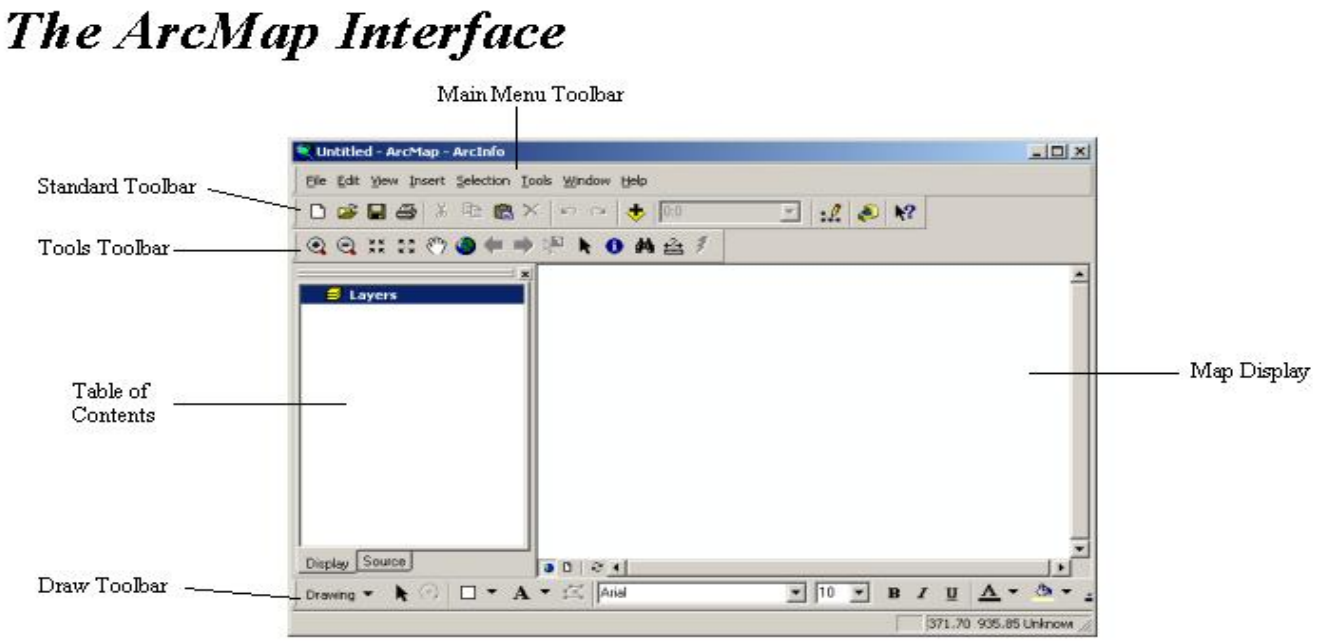
Working with Census Data in ArcMap 9.x

Open ArcMap 9.x. Go to Start > Programs > ArcGIS > ArcMap**.** The program opens up to a screen that looks like this:



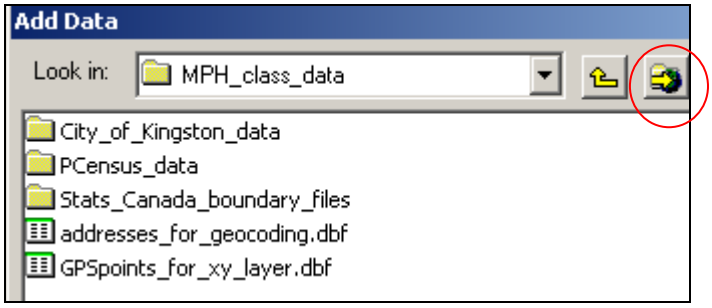
Open the View menu at the top. Make sure Status Bar is turned on. Then select Toolbars and turn on Draw, Layout, Main Menu, Standard, Tools. The toolbars float but you can dock them anywhere you like.

Standard Toolbar: tools to navigate directories, Add Data, among others

Tools Toolbar: tools to navigate in the View canvas as well as finding, identifying and measuring distance

Draw Toolbar: similar to that found in most Windows applications

Click on the  Add Data button (Black 'plus' sign with yellow behind it) on the Standard bar and use the  Connect to Folder button to navigate to the **MPH_class_data** folder.



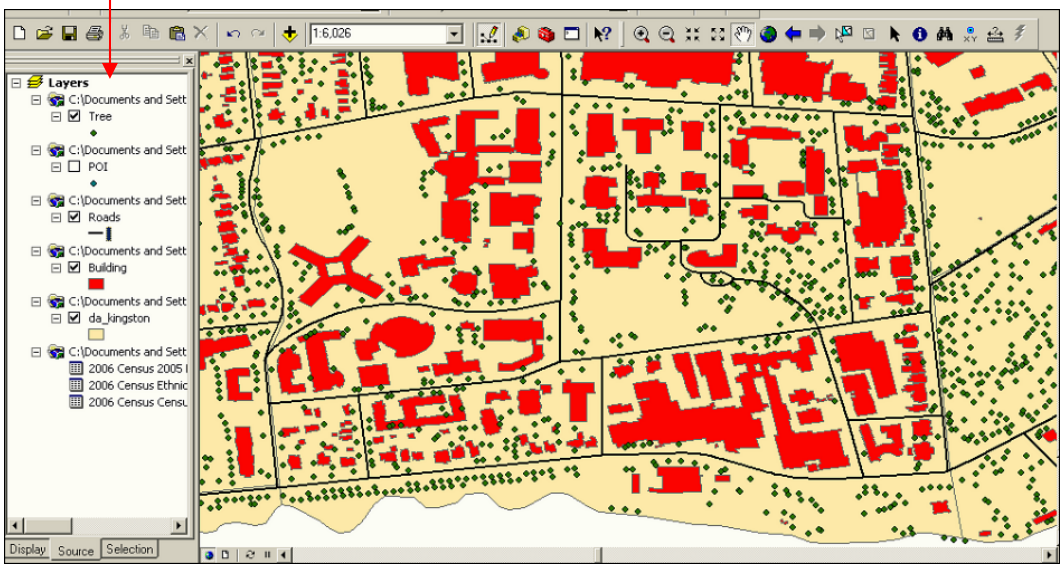
PCensus_data folder: Add *Census Snapshot.dbf. (*Later, explore mapping with *2006 Household Income.dbf* and *Ethnic Origin Summary.dbf*, if there is time*)*
Stats-Canada_boundary folder: Add *da_Kingston.shp;*
City_of_Kingston folder: Add *Buildings; POI; Roads: Tree; Water* (you will need to drill down)

