Geographically Weighted Regression in Health: An Application to Cardiac Catheterization in Calgary

Oct 09, 2007

ESRI Health GIS Conference

Rizwan Shahid, *GISP, PhD (c)*

Senior Research Analyst

Calgary Health Region

APPROACH
Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease

GEOIDE

UNIVERSITY OF CALGARY

The Department of GEOGRAPHY

Outline

- Introduction
- Objectives
- Datasets
- Methodology
- Results
- Comparison of Models
- Conclusion

Cardiac Catheterization

- Heart disease is currently the leading cause of death in North America (Barnes, 2005).
- Some heart patients need to go through cardiac catheterization.
- Cardiac Catheterization involves the passage of a catheter, a thin flexible tube, into the right or left side of the heart.
- Performed to obtain diagnostic information about the heart or its blood vessels or to provide treatment in certain types of heart condition (Keller, 2004).
- There is only one hospital in Calgary Health Region which provides the facility of this procedure.

Spatial Analysis

- Spatial Dependence
- Spatial Non-Stationarity

Spatial Analysis

- Spatial Dependence
  - Data for particular spatial units are related and similar to data for other nearby spatial units in a spatially identifiable way (Getis, 1994).
  - An analysis that does not take spatial dependence into account may give spurious precision and potentially biased estimates (Elliot and Wakefield, 2000).
  - Spatial Auto Regression models deals with Spatial Dependence

Spatial Analysis *(Cont.)*

- Spatial Non-Stationarity
  - When the same stimulus provokes a different response in different parts of the study region.
  - Variation in relationships over space (Fotheringham et al. 2002).
  - Geographically Weighted Regression (GWR) takes into account Spatial Non-Stationarity.

Geographically Weighted Regression (GWR)

- Global regression model identifies the nature of relationships among variables under the assumption that the relation is universal across the study region.

- GWR extends the global model by allowing local parameters to be estimated. A weight matrix is created in which it is assumed that the observed data near a certain point, say *i*, exerts more influence in the estimation of parameters at location *i* than those further away.

Objectives

- To apply GWR to cardiac catheterization cases and socioeconomic data to analyze spatial variation of cardiac catheterization cases and probe the relationship between catheterization cases and socioeconomic variables.

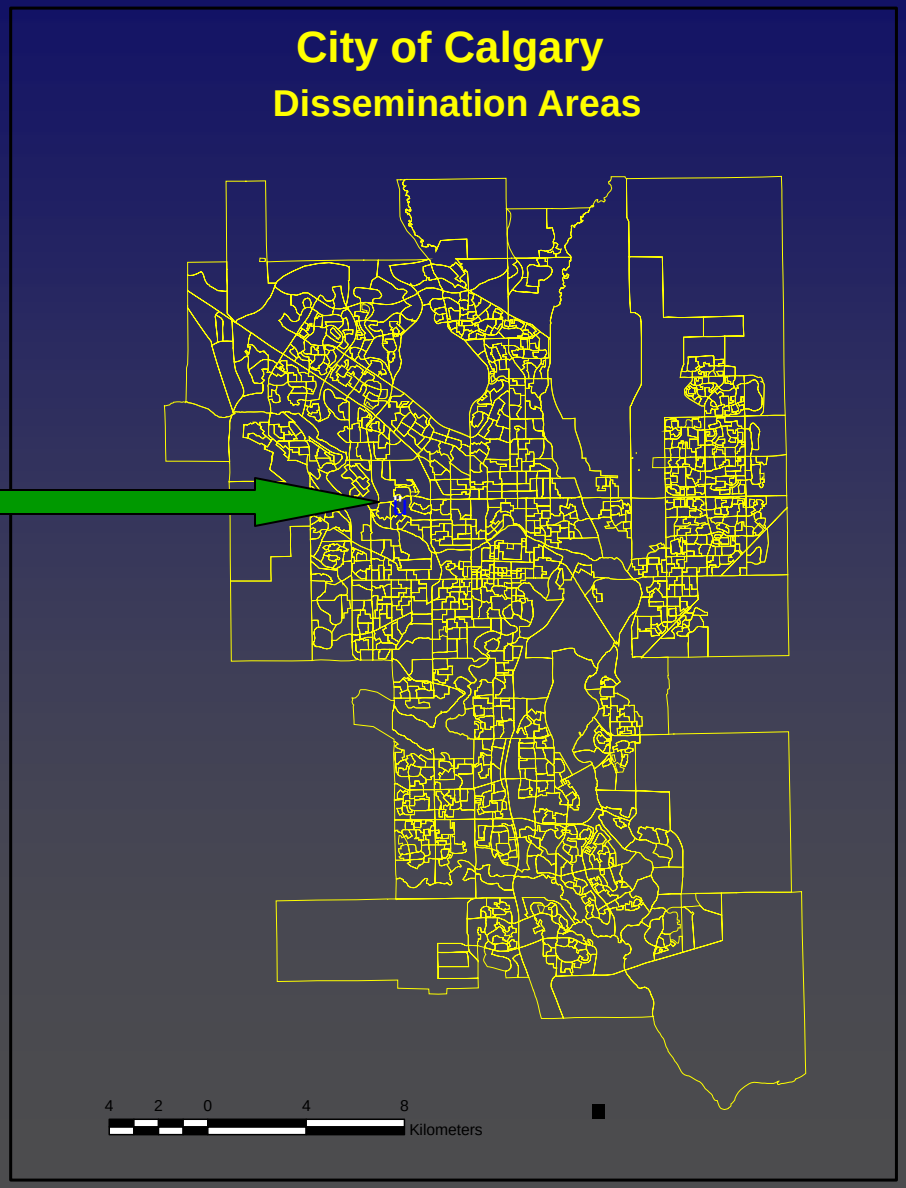
- To compare the GWR technique with standard and spatial regression in order to propose the better model, as well as to investigate the robustness of GWR in the presence of spatial dependence.

Datasets

- Cardiac Catheterization patient information
  - *Obtained from APPROACH (Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease)*
  - *Data at postal code level*
  - *Year 1998 – 2002*
- Socioeconomic Data
  - *Acquired from Statistics Canada*
  - *Data at Dissemination Area Level*
  - *Year 2001*

