Using GIS to Analyze School Exam Scores

Lixin Huang, PhD, GISP
GIS Analyst III
Brevard County, Board of County Commissioners
2015 Esri International User Conference
July 21, 2015

Model School Exam Scores

- Commuting Style
- Marital Status
- Education Level
- Income
- Unemployment Rate
- Poverty Status

Data Sources

- Socio-economic Factors: Latest Census ACS
- School Attendance Boundary: County School Board
- Exam Scores: State Department of Education

Model Process

- Prepare census data.
- Prepare school attendance boundary and exam scores.
- Overlay school attendance boundary, census data, and exam scores.
- Model exam scores.

Prepare Census Data

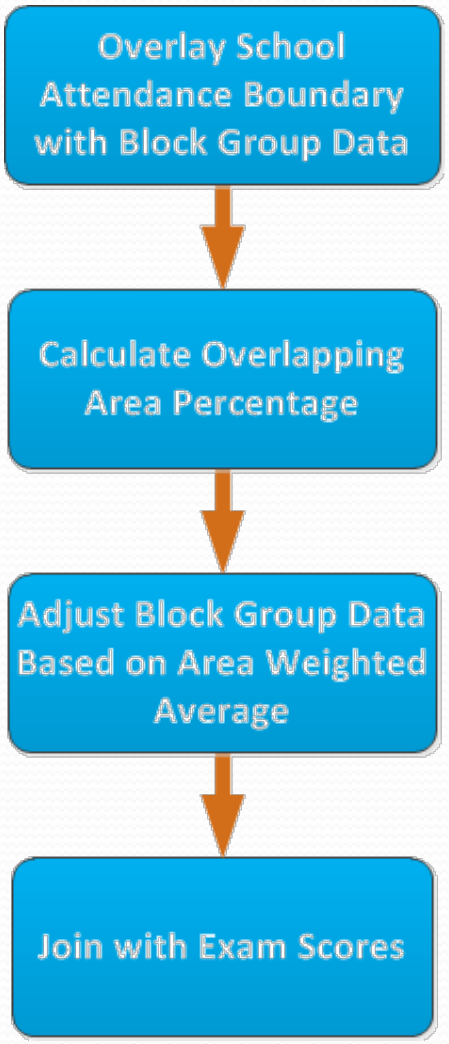
Census Data Selected

- Means Of Transportation To Work (Percentage)
- Travel Time To Work (Percentage)
- Household Type (Percentage)
- Educational Attainment (Percentage)
- Poverty Status (Percentage)
- Median Household Income (Number)
- Employment Status (Percentage)

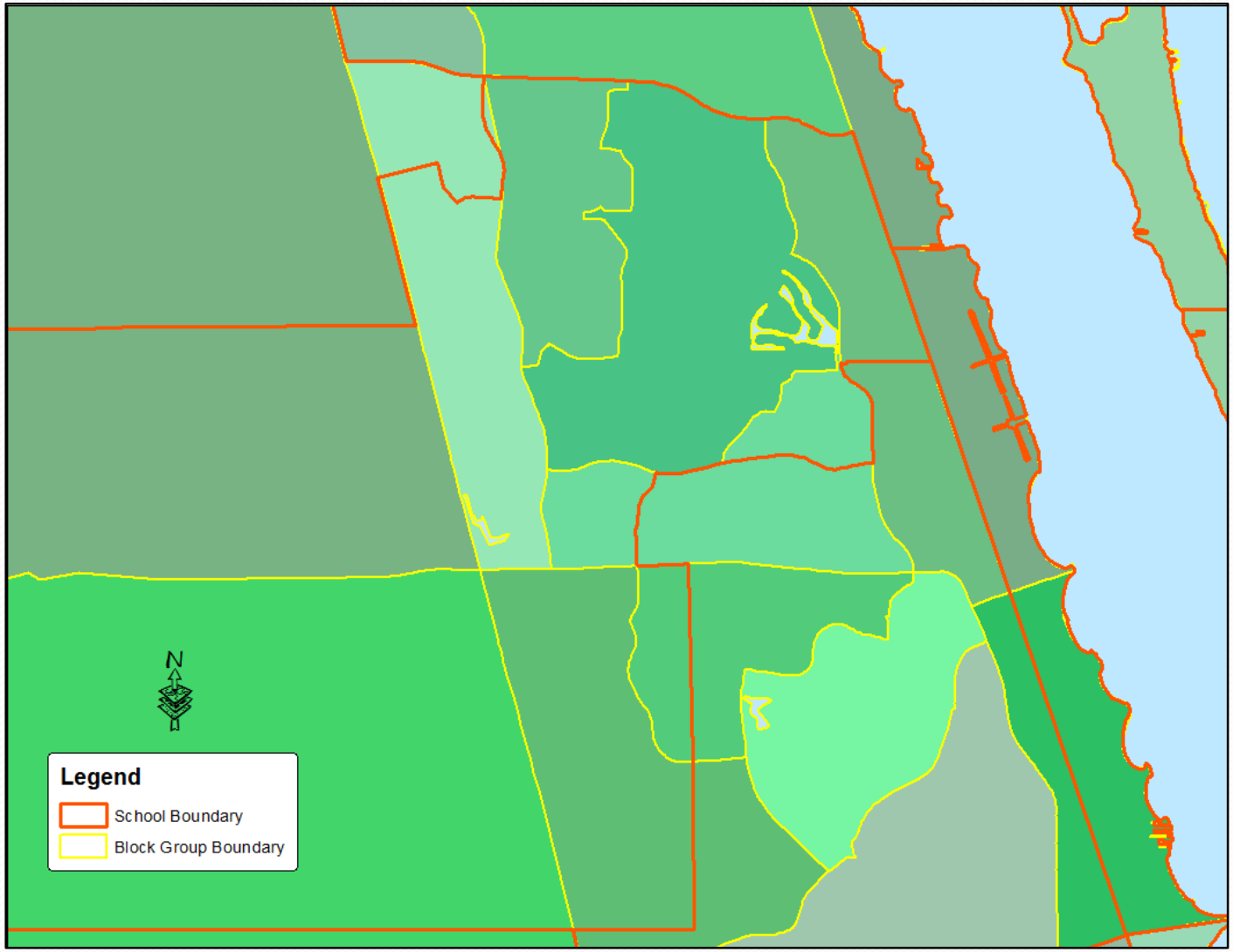
Prepare School Attendance Boundary and Exam Scores