These are Census population and housing data files and boundary files for Dissemination Areas that have been selected and downloaded for this class. Some extra data prepared by the City of Kingston is included. To obtain similar census data files yourself, consult *Extract Census data in PCensus to use in ArcMap 9.x and Join PCensus data with Census geography files (in ArcMap 9.x)* *** Abbreviations for the data from PCensus are explained in the metadata file *Export Log File.txt*, also named *Study Area Export Information,* stored in the folder. See the Statistics Canada Census Dictionary for complete information about each data element.

To obtain other geographic and statistical data (including health statistics), consult MADGIC, the Map, Data, and Government Information Centre in Stauffer Library.

After adding data, your map may look something like this:

Table of Contents



Renaming files: Notice the filenames that appear in the Table of Contents on the left. You can rename these to better describe content. Renaming is especially useful when working with many layers with confusing or similar labels.

To rename: Right click on a filename and choose Properties. In the Layer Properties dialog box click on the General tab and type in a name for that file – for instance Points_**of_interest** for **POI**. You won't be able to change the names of the database files.

Turn layers on and off by clicking on the check mark in the box to the left of the layer name.

Zoom and pan by using the navigational tools on the Tools toolbar. If you need to return to the full extent, use the icon that looks like a blue globe. Notice the scale bar.

Three tabs at the bottom of the Table of Contents: **Display, Source,** and **Selection.**

When **Display** is active, you can change the order in which layers display by dragging the names. The bottom layer plots (or draws) first and is covered by subsequent layers. For instance, zoom in quite close and drag the **Building** layer on top of the **Roads** layer and then put the **da_kingston** layer on top of everything.

The **Source** tab shows the layers in each data frame with the layers organized by the folders or databases in which the data sources referenced by the layers can be found. This view will also list tables that have been added to the map document as data.

The **Selection** tab shows a list of layers in the active data frame and lets you check which ones will be selectable.

Change colours:

Turn off the **Tree** and **Points_of_interest** layers.

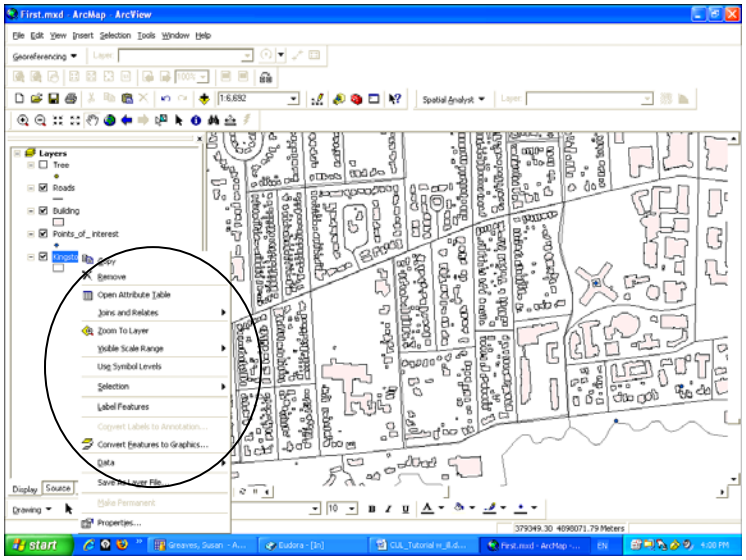
Make sure that the **Building** layer is visible and click once in the color box below the name. In the Symbol Selector menu, select another Fill Color or Pattern.
Again, make sure the **da_Kington** layer is visible. Click on the color box below **da_Kington** and make the Outline Width thicker, the Fill Color different. Notice the top area for No Color under Options > Fill Color.

Labeling:

Try labeling features. Make sure that the **Roads** layer is visible. Right click on the Layer name and select Properties. Click on the Labels tab. In the Text String roll-down, select STREET_CP. Choose a bold size 10 font in the Text Symbol menu. In the Scale Range menu, choose same scale as the feature layer. Click OK and OK. Then, Right Click on the Roads layer and select Label Features. (To remove the labels, check Label Features off.)

Layer properties in ArcMap 9.x

Right click on any shapefile and an option box opens.



Copy – make a copy for pasting.

Remove – delete from the Map Display and the Table of Contents.

Open Attribute Table – all shapefiles have tabular data associated with them.

Joins & Relates -- Joining two tables appends the attributes from one onto the other based on a field common to both tables. Relating tables defines a relationship between two tables—also based on a common field—but doesn't append the attributes of one to the other. Instead, you can access the related data when necessary.