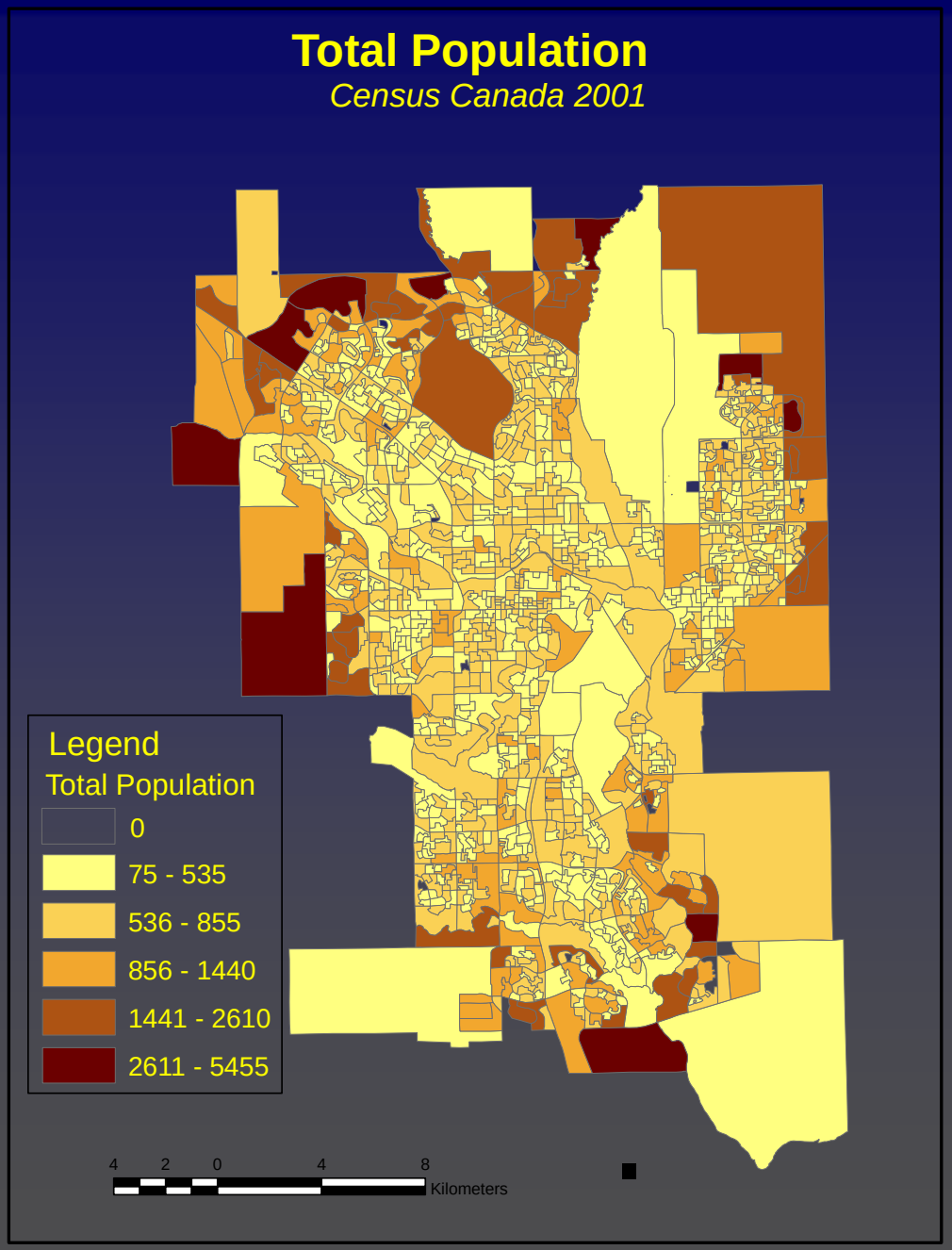
Study Area

Foothills Hospital



Total Population

Census Canada 2001



Methodology

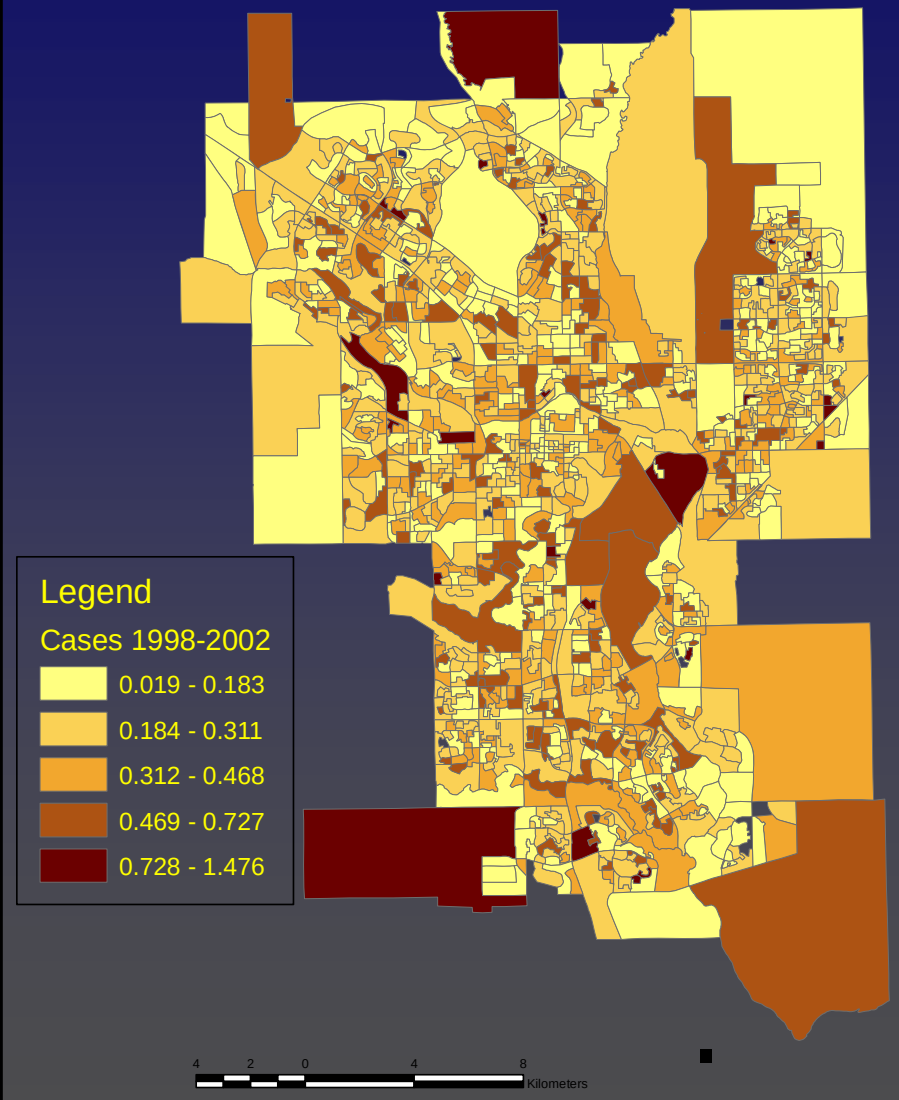
- Cardiac Catheterization patient information converted from postal code level to dissemination area level
- Variables were normalized by per 100 individuals for each dissemination area
- Correlation and cross correlation were checked and non-correlated variables were selected.
- A set of 8 significant variables was chosen

Dependent Variables

- Age 55 to 84
- Male to female Ratio
- Personals holding trade certificate
- Married and in common law
- Single parent
- Individuals having education Grade 1-13
- Individual median income
- Personals having university education with degree

Cardiac Catheterization Cases 1998-2002

Normalized by Population per Dissemination Area



Standard Linear Regression

Parameter	*Estimate*	*Std. Error*	*t*
Age 55-84	0.0091079	0.0003574	25.4804
MF-Ratio	0.1839129	0.0246775	7.4527
Trade-Certificate	0.0056455	0.0009058	6.2327
Married-Com. Law	0.0026099	0.0005455	4.7841
Single-Parent	0.0060093	0.0014476	4.1512
Ind-Med-Income	-0.0000022	0.0000006	-3.8657
Grade 1-13	0.0019050	0.0005850	3.2563
Univ-With-Degree	0.0011726	0.0004892	2.3971

Diagnostic Information:

Residual Sum of Squares : 23.078

Akaike Information Criterion : -1610.656

Coefficient of Determination : 0.388

Adjusted R – Square : 0.384

Spatial AutoRegression (SAR)

Parameter	Estimate	Std. Error	t
Age 5-84	0.0091059	0.0003570	25.5066
MF-Ratio	0.1833270	0.0246496	7.4373
Trade-Certificate	0.0056463	0.0009047	6.2408
Married-Com. Law	0.0026084	0.0005449	4.7868
Single-Parent	0.0060015	0.0014464	4.1491
Ind-Med-Income	-0.0000023	0.0000006	-3.8867
Grade 1-13	0.0019193	0.0005844	3.2845
Univ-With-Degree	0.0011687	0.0004886	2.3919

Diagnostic Information:

Residual Sum of Squares : 23.023

Akaike Information Criterion : -1619.01

Coefficient of Determination : 0.389

Adjusted R – Square : 0.385

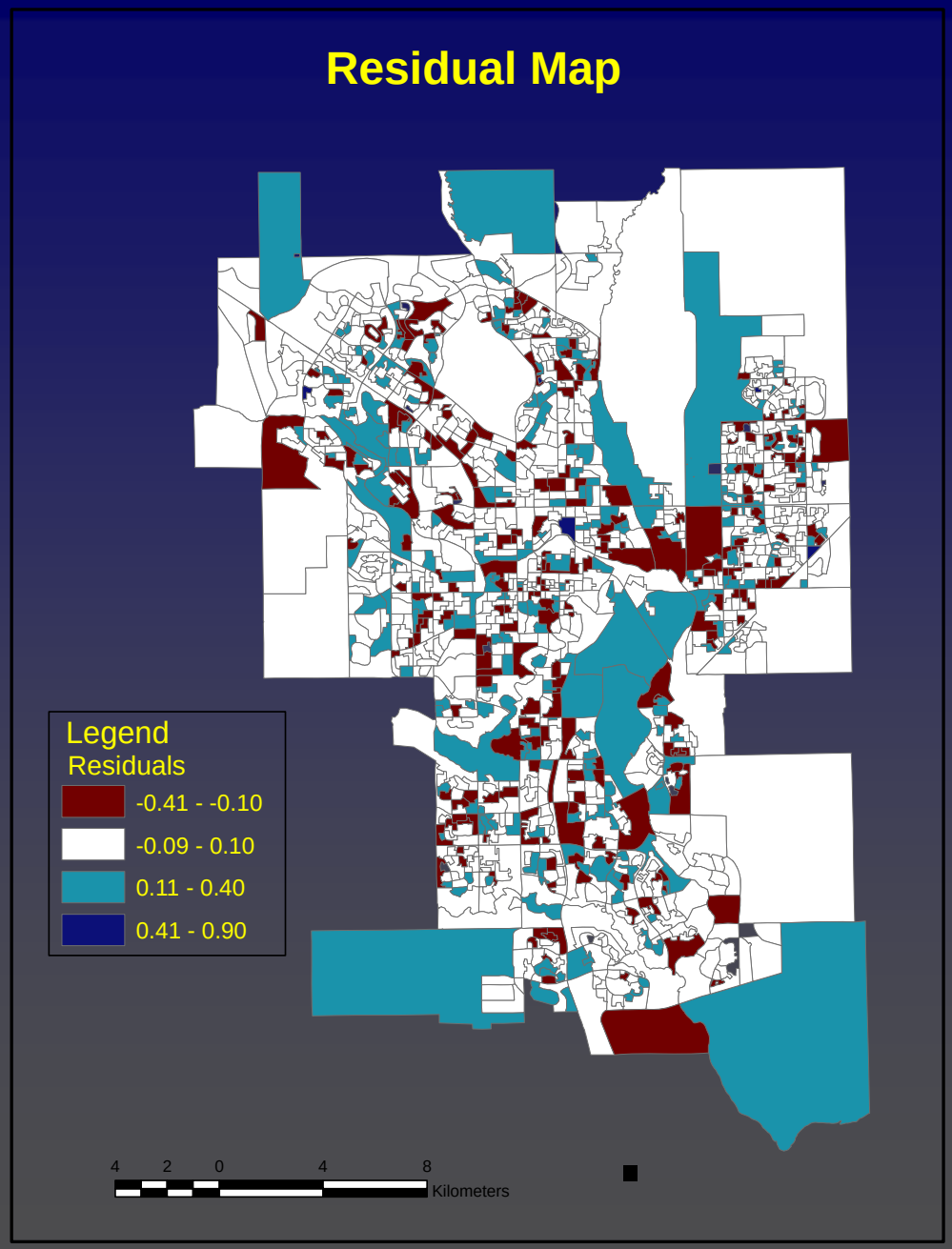
GWR – Parameter 5 Number Summaries

Label	Minimum	Lower Quartile	Median	Upper Quartile	Maximum
Age 55-84	0.006237	0.007948	0.009626	0.010505	0.012708
MF-Ratio	-0.161990	0.083884	0.150172	0.206647	0.387371
Trade-Certificate	-0.002386	0.002595	0.004766	0.007564	0.011498
Married-Com. Law	-0.002359	0.000556	0.002227	0.003337	0.007683
Single-Parent	-0.003528	0.000800	0.004983	0.007914	0.019415
Ind-Med-Income	-0.000010	-0.000002	-0.000001	0.000000	0.000002
Grade 1-13	-0.003424	0.000806	0.002779	0.003527	0.006294
Univ-With-Degree	-0.002296	-0.000499	0.000694	0.002508	0.003668

Diagnostic Information:

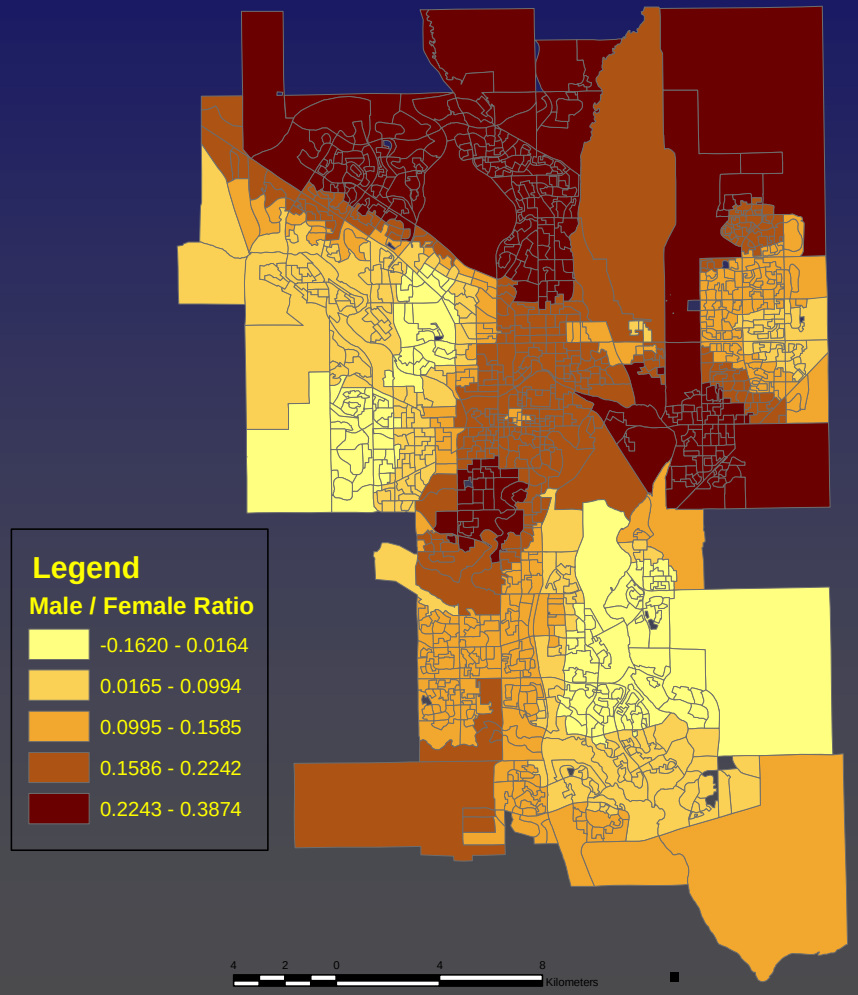
Residual Sum of Squares	:	18.842
Akaike Information Criterion	:	-1650.034
Coefficient of Determination	:	0.501
Adjusted R – Square	:	0.454

