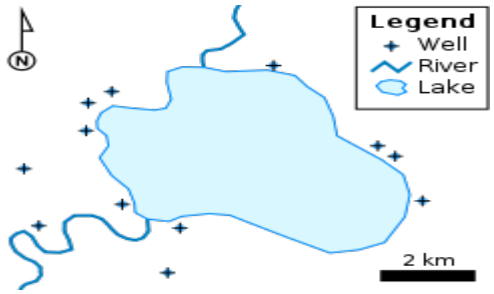
Tutorial 3: Analysing Vector Data[1]

Objective	After completing this tutorial you will be able to select and clip features of a map, join attribute tables, analyse density, and use time to represent land use change.

Introduction Like any urban centre, the GTA has developed and expanded over time. These changes in land use have brought upon various threats to the natural habitat that at one point dominated the area. We will analyse the urban growth of the GTA over periods of time. The rural to urban gradients of the GTA, can be described as a change in industry, population and social dynamics. We will analyse the latter by using the density of coffee shops as a base for determining urbanity.

Data You will analyse data provided covering different aspects of the GTA (municipalities, business locations, etc.). Overall, looking at different types of vector data and how they can be used as a descriptive source for landscape change. Some data is pre-prepared for you, however it is expected that you utilize the reference card, ESRI help tutorials, and lessons in ArcGIS from previous tutorials to supplement the outlines given to solve the tasks at hand.

A key aspect will be to use the different types of vector data in various analyses (point, line and polygon data). Point, line and polygon features are all types of discrete features within vector data models. (See Terminology)



http://en.wikipedia.org/wiki/Geographic_information_system#Vector

Important: To ensure your work is easily accessible and well maintained, be sure to organise and properly label each file you use and create. Assemble your final results for each question in an appropriate location – check that file names are representative of the product and remove files that are no longer needed.

For your convenience, links to ESRI help files are provided. Follow these links to find further information about the procedures used in this tutorial.

[1] This lab outline was jointly developed by A. DiFederico and H. Wagner, UTM Biology.

Data Sets

File	Data Type	Description
ONmun.shp	Shapefile - Polygon	Ontario municipalities
GTADwellings.shp	Shapefile - Polygon	GTA wards with census tract data on private dwellings
GTADwellingsMaxPer.dbf	dBase – Attribute Features	Maximum dwelling growth data
Tim_Hortons.shp	Shapefile - Point	GTA Tim Horton's locations
Starbucks.shp	Shapefile - Point	GTA Starbucks locations
Starbucks_and_TimHortons.shp	Shapefile - Point	GTA compiled locations
GTA_highways.shp	Shapefile - Line	Highways in the GTA
GTA_rivers.shp	Shapefile – Line	Rivers in the GTA

Resources

GIS Introductory Book: Chang, Kang-Tsung. *Introduction to Geographic Information Systems*. 3rd ed. Boston: McGraw-Hill Higher Education, 2006.

ArcGIS Help online guide:
http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=welcome

GIS Tutorial Guide: Gorr, Wilpen L., and Kristen Seamens Kurland. *GIS Tutorial: Workbook for ArcView 9*. 3rd ed. Redlands, CA: ESRI, 2008.

Additional Data: http://maps.chass.utoronto.ca/cgi-bin/search.pl?keyword=dmti

Information about census tracts (definition, naming convention):
http://www.library.yorku.ca/ccm/LibraryDataServices/Courses/censustracts.htm

Concepts

Displaying time series:

- Visualizing time steps vs. change between time steps
- Making data comparable
- Symbology

Working with attribute tables:

- Displaying table
- Sorting
- Selecting features from map and from table
- Exporting selection as a new file
- Creating new fields (variables)
- Calculating fields
- Joining tables

Questions A. How did development occur throughout the GTA over time?

B. Identify gradients of urbanity within the GTA.

Overview of tasks

1. Limit map display to GTA.
    a. Select GTA municipalities from all Ontario municipalities.
    b. Export GTA municipalities to new file.
    c. Set map extent to new GTA file.
2. Evaluate development within GTA over time.
    a. For each period, display the total number of new private dwellings in each census tract.
    b. Display the proportion of dwellings constructed within a period.
    c. Display for each census tract the period with the highest development.
    d. Reflect on data quality.
3. Identify gradients of urbanity based on density of coffee shops.
    a. Display coffee shops (Tim Hortons and Starbucks) in the GTA.
    b. Count the number of coffee shops in each municipality.
    c. Calculate the number of coffee shops per 1000 inhabitants.
    d. Rank municipalities by coffee shop density.
    e. Reflect on appropriateness of data.