Zoom to Layer – that Layer becomes the focus in the Map Display.

Visible Scale Range – you can set the scale where layers are or are not visible.

Use Symbol Levels -- symbol level drawing allows you to achieve special cartographic effects.

Selection – allows selection of certain parts of a layer.

Label Features – to set the label properties, open the Layer Properties window.

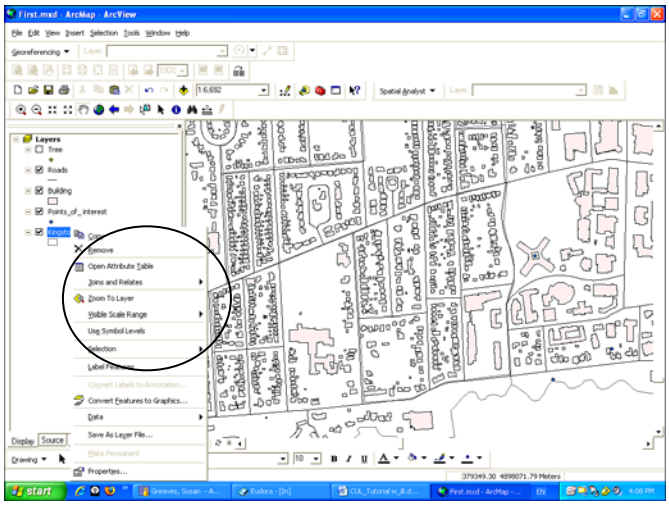
Convert Features to Graphics –converts the features in a layer to graphics that can be moved, resized, and edited on the map. This is useful if you want to change the location of features relative to each other for cartographic purposes (such as generalization), but you don't want to edit the source data that your layer represents.

Data – allows you to export the data in the current layer to another location.

Save as Layer File – allows you to save the current file with all of your adjustments as a separate layer file in the map document.

Properties – used the most often to perform many different operations on the layer.

Right clicking on a database file offers different options.



Copy Records – to a new table

Remove – this table from the project

Open – opens the Attributes Table in a new window

Joins and Relates – same as above

Data – allows exporting tabular data to another location or another format (i.e. .csv to .dbf)

Geocode Addresses – with sufficient data to work with, provides latitude and longitude values for a list of street addresses.

Display Route Events -- route event tables are typically organized around a common theme. For example, an event table for highways might include speed limits, year of resurfacing, present condition, and accidents.

Display XY Data – displays latitude / longitude points from the table on the map.

Properties – set names for fields, build queries, perform joins

Join spatial and attribute data in ArcView 9.x

Make the Source tab is active in the Table of Contents so that all the available data sets are visible.

Open up and view the Attribute Tables for the *da_kingston* shapefile and the Census Snapshot database file. Notice the headers, especially the ones labeled DAUID and CODE. These are the fields that we will use to make a join between the two tables.

Attributes of da_kingston

FID	Shape *	DAUID	CSDUID	CCSUID	CDUID	ERUID	PRUID	CTUID	CMAUID
0	Polygon	35100029	3510010	3510010	3510	3515	35	5210011.02	521
1	Polygon	35100030	3510010	3510010	3510	3515	35	5210011.01	521
2	Polygon	35100031	3510010	3510010	3510	3515	35	5210011.01	521
3	Polygon	35100032	3510010	3510010	3510	3515	35	5210011.02	521

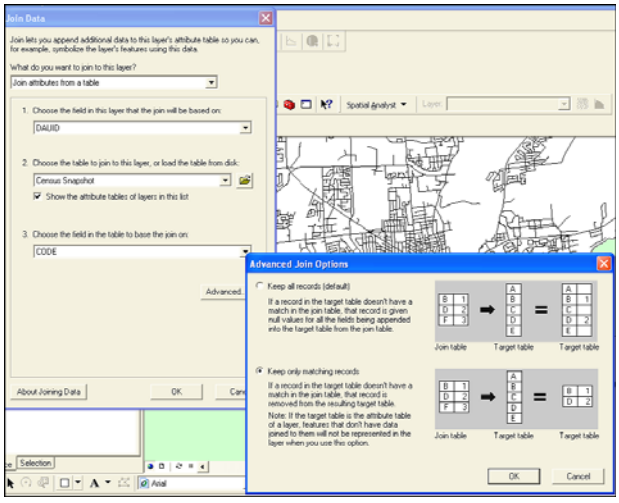
Attributes of 2006 Census Census Snapshot

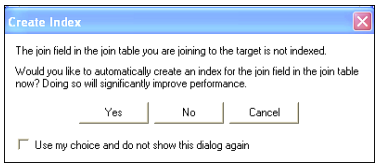
OID	AREA_NAME	CODE	XCOORD	YCOORD	TOTPOP	MALES	FEMALES	POPAGE0004
0	35100193	35100193	-76.48832	44.2311	420	200	220	5
1	35100194	35100194	-76.48385	44.22946	465	235	230	5

Turn on *da_kingston* layer by putting a check mark in the box.

Right click on the file name and highlight Joins and Relates. Select Join. The Join Data dialog box opens. For this class, fill in the following and click OK when done.

What do you want to join to this layer? ***Join attributes from a table.***
Choose the field in this layer that the join will be based on. ***DAUID***
Choose the table to join to this layer. ***Census Snapshot***
Choose the field in the table to base the join on. ***CODE***
Click the **Advanced** button and select **Keep only matching records.**
Click OK.





When prompted to create an index, you can say No.