Residual Map

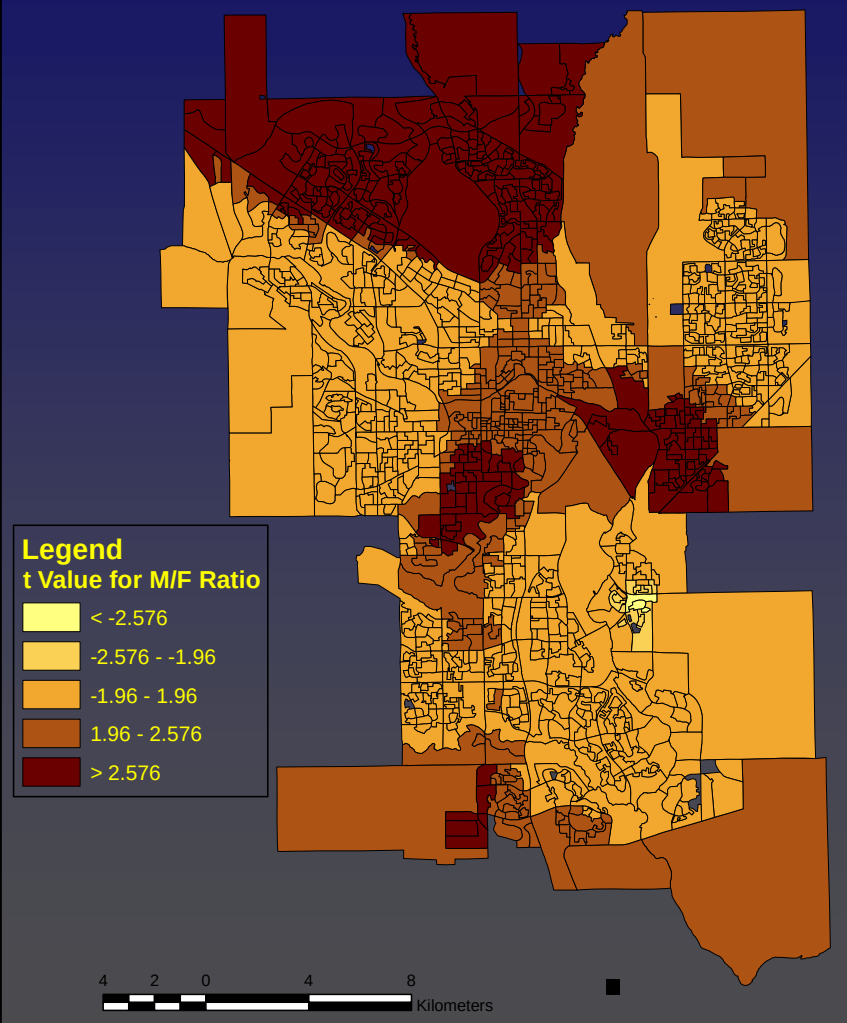


Local Parameter Estimates for Male to Female Ratio

Geographically Weighted Regression by using Adaptive Kernel with 255 Nearest Neighbors

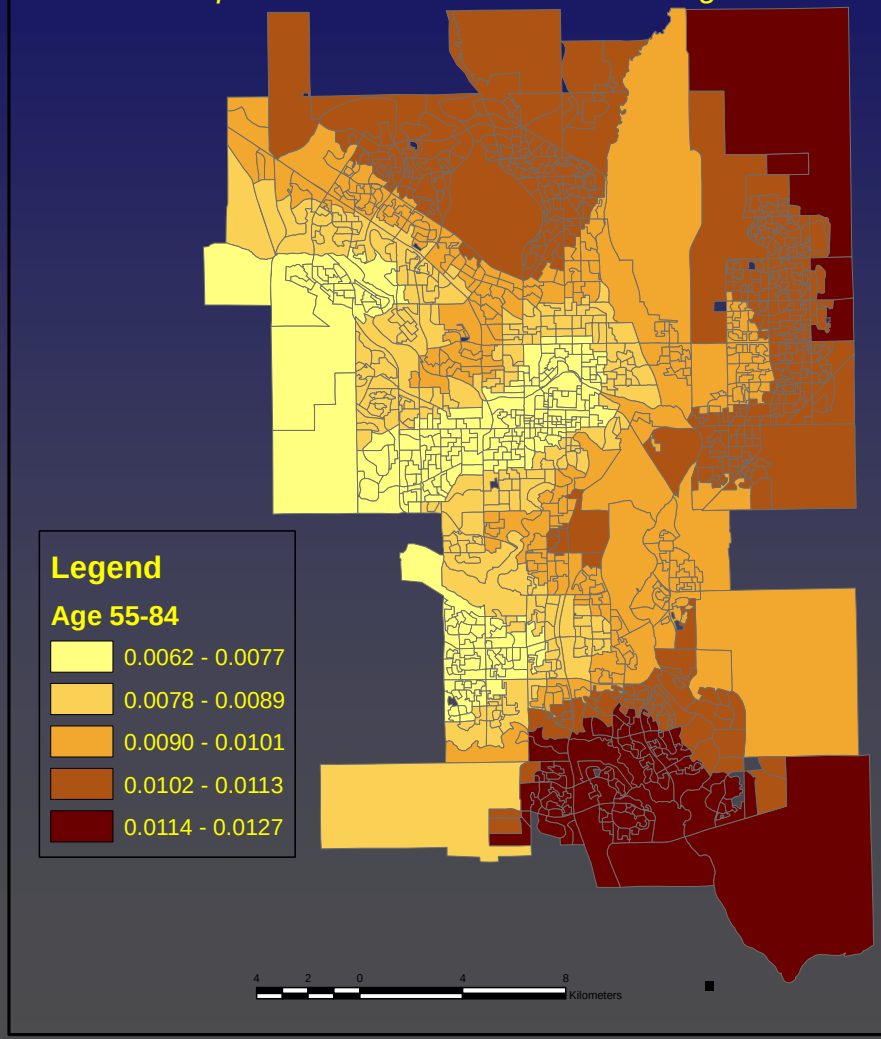


t-Value Map for Male to Female Ratio Parameter

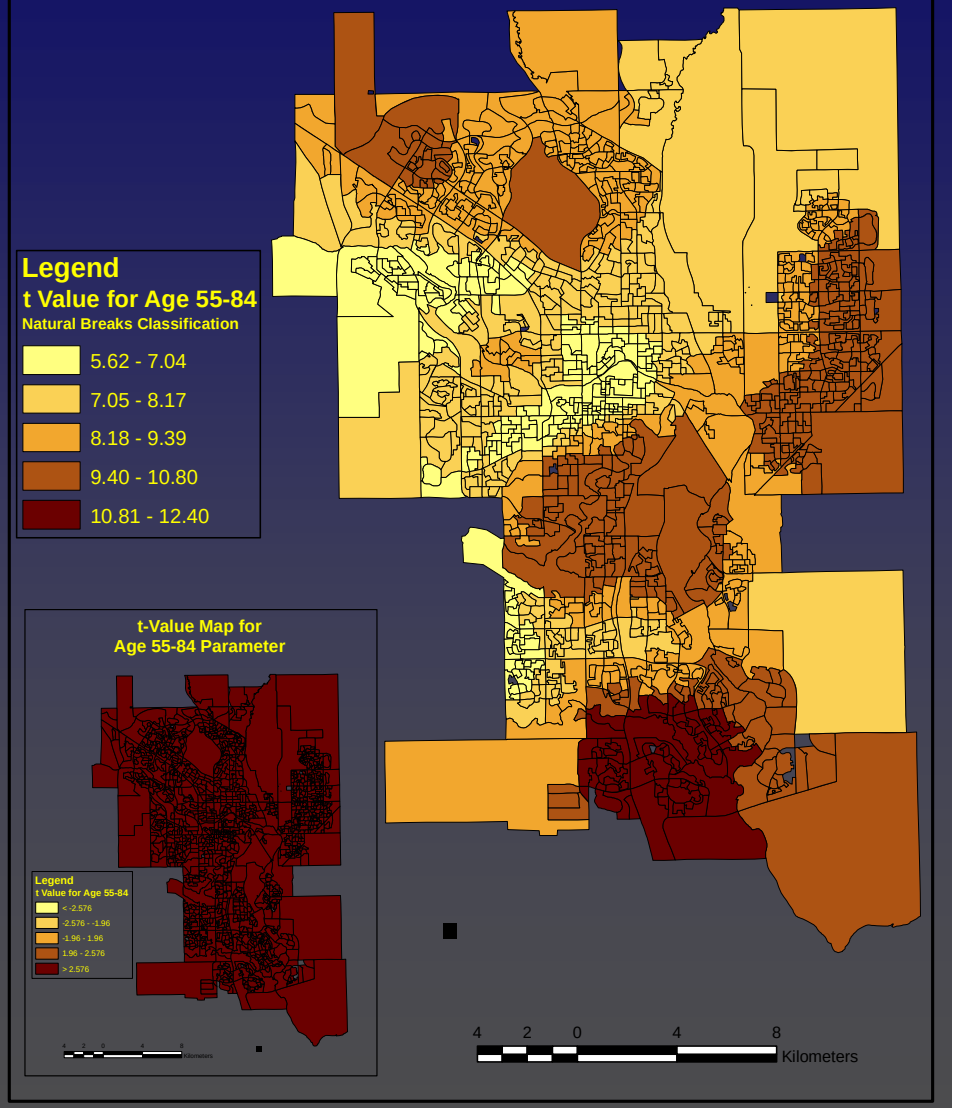


Local Parameter Estimates for Age 55-84

Geographically Weighted Regression by using Adaptive Kernel with 255 Nearest Neighbors