- School attendance boundary data format is Feature Class.
- Original exam scores include three subjects.
  - Math
  - Reading
  - Writing
- Original exam scores format is Excel Spreadsheet.
- Convert exam scores to standalone tables.

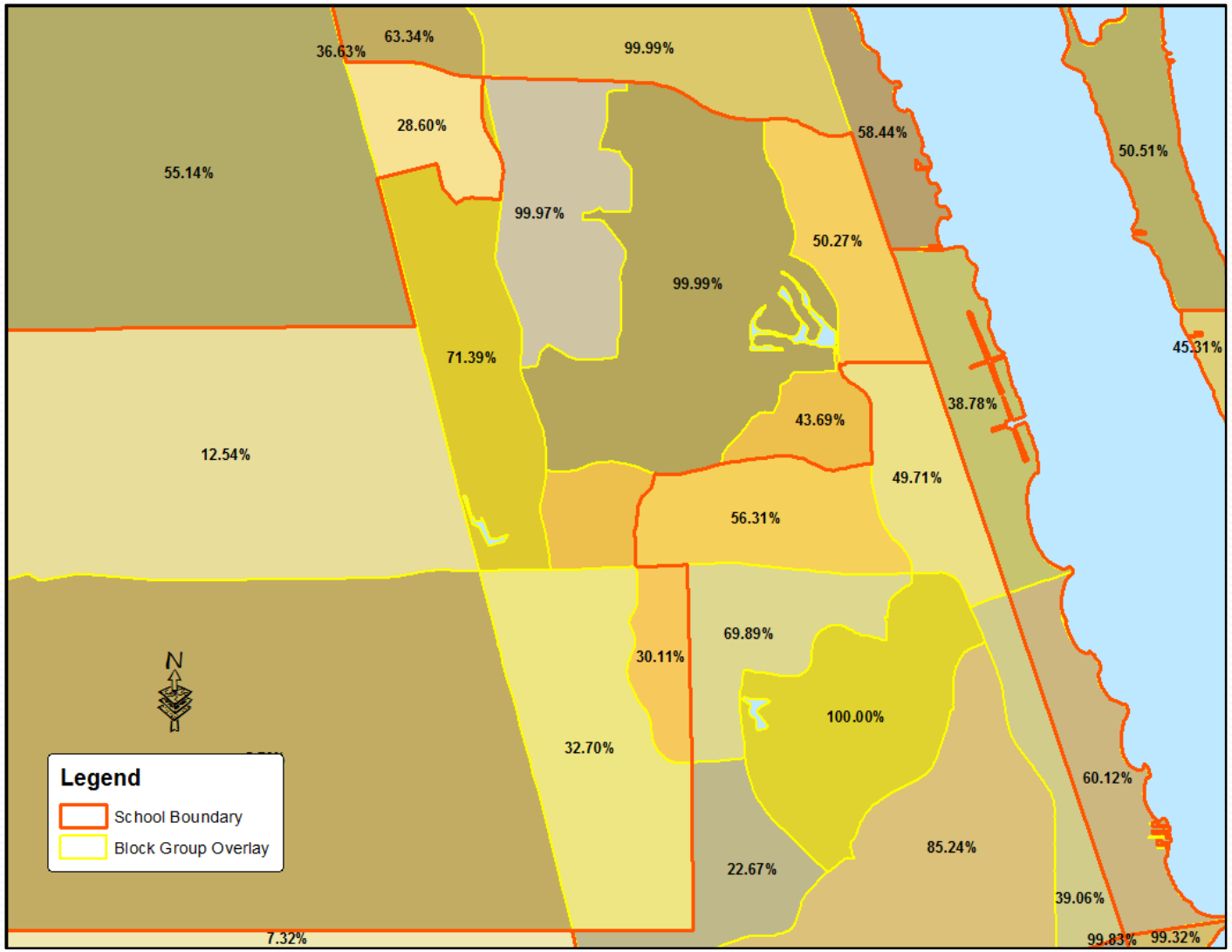
Overlay School Attendance Boundary with Census Data, and Exam Scores