Now, right click on the *da_kingston* layer and look at the attribute table. You will notice that census attribute data had been joined to the Dissemination Area geographic data. Also, the headers have changed. They look something like *da_kingston.CMAUID* or *2006 Census Snapshot.AREA_NAME* where the two table headers are separated by a period.

da_kingston.CMAUID	2006 Census Census Snapshot.OID	2006 Census Census Snapshot.AREA_NAME
521	72	35100029
521	62	35100030
521	63	35100031

Clean this up to make the data easier to refer to and to work with in the next exercises.

Right click on the d*a_kingston* file name, highlight Data and select Export Data.

Export all features.

Use the same coordinate system as this layer's source data.

Output the shapefile to the same folder you are using.

When asked about adding the exported data as a layer, say yes.

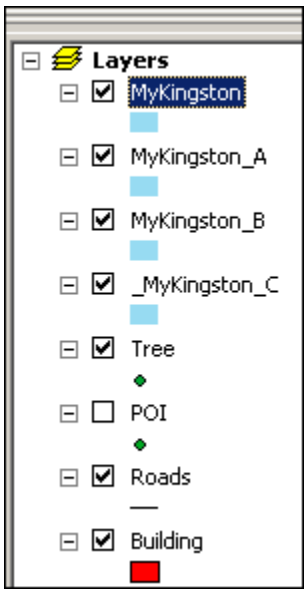
Give the new file a useful name, such as MyKingston.

Right click on the layer and look at the attribute table.

Attributes of MyKingston

FID	Shape *	DAUID	CSDUID	CCSUID	CDUID	ERUID	PRUID	CTUID	CMAUID	OID_	AREA_NAME
0	Polygon	35100029	3510010	3510010	3510	3515	35	5210011.02	521	72	35100029
1	Polygon	35100030	3510010	3510010	3510	3515	35	5210011.01	521	62	35100030
2	Polygon	35100031	3510010	3510010	3510	3515	35	5210011.01	521	63	35100031
3	Polygon	35100032	3510010	3510010	3510	3515	35	5210011.02	521	73	35100032
4	Polygon	35100033	3510010	3510010	3510	3515	35	5210011.02	521	74	35100033

Make duplicate layers. (Why? Just to have material handy to use in the exercises ahead.) Right Click on MyKingston > Copy. Then Edit > Paste. Repeat this process a couple of times to build up layers of themes which have all of your extracted data in their attributes tables. You can rename each one.



At this point, give your ArcMap Document (*.mxd) a name and save it in the data folder.

Create thematic maps in ArcMap 9.x

Types of thematic maps in *ArcMap*

This section is a brief description of the types of maps available for use in *ArcMap*. Similar theory can be found in most introductory cartography or GIS books.

Single Symbol All the features in the theme are displayed with the same colors and symbols. This is useful when you only need to show where a theme's features are located. An example is the boundaries of states, roads, and a single symbol for cities.

Unique Value Each unique value in a theme is represented with a unique symbol. The symbol for one value is no more or less prominent than the symbols for any other value.

> **Is your data categorical, like names of places or types of roads?** Use the Unique Values Legend Editor.

Graduated Color (often called **choropleth**) Features are displayed with the same symbol type, but the colors represent the progression of values for a data attribute you specify.

> **Is your data normalized, like percentage data or population per unit of area?** For polygon data, your first choice of legends should be Graduated Colors. If you have point or line data, then use a Graduated Symbol legend. If you need to show several normalized attributes, consider using the Chart legend.

Graduated Symbol The features are displayed with the same colors and symbols, representing a progression of values. This is the best way to symbolize data that expresses size or magnitude. Graduated symbol is only available for point and line data.

> **Is your data ranked, like best sales territory to worst sales territory?** Use either Graduated Symbols or Graduated Colors as these types of legends are best suited to show a progression of data values. However, if your data is qualitative, Unique Values may also be appropriate.

Dot Density The features of a polygon theme are displayed with a number of dots corresponding to a value. This method is good for showing how particular things are distributed throughout an area. For instance, a dot map depicting population will most likely have the strongest concentrations of dots along rivers and near coastlines.

> **Are your data raw counts, like population or barrels of crude oil?** For polygon data, your first choice of legends should be Dot Density .If you have point or line data, then use a Graduated Symbols legend. If you use a Graduated Symbol or a Chart legend you may also want to consider normalizing your data's values to make your legend easier to read.

Chart The features are displayed with a chart. The components of the chart correspond to data attributes you specify, and the size of each part in a chart is determined by the value of each data attribute. You can specify whether the charts are pie charts or column charts. This is a good method for displaying the values of many attributes.

Classification methods in *ArcMap*

ArcMap provides seven classification methods to display data; Manual: Equal Interval, Defined Interval, Quantile, Natural Breaks (Jenks), Geometric Interval, Standard Deviation. For more details on each method, see the help topic "Classification Schemes" within the Help function in the program. (This help material is definitely worth reading!)

ArcMap exercises

1. What is the median age (MEDAGE), both sexes, by dissemination area (da)? What is the average household income (IHTAV), by da?

PCensus Census Snapshot data contains values for these attributes. These values can be displayed in a graduated color map.

Right Click on MyKingston. Select Properties.

In the Layer Properties box, click the Symbology tab.

In the Show box, select Quantities > Graduated Colors.

In the Fields > Value drop-down box select *MEDAGE*.(or IHTAV).

