



# Analyzing Student Enrollment at Clark State Community College

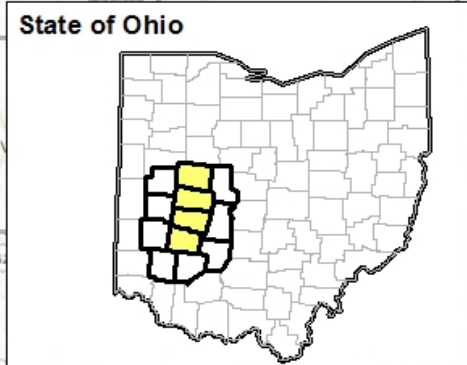
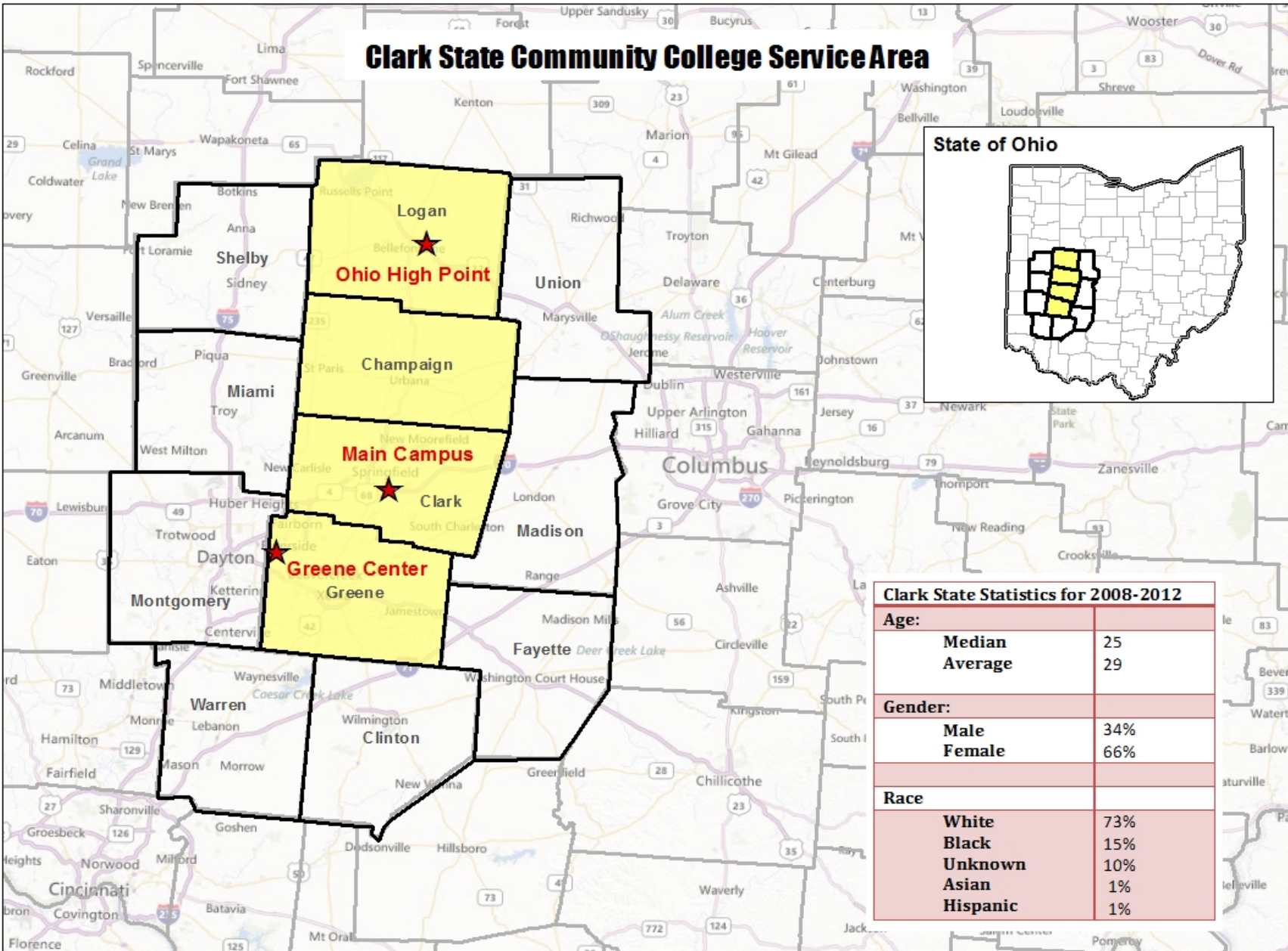
Aimee Belanger-Haas, GISP

Advisor: Stephen A. Matthews

# Outline for the session

- Background information on CSCC and motivation for the study
- Investigation of student & census data
- Exploratory spatial data analysis (ESDA) and spatial structure considerations
- Limitations & next steps

# Clark State Community College Service Area



<b>Age:</b>	
Median	25
Average	29
<b>Gender:</b>	
Male	34%
Female	66%
<b>Race</b>	
White	73%
Black	15%
Unknown	10%
Asian	1%
Hispanic	1%

# Demographic Profile of Service Area and Study Area

	Service Area (4 counties)	Study Area (12 counties)
Percent of Population 20-29	13.4%	12.4%
Median Age	39.4	38.7
Median Income	\$32,636.50	\$33,952.00
Percent White	88.3%	84.5%
Percent with no College	59.86%	62.16%
Percent Unemployment	9.54%	9.99%

# College Financial Reality

- identifying “where” to recruit students from can be an important financial strategy
- A large part of institutional revenue is generated via student tuition
- Other sources include:
  - Alumni society
  - Fundraising
  - Government funding (changing and reducing)

# Research Questions

- Based on five years of registration data at Clark State Community College:
  - Where are CSCC enrollment rates the highest and lowest?
  - What ecological factors potentially explain CSCC enrollment?