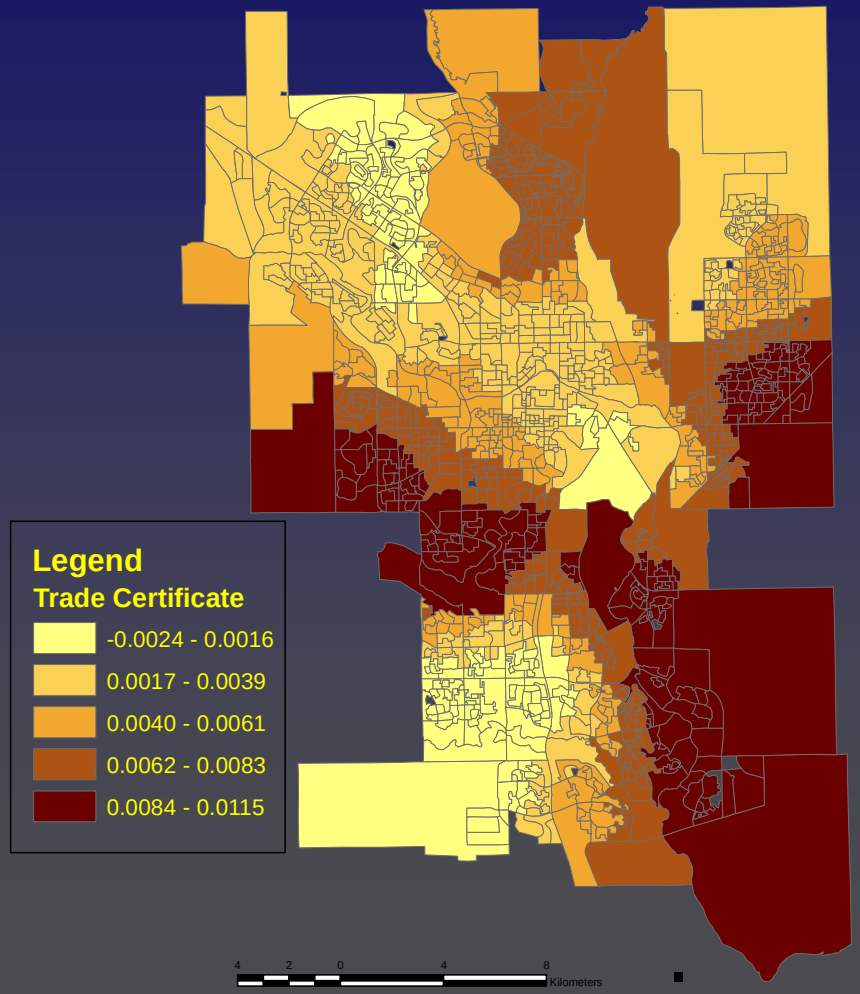


t-Value Map for Age 55-84 Parameter

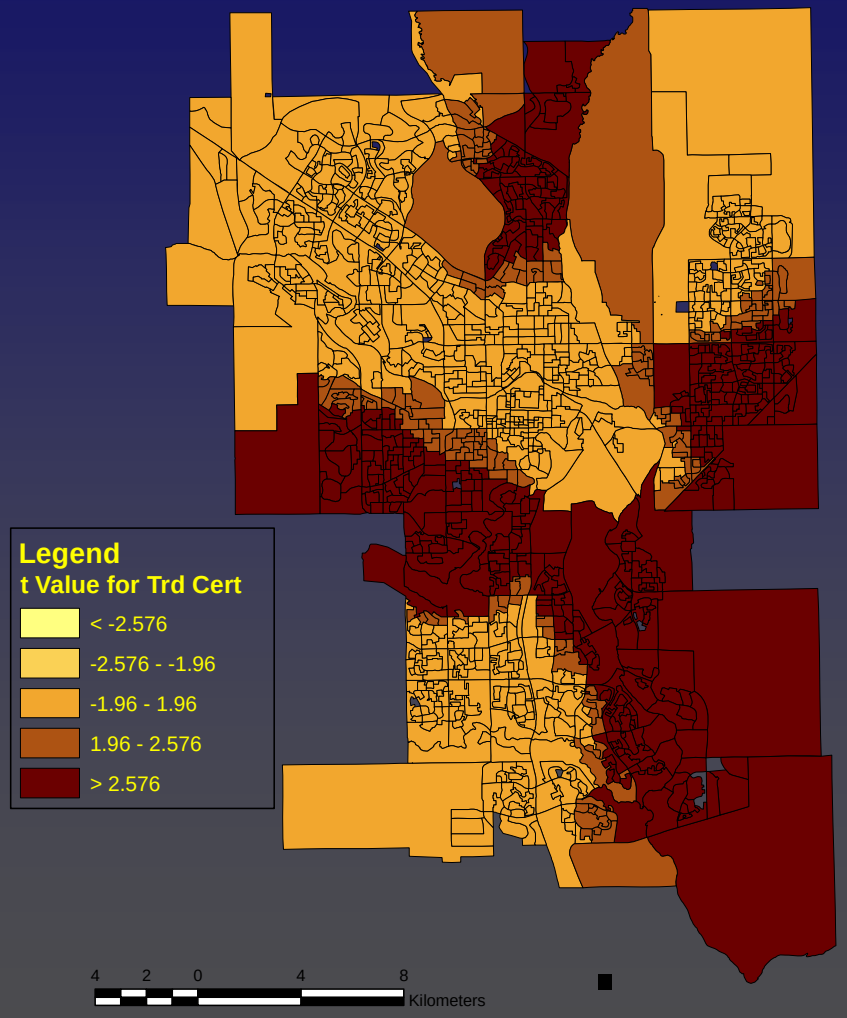


Local Parameter Estimates for Education (Trade Certificate)

Geographically Weighted Regression by using Adaptive Kernel with 255 Nearest Neighbors

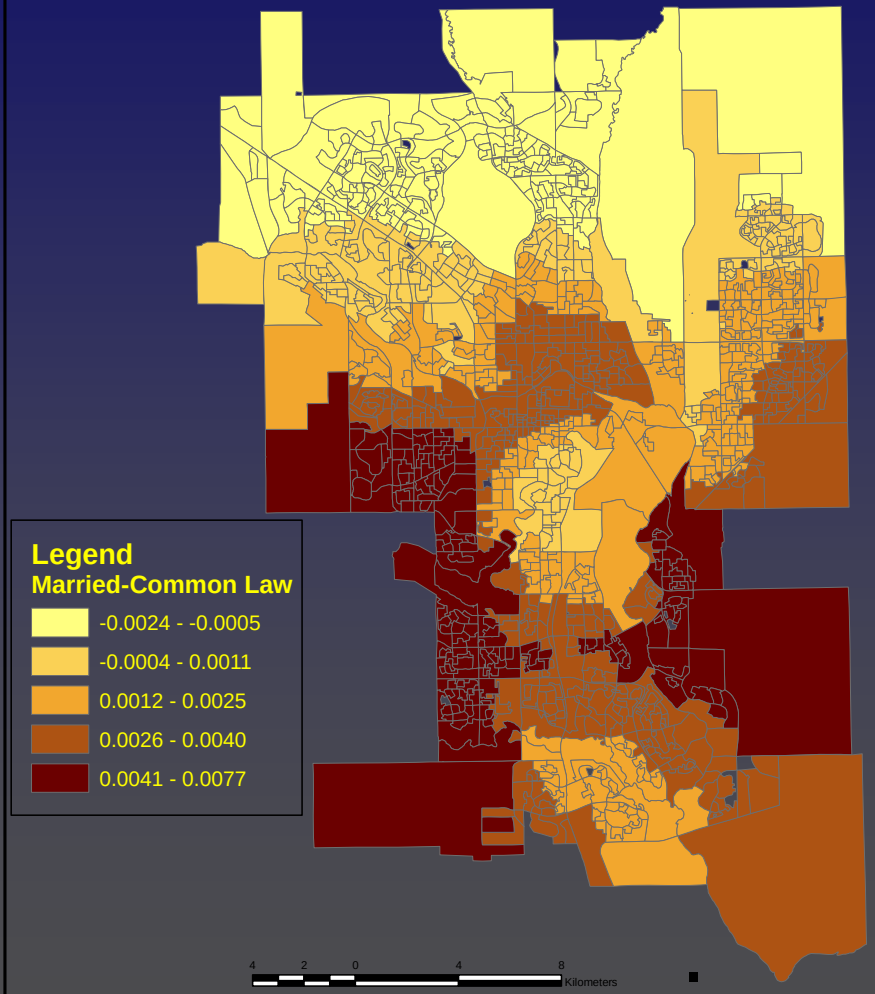


t-Value Map for Education (Trade Certificate) Parameter

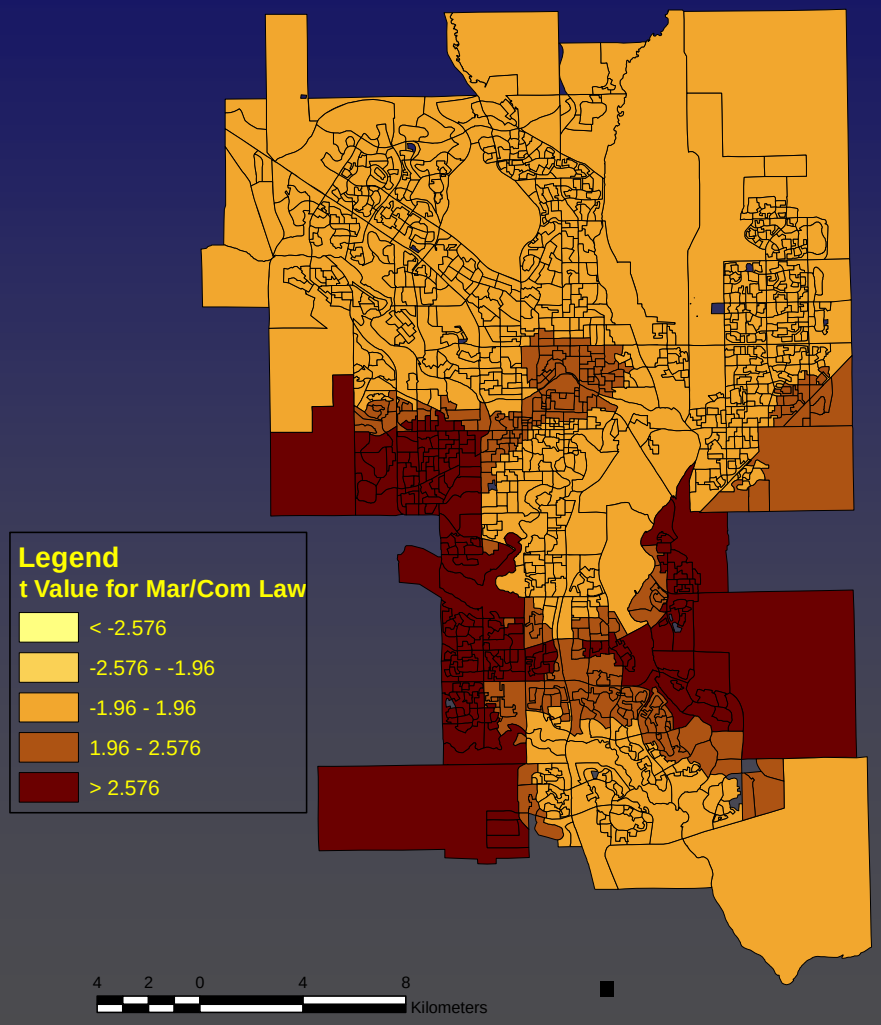


Local Parameter Estimates for Married and Common Law

Geographically Weighted Regression by using Adaptive Kernel with 255 Nearest Neighbors