In the Fields > Normalization drop-down box, select **<**NONE**>.**

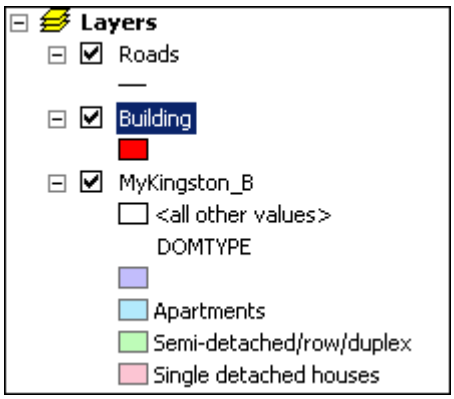
In the Classification > Natural Breaks drop-down box, choose 5.

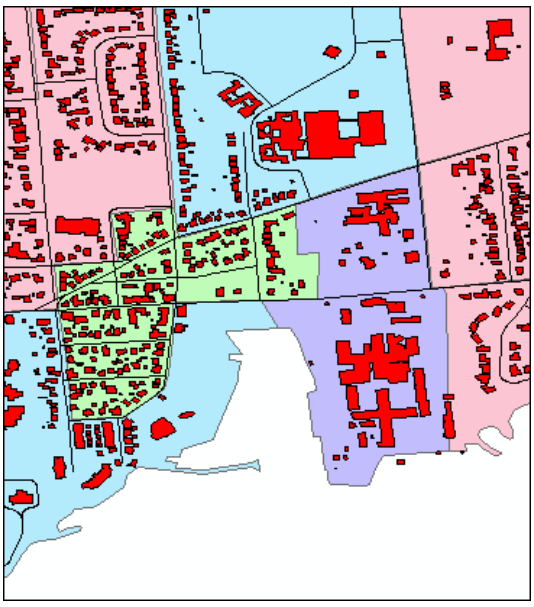
Pick a color ramp > click OK .

Go back and change colors and number of breaks to change the display.

2. Where are the Dominant Type of Housing (DOMTYPE) single detached houses, semi-detached or row houses, apartments or other in the Kingston CMA?

PCensus Census Snapshot has a field called DOMTYPE which lists this data. Show the categories of housing using Symbology → Show Categories → Unique Values → Add all values.





A few places have no dominant housing type assigned. Why is this?

3. Where do concentrations of people aged 20 – 24 live? People over 85?

In the data are rounded counts by DA for the total numbers of persons aged 20 through 24 (POPAGE2024), 85 and over (POPAGE85PL) and total population (TOTPOP).

Method 1. Make a **dot density** map.

Right click on MyKingston . Select Properties. In the Layer Properties box, click the Symbology tab.

In the Show box, select Quantities > Dot density.

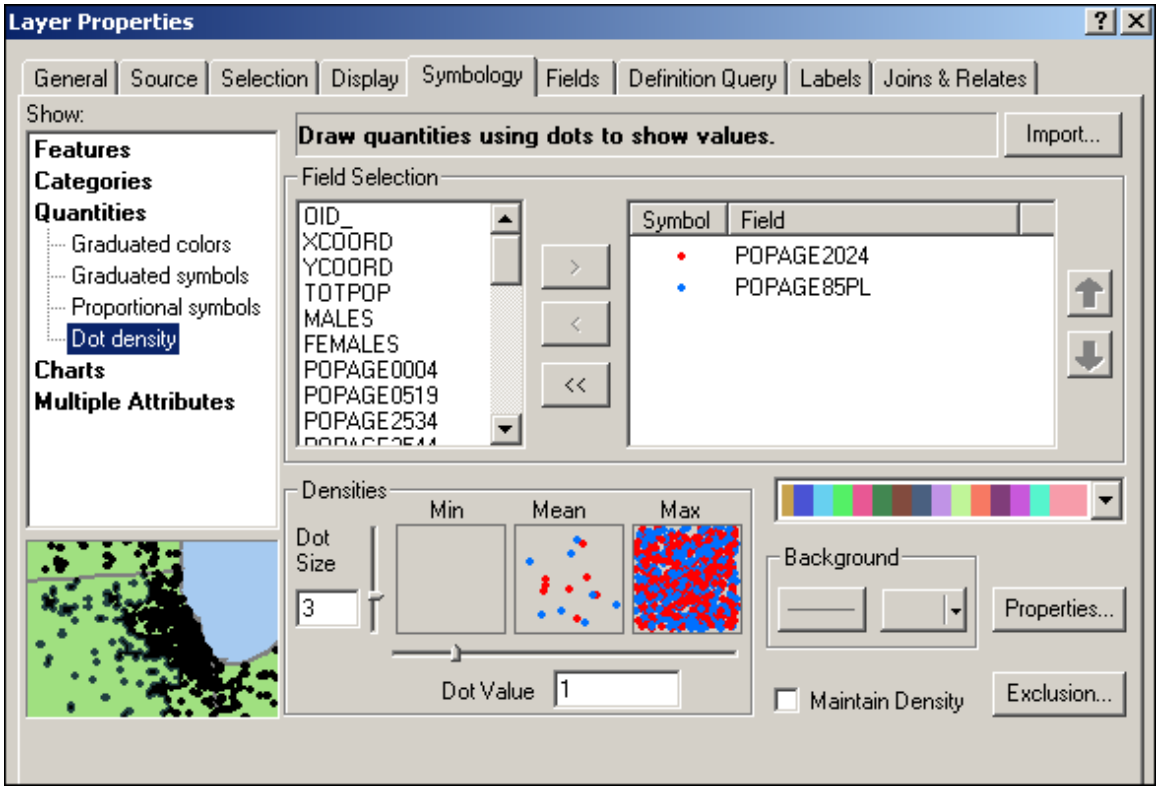
Choose Field selection: *POPAGE2024* and move it into the Symbol and Field box.

