Background Information

Geocoding is the process of assigning a geographic location to point data based on a description. The description usually comes in the form of street addresses, postal/zip codes, or city names. The geocoding service then converts this descriptive information into a point feature on a map with precise location coordinates. In order to create the point feature, a **reference layer** such as an addressed street file is required.

A **Geocoding Service** defines the process for converting these descriptions to points on a map by setting the parameters of the transformation. It is possible to rerun the geocoding service in order to match unmatched points interactively and thereby increase the percentage of matched points.

It is important to note that geocoding is not an exact science. Each point that is inputted into the geocoding service is compared with potential candidates in a **Reference Table**. The points are then assigned a score based on their sameness to points in the reference table. Scores that exceed a user defined percentage are automatically matched. Points that fall below the designated grade are not matched but can be rematched interactively.

Once the points have been geocoded a new output table containing 4 new auto generated columns will appear in your map view.

- The **status** field simply indicates whether the record was matched or not. The value **"M"** for matched addresses and **"U"** designates unmatched addresses.
- The **score** field lists the score of the candidate to which the address was matched as a percentage.
- The **side** field lists the side of the street that the address was matched on if that information is available. **"L"** means that the address is located on the left side of the street and **"R"** stands for right side of the street.
- **Arc_Street** contains the address information that was geocoded.

Overview

In this tutorial, you will:

- Create a geocoding service in ArcCatalog
- Add the geocoding service in ArcMap
- Geocode addresses in ArcMap
- Interactively rematch addresses
- Conduct a series of queries on the newly geocoded data

Note: To complete this exercise, you will need to download the dataset designed to be used in this tutorial. Download the data files by clicking either the **WinZip** or the **Self Extracting** link located just below this tutorial. The **Self Extracting** link will automatically extract the files to a folder called **Geocode** on your **c: drive**.

Data for this tutorial has been provided by DMTI Spatial. If you are interested in similar data for your school area, ESRI Canada and DMTI Spatial have partnered to offer Canadian K-12 schools a Local Dataset. The Local Dataset is a complete, custom designed set of local data surrounding a school and includes a wealth of census data at the enumeration area level, a comprehensive street network file with address ranges, points of interest, land use and building footprints (where applicable), topographic features, and much more. Please contact k-12@esricanada.com for more information.

12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com

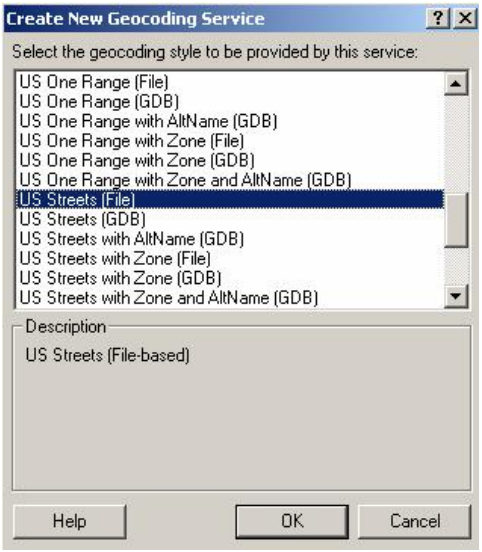
Scenario

In this tutorial you are given a table of data that consists of 26 banks in the Oakville, Ontario area, a shapefile with enumeration area data and a roads shapefile. You are going to geocode the banks using the roads file as a reference layer. Once the banks have been geocoded, then you will query the data to conduct some spatial analysis.

Creating a Geocoding Service in ArcCatalog

In order to geocode files in **ArcMap**, you must first create a **New Geocoding service** in **ArcCatalog**.