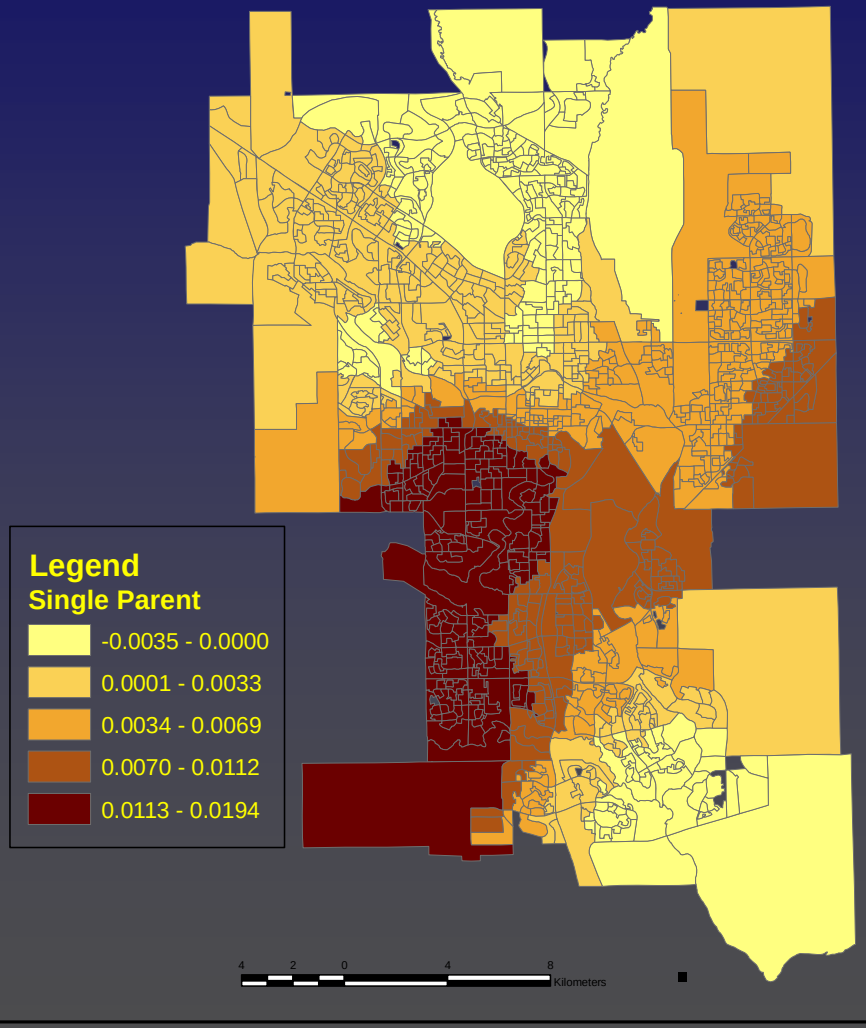


t-Value Map for Married and Common Law Parameter

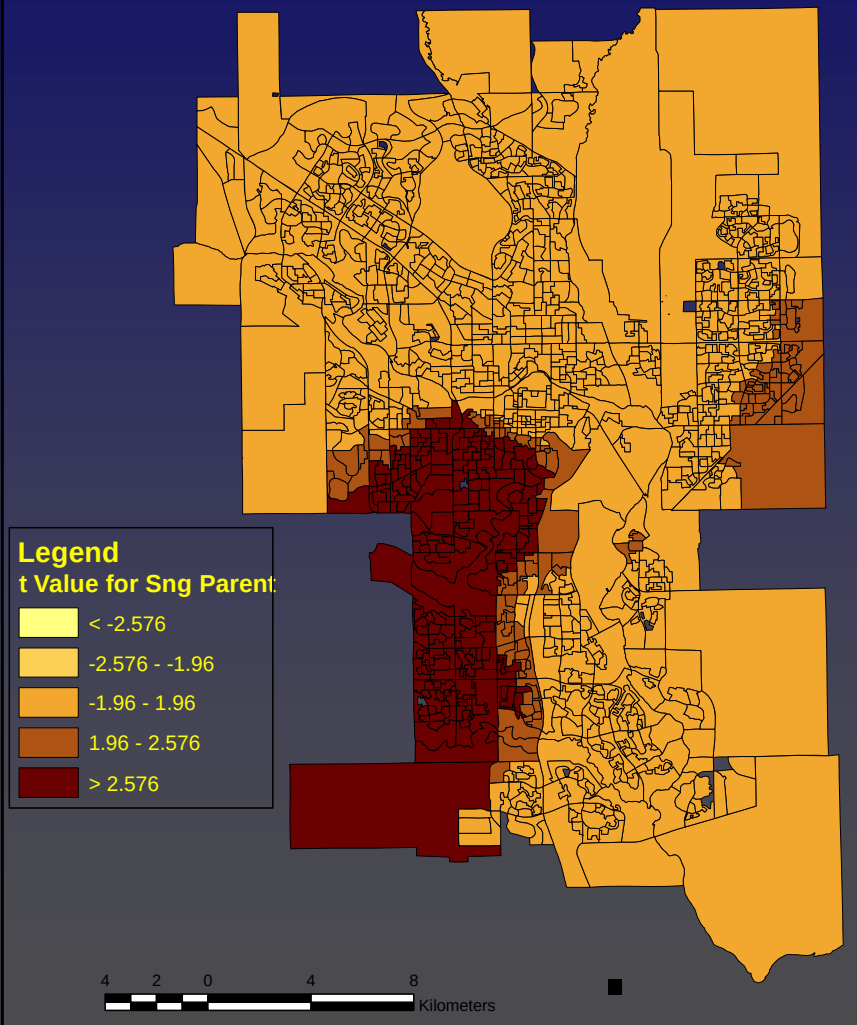


Local Parameter Estimates for Single Parent

Geographically Weighted Regression by using Adaptive Kernel with 255 Nearest Neighbors

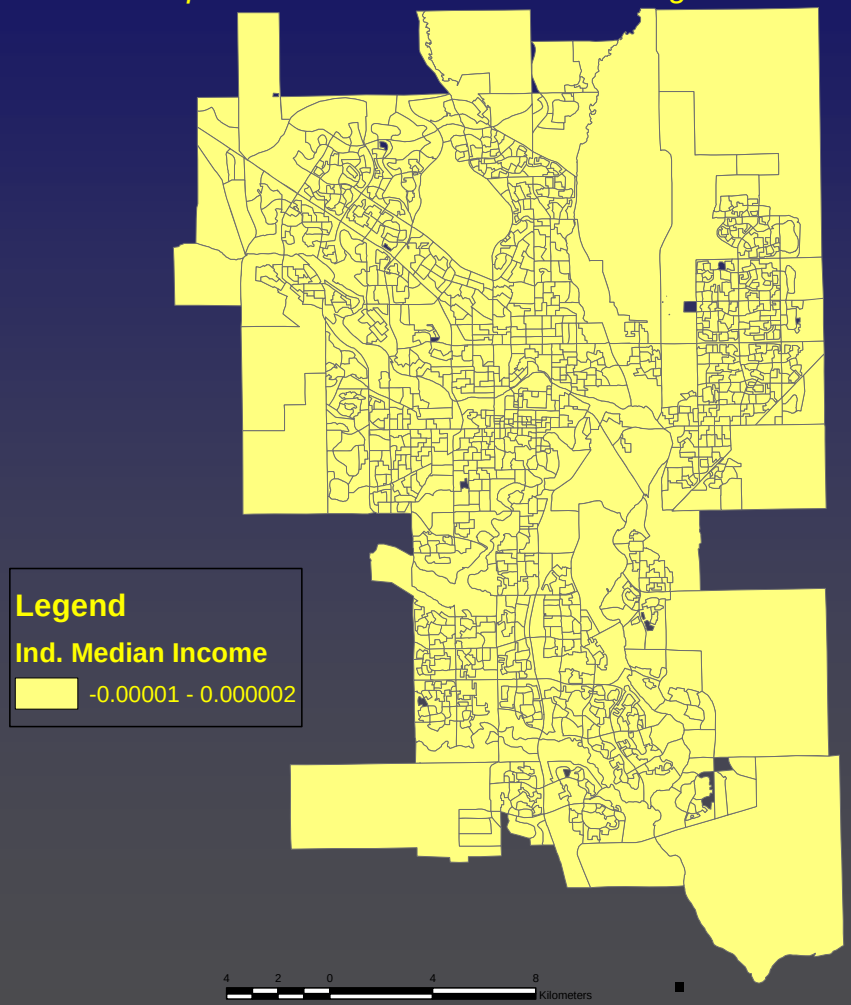


t-Value Map for Single Parent Parameter

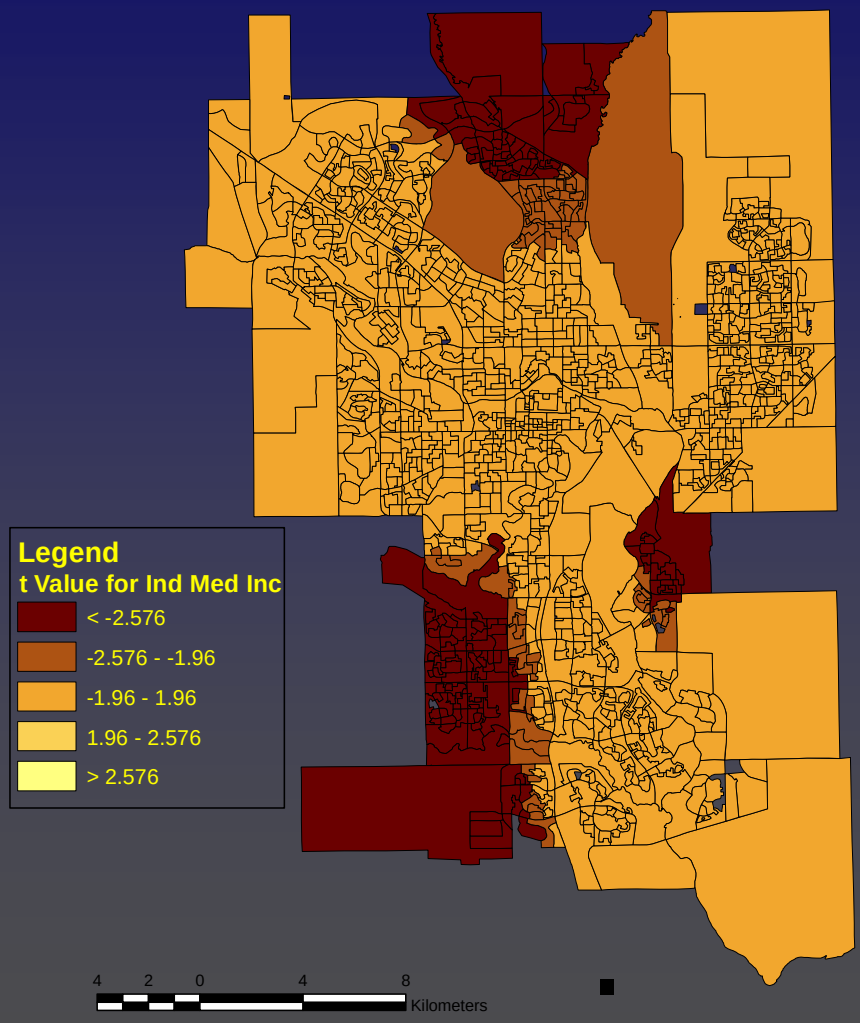


Local Parameter Estimates for Individual Median Income

Geographically Weighted Regression by using Adaptive Kernel with 255 Nearest Neighbors