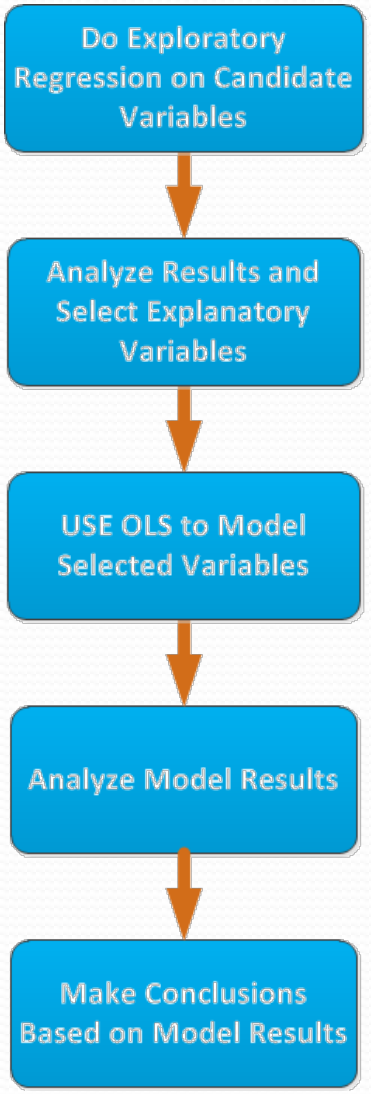
School Attendance and Block Group Boundaries