Notice Dot value and Dot size. *ArcMap* will calculate a default number for each dot representing values in this particular field. This default will vary, depending on the values in the field, the size of the area that these values apply to, the scale of the map on the screen. You can always modify this number and will may need to do so, in order to preserve comparability. Click on the Symbol and choose a color, shape, size, etc. for this Density Field. Click OK to view this theme plotted on the map.

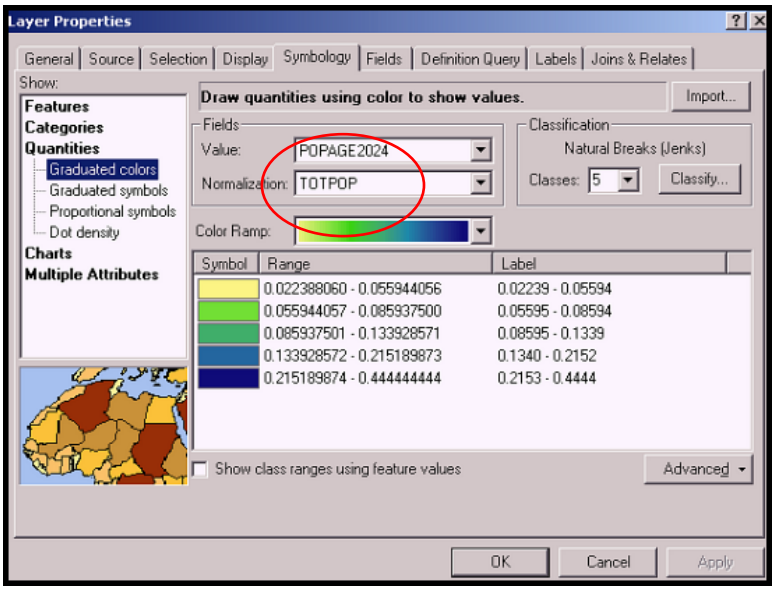
Repeat the process, choosing *POPAGE85PL* in the Field Selection box. Select contrasting colors in the Symbol and Field box. Select the order of plotting by rearranging the terms in the FIELD box.

Experiment with Dot Size and Dot Value. Zoom in and out on the map.

If you have questions about why there are anomalies or gaps in the data display, it is always possible to query the data for a particular area by using the Information Button .



Method 2. Use the **graduated color** method from exercise 1. to show a progression of percentage values. This time, use the count for each age range **normalized** by the count for total population (TOTPOP). Using ratios or normalizing the data is one method for mapping quantitative data to minimize differences based on the size of areas or number of features in each area. Dividing the 20 to 24-year-old population by the total population yields the percentage of people aged 20–24. Similarly, dividing a value by the area of the feature yields a value per unit area, or density.



Change the Label that appears in the Table of Contents (and eventually in the legend) by highlighting the Label text and typing OR right clicking in the same area and choosing Format Labels. Change the value Range by typing in the higher end of the desired range – i. e. for ."001 - .02", type ".02". (Be careful that you choose appropriate ranges for the subject.)

Change the Classification method to examine your data in various ways.

4. Compare the ratio of owned dwellings to rented dwelling.

PCensus Census Snapshot data contains values for owned (DWOWN) and rented (DWRENT).

*One way to show ratios is to use **pie charts**:*

Right Click on another MyKingston . Select Properties, General, and change the Layer Name.

Then click the Symbology tab.

In the Show box, select Charts> Pie.

In the Field Selection box, highlight the fields *DWOWN and DWRENT* and select them.

Look at the Background Symbol choices.

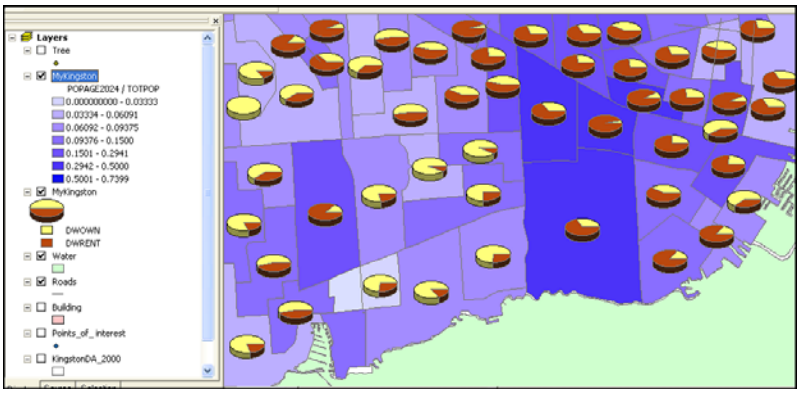
Look at the Properties dialog.

Click on Apply to view this theme plotted on the map. You may need to go back and change a few things, depending on the scale of your map.

While the new layer is active, go in and change its Layer Name to something descriptive for the Table of Contents.

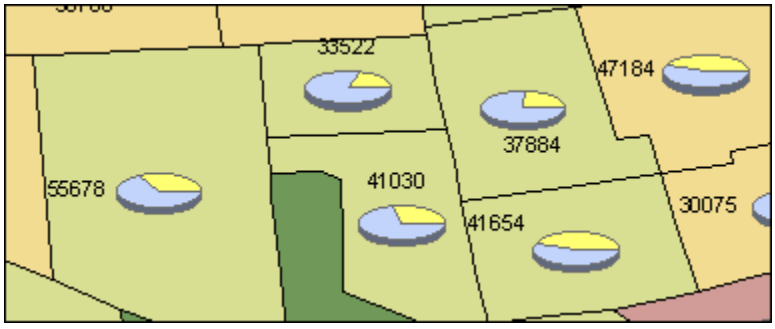
4.a) Which housing is more prevalent where college age people live? People over 85?

Make the map display the ownership/rental pie charts on top of a graduated color layer showing concentrations of persons of a certain age group.



4.b) How does average household income (IHTAV) compare with the prevalence of owned or rented housing?

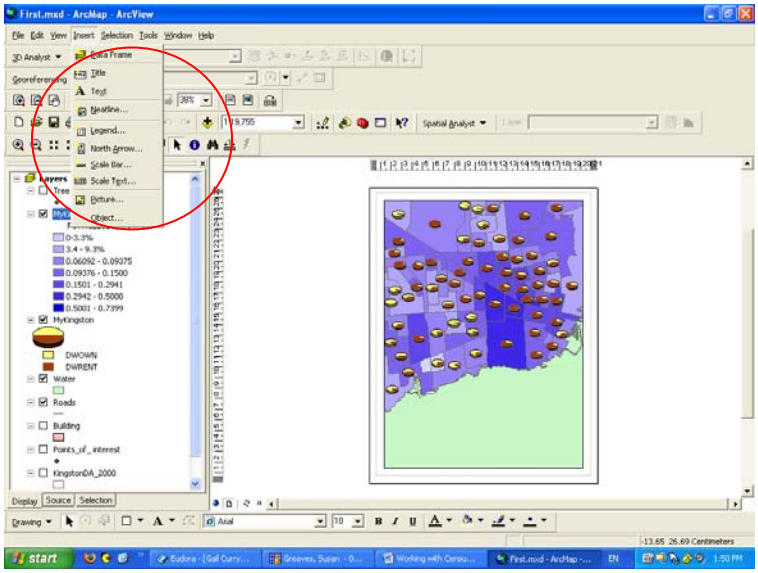
On top of the layer showing age in the map created above, go to Properties and change the Label field to be IHTAV. Then, right click on the age layer and turn the 'label' box on. Now each Dissemination Area polygon will be labeled to show the average household income in the area, in addition to whatever you have added concerning age or home ownership.



This can get complicated! Remember that all of this data can also be queried and selected by condition (for instance, median income > $ 100,000 AND average age > 50). There are many, many possibilities with the software. First of all, it is really important to get to know your data.

5. Create a map layout

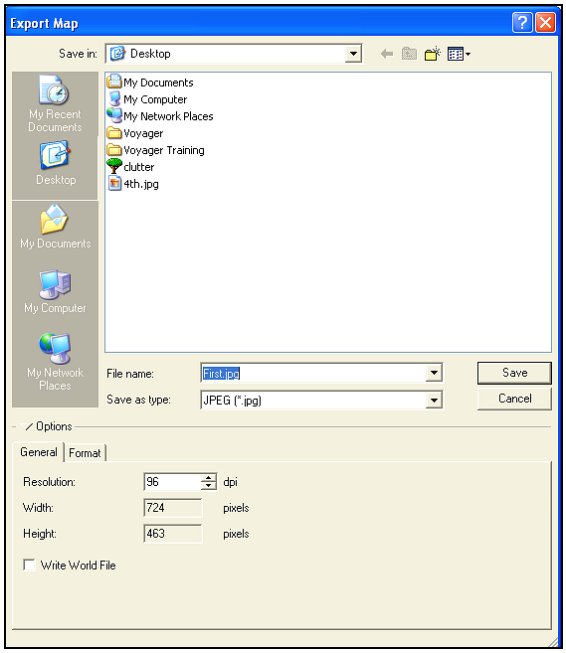
Press the Map Layout button: 



Map layout options can be inserted here. At this point, it is really helpful to have descriptive names for your layers.

6. Insert your finished map into a general document type such as Word or PowerPoint

By doing the following, you will export a flat graphic image file that will not have the layer capabilities of the GIS data files you have been working with.



Choose file … Export map. In the Dialogue box select the destination, a file name, and a file type. JPG is excellent for Word, PowerPoint, and web pages. There are some options for resolution and colour format. After the image file is saved, try opening Word and insert your map as a picture.

Bonus exercise: Add points defined by x,y coordinates to your map

In addition to data sources such as a shapefile, you can also add tabular data that contains geographic locations in the form of x,y coordinates to your map.

X,Y coordinates describe discrete locations on the earth's surface such as the location of fire hydrants in a city or the points where soil samples were collected. You can easily collect x,y coordinate data using a global positioning system (GPS) device.

In order to add a table of x,y coordinates to your map, the table must contain two fields—one for the x-coordinate and one for the y-coordinate. The values in the fields may represent any coordinate system and units such as latitude and longitude or meters.

Your class date includes *GPSpoints_for_xy_layer.dbf*, an example of tabular data containing x,y, point locations:

LOCATION	Y(long)	X (lat)	CONDITION
1	44.22	-76.53	good
2	44.22	-76.52	poor
3	44.24	-76.51	excellent
4	44.24	-76.53	poor
5	44.25	-76.54	poor
6	44.26	-76.54	good
7	44.26	-76.5,	excellent
8	44.25	-76.52	poor
9	44.25	-76.54	poor
10	44.25	-76.5	good
11	44.24	-76.5	excellent

Open a new map. Start by adding *da_kingston.shp*. Add *GPSpoints_for_xy_layer.dbf*. (At this point you can add *roads.shp* to provide context, but in order to project properly, start with *da_kingston.shp*.)