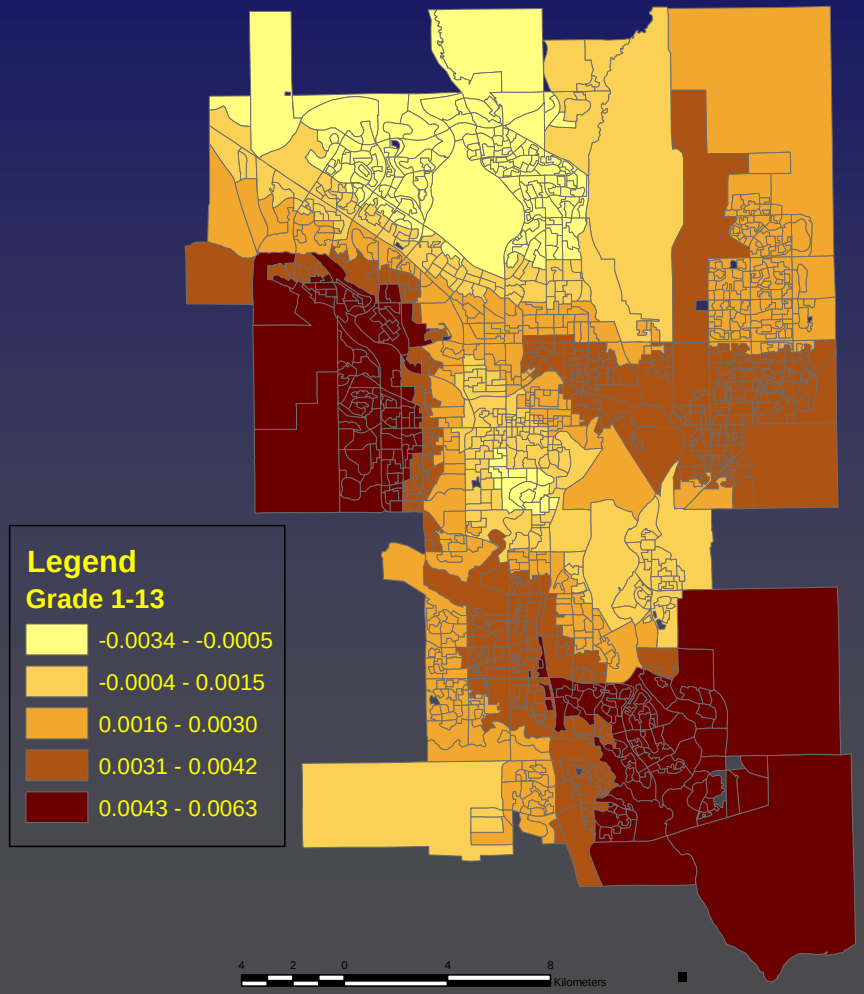


t-Value Map for Individual Median Income Parameter

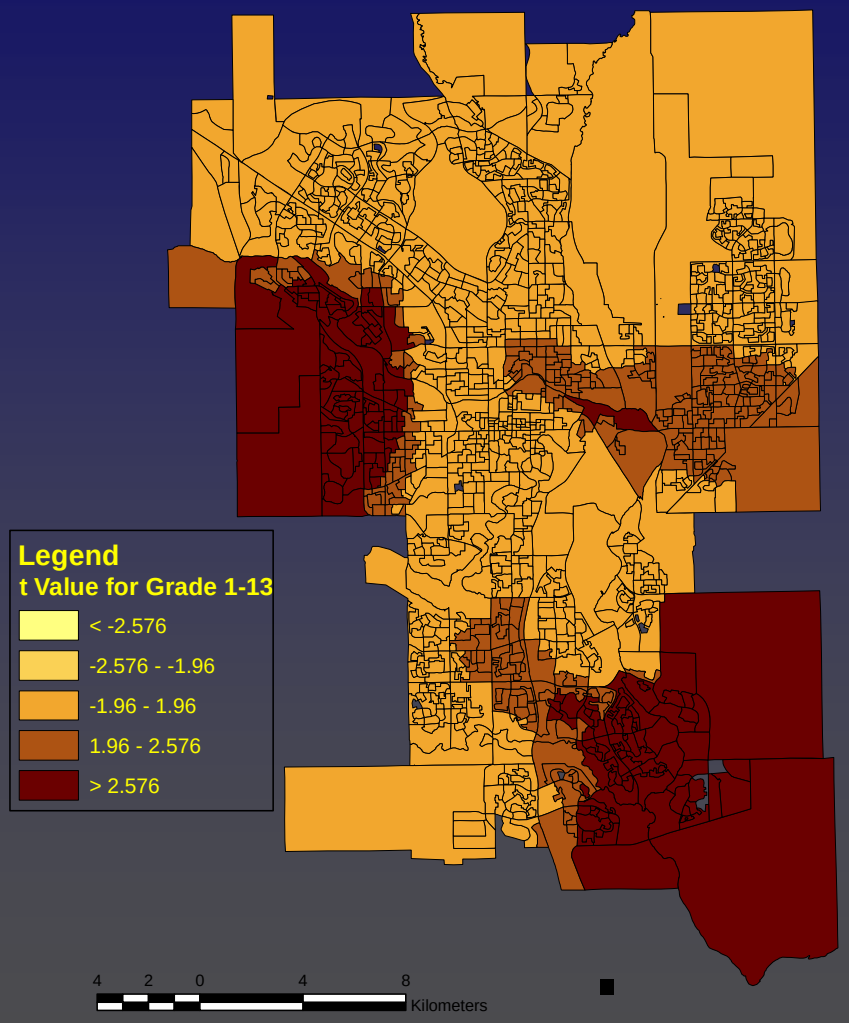


Local Parameter Estimates for Education (Grade 1-13)

Geographically Weighted Regression by using Adaptive Kernel with 255 Nearest Neighbors