School Attendance and Block Group Boundaries Overlay



Model Exam Scores

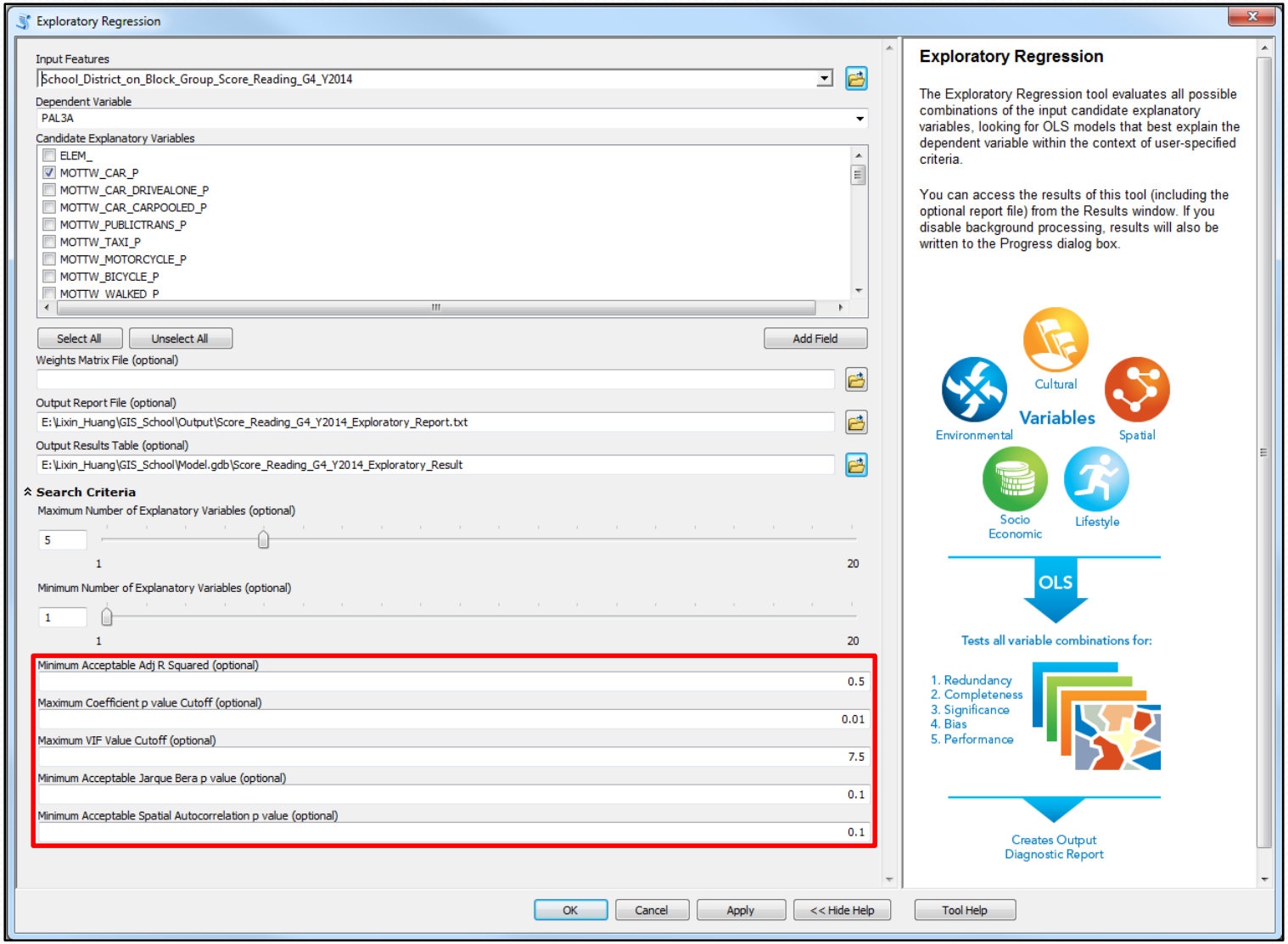


Candidate Explanatory Variables

- Means Of Transportation To Work (Passenger Car)
- Educational Attainment (Bachelor and Above)
- Travel Time To Work (Less Than 15 Minutes)
- Poverty Status (Income Below)
- Travel Time To Work (Between 15 and 30 Minutes)
- Median Household Income
- Household Type (Married Couple)
- Employment Status (Unemployed)

Dependent Variable

Exploratory Regression Criteria



Exploratory Regression Result

********** Exploratory Regression Global Summary (PAL3A) **********

Percentage of Search Criteria Passed

Search Criterion	Cutoff	Trials	# Passed	% Passed
Min Adjusted R-Squared	> 0.50	218	168	77.06
Max Coefficient p-value	< 0.01	218	10	4.59
Max VIF Value	< 7.50	218	218	100.00
Min Jarque-Bera p-value	> 0.10	218	217	99.54
Min Spatial Autocorrelation p-value	> 0.10	17	17	100.00

Summary of Variable Significance

Variable	% Significant	% Negative	% Positive
EAO25Y_BACHELORANDABOVE_P	68.69	0.00	100.00
MHIP12M	57.58	0.00	100.00
ESO16Y_UNEMPLOYED_P	46.46	100.00	0.00
PSP12M_INCOMEBELOW_FAMILYHH_P	35.35	100.00	0.00
TTTW_LT15MIN_P	13.13	94.95	5.05
HT_FAMILYHH_MARRIEDCOUPLE_P	11.11	39.39	60.61
MOTTW_CAR_P	0.00	47.47	52.53
TTTW_15TO30MIN_P	0.00	63.64	36.36

Summary of Multicollinearity

Variable	VIF	Violations	Covariates
MOTTW_CAR_P	1.69	0	--------
TTTW_LT15MIN_P	3.70	0	--------
TTTW_15TO30MIN_P	2.07	0	--------
HT_FAMILYHH_MARRIEDCOUPLE_P	3.90	0	--------
EAO25Y_BACHELORANDABOVE_P	5.46	0	--------
PSP12M_INCOMEBELOW_FAMILYHH_P	1.86	0	--------
MHIP12M	6.36	0	--------
ESO16Y_UNEMPLOYED_P	1.87	0	--------

OLS Parameters

Ordinary Least Squares

Input Feature Class
School_District_on_Block_Group_Score_Reading_G4_Y2014

Unique ID Field
UNIQUE_ID

Output Feature Class
E:\Lixin_Huang\GIS_School\Model.gdb\Score_Reading_Case01_G4_Y2014_OLS

Dependent Variable
PAL3A

Explanatory Variables
- ELEM_
- MOTTW_CAR_P
- MOTTW_CAR_DRIVEALONE_P
- MOTTW_CAR_CARPOOLED_P
- MOTTW_PUBLICTRANS_P
- MOTTW_TAXI_P
- MOTTW_MOTORCYCLE_P
- MOTTW_BICYCLE_P
- MOTTW_WALKED_P

Select All | Unselect All | Add Field

Output Report File (optional)
E:\Lixin_Huang\GIS_School\Output\Score_Reading_Case01_G4_Y2014_OLS_Report.pdf

Additional Options

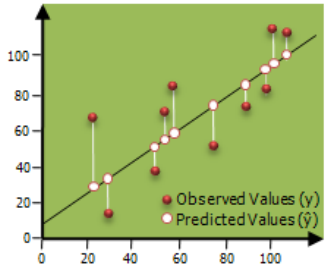
Coefficient Output Table (optional)
lel.gdb\Score_Reading_Case01_G4_Y2014_OLS_Coefficient

Diagnostic Output Table (optional)
del.gdb\Score_Reading_Case01_G4_Y2014_OLS_Diagnostic

Ordinary Least Squares

Performs global Ordinary Least Squares (OLS) linear regression to generate predictions or to model a dependent variable in terms of its relationships to a set of explanatory variables.