Click the Tools menu on the Standard toolbar and click Add XY Data. Click the table dropdown arrow and locate the table *GPSpoints_for_xy_layer.dbf*. If the table is not on the map, click the Browse button to access it from disk.

Click the X Field dropdown arrow and click the X field containing x-coordinate values.

Click the Y Field dropdown arrow and click the Y field containing y-coordinate values.

If you know the coordinate system used to collect the points, then click Edit to define the coordinate system and units represented in the x and y fields. The x,y coordinates will be automatically transformed to match the coordinate system of the data frame. In this exercise, just click OK.

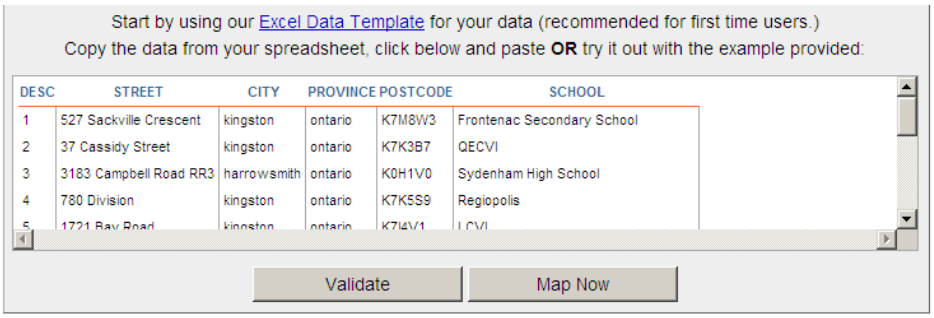
Press OK. A new Events theme called *GPSpoints_for_xy_layerEvents* is added containing all of the points defined by the data.

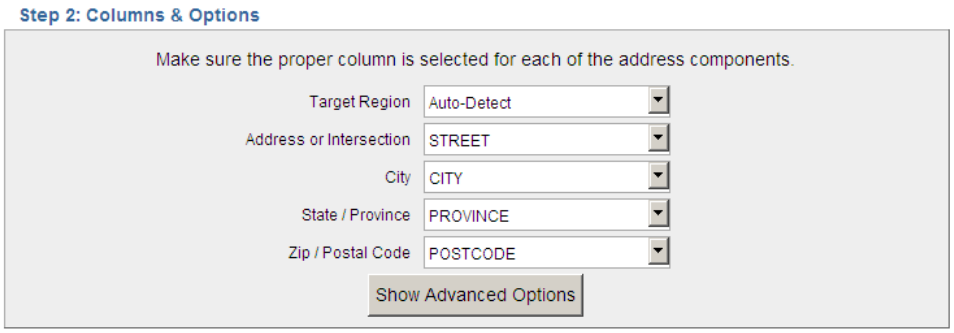
Label the points by right click on the Layer name and select Properties. Click on the Labels tab. In the Text String roll-down, select CONDITION. Choose a bold size 10 font in the Text Symbol menu. In the Scale Range menu, choose same scale as the feature layer. Click OK and OK. Then, right click on the Layer name and select Label Features.

Bonus exercise: Geocoding from street addresses

Geocoding is the process of finding associated geographic coordinates (often expressed as latitude and longitude) from other geographic data, such as street addresses, or postal codes. This is a useful guide: A Geocoding Best Practices Guide

Although ArcMap has built-in geocoding capability, there are other options. I think this is a good one. **Go to http://www.batchgeocode.com/** and follow the directions. Copy our sample file ***addresses_for_geocoding.xls*** into the Source Data box. Validate it.





Run the geocoder.
Show the Geocode data.
Copy the data into another Excel file and then save it as ***geocoded_addresses.csv.*** (CSV Comma Delimited)

Notice the new columns that have values for latitude and longitude. (You also will see the data plotted in GoogleMaps.) There is also a column that describes the geocoding method used. These result are not as accurate as other methods might be, but they also might be "good enough". The Good Practices document mentioned above goes further into discussions of accuracy and methodology.

DESC	STREET	CITY	PROVINCE	POSTCODE	SCHOOL	bg_lat	bg_long	bg_postal	bg_accuracy
1	527 Sackville Cresc	kingston	ontario	K7M8W3	Frontenac	44.2333783	-76.599538	K7M	RANGE_INTERPOLATED
2	37 Cassidy Street	kingston	ontario	K7K3B7	QECVI	44.2552882	-76.485151	K7K	RANGE_INTERPOLATED
3	3183 Campbell Roa	harrowsm	ontario	K0H1V0	Sydenham	44.405705	-76.665573	K0H 1V0	APPROXIMATE
4	780 Division	kingston	ontario	K7K5S9	Regiopoli	44.2548837	-76.49692	K7K	RANGE_INTERPOLATED
5	1721 Bay Road	kingston	ontario	K7I4V1	LCVI	44.2649064	-76.371888	K7L	RANGE_INTERPOLATED

Back in ArcMap, (ok to start from the same map as in the above Bonus Exercise) add the .csv file to your existing project. Change it to a .dbf format by right clicking on its name in the Table of Contents and export the data to a .dbf file. Keep the new .dbf file with the rest of the project data.

Now you are ready to add the new data as a XY data, as done in the Bonus Exercise above. (Hint: X is 'bg_long'; Y is 'bg_lat')