551	0.124004	0.085841	60391	0.073611	0.000201	1	
1061	Census Tract 1061	05020	S	962329	0	-38.6693860	-90.281551	2901106100	0.687796	0.293283	0.384513	1.345161	0.070215	27	1747	10	0	1784	0.915135	0.07926	2901106100	0.489916	0.779272	0.352018	0.132287	19560	0.044206	0.0001	1	
2131.02	Census Tract 2131.02	05020	S	16094099	1208882	-38.7720834	-90.4989770	29189213102	0.729817	0.078982	0.201635	2.588235	0.33664	234	31	14	20	979	0.838781	0.111111	29189213102	0.591438	1.060403	0.121984	0.017118	35643	0.198774	0.001793	1	
2136	Census Tract 2136	05020	S	2977292	0	-38.7140994	-90.4444471	29189213600	0.450504	0.295191	0.163074	0.552542	0.662693	3190	1624	235	110	5076	0.804825	0.037098	29189213600	0.276510	1.097635	0.086247	0.110471	47201	0.084455	0.000302	1	
2138	Census Tract 2138	05020	S	2397482	0	-38.6890498	-90.287482	29189213800	0.805943	0.33024	0.289703	0.802713	0.154635	144	4724	35	48	4872	0.839882	0.859121	29189213800	0.302187	0.858683	0.255287	0.135861	32309	0.085221	0.000246	1	
1055	Census Tract 1055	05020	S	1195498	0	-38.6600293	-90.2789757	2901105500	0.324162	0.115558	0.209105	0.817391	0.116338	56	2553	21	0	2647	0.821156	0.964488	2901105500	0.460034	1.179526	0.231499	0.267895	4019	0.057659	0.000124	1	
2139	Census Tract 2139	05020	S	2345344	0	-38.6745403	-90.2935362	29189213900	0.605	0.449	0.176	0.359916	0.126959	29	1603	6	27	1665	0.917417	0.962763	29189213900	0.370741	0.920653	0.454655	0.155135	23304	0.032572	0.000243	1	
1052	Census Tract 1052	05020	S	830997	0	-38.6555680	-90.2891124	2901105200	0.355941	0.136139	0.219862	1.814445	0.723127	1009	1319	87	467	183	2739	0.362383	0.481863	2901105200	0.358569	1.889231	0.104396	0.056225	52746	0.040541	0.00006	1
1053	Census Tract 1053	05020	S	960234	0	-38.6575904	-90.2803321	2901105300	0.62809	0.410120	0.217851	0.53	0.411098	445	2202	17	272	7374	0.160418	0.7636	2901105300	0.460755	1.410583	0.213603	0.092795	24484	0.048822	0.000099	1	
1054	Census Tract 1054	05020	S	721522	0	-38.6650530	-90.2895436	2901105400	0.454405	0.332544	0.222032	0.367083	0.521737	245	1645	165	165	2205	0.111111	0.746032	2901105400	0.51287	0.967374	0.328168	0.165451	26731	0.036767	0.000075	1	
2159	Census Tract 2159	05020	S	3429775	0	-38.6718259	-90.328142	29189215900	0.487705	0.21648	0.271212	1.252784	0.489638	1364	4452	168	107	6171	0.721234	0.721439	29189215900	0.324545	1.386258	0.135055	0.01392	51678	0.09116	0.000355	1	
1057	Census Tract 1057	05020	S	130529	0	-38.6577053	-90.3740494	2901105700	0.244483	0.133708	0.110775	0.628478	0.730789	3452	1182	397	327	5782	0.600159	0.191586	2901105700	0.255466	2.305436	0.388533	0.043105	30329	0.056812	0.000172	1	
1042	Census Tract 1042	05020	S	1130321	0	-38.6283945	-90.3001719	2901104200	0.231966	0.079145	0.15275	1.93	0.46562	2509	142	91	3113	0.803405	0.035014	2901104200	0.376876	1.97665	0.141825	0.050886	55000	0.05085	0.000117	1		
1036	Census Tract 1036	05020	S	1372430	0	-38.6188916	-90.2893166	2901103600	0.556306	0.36306	0.195466	0.54375	0.260281	1248	52	4	66	1382	0.903039	0.037827	2901103600	0.287786	1.567782	0.171481	0.081766	51389	0.057982	0.000142	1	
1051.98	Census Tract 1051.98	05020	S	1175837	0	-38.6495453	-90.2997487	2901105198	0.368025	0.100088	0.267397	2.60515	0.653458	2078	452	109	93													

1. Clean data
2. Export shapefile

Saving Data

Look in: gis

Name	Type
stl_ct.shp	Shapefile

Name:

Save as type:

NAME		NAMLSAD	MTFCC	FUNCTCAT	ALAND	AWATER	INTPTLAT	INTPTLON	FIPS	dp	cdp	sdp	ai	e	wht	blk	lat	oth	tot	pwht	pbik	FIPS_1	thel	edtot	pov	nhi	mhi	Shape_Length	Shape_Area	keep
9800	Census Tract 9800	05020	S	75040204	57454149	+38.6934150	-90.7426541	29183980000	0.825	0.825	0	0	0	0	13	0	0	0	13	1	0	29183980000	0	1.5	0	0	0	458916	0.008423	1
2131.02	Census Tract 2131.02	05020	S	16040409	1208882	+38.7720834	-90.4498770	29189213102	0.279817	0.279817	0.201835	2.588235	0.33684	234	31	14	0	279	0.83871	0.111111	0.29189213102	0.591438	1.008403	0.121864	0.179211	36563	0.198774	0.001793	1	
1115	Census Tract 1115	05020	S	780804		+38.6506865	-90.2251410	29510111500	0.552279	0.383378	0.168901	0.440559	0.10784	4	558	0	17	579	0.006908	0.963731	29510111500	0.276659	1.108359	0.32872	0.096718	35337	0.04408	0.000082	1	
4019.04	Census Tract 4019.04	05020	S	2333611	177215	+38.7594499	-90.0927729	17119401904	0.069811	0.193396	0.476415	2.463415	0.09196	687	0	17	708	0.970339	0	0.1719401904	0.178437	1.011429	0.072034	0.04096	85243	0.115062	0.00026	1		
1105	Census Tract 1105	05020	S	1247137		+38.6659536	-90.2176533	29510110500	0.372404	0.139466	0.232936	1.670213	0.077201	25	900	0	0	925	0.027027	0.972973	29510110500	0.262562	0.845662	0.315676	0.095135	22356	0.052234	0.000129	1	
1112	Census Tract 1112	05020	S	1047490		+38.6633743	-90.2505543	29510111200	0.350778	0.164074	0.186704	1.137931	0.241808	63	863	25	0	955	0.065969	0.903665	29510111200	0.268138	1.029605	0.467593	0.262827	26705	0.063509	0.000108	1	
5027	Census Tract 5027	05020	S	1933037		+38.5794683	-90.1272292	17163502700	0.094043	0.208857	0.351786	0.640205	0.139225	9	1007	0	45	1061	0.000403	0.949105	17163502700	0.535554	0.853774	0.048961	0.186816	14474	0.0598	0.0002	1	
1181	Census Tract 1181	05020	S	1563677		+38.6235565	-90.2538984	29510118100	0.20524	0.145197	0.060044	0.413534	0.622791	532	460	95	10	1104	0.481884	0.168667	29510118100	0.493465	1.526627	0.24116	0.174819	41786	0.059353	0.00162	1	
1267	Census Tract 1267	05020	S	2224577	383318	+38.6626816	-90.1929647	29510126700	0.327751	0.229665	0.090806	0.427083	0.408208	242	837	6	25	1110	0.218018	0.754654	29510126700	0.346207	1.058215	0.327027	0.176577	27206	0.068279	0.00027	1	
5025	Census Tract 5025	05020	S	4079376	66329	+38.5857898	-90.1164529	17163502500	0.458333	0.291687	0.168687	0.571429	0.101779	15	1120	17	0	1155	0.012987	0.969697	17163502500	0.524793	0.947514	0.333916	0.122844	25625	0.124541	0.000429	1	
5038	Census Tract 5038	05020	S	10165498	71982	+38.5429744	-90.8521267	17163503800	0.553006	0.553006	0	0	0.533498	902	105	45	93	1164	0.761824	0.088682	17163503800	0.260959	2.181524	0.064035	0.017078	95938	0.192101	0.001058	1	
1122	Census Tract 1122	05020	S	695271		+38.6550030	-90.2689147	29510112200	0.639281	0.367137	0.272144	0.741259	0.282017	110	1124	8	35	1277	0.086139	0.880188	29510112200	0.470001	1.150617	0.278956	0.131369	30125	0.054173	0.000094	1	
1114	Census Tract 1114	05020	S	911525		+38.6547267	-90.2316895	29510111400	0.44009	0.260034	0.150056	0.517375	0.079459	29	1248	0	0	1286	0.022551	0.968986	29510111400	0.218424	0.806913	0.254489	0.103444	33214	0.043476	0.000094	1	
2160	Census Tract 2160	05020	S	675447		+38.6661799	-90.3054999	29189216000	0.431052	0.245303	0.190599	1.076281	0.522947	426	725	45	98	1294	0.235212	0.580278	29189216000	0.285775	1.257656	0.176188	0.165379	39635	0.039887	0.00007	1	
9004.01	Census Tract 9004.01	05020	S	20849594	665974	+38.6463037	-90.2016216	17027900401	0.567147	0.26908	0.23075	0.736244	0	17027900401	0	62	0	1347	0.852444	0	17027900401	0.995048	1.046187	0.046154	0.046154	43608	0.786429	0.021641	1	
1062	Census Tract 1062	05020	S	1078479		+38.6802851	-90.2780862	29510106200	0.916104	0.661303	0.25072	0.379085	0.178262	37	1251	19	20	1327	0.028072	0.942728	29510106200	0.312087	0.729013	0.503058	0.091937	39000	0.046329	0.00112	1	
1072	Census Tract 1072	05020	S	655397		+38.7006180	-90.2534537	29510107200	0.492721	0.361702	0.131019	0.362229	0.054437	0	1313	0	13	1333	0	0.904996	29510107200	0.395637	0.836502	0.452075	0.14178	22361	0.040729	0.000068	1	
2141	Census Tract 2141	05020	S	1105485		+38.6770793	-90.3044709	29189214100	0.615846	0.411765	0.204082	0.495627	0.15081	48	1274	0	24	1346	0.035681	0.946508	29189214100	0.275309	1.015819	0.223626	0.145617	32098	0.044721	0.000114	1	
1202	Census Tract 1202	05020	S	579422		+38.6596094	-90.2086099	29510120200	0.631391	0.420927	0.210464	0.5	0.255608	67	1231	0	71	1372	0.048834	0.89723	29510120200	0.405424	1.002751	0.499239	0.111111	51393	0.035821	0.00006	1	
1036	Census Tract 1036	05020	S	1372430		+38.6168916	-90.2893168	29510103600	0.556306	0.36036	0.195946	0.54375	0.260281	1248	52	4	66	1382	0.903039	0.037627	29510103600	0.297786	1.567762	0.171491	0.081766	51389	0.057992	0.00142	1	
1113	Census Tract 1113	05020	S	951117		+38.6592135	-90.2403939	29510111300	0.602228	0.178815	0.421412	2.356868	0.039337	4	1393	0	8	1405	0.028487	0.991459	29510111300	0.470656	1.030691	0.325	0.052857	19107	0.040881	0.000098	1	
4024	Census Tract 4024	05020	S	1017115		+38.6718384	-90.2714789	29510108300	0.687062	0.408768	0.258294	0.831884	0.102371	83	1047	0	13	1627	0.007818	0.971571	29510108300	0.295887	1.404725	0.038866	0.140725	28125	0.047879	0.00007	1	
5004	Census Tract 5004	05020	S	2964324		+38.6303831	-90.1198244	17163500400	0.602228	0.178815	0.421412	2.356868	0.039337	4	1393	0	8	1405	0.028487	0.991459	29510111300	0.470656	1.030691	0.325	0.052857	19107	0.040881	0.000098	1	
1184	Census Tract 1184	05020	S	1448164		+38.6286783	-90.2267576	29510118400	0.055052	0	0.055052	0.761889	0.14	368	96	54	1514	0.537849	0.243065	29510118400	0.567579	0.916244	0.567896	0.028662	11769	0.055837	0.00015	1		
5011	Census Tract 5011	05020	S	2234577		+38.6023821	-90.1353464	17163501100	0.411163	0.110688	0.300465	2.714286	0.051306	11	1496	0	10	1517	0.007251	0.986157	17163501100	0.245714	1.000831	0.23797	0.182967	32414	0.074574	0.000231	1	
105	Census Tract 105	05020	S	131598028	11890923	+38.6018088	-90.5059765	17083010500	0.626993	0.250797	0.378196	1.5	0.221118	1417	44	32	33	1531	0.922539	0.028739	17083010500	0.381996	1.232727	0.196249	0.020764	51136	0.049698	0.014924	1	
1066	Census Tract 1066	05020	S	845854		+38.6627121	-90.2864515	29510106600	0.491012	0.266793	0.224219	0.840426	0.289934	35	1413	17	61	1576	0.022028	0.896574	29510106600	0.494123	0.823264	0.298325	0.198454	22188	0.051686	0.000088	1	
1123	Census Tract 1123	05020	S	851630		+38.6549037	-90.2572398	29510112300	0.474608	0.240897	0.23361	0.989349	0.249485	84	1433	0	80	1597	0.052599	0.897307	29510112300	0.458586	1.12312	0.418284	0.167815	20552	0.038656	0.000088	1	
4024	Census Tract 4024	05020	S	2603209		+38.6805394	-90.1598639	17119402400	0.579912	0.221359	0.355252	1.619421	0.473693	1121	437	11	58	1627	0.008988	0.980983	17119402400	0.541917	1.188427	0.351177	0.066131	23333	0.194621	0.000907	1	
4012	Census Tract 4012	05020	S	8112710	917763	+38.6571908	-90.1043204	17119401200	0.622666	0.347763	0.274677	0.790589	0.202515	1526	75	38	10	1647	0.028533	0.945537	17119401200	0.454201	0.919794	0.390873	0.177225	18205	0.079008	0.000307	1	
2139	Census Tract 2139	05020	S	2345344		+38.6745045	-90.2935382	29189213900	0.665	0.489	0.176	0.359918	0.120659	29	1603	6	27	1665	0.019417	0.962763	29189213900	0.370741	0.920653	0.454655	0.135135	22304	0.082572	0.000243	1	
1171	Census Tract 1171	05020	S	1083050		+38.6146101	-90.2620141	29510111700	0.289074	0.106404	0.10686	0.5875	0.631452	1137	317	76	106	1699	0.699217	0.18658	29510111700	0.284185	1.783282	0.205415	0.086884	21784	0.043257	0.000112	1	
5024.01	Census Tract 5024.01	05020	S	1862913		+38.6219122	-90.0990002	17163502401	0.46588	0.288964	0.178917	0.612245	0.095565	0	1678	0	62	1740	0	0.964368	17163502401	0.362509	0.827358	0.51092	0.056322	26599	0.067928	0.000205	1	
1271	Census Tract 1271	05020	S	1741684		+38.6505842	-90.2110922	29510127100	0.72434	0.346041	0.378299	1.09322	0.170352	96	1649	12	5	1764	0.055556	0.934807	29510127100	0.594012	0.860215	0.53066	0.145407	17183	0.054559	0.00018	1	
1076	Census Tract 1076	05020	S	1168198		+38.6848910	-90.2387722	29510107600	0.347826	0.130863	0.209764	1.519337	0.039087	13	1749	0	5	1767	0.070357	0.989813	29510107600	0.310405	0.945882	0.25863	0.139219	24554	0.05442	0.000121	1	
1051	Census Tract 1051	05020	S	862329		+38.6663968	-90.2811551</																							

STATA code to read shapefile

First, we need to export shapefile out the geodatabase

Second, install *shp2dta* if it is not installed

Third, import shapefile into geodatabase

```
clear  
set more off  
cd U:\spatdemo20\data\gis  
shp2dta using stl_ct, database(geo_stl) coordinates(uscd))
```

Run STATA code to merge & produce final.xlsx

```
clear
set more off
```

```
use "U:\spatdemozo\data\gis\geo_stl.dta"
```

```
sort FIPS
```

```
save "U:\spatdemozo\data\gis\geo_stl01.dta", replace
```

```
use "U:\spatdemozo\data\labo2\stl_parto2.dta"
```

```
sort FIPS
```

```
merge 1:1 FIPS using "U:\spatdemozo\data\gis\geo_stl01.dta"
```

```
keep if keep==1
```

```
*keep if tot17>15
```

```
*Data Normalization =x-MIN/MAX/MIN replace min and max values.
```

```
*gen ed_sc15(ED_TOT-min)/((max-min)
```

```
*gen mhi_sc15=(log(mhi15)-log(min))/(log(max)-log(min))
```

```
*Income Inequality Scale - High values are bad and low values are good
```

```
*gen inc_sc15=(theil15-min)/((max-min)
```

```
*Recode income Inequality Scale - High values are good and low values are bad
```

```
*gen rinc_sc15=1-inc_sc15
```

```
*gen index01=(ed_sc15+rinc_sc15+mhi_sc15)/3
```

```
summarize theil edtot mhi
```

* Variable	Obs	Mean	Std. Dev.	Min	Max
*-----+-----					
* theil	614	.3002117	.0847341	.1283958	.6564058
* edtot	614	1.353836	.3409433	.5843374	2.449892
* mhi	614	63109.49	28748.2	11769	209096

```
*Education Scale - High values are good and low values are bad
```

```
gen ed_sc=(edtot-.5843374)/(2.449892-.5843374)
```

```
*Income Inequality Scale - High values are bad and low values are good
```

```
gen inc_sc=(theil-.128396)/(.6564058-.128396)
```

```
*Recode income Inequality Scale - High values are good and low values are bad
```

```
gen rinc_sc=1-inc_sc
```

```
*Median Income - High is good and low is bad
```

```
gen mhi_sc=(log(mhi)-log(11769))/(log(209096)-log(11769))
```

```
*High is good and low is bad - all variables are equal
```

```
gen index01=(ed_sc+rinc_sc+mhi_sc)/3
```

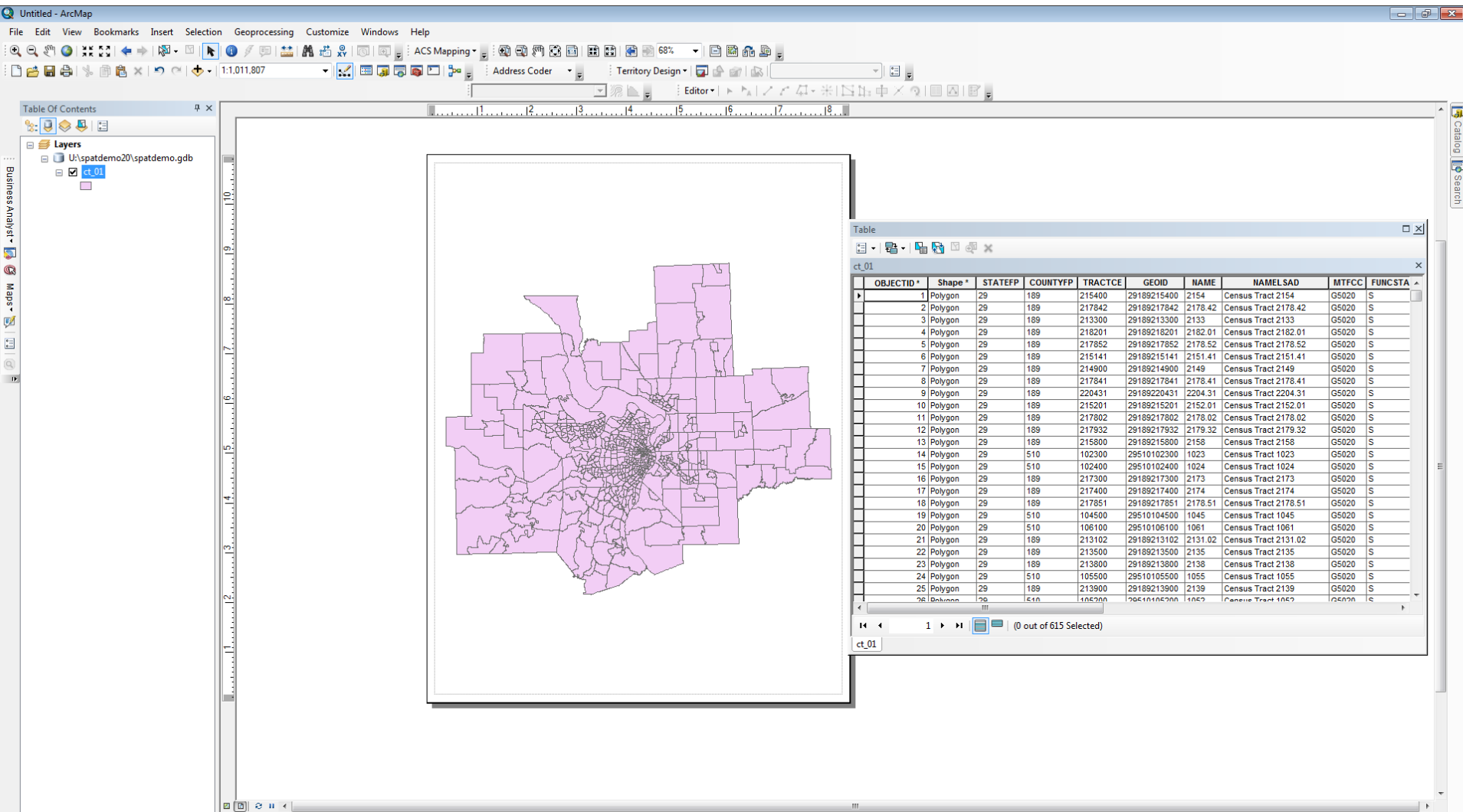
```
*High is good and low is bad - education is 50%, inequality is 25% and income is 25%
```

```
gen index02=(ed_sc*.5)+(rinc_sc*.25)+(mhi_sc*.25)
```

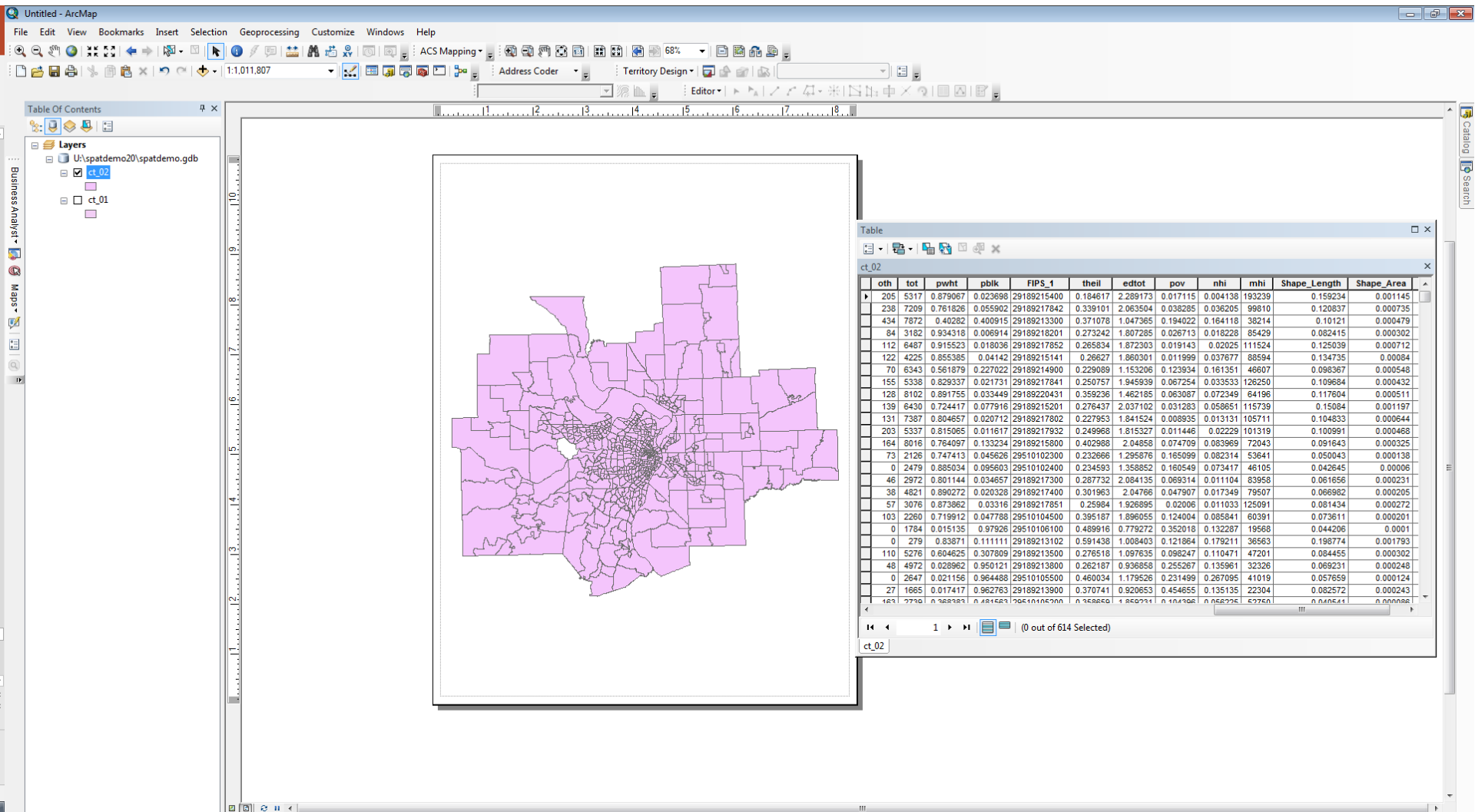
```
export excel FIPS index01 index02 using "U:\spatdemozo\data\labo2\final.xls", firstrow(variables) nolabel replace,
```

```
save "U:\spatdemozo\data\labo2\mas_stl.dta", replace
```


The shapefile - ct_01 has part 1 of the data



Join Part2 to ct_01 and make a permanent file called ct_02



Join shapefile ct_o2 to final and make a permanent file called ct_o3

Untitled - ArcMap

File Edit View Bookmarks Insert Selection Geoprocessing Customize Windows Help

ACS Mapping 68% Address Code Territory Design Editor

1:1,011,807

Table Of Contents

Layers

- U:\spatdemo20\spatdemo.gdb
 - ct_o3
 - ct_o2
 - ct_o1

Table

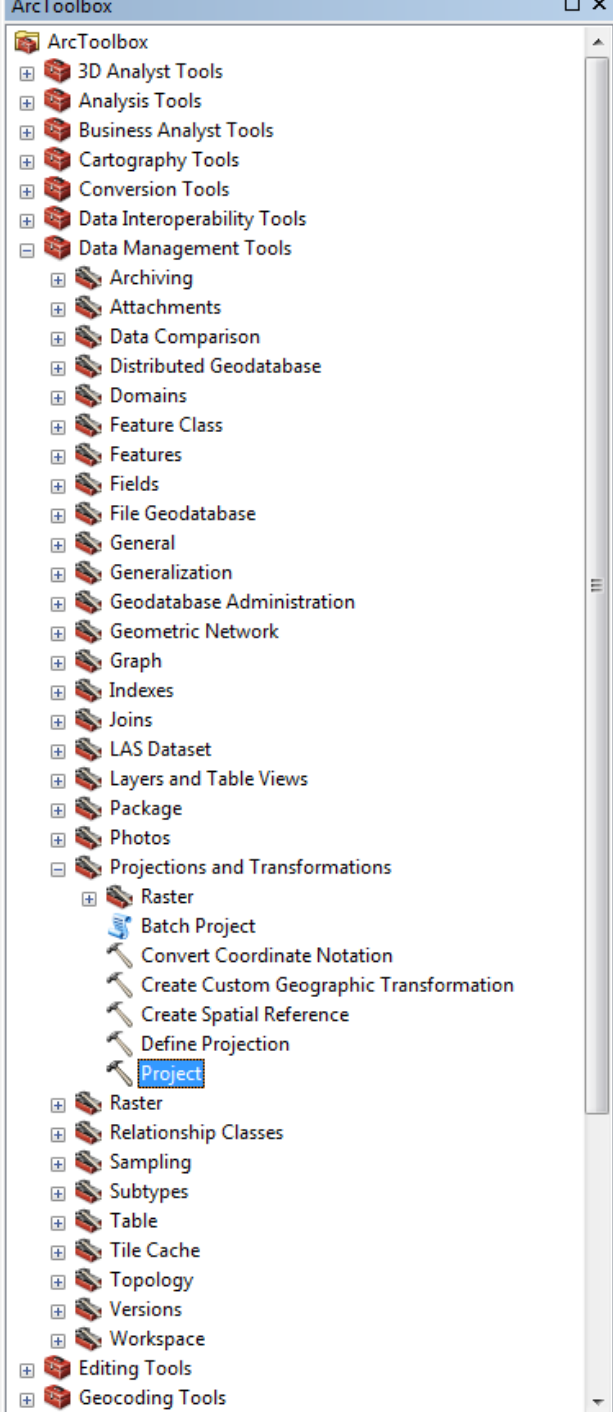
ct_o3

theil	edtot	pov	nhi	mhi	keep	FIPS_12	index01	index02	Shape_Length	Shape_Area
0.184617	2.289173	0.017115	0.004138	193239	1	29189215400	0.926854	0.923453	0.159234	0.001145
0.339101	2.063594	0.038285	0.036205	99810	1	29189217842	0.71227	0.732423	0.120837	0.000735
0.371078	1.047365	0.194022	0.164118	38214	1	29189213300	0.399299	0.361524	0.10121	0.000479
0.273242	1.807285	0.026713	0.018228	85429	1	29189218201	0.690042	0.681417	0.082415	0.000302
0.265834	1.872303	0.019143	0.02025	111524	1	29189217852	0.737216	0.72551	0.125039	0.000712
0.26627	1.860301	0.011999	0.037677	88594	1	29189215141	0.708131	0.702088	0.134735	0.000584
0.229089	1.153206	0.123834	0.161351	46807	1	29189214900	0.53085	0.474371	0.098367	0.000548
0.250757	1.945939	0.067254	0.033533	126250	1	29189217841	0.774259	0.78316	0.109884	0.000432
0.359236	1.462185	0.063087	0.072349	64196	1	29189220431	0.540989	0.523381	0.117604	0.000511
0.276437	2.037102	0.031283	0.058651	115739	1	29189215201	0.764266	0.767882	0.15084	0.001197
0.227953	1.841524	0.008935	0.013131	105711	1	29189217802	0.749429	0.730545	0.104833	0.000644
0.249968	1.815327	0.011446	0.02229	101319	1	29189217932	0.725935	0.709414	0.100991	0.000468
0.402988	2.04858	0.074709	0.083969	72043	1	29189215800	0.631504	0.669849	0.091643	0.000325
0.232866	1.295876	0.165099	0.082314	53641	1	29510102300	0.570368	0.523128	0.050043	0.000138
0.234593	1.358852	0.160549	0.073417	46105	1	29510102400	0.562865	0.52594	0.042645	0.00006
0.287732	2.084135	0.069314	0.011104	83958	1	29189217300	0.72835	0.747248	0.061656	0.000231
0.301963	2.04766	0.047907	0.017349	79507	1	29189217400	0.708538	0.726001	0.066982	0.000205
0.25984	1.926895	0.02006	0.011033	125091	1	29189217851	0.764054	0.752955	0.081434	0.000272
0.395187	1.896055	0.124004	0.085841	60391	1	29510104500	0.588738	0.617335	0.073611	0.000201
0.489916	0.779272	0.352018	0.132287	19568	1	29510106100	0.198836	0.17525	0.044206	0.0001
0.591438	1.008403	0.121864	0.179211	38563	1	29189213102	0.248107	0.242909	0.198774	0.001793
0.276518	1.097635	0.098247	0.110471	47201	1	29189213500	0.492446	0.43812	0.084455	0.000302
0.262187	0.936858	0.255267	0.135961	32326	1	29189213800	0.428912	0.368925	0.069231	0.000248
0.460034	1.179526	0.231499	0.267095	41019	1	29510105500	0.374961	0.360981	0.057659	0.000124
0.370741	0.920653	0.454655	0.135135	22304	1	29189213900	0.314495	0.28094	0.082572	0.000243
0.369860	1.840731	0.104306	0.056775	67760	1	29510105700	0.638647	0.613006	0.040641	0.000096

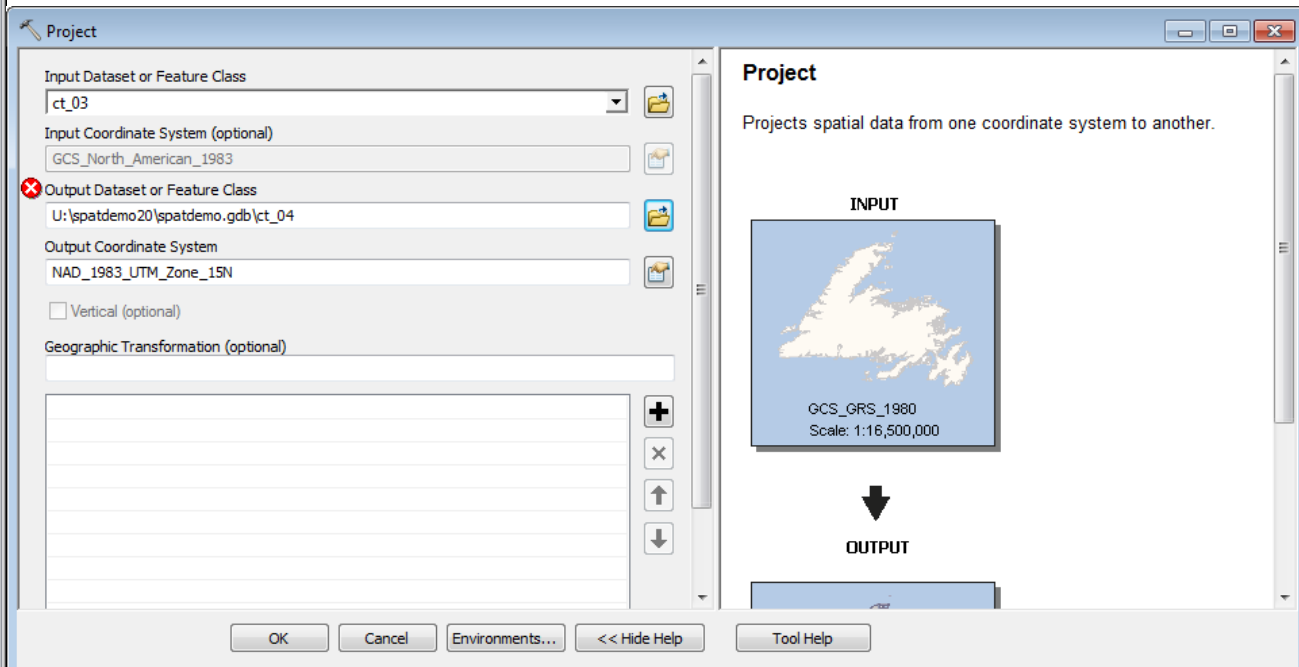
1 (0 out of 614 Selected)

ct_o3

-355 412 Inches



Project the shapefile ct_03 to ct_04 TM Zone 15



Change Data Frame Property to match projection