You can access the results of this tool (including the optional report file) from the Results window. If you disable background processing, results will also be written to the Progress dialog box.



OK | Cancel | Apply | << Hide Help | Tool Help

OLS Result – Model Variables

Summary of OLS Results - Model Variables

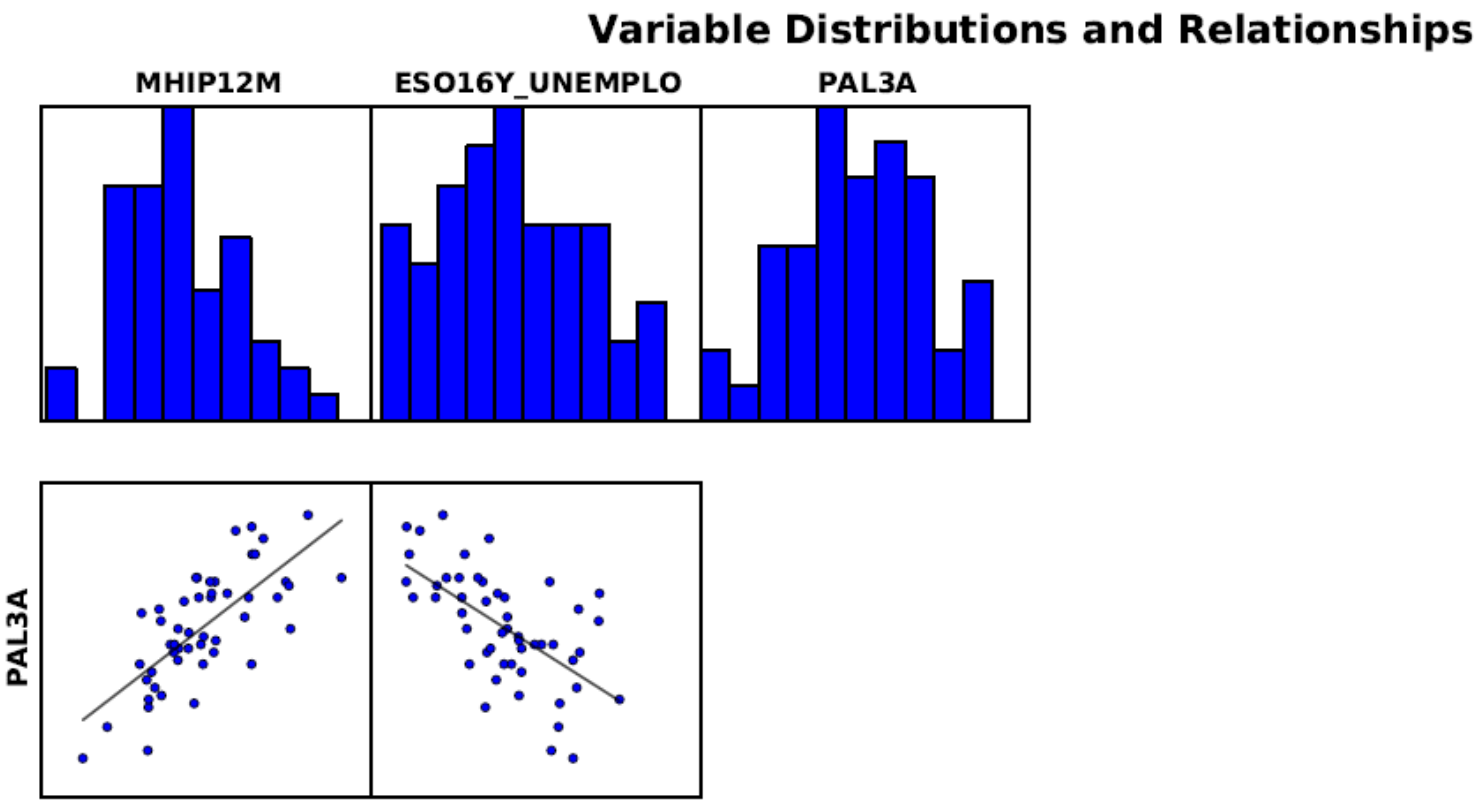
Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	43.595465	11.141898	3.912750	0.000294*	11.180687	3.899176	0.000306*	--------
MHIP12M	0.000549	0.000110	4.977658	0.000009*	0.000118	4.656206	0.000027*	1.436911
ESO16Y_UNEMP	-225.055200	80.015332	-2.812651	0.007149*	79.339979	-2.836593	0.006708*	1.436911

