

Project Tutorial

1. Topic

Demographic characteristics of the School of International Studies 9th Grade class and their success their first semester.

2. Research Question/Problem

For the 9th Grade class at The School of International Studies for the school year 2006 – 2007, is there a correlation among the Middle School attended, median income of the area of the county they reside in, and the success or lack of success in classes their first semester of 9th Grade?

Criteria:

Success – Passing all classes

Lack of success – Failing one or more classes

3. Data Availability

❖ San Diego Unified School District:

- Student Home Address
- Student Middle School
- Student Number of F's in 9th Grade – First Semester

❖ SanDag:

- Location of Public and Private Schools
- Census data (socioeconomics) for the county – 2006 estimates
- County boundaries & roads
- Zip code boundaries

❖ SanGIS:

- Roads_All – Geocoding reference data for San Diego County

4. New Information to be produced from the data

The new information that will be produced will be correlational data based on the relationship between different data sets. The new information will be the identification of possible locations/areas that tend to generate particular trends in success rates and the environmental factors that may be contributing to the success or lack of success of students entering the 9th grade at our school.

5. Data Analysis

❖ Tools:

- Geocoding
 - Addresses of students to spatial coordinates
- Join &/or Relate

- Link data sets for analysis
- Data Classification
 - Layer Properties
 - Display data to allow for interpretation
- Spatial Autocorrelation
 - Identifying probability that value clustering is random.

6. Process

Program

Task

Procedure

In ArcCatalog:

Create a new address locator

Style: US Streets with Zone [File]

Reference Data: Roads_All

In ArcMap:

Add Data

County.shp (County Boundaries)

Schools.shp (County School Locations)

Zip.shp (Zip Code Boundaries)

Student Information Database (Student Addresses, Middle School and Grade Data)

Batch Geocode student addresses

Source tab in Table of Contents

Right click on Student Information Database

Geocode Addresses

Geocoded addresses and Student Information is added as a shapefile

Relate School.shp to Geocode results

Right click on School.shp

Join or Relate

Relate

Field: School

Relate File: Geocode_Results

Field: Middle_school

Display Geocode Results

Layer: Geocode Results

Layer properties

Categories

Unique Values

Middle_school

Home addresses are color coded by Middle School attended.

Display number of F's earned by students in semester 1

Layer: Geocode Results

Layer properties

Symbology

Multiple Attributes

Quantities by category

Value Field: Middle_school

Symbol size: 9 classes based on 9th grade Fs

Symbol Size

7 – 24

In Microsoft Excel:

Convert data to a form that can be imported into ArcMap.

Open Census Data .csv file from the SanDag site containing zip code, ethnicities & family incomes.

Open a new Excel spreadsheet.

Copy and paste from the original spreadsheet to the new spreadsheet:

Zip code field

Median household income field

Name fields to contain no spaces

Save as Income.csv file

In Windows Explorer:

Change the extension from .csv to .txt

In ArcMap:

Join the zip code shape file to the Income.txt file

Right click on Zip.shp

Join or Relate

Join

Field: Zip

Join: Income.txt

Field: Zip

Display median family income by zip code:

Layer: Zip.shp

Layer properties

Symbology

Quantities

Graduated Colors

Field: Median Household Income

15 Classes

Zip codes are color coded by median household income.

Save the School data of interest as it's own layer

Open Attribute Table of Geocode_Results
Open Attribute Table of Schools.shp
 Select schools in Schools.shp that are present in Geocode_Results
 Display selected records
 Save as layer file
Only displays the 32 schools of interest on the map.

Revise display of median family income by zip code:

Layer: Zip.shp
 Layer properties
 Symbology
 Quantities
 Graduated Colors
 Field: Median Household Income
 Classify
 Manual
 32 Classes

Zip codes are color coded by median household income based on user defined classes. Incomes \$0 - \$100,000 in \$5000 Breaks, \$100,000 - \$220,000 in \$10,000 Breaks.

Color Ramp
Green

The zip code areas are now light green for low median household income in a zip code and dark green for high median household income.

Display Zip Code and Median Household income on map:

Layer: Zip.shp
 Label Features

 Layer Properties
 Labels
 Label Field: Zip
Zip Code displayed on map.
 Layer Properties
 Labels
 Label Field: MED_HH_INC

Median Household income displayed on map.

 Layer Properties
 Labels
 Label Field – Expression
 [ZIP] & “ “ & “\$” & [MED_HH_INC]
Both Zip Code and Median Household Income displayed on map.

Correlate number of F's with clustering:

Toolbox

Spatial Statistics Toolbox

Analyzing Patterns Toolset

Spatial Autocorrelation (Morans I)

Input Feature Class: Geocode_Results

Input Field: 9Num_F (Number of F's)

Check Box: Display Data Graphically

OK

The level of clustering (Dispersed – Clustered) is displayed with a Morans index and Z score. The significance level and probability of the distribution being random is also displayed.

7. Results

The Results are Qualitative in nature (Quantitative data is analyzed and manipulated, but due to the small sample size, a qualitative correlation is made.), and most of the Middle Schools (n=31) had so few students (1 - 3) (0.7% - 2.1% of enrollment) from them that the group would be statistically insignificant. Only 3 schools had 11 (7.6%) or more student from them.

Middle School & Academic Success

❖ Memorial Academy

- 10 of 11 students (91%) had one or more F's.
 - (Mean = 3.8, Median = 3.5, Mode = 2, 3, 5, Range 1 - 7)
 - [Only Students receiving 1 or more F's at the semester.]
 - (Mean = 3.5, Median = 3, Mode = 2, 3, 5, Range 0 - 7)
 - [All Students attending from this school.]

❖ Roosevelt

- 22 of 45 students (49%) had one or more F's.
 - (Mean = 3.2, Median = 3, Mode = 2,3, Range 1 - 8)
 - [Only Students receiving 1 or more F's at the semester.]
 - (Mean = 1.6, Median = 0, Mode = 0, Range 0 - 8)
 - [All Students attending from this school.]

❖ Language Academy

- 7 of 28 students (25%) had one or more F's.
 - (Mean = 2, Median = 2, Mode = 2, Range 1 - 4)
 - [Only Students receiving 1 or more F's at the semester.]
 - (Mean = .5, Median = 0, Mode = 0, Range 0 - 4)
 - [All Students attending from this school.]

❖ All other Middle Schools

- 14 of 59 students (24%) had one or more F's.
 - (Mean = 2.2, Median = 2, Mode = 2, Range 1 - 4)
 - [Only Students receiving 1 or more F's at the semester.]

- (Mean = .5, Median = 0, Mode = 0, Range 0 - 4)
 - [All Students attending from this school.]

Median Household Income & Academic Success

- ❖ Most of our students are from areas with Median Household Incomes <\$45,000.
 - The largest proportion of these are from areas with Med_HH_Inc <\$35,000.
 - Very few are from areas with Med_HH_Inc >\$45,000
- ❖ There is no direct correlation between Median Household Income and academic success based on this data.
- ❖ Since most of our students are from low income areas and there is no individual income data available, the results have a low reliability and validity.

8. Conclusion and Extensions

- ❖ Conclusion:
 - The highest correlation with academic success in the first semester of 9th grade is with the Middle School that the student attended.
 - The Autocorrelation tool indicated that there is a high likelihood that the clustering (around home schools) is not random. The correlation with Median Household income appears to be a random distribution.
- ❖ Extensions & Further Experimentation:
 - Gather individual household income data to evaluate actual income to academic success.
 - Gather equivalent data from schools with similar API scores to compare.
 - Gather equivalent data from other schools in the San Diego High Educational Complex to compare.