# Education is like a Business

- The retail sector embraces the use of geodemographics to increase business and profits by being more efficient in identifying potential customers
- This same methodology can be applied to Higher Educational institutions as they both have customers (students) with addresses that can be geocoded that can help uncover varying themes through their geodemographic profile

# Previous Research

- Studies have focused on higher education institutions.
- However, most have been at 4-year universities who recruit straight out of high school; this is different from CSCC
- Most institutions analyze enrollment data but do not reveal all their methods

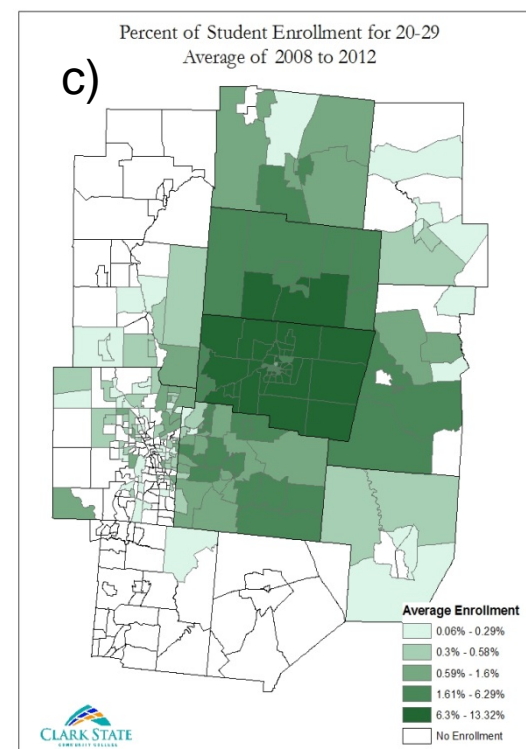
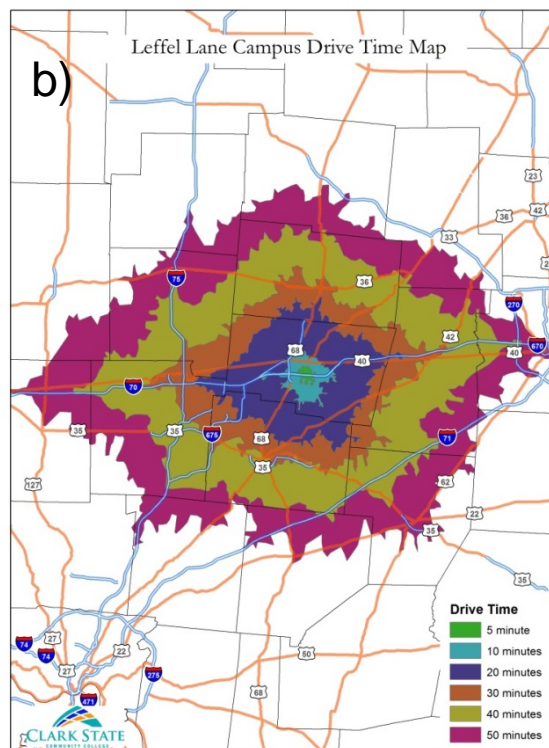
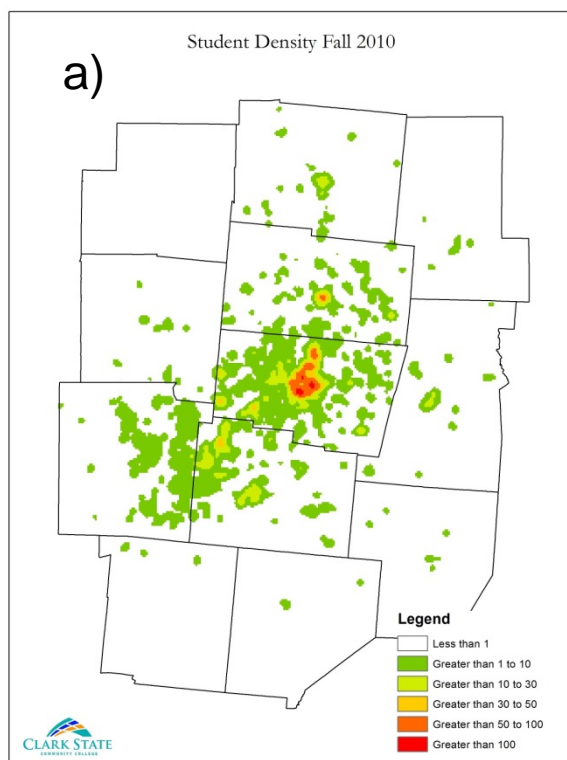