OLS Result – Diagnostics

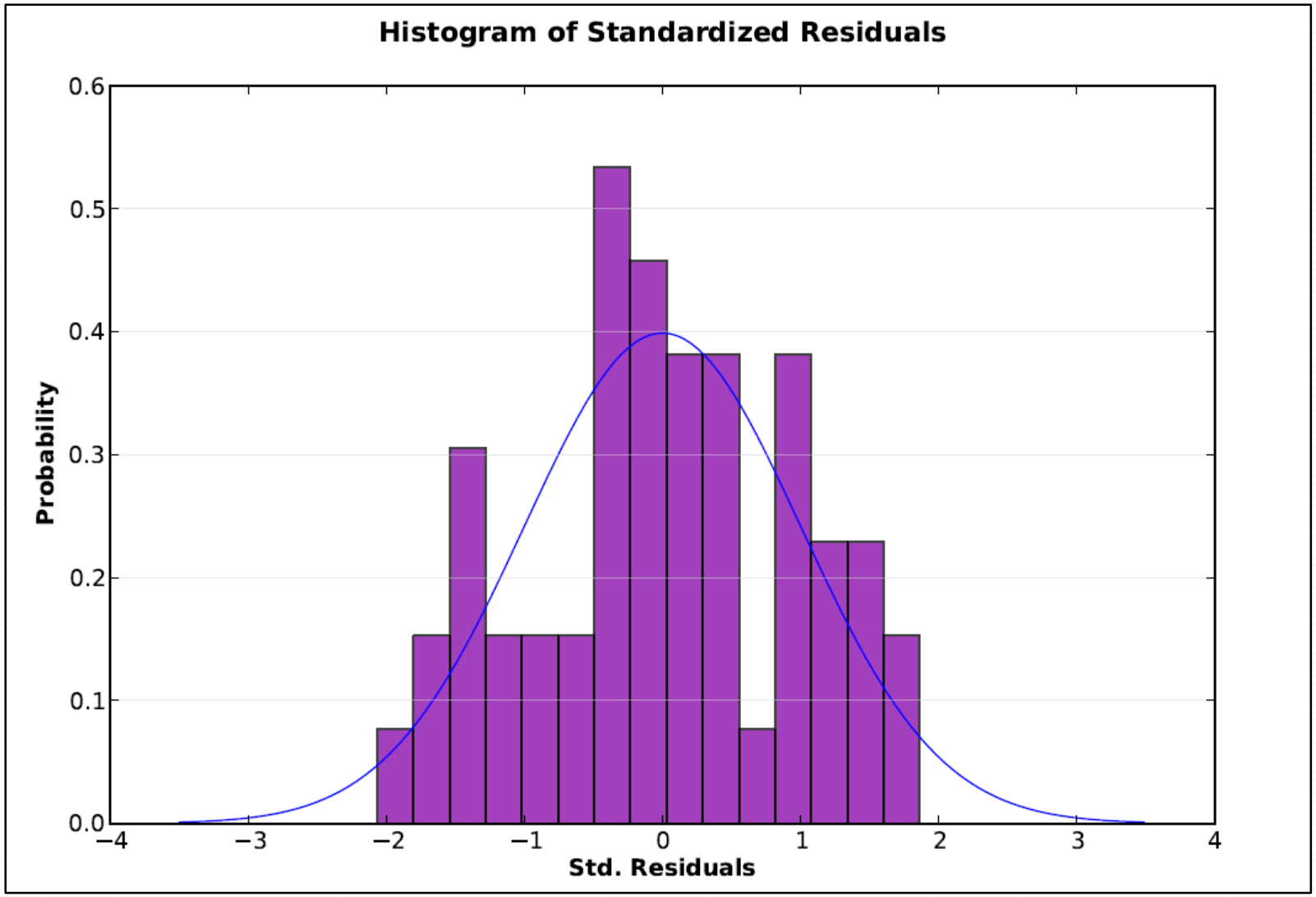
OLS Diagnostics

Input Features:	School_District_on_Bloc	Dependent Variable:	PAL3A
Number of Observations:	50	Akaike's Information Criterion (AICc) [d]:	370.560057
Multiple R-Squared [d]:	0.595372	Adjusted R-Squared [d]:	0.578154
Joint F-Statistic [e]:	34.578031	Prob(>F), (2,47) degrees of freedom:	0.000000*
Joint Wald Statistic [e]:	60.269095	Prob(>chi-squared), (2) degrees of freedom:	0.000000*
Koenker (BP) Statistic [f]:	1.923184	Prob(>chi-squared), (2) degrees of freedom:	0.382284
Jarque-Bera Statistic [g]:	1.102379	Prob(>chi-squared), (2) degrees of freedom:	0.576264