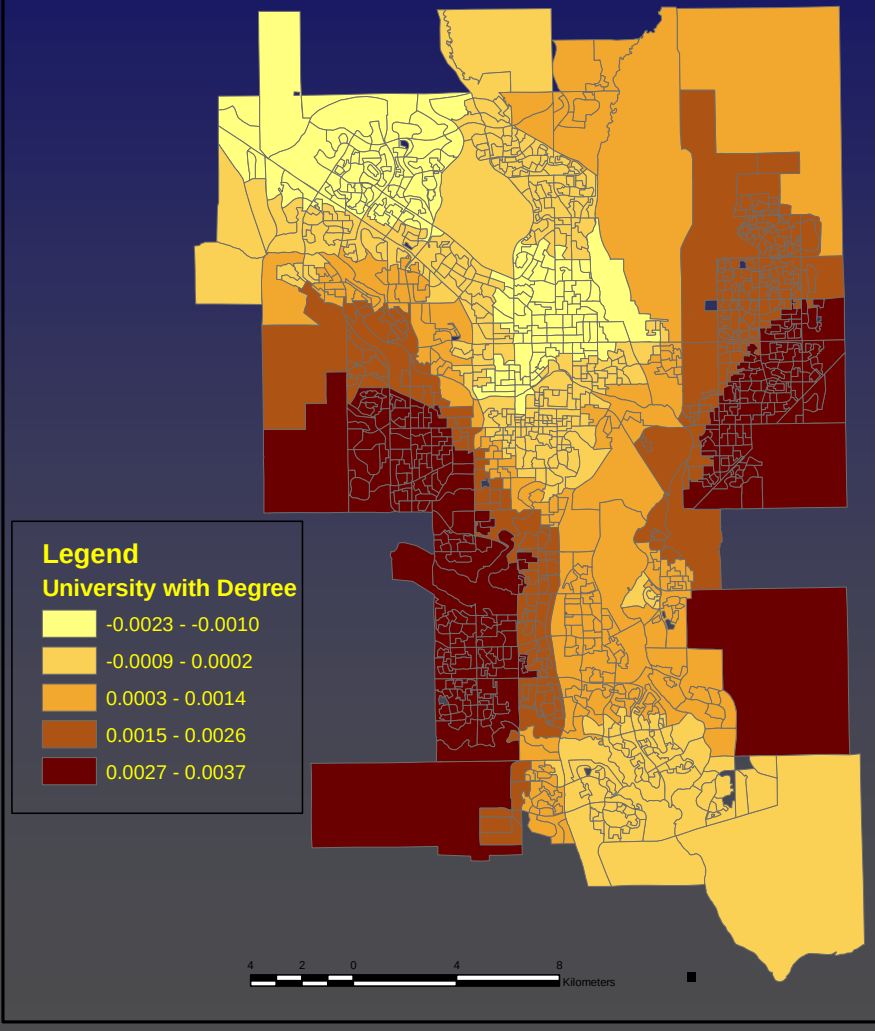


t-Value Map for Education (Grade 1-13) Parameter

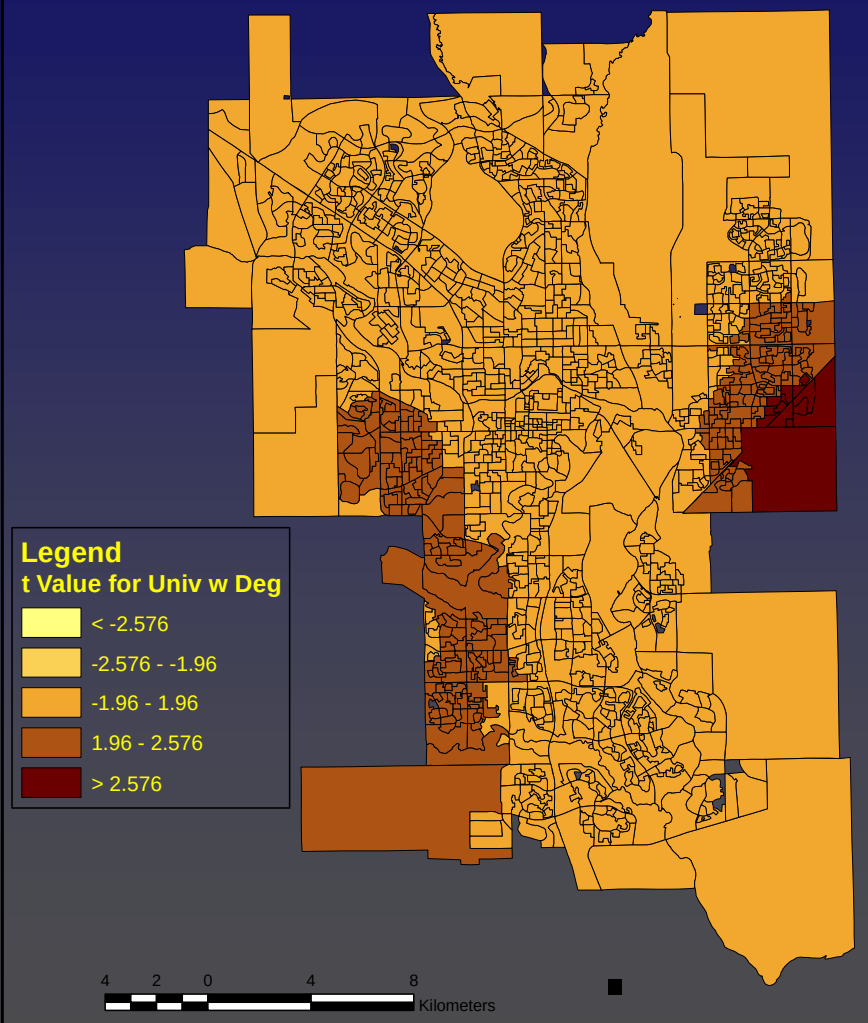


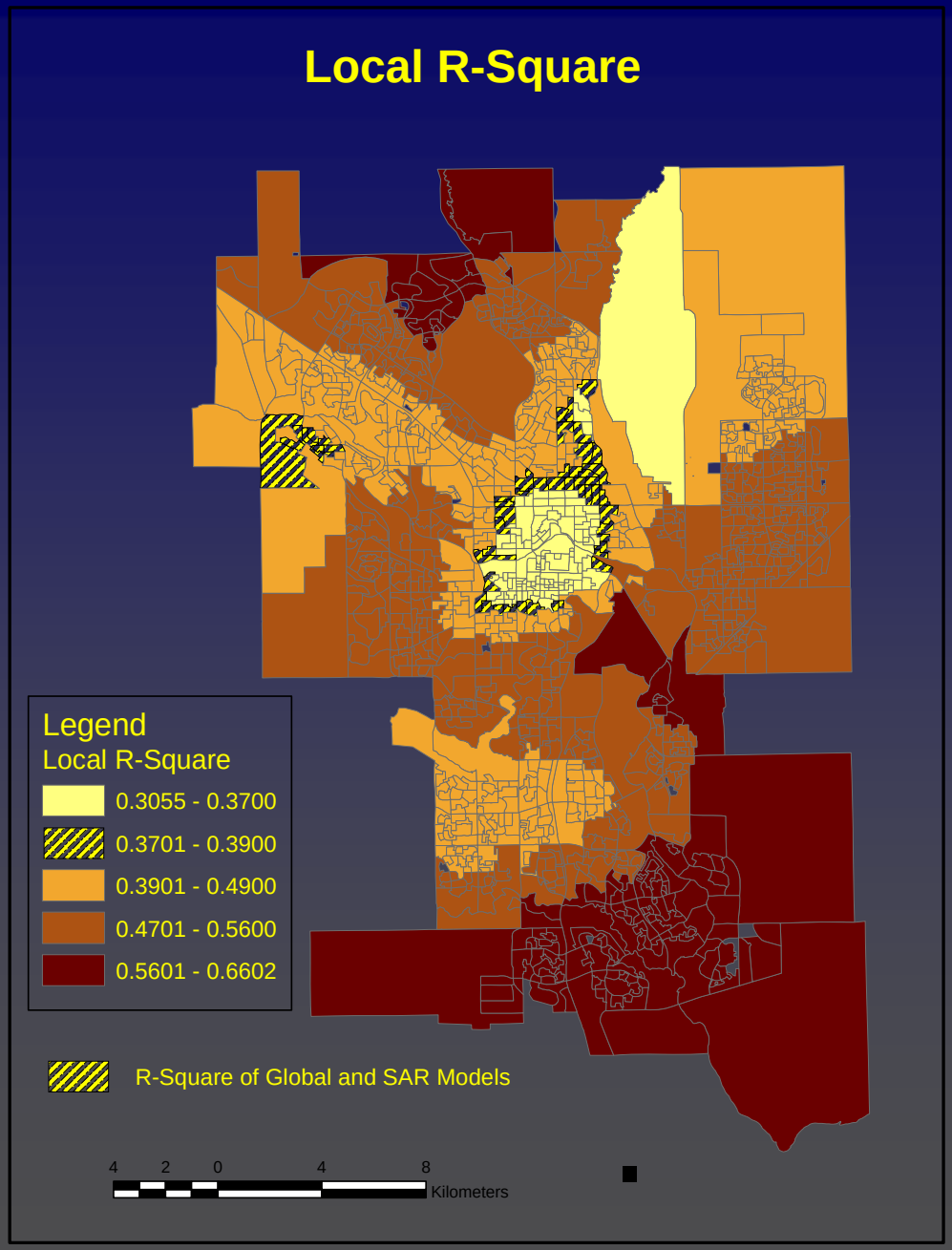
Local Parameter Estimates for Education (University with Degree)

Geographically Weighted Regression by using Adaptive Kernel with 255 Nearest Neighbors

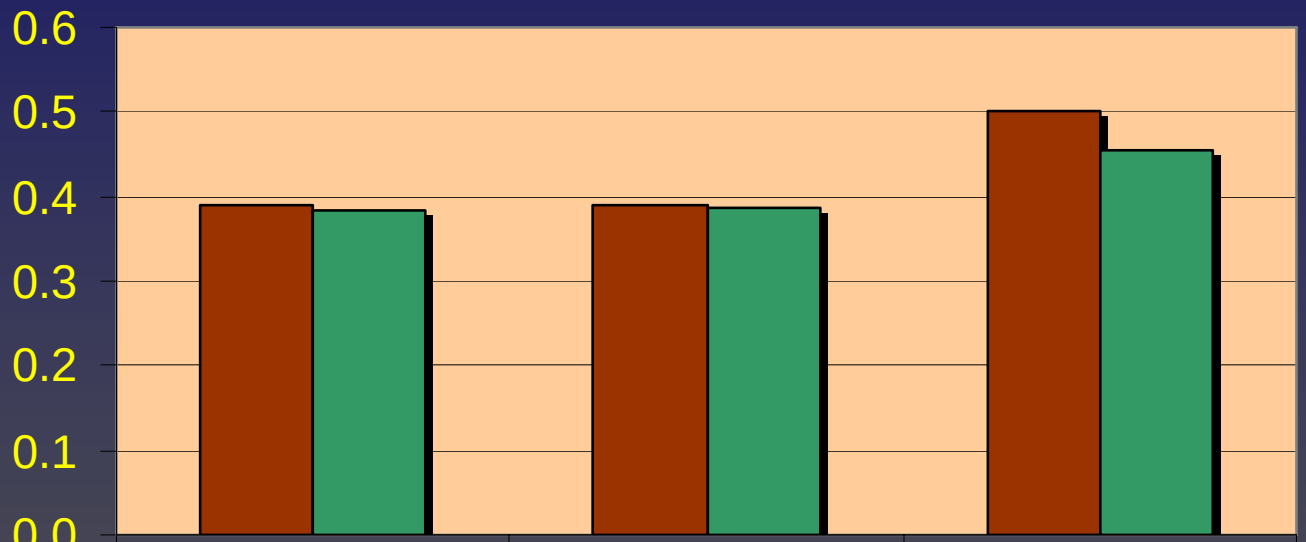


t-Value Map for Education (University with Degree) Parameter

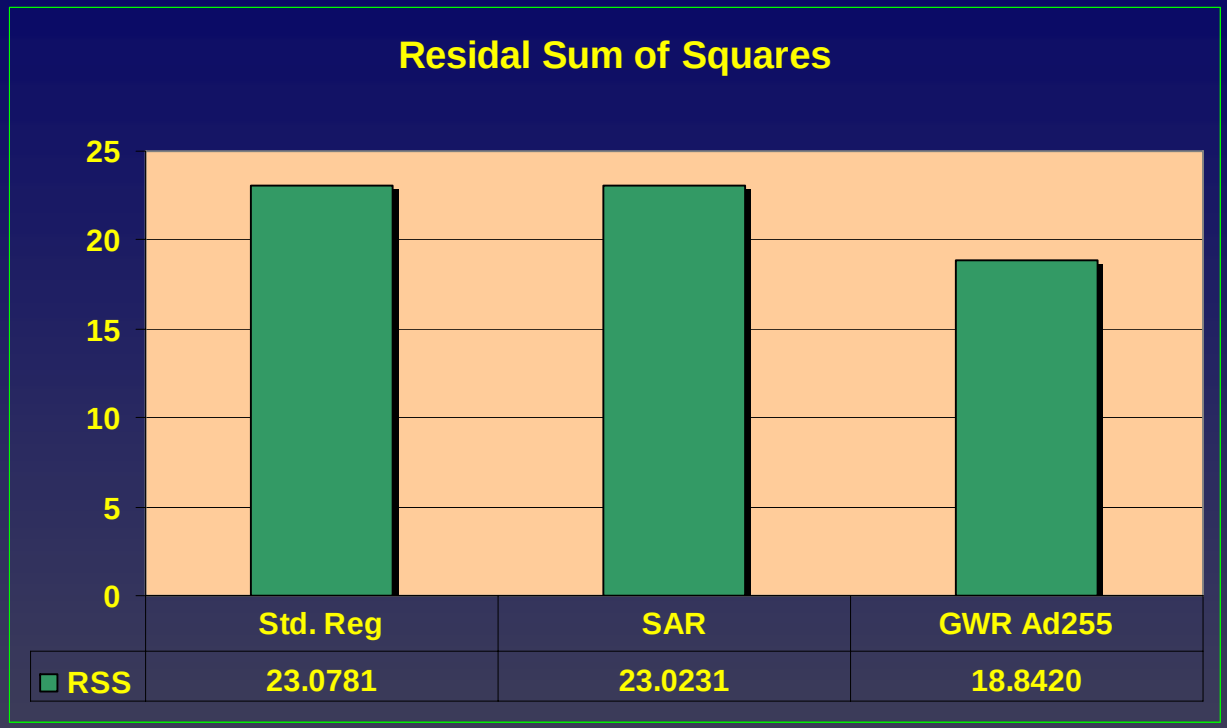


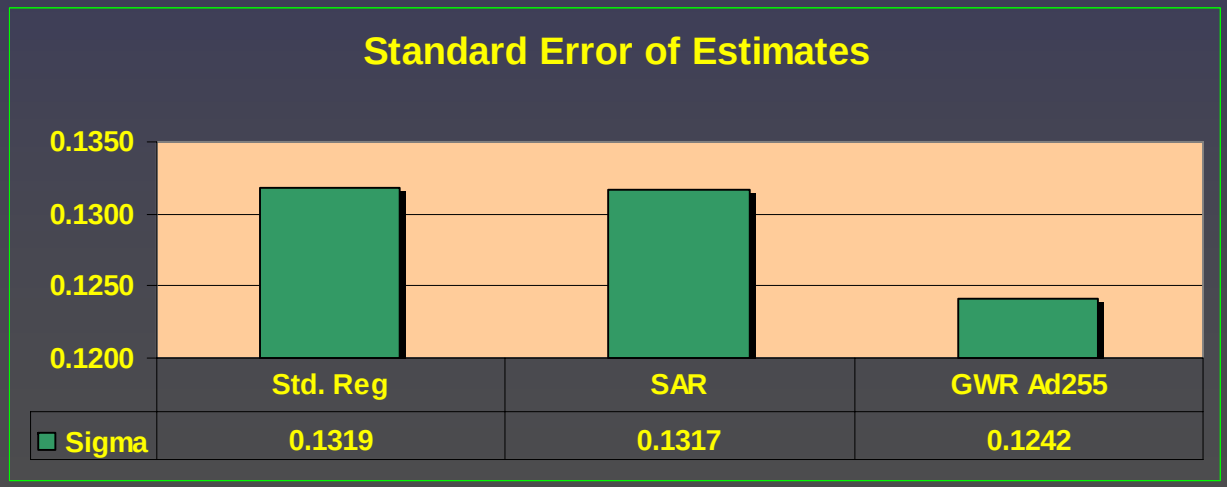

Local R-Square
Legend
Local R-Square
0.3055 - 0.3700
0.3701 - 0.3900
0.3901 - 0.4900
0.4701 - 0.5600
0.5601 - 0.6602
R-Square of Global and SAR Models
4 2 0 4 8
Kilometers