# Brief summary of packages and data-related tasks completed in this project

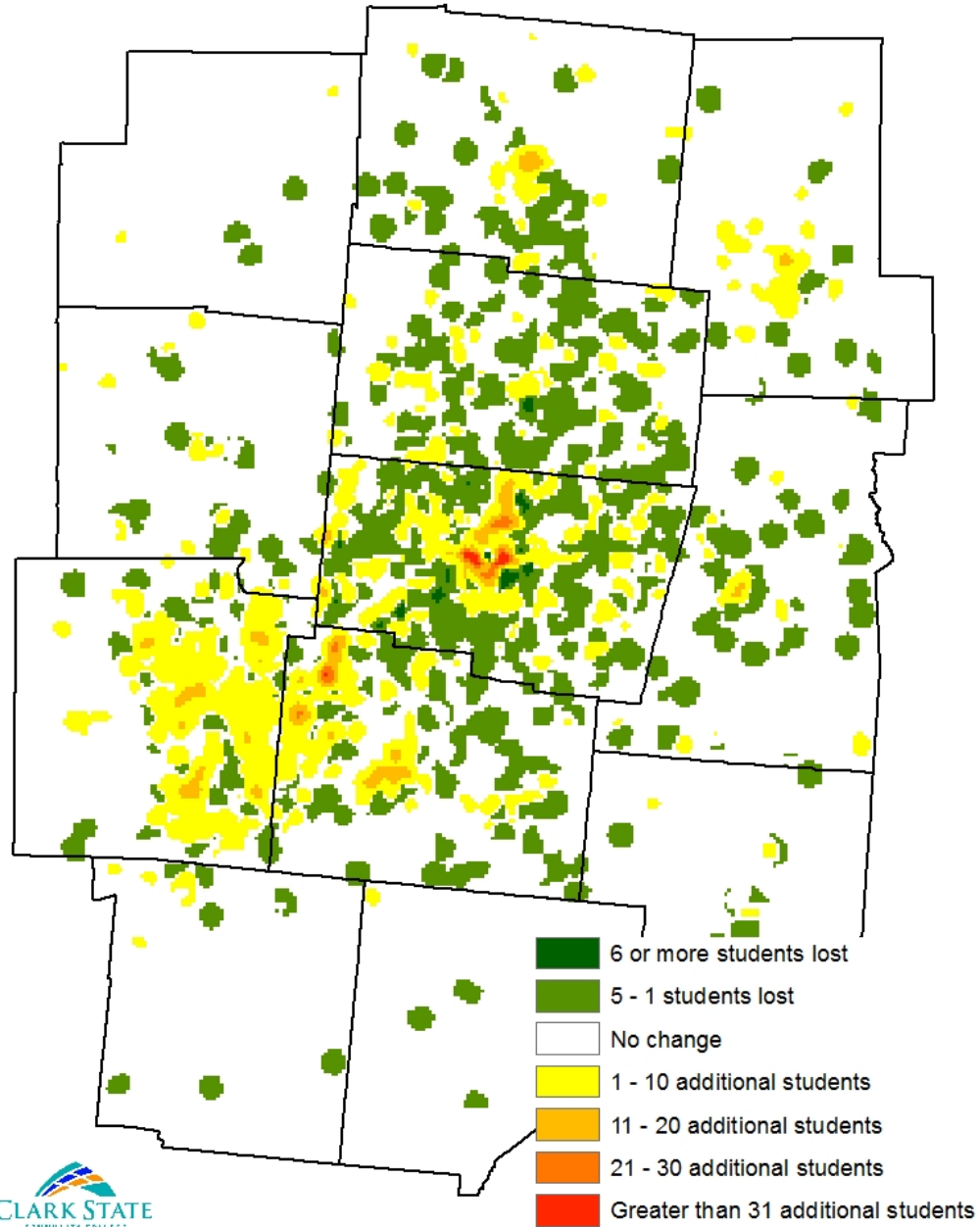
Software	Task
ArcGIS	Geocoding, locator and standard choropleth maps, and density mapping.
Excel	Data cleaning and manipulation
SPSS	Individual data analysis (summary statistics) Tract level data: univariate, bivariate (correlations, difference of means), and some OLS regression
GeoDA	Moran's I, LISA maps, OLS and spatial lag regression

# Enrollment Data, Geospatial Links and Data Aggregation

Having individual address data on the enrollees enables us to a) map distributions (density surfaces), b) generate drive-time maps and c) generate aggregate enrollment data at the census tract level.