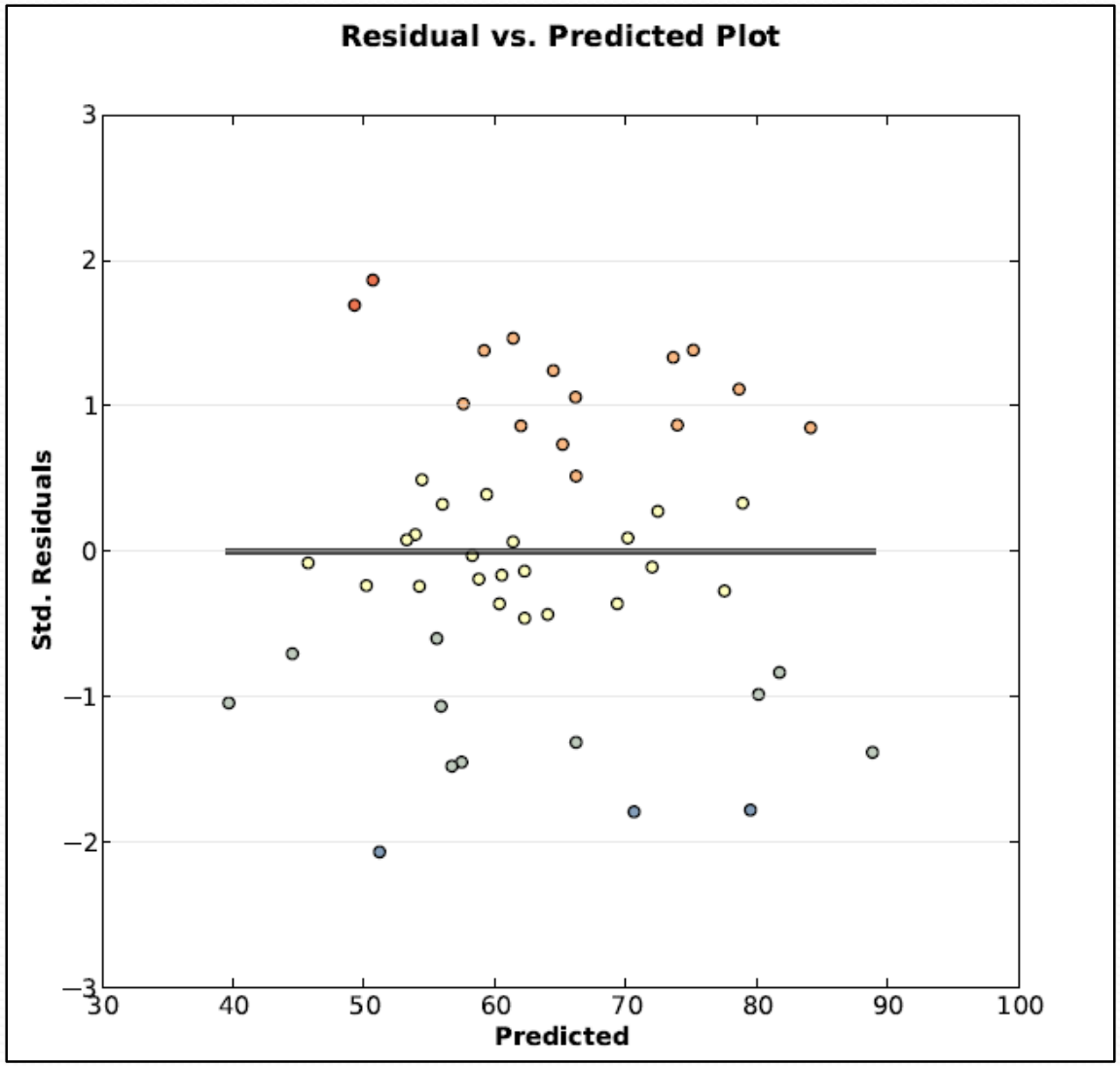
OLS Result – Variable Distributions and Relationships