Task description

1. Limit map display to GTA.

Goal: Create a clip file of the GTA from the ONmun.shp shapefile (Ontario municipalities), to be used throughout tutorials 3 and 4 for display and analysis.

http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Clip%20(Analysis)

- Look up "Clip" under "Procedures" at the end of this tutorial.
- Follow steps to select GTA municipalities.
- Follow steps to save selected features in a new file. Name it appropriately.
- Add the new file to the map.
- Look up "Define Map Extent" under "Procedures" and follow steps to change map extent to the new GTA clip file.

2. Evaluate development within GTA over time.

2.1. With the class, derive competing hypotheses about the spatio-temporal process of urban growth within the GTA. Where did development first occur? And over time where did it expand to? Name three hypotheses:

_____

_____

_____

2.2. Analyse the attribute table for the GTADwellings feature (right click on layer and select "Open attribute table"). Use the dwellings metadata Excel spreadsheet to determine the meanings of column headings.

http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=An_overview_of_tables_and_attribute_information

Question:

- When and where do you expect to find the greatest increases in dwellings within the GTA?

2.3. Visualize the number of new dwellings per census tract for each period from before 1946 until 2006. Add highways and/or rivers to facilitate orientation.

Hint: Open the Symbology tab under Properties. Under Quantities, Graduated Symbols, select the appropriate Field, then click "Apply" instead of "OK". The dialog box will remain open and you can cycle through all periods.

2.4. The absolute numbers of dwellings may be misleading to compare between census tracts, as one may be larger or more populated than another in the first place. We can standardize the data by dividing the number of new dwellings in a period by the total number of dwellings.

- In the Symbology tab, click on "Normalization" and select the column with the total number of private dwellings. Click "Apply".
- How does this affect the map?

2.5. To summarize the analysis of the spatio-temporal process of development in the GTA, an analysis was performed to determine for each census tract the period with the highest number of new dwellings (MaxPeriod). Before we can visualize these data, they need to be combined (joined) with the GTADwellings attribute table:

http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=Joining_tables

- Add the data base file GTADwellingsMaxPer.dbf.
- Check "Joining Attributes From Two Tables" under "Procedures.
- Follow steps to join GTADwellingsMaxPer.dbf to GTADwellings.shp.
- GTADwellings.shp > Symbology Tab > Quantities > Graduated Colours > Values = MaxPeriod. Select a suitable color ramp.

Questions:

- Which hypothesis about the process of urban growth in the GTA was supported by your findings?
- Which representation is most helpful for answering this question? Why?
- Where do you think further development will take place over the next 15 – 20 years? See answer sheet.

3. Identify gradients of urbanity based on density of coffee shops.

3.1. Before analysing gradients of urbanity take a look at the following predictions concerning the social aspect of urbanization.

a) The number of coffee shops is proportional to the population.
b) Urbanites visit coffee shops more often than those who live outside of urban centres; hence there would be higher number of coffee shops per 1000 inhabitants in urban centres.
c) Urbanites prefer Starbucks to Tim Horton's, hence there would be a higher ratio of Starbucks : Tim Horton's in urban centres.

3.2. Count the number of coffee shops (Tim Horton's and Starbucks) in each municipality within the GTA.

- Add "Starbucks_and_TimHortons.shp" to your map.
- Check "Counting Points Within Polygons" under "Procedures".
- Follow steps to count the number of coffee shops within each municipality.
- Save the results and add them to the map.

3.3. Is the density of coffee shops proportional to population?

- Display the total number of coffee shops in each municipality.
- Use "Normalization" (Symbology tab) to account for population in 2001.
- Note: This normalization only changes the display, not the data. Hence it is not easy to determine the exact value for each municipality.

Questions:

- Is the density of coffee shops proportional to population or are there gradients of urbanity?
- If there are gradients, how would you describe them?

3.4. Create a new variable that expresses the number of coffee shops per 1000 inhabitants.

- In the attribute table for the GTA municipalities with the newly created column for the coffee shop count, add a new Field (variable) "Coff_pop" (under Options).
- Use the Field calculator (right click).
- Calculate the variable as 1000 * Coffee shop count / Population in 2001 (select appropriate Fields).
- Sort the attribute table by the new variable "Coff_pop".

Question:

- Which municipalities have the most coffee shops per 1000 inhabitants? See answer sheet.

3.5. Challenge: How would you go about calculating the ratio between the number of Starbucks and the number of Tim Horton's coffee shops?

3.6. What other processes could affect the density of coffee shops?

- Display the actual locations of the coffee shops.
- Add major roads (GTARoads.shp) to facilitate interpretation.

Question:

- What other processes beyond degree of urbanity could lead to the observed pattern? Think about factors that we did not previously consider.
- In what types of locations are coffee shops typically concentrated?