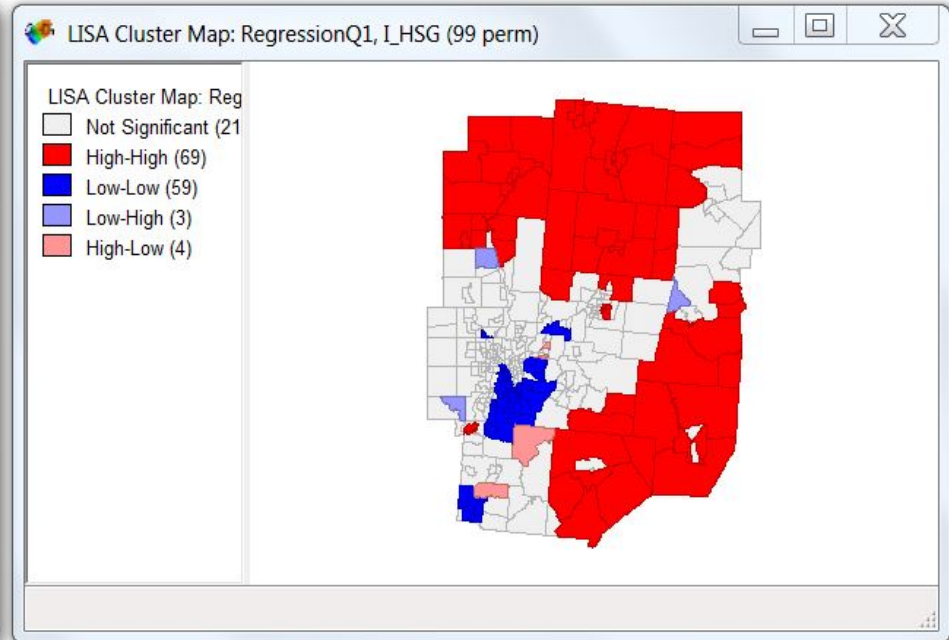
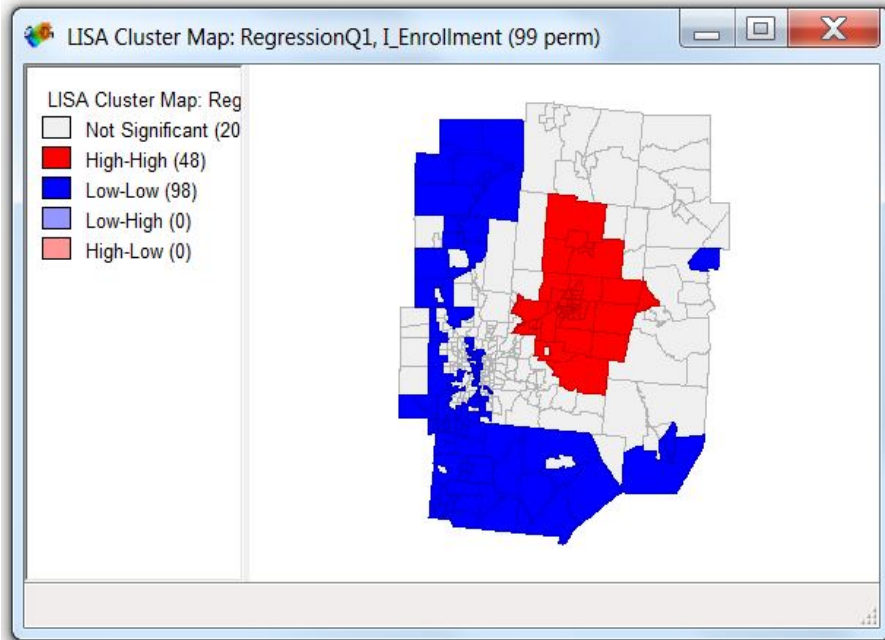
# Change in number of students Fall 2008 to Fall 2012



# Socio-economic variables and the location of students

- What kinds of relationships exist among CSCC enrollment rates and tract level factors such as
  - Educational attainment
  - Employment status
  - Median earnings
  - Poverty level
  - Race

# ESDA and Spatial Structure



Variable	Mean	Standard Deviation	Moran's I
Enrollment	0.346	0.030	0.90508
Employment	57.007	0.523	0.41538
9 to 12 no diploma	9.404	0.323	0.45234
High School Graduate	34.441	0.560	0.58251
Some College	22.501	0.303	0.27041
Median Earnings	33,952.475	543.087	0.46308
WhiteTot	81.860	1.301	0.84991
Closest Distance	68,193.960	2,493.286	0.93195
Average Distance	110,361.437	2,834.789	0.94367

# Correlation Matrix

	Enrollment	Employment	9th to 12th (no diploma)	High School Graduate	Some College	Median Earnings	Closest Distance	White Percent	Average Distance
Enrollment	1								
Employment	-0.11	1							
9th to 12th (no diploma)	0.12	-0.59	1						
High School Graduate	0.15	-0.11	0.43	1					
Some College	-0.01	-0.06	-0.09	-0.19	1				
Median Earnings	-0.15	0.51	-0.70	-0.53	-0.21	1			
Closest Distance	-0.41	0.29	-0.05	0.22	-0.32	0.20	1		
White Percent	0.00	0.52	-0.34	0.14	-0.36	0.37	0.25	1	
Average Distance	-0.40	0.27	-0.03	0.32	-0.37	0.18	0.81	0.25	1.00

Correlation is significant at the 0.05 level

Correlation is significant at the 0.01 level

# Difference of Means/ANOVA/Tukey HSD

	1	2	3	4	5	
	Mean	Mean	Mean	Mean	Mean	
Employment Rate		58.79	56.95	56.24	57.36	55.25
Unemployment Rate		8.89	11.19	11.19	9.27	10.34
Population Over 25 with Less than 9th grade		2.94	3.44	3.79	2.88	4.18
Population Over 25 with 9-12 (no diploma)		8.78	9.99	9.60	7.74	10.62
Population Over 25 that are High School Graduates		34.34	34.30	33.73	33.60	37.91
Population Over 25 with Some College**		20.46	24.40	24.89	22.33	22.31
Population Over 25 with an Associate's Degree		8.02	8.01	8.23	8.88	8.05
Population Over 25 with a Bachelor's Degree**		15.41	12.43	12.52	14.28	10.73
Population Over 25 with a Professional or Graduate Degree*		10.04	7.47	7.26	10.27	6.21
Population Over 18 Enrolled in College*		7.30	8.46	8.86	12.73	8.74
No Vehicle Available Rate		7.74	8.19	9.48	6.14	8.25
One Vehicle Available Rate**		29.22	35.20	36.61	31.49	34.67
Two Vehicles Available*		40.07	36.43	36.53	39.34	34.95
Three Or More Vehicles Available*		22.96	20.16	17.37	23.03	22.13
Median Earnings Population Over 25**		\$ 36,558.22	\$ 32,540.85	\$ 31,819.39	\$ 36,344.41	\$ 30,377.84
Percent of Population between ages of 18-64 that are Below Poverty Level**		12.06	14.64	15.98	12.11	17.37
Average Tract Distance (miles)**		25.35	21.55	18.99	18.39	12.82
Percent White**		89.11	77.29	71.67	81.25	85.25
Percent Enrollment**		.03264378	.06742010	.11976648	.32334264	1.46458567
* Significant at 95%						
** Significant at 99%						
Difference between quartiles 95%						
Difference between quartiles at 99%						