OLS Result – Histogram of Standardized Residual

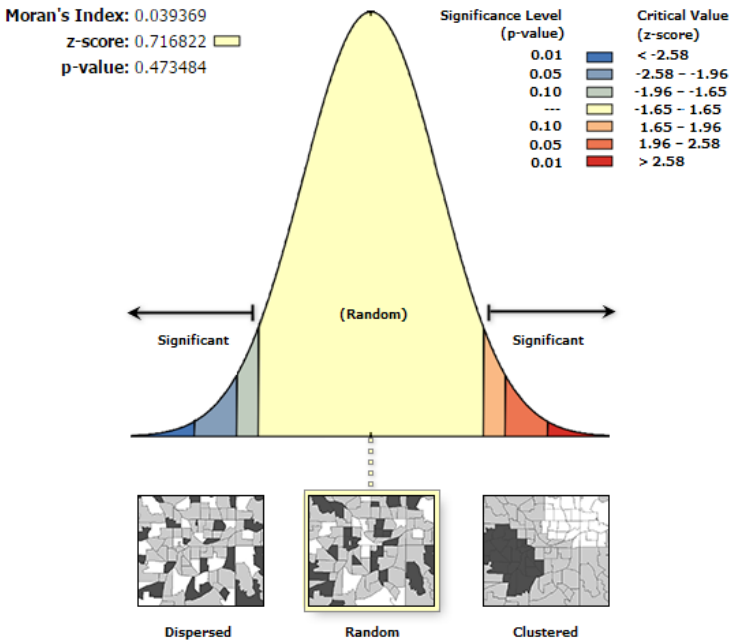


OLS Result – Residual vs. Predicted Plot



OLS Result – Spatial Autocorrelation

Spatial Autocorrelation Report

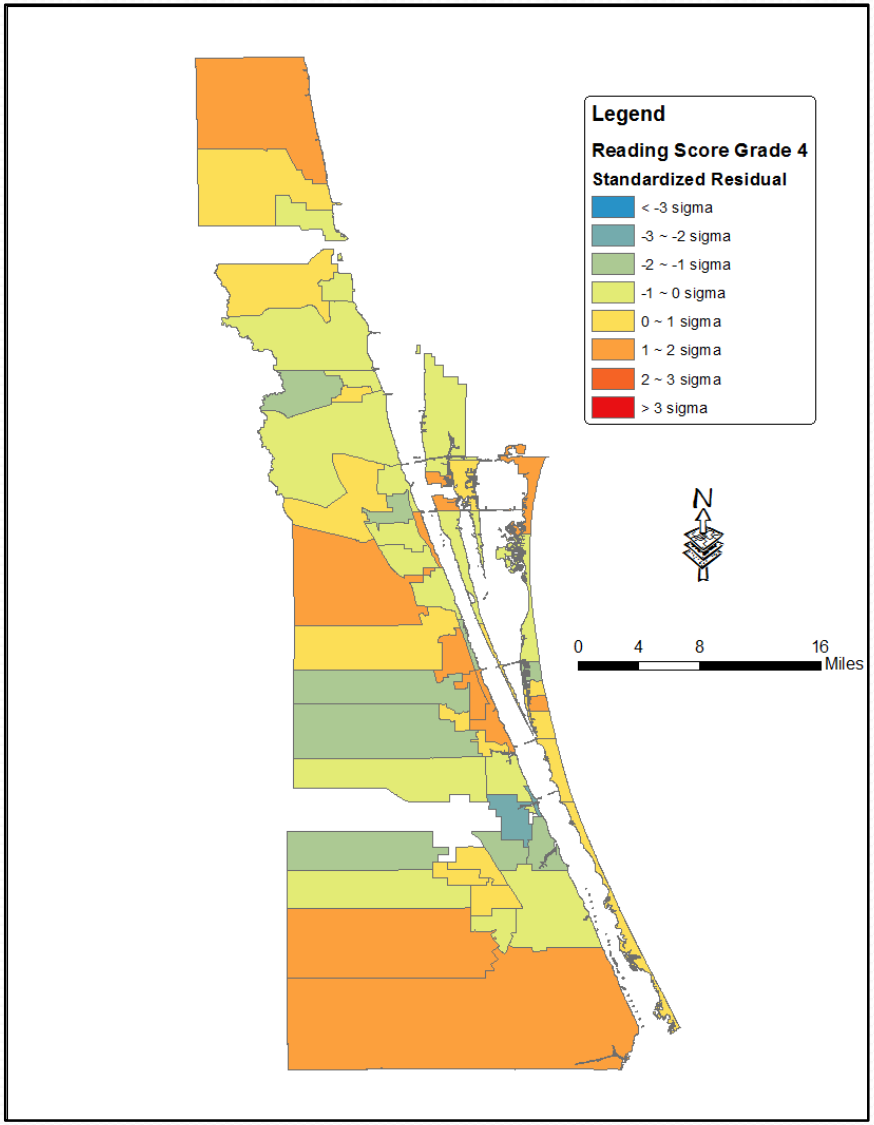


Given the z-score of 0.72, the pattern does not appear to be significantly different than random.

Global Moran's I Summary

Moran's Index:	0.039369
Expected Index:	-0.020408
Variance:	0.006954
z-score:	0.716822
p-value:	0.473484

OLS Result – Standardized Residual



Model Result – Math

Percentage of Exam Passing Score (Math) = 10.97 + 64.45 Household Type (Married Couple) + 67.75 Educational Attainment (Bachelor and Above)

Adjusted $R^2 = 0.444$

Model Result – Reading

Percentage of Exam Passing Score (Reading) = 43.60 + 0.000549 Median Household Income - 255.06 Employment Status (Unemployed)

Adjusted $R^2 = 0.578$

Model Result – Writing

Percentage of Exam Passing Score (Writing) = 27.87 + 73.42 Educational Attainment (Bachelor and Above)

Adjusted $R^2 = 0.145$

Conclusions

- Math exam score was affected by factors:
  - Household with married couple
  - Educational attainment with bachelor degree and above
- Reading exam score was affected by factors:
  - Median Household Income
  - Unemployment
- Weak relationship between Writing exam score and selected socioeconomic factors was found.
- Other important factors may be missing, which may improve model performance.

Questions