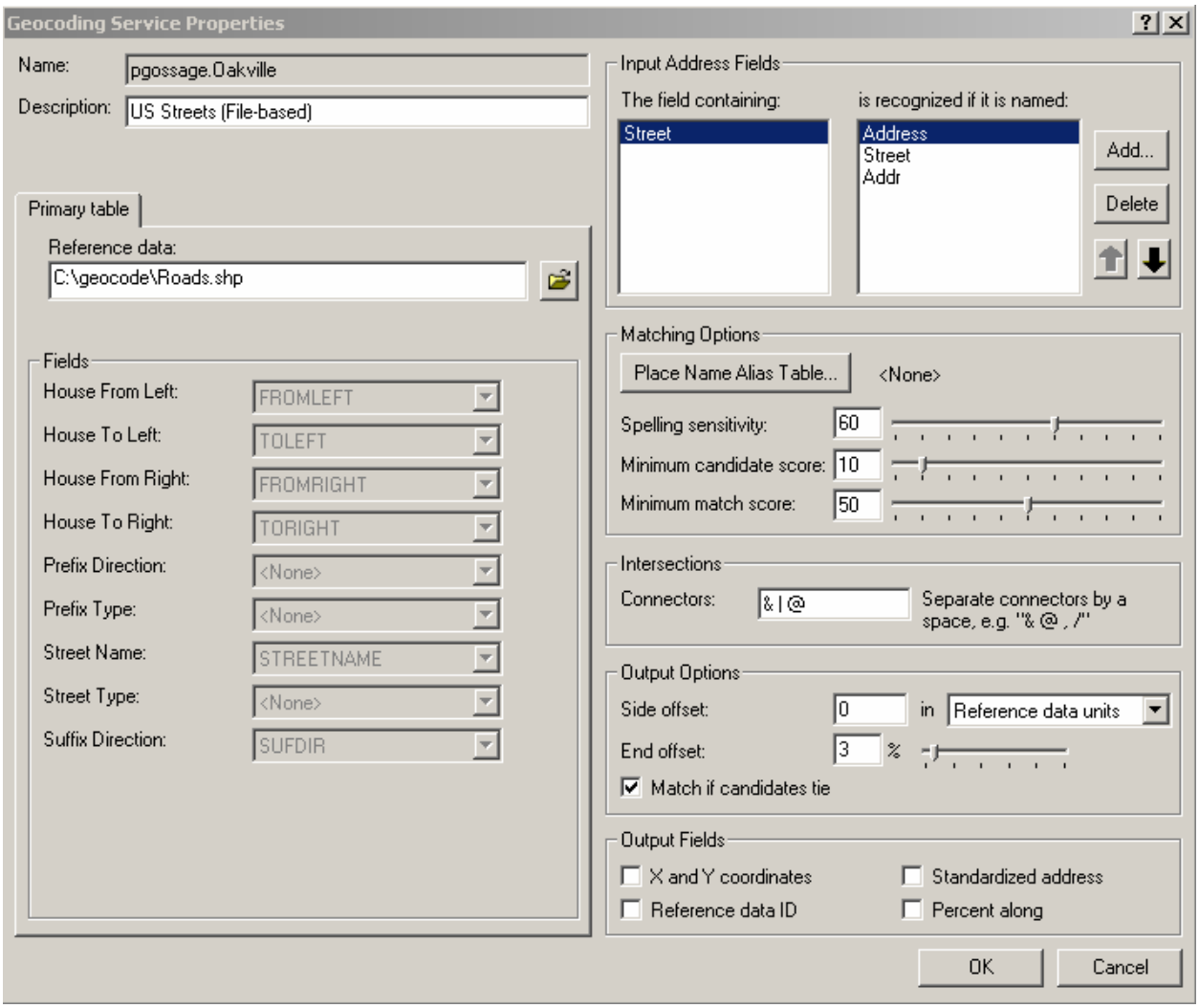
1. Open **ArcCatalog** .
2. Scroll down in the table of contents to the **Geocoding Services** tree and expand it by clicking the **+** sign next to it.
3. Double click on **Create New Geocoding Service** .
4. In the resulting dialog box, scroll down, select **US Streets (File)** and click **OK**.



5. In the **New US Streets (File)** dialog box, type in **Oakville** as the name of the service. In the **Reference Data** window, click the Browse button and browse to your **Geocode** folder select **Roads.shp,** and click **Add.** Decrease the **Spelling sensitivity** to **60** and the **Minimum match score** to **50.** Click **OK.**

Note: Reducing the spelling sensitivity and the minimum match score is important. Data in your table will sometimes have mistakes and data in the reference table could have mistakes or could be recorded in a different manner. If you set both scores equal to 100, the geocoding service would only match addresses that were identical in both tables. Reducing the scores gives the geocoding service the ability to pick up on subtle differences in spellings and the way addresses are recorded.

12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com



This will create the **Geocoding Service** that you can reference in **ArcMap** to geocode points. It is possible to change the **Matching Option** settings while you are geocoding in **ArcMap,** however it is important to note that those changes are not permanent. Those temporary changes are most often used when rematching unmatched records. In order to permanently change the geocoding service you