Procedures

Clip

In ArcMap open the Ontario Municipalities layer (ONmun.shp) >>
Open the Attribute Table (right click) >>
Sort by attribute (here: Name) (right click on column header, sort ascending) >>
Using the CRTL key, select the municipalities that represent the GTA (as listed below; do NOT select the lines with Feature = Waterbody) (click in left-most field to select a line) >>
Right-click the layer and select Data >> Export Data >> Save as GTAClip shapefile.

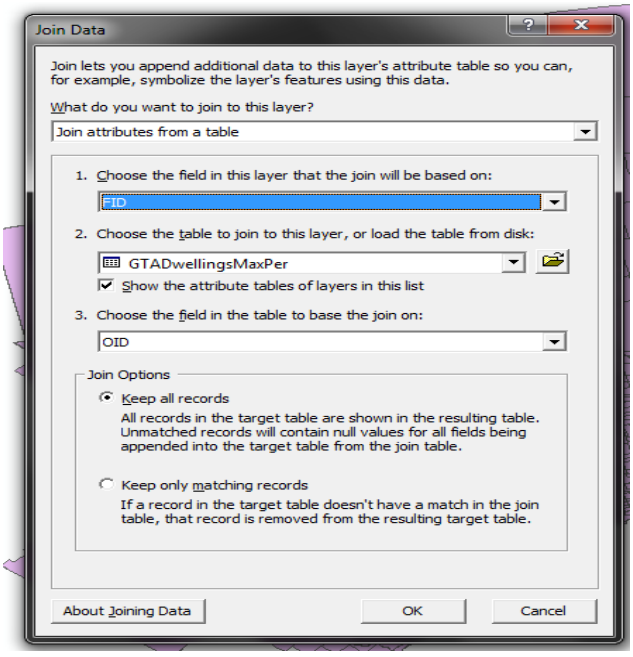
GTA: Ajax, Aurora, Brampton, Burlington, Caledon, East Gwillimbury, Georgina, Halton Hills, King, Markham, Milton, Mississauga, Newmarket, Oakville, Oshawa, Pickering, Richomnd Hill, Scugog, Toronto, Uxbridge, Vaughan, Whitby, Whitchurch-Stouffville.

Define Map Extent

In the ArcMap menu, go to View, Data Frame Properties >>
Data Frame >> Clip To Shape >> Enable >> Specify Shape >>
Data Frame Clipping: Outline of Featuers >> select GTAClip file.
Features: All >> OK.

Joining Attributes From a Table

Add GTADwellings and GTADwellingsMaxPer >>
Right-click the shapefile (GTADwellings.shp) >>
Joins and Relates >> Join >>
Join attributes from a table >>
Choose the field in this layer that the join will be based on: here FID >>
Choose the table to join to this layer: here GTADwellingsMaxPer.dbf >>
Choose the field in the table to base the join on: here OID >> Keep all records
Click OK.



Counting Points Within Polygons

First, create a variable "count" that has the value 1 for each line (i.e., each coffee shop). Note: in this example, the variable "count" or "count_tot" may already have been created for you, check.

Once the point data has been added >>
Open the Attribute Table for the point data >>
Options >> Add a Field (column); "Count" >>
Right-click on "Count" >> Field Calc. >> Count = 1 >>
Close Attibute Table.

Then, count the points (coffee shops) within polygons (municipalities):

Right-click GTAClip >> Join and Relates >> Joins >>
Join Data from another layer based on spatial location >> Select file with coffee shops
Select point data >> Select "SUM" >>
Save output file as GTACoffeeJoin.

Terminology

Vector Data Model: uses points and their associated *x*- and *y*- coordinates to construct spatial features of points, lines and areas. No measurements are possible with points; also points can be small scale representatives of larger polygons. Lines can also be used to represent polygons. With polygons both area and perimeter can be measured. Each geometric feature has certain attributes (for a lake polygon – depth, water quality and pollution) and theses attributes can lead to higher definition of mapping features (e.g. colour gradients). Different features can also be compared (points and polygons – wells and lakes).

Joining Spatial and Attribute Data: the two data types are linked by the feature IDs. These IDs are a type of key (a common field whose values can uniquely identify a record in a table).

Wards: is a sub-district usually within municipalities to distinguish electoral voting areas.

MaxPeriod: is a numerical reference to a time period of when the largest increase in dwellings occurred for a ward – for which numbers relate to the time periods see metadata spreadsheet.

Answer Sheet

Full name: _____

Fill in the required files:

File name	Data Type	Description
		GTA clip from Task 2.
		File with joined data from Task 4. (Joined Coffee Shops to GTAClip)

Answer the following:

1. Where do you expect the GTA to develop in the next 10 – 20 years?

   _____

   _____

2. Rank the top 7 GTA municipalities by level of urbanity measured by coffee shop density.

   1. _____
   2. _____
   3. _____
   4. _____
   5. _____
   6. _____
   7. _____