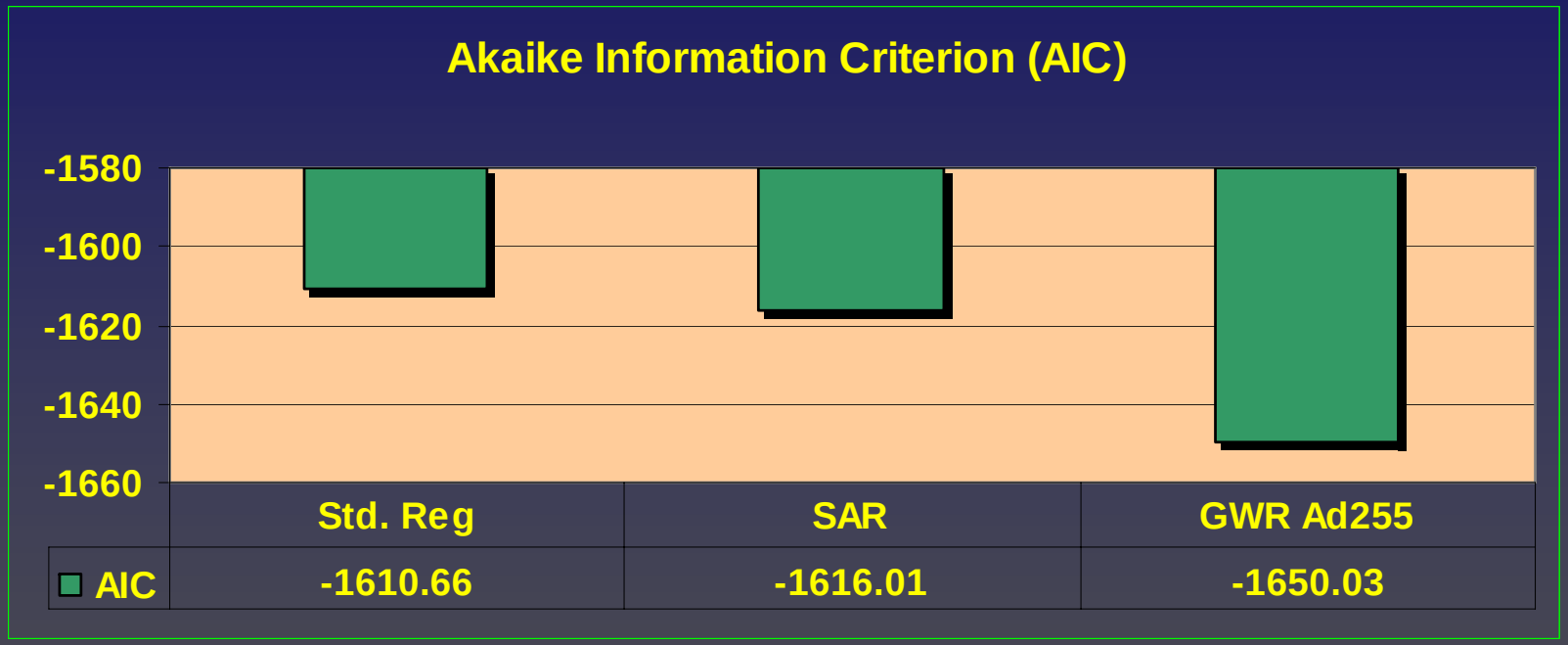
Coefficient of determination (R^2 and Adj. R^2)



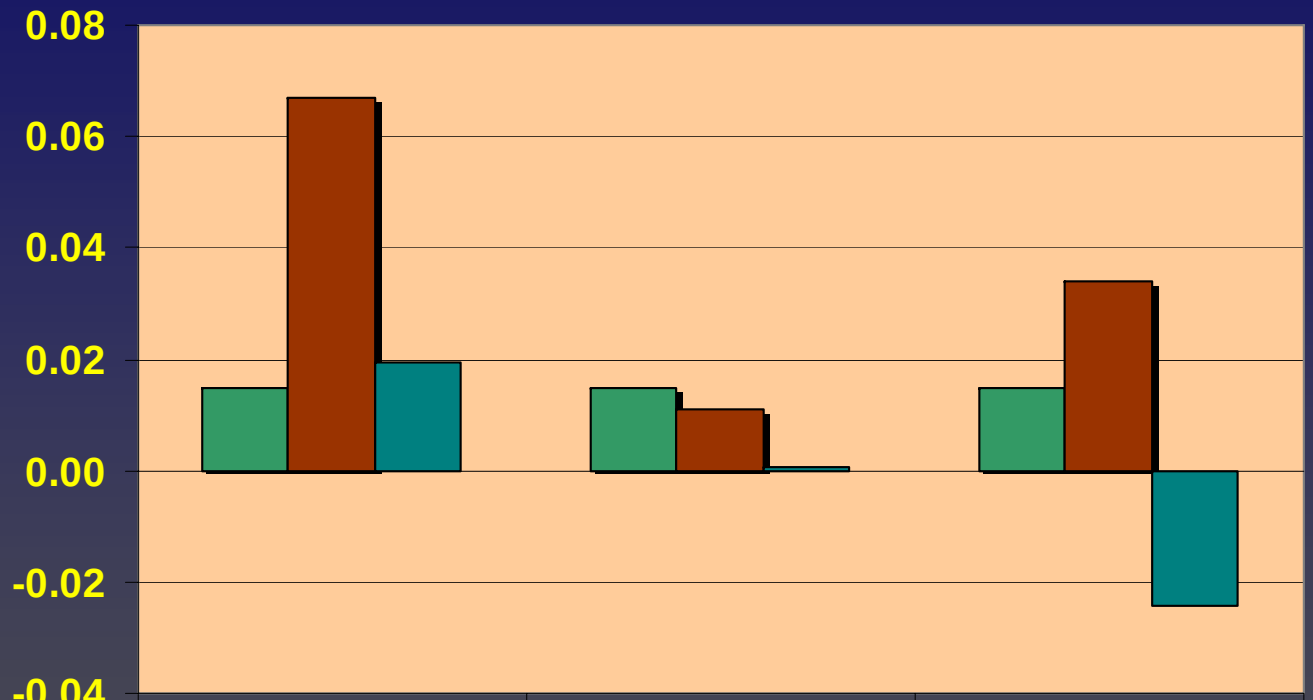
	Std. Reg	SAR	GWR Ad255
R-Sq	0.3887	0.3894	0.5009
Adj. R-Sq	0.3845	0.3858	0.4543


Residal Sum of Squares
25
20
15
10
5
0
Std. Reg
SAR
GWR Ad255
RSS
23.0781
23.0231
18.8420


Standard Error of Estimates
0.1350
0.1300
0.1250
0.1200
Std. Reg
SAR
GWR Ad255
Sigma
0.1319
0.1317
0.1242


Akaike Information Criterion (AIC)
-1580
-1600
-1620
-1640
-1660
Std. Reg
SAR
GWR Ad255
AIC
-1610.66
-1616.01
-1650.03

Moran's I on Observed, Predicted and Residuals



	Std. Reg	SAR	GWR Ad255
Observed	0.0150	0.0150	0.0150
Predisted	0.0668	0.0109	0.0340
Residuals	0.0196	0.0008	-0.0244

Conclusion

- The relationship between cardiac catheterization cases and independent variables displays a spatial variation which can be modeled efficiently by GWR.

- Age 55-84 is the most important variable as it has the highest impact on cardiac catheterization cases.

- Spatial variation in local parameter estimates and t-value maps both in magnitude and direction depict the short coming of using only one parameter estimate (Standard Regression, SAR).

Conclusion *(Cont.)*

- GWR outperformed the global and SAR models based on the standard error of estimates, residual sum of squares, adjusted R^2 and AIC.

- However, GWR did not accommodate spatial dependence.

- It is concluded that GWR can be used effectively in spatial epidemiology as it may provide a great local insight to the varying relationships, and can be used as an excellent exploratory data analysis tool.

References

- Barnes, J., 2005, Minneapolis Fire and MMRF Researchers Collaborate to Test Two New COR Devices, Minneapolis Medical Research Foundation, URL: http://www.mmrfweb.org/aboutus/resources/press.html

- Elliot, P., J.C. Wakefield, 2000, Bias and confounding in spatial epidemiology, In: Elliot, P., J.C. Wakefield, N.G. Best, and D.J. Briggs (eds) Spatial epidemiology: methods and application, Oxford University Press, Oxford, pp. 68-84.

- Fotheringham, A.S., C. Brunsdon, and M.E. Charlton, 2002, Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. West Susses: John Wiley and Sons.

- Getis, A., 1994, Spatial dependence and heterogeneity and proximal databases, In: Fotheringham, S., and P.A. Rogerson, (eds) Spatial Analysis and GIS, Taylor & Francis, pp. 105-120.

- Keller, S., 2004, Cardiac Catheterization, Medical Encyclopaedia, U.S. National Library of Medicine, National Institute of Health, URL: http://www.nlm.nih.gov/medlineplus/encyarticle.003419.htm

THANK YOU

- Dr. Stefania Bertazzon, Associate Professor and Graduate Program Coordinator, University of Calgary
- Dr. Nigel Waters, Director of Mason's Center of Excellence for Geographic Information Science. (Former Director for MGIS program, University of Calgary)
- Dr. William Ghali, Research Director, Community Health Sciences, University of Calgary
- James Norris, MGIS Candidate
- Dr. David Strong and Tracy Halbert, Calgary Health Region