# Spatial Regression

- Ordinary Least Squares (OLS) can serve to determine if a relationship exists between CSCC enrollment and theoretically relevant predictor variables
- As tracts are not independent observations (and do not satisfy the i.i.d assumptions) – see the map of our outcome variable (enrollment) – we might also anticipate the need for a spatial regression model.
- This is tested this in GeoDa (<http://geodacenter.asu.edu>)



# Regression Analysis (n=354)

Variable	T-Statistic	Significance
Constant	1.262	0.205
Average Distance	-11.07	0.000
High School Graduate	5.90	0.000
Some College	-2.46	0.014
Earnings	2.16	0.031
% White	-0.16	0.874
Adjusted R-square	0.27	
F-Test	27.46	
Log Likelihood	-245.31	
AIC	502.83	

**OLS**

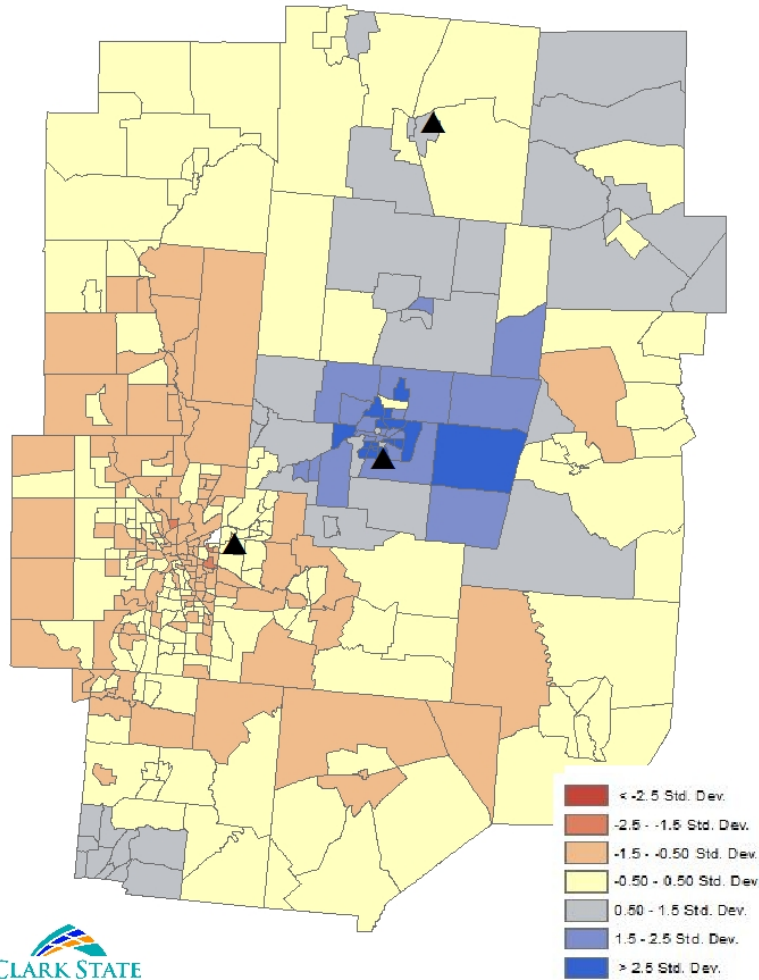
Variable	Z-Value	Probability
Constant	-1.43	0.154
Average Distance	-3.70	0.001
High School Graduate	3.292	0.001
Some College	0.642	0.521
Earnings	2.27	0.023
% White	-0.697	0.486
W_Enrollment	53.69	0.000
Adjusted R-square	0.922	
Log Likelihood	98.39	
AIC	-182.79	

**Spatial Lag**

# Ordinary Least Squares Residuals

Moran's I = 0.794

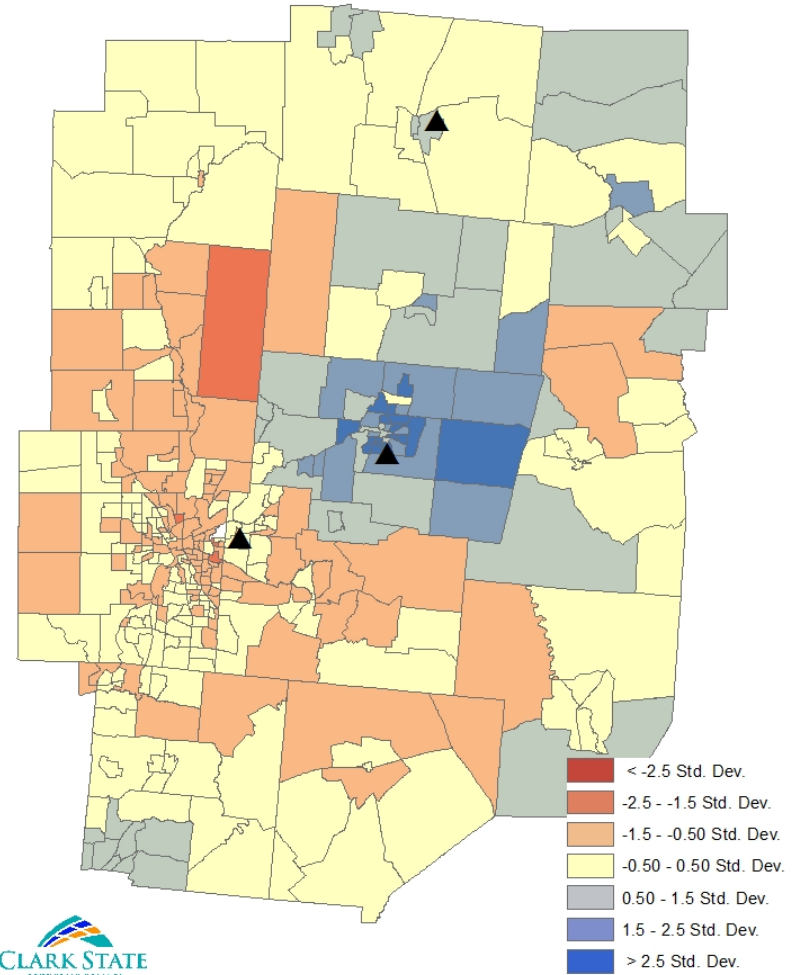
p-value: 0.561



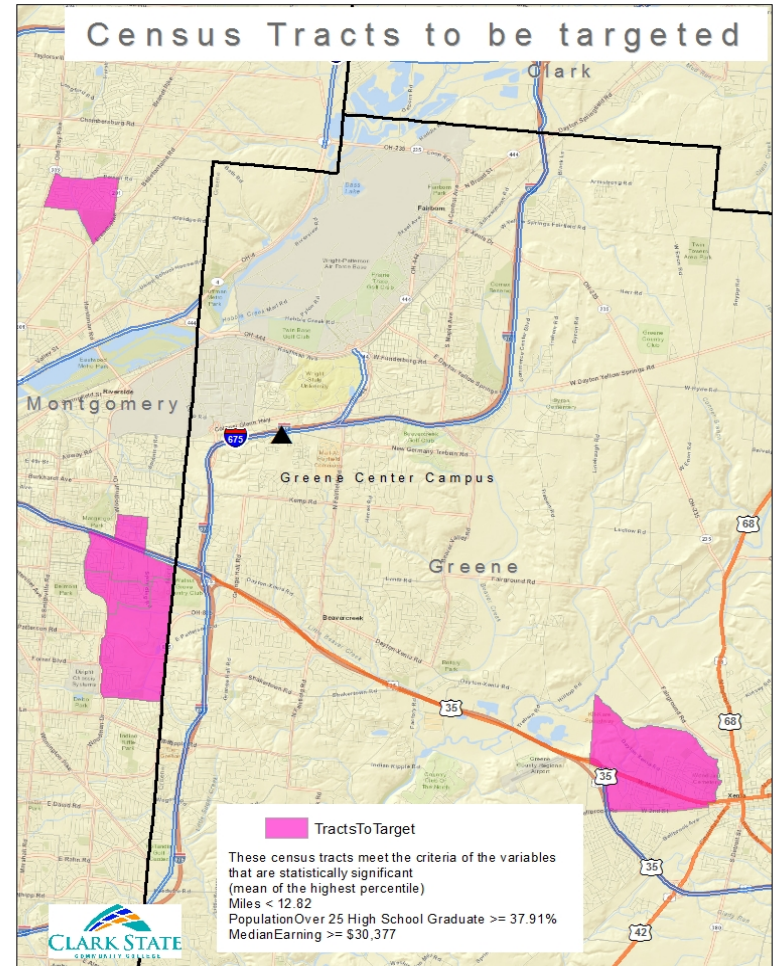
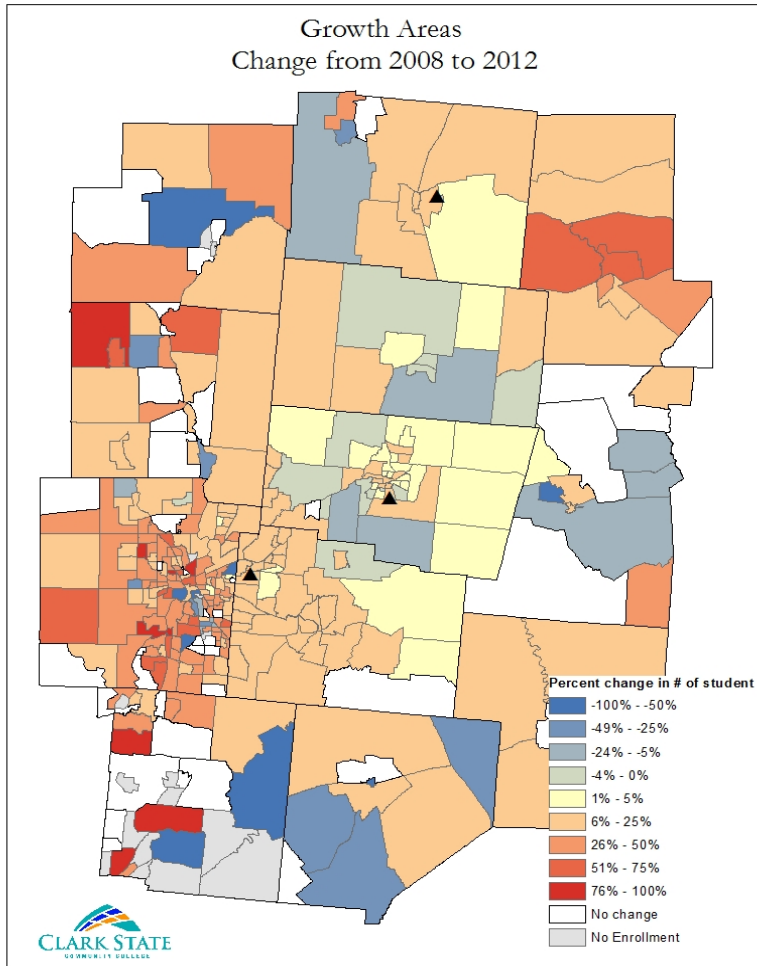
# Spatial Lag Model

Moran's I = 0.140

p-value: 0.000

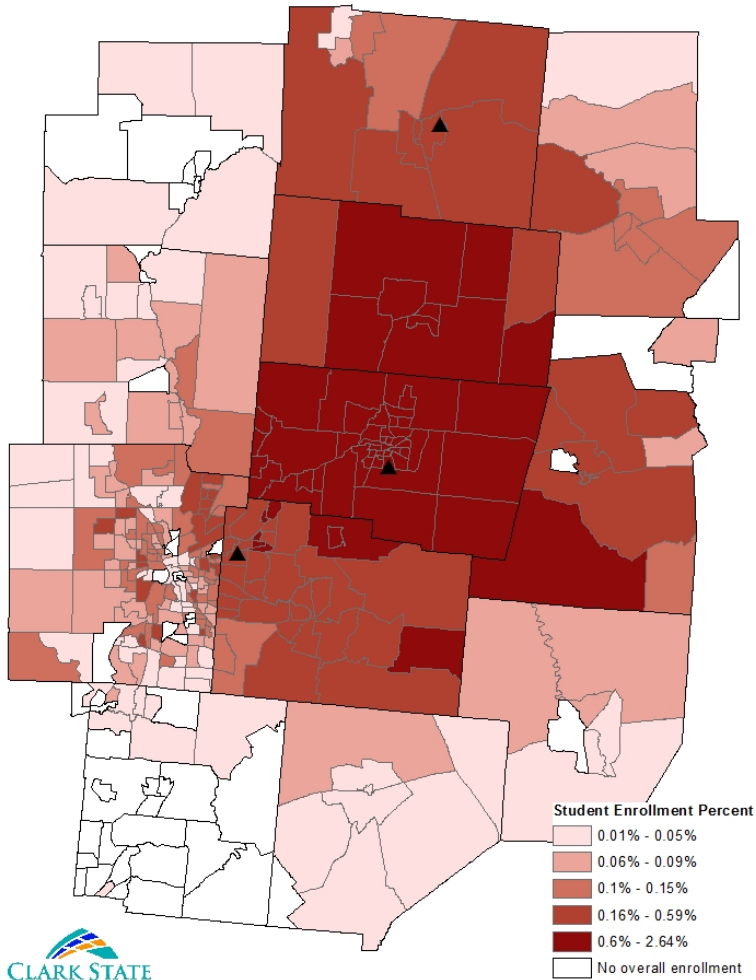


# Results

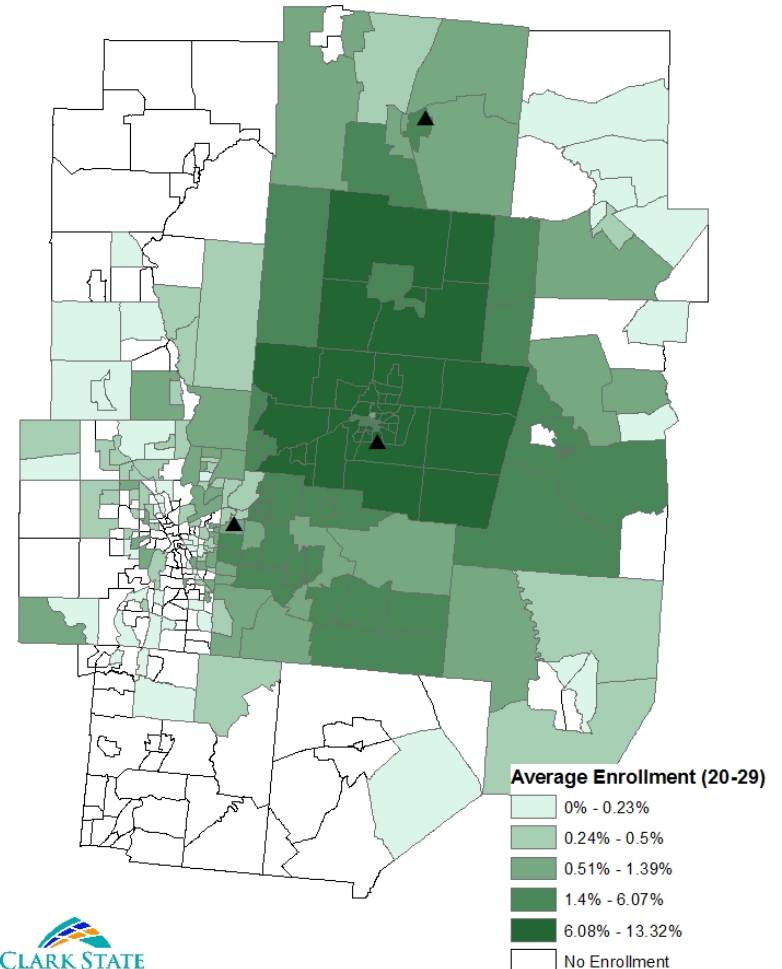


# Market Penetration

Percent of Student Enrollment  
Average of 2008 to 2012



Percent of Student Enrollment for Age Group 20-29  
Average for 2008 to 2012



# Limitations

- Analysis was limited due to the fact that the college doesn't maintain the "home" campus of each student.
- I wasn't able to acquire all variables requested from CSCC.
- No marketing information was available to correlate the success of past marketing efforts.
- The data relationships were not linear and therefore regression techniques in GIS software wasn't possible.

# Limitations #2

On the ecologic side there are limitations, including:

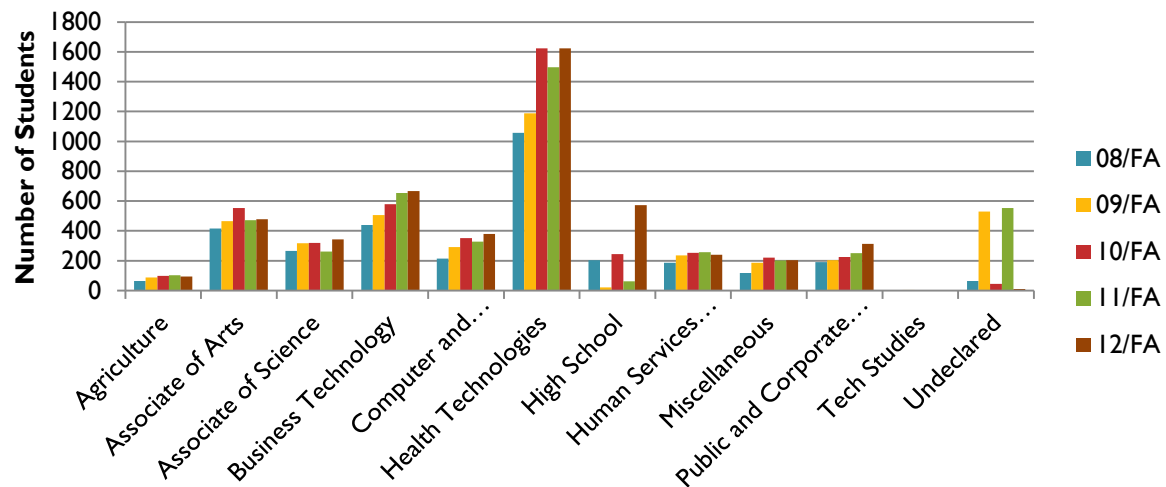
- The ACS data are an estimate and the margin of error (MOE) associated with small 'areas' can be problematic
- The assumption was made that census tracts was the relevant ecologic unit for my analysis. This is standard practice but does not preclude the possibility that MAUP issues may exist, and that other geographical units may be more relevant to the study of enrollment at CSCC.



# Next Steps/Additional Studies

- Do more analysis based on the rich CSCC enrollment date (e.g., more closely examine variation by gender, ethnicity, and major).

**Major per Academic Year**



- Also undertake analysis that also include information on other sources of education in the study area (i.e., the competition).
- Continue with regression analysis with non-linear models

# Acknowledgments

- Would like to acknowledge the following people:
  - **Advisor: Stephen Matthews** Associate Professor of Sociology, Anthropology & Demography (Courtesy Geography) at Penn State Director, Graduate Program in Demography (Dual Degree)



# References

- Adnan, M., Longley, P., Singleton, A. and Brunson, C. (2010). Towards Real Time Geodemographics: Clustering Algorithm Performance for Large Multidimensional Spatial Databases. *Transactions in GIS* , 283-297.
- Batey, P. (1999). Participation in higher education: A geodemocratic perspective on the potential for further expansion in student numbers. *Journal of Geographical Systems* , 277-303.
- Crosta, P., Leinbach, T. et al (2006). Using Census Data to Classify Community College Students by Socioeconomic Status and Community College Characteristics. *Community College Resource Center Research Tools* , 1-12.
- DesJardins, S. L. (2002). An Analytical Strategy to Assist Institutional Recruitment and Marketing Efforts. *Research in Higher Education* , 531-553.
- Hanewicz, D. C. (2012, 02 28). *Geographic Information Systems and the Political Process*. Retrieved 2012 17-10 from wpsa.research.pdx: <http://wpsa.research.pdx.edu/meet/2012/hanewicz.pdf>
- Krestle, J. (2004). Geodemographic target clusters: A case study. *Monday Report on Retailers* , 2-4.
- Lane, J. (2003). Studying Community Colleges and Their Students: Context and Research Issues. *New Directions for Institutional Research* , 51-67.
- Livingson, A. (2000, 07 16). Colleges search for applicants, with glitz and geodemographics. *The Associated Press* . New York.
- Marble, D. (1997, 07). *A Model for the Use of GIS Technology in College and University Admissions Planning*. Retrieved 2012 йил 01-10 from ESRI User Conference Proceedings: <http://proceedings.esri.com/library/userconf/proc97/proc97/to250/pap218/p218.htm>
- Marble, D. and Herries, J. (2001). A Model for the Use of GIS Technology in College and University Admissions Planning. *ESRI User Conference*, San Diego California.
- Marble, D. (1995, 07). *Applying GIS Technology to the Freshman Admissions Process*. Retrieved 2012 йил 01-10 from ESRI User Conference Proceedings: <http://proceedings.esri.com/library/userconf/proc95/to200/p182.html>
- Mora, V. (2003). Applications of GIS in Admissions and Targeting Recruiting Efforts. *New Directions for Institutional Research* , 15-21.
- Read, P., Higgs, G and Taylor, G. (2005). The potential and barriers to the use of geographical information systems for marketing applications in higher education institutions. *Marketing Intelligence & Planning* , Volume 23, 30-42.

# Questions?

Please feel free to contact me at

[belanger-haasa@clarkstate.edu](mailto:belanger-haasa@clarkstate.edu)

Aimee Belanger-Haas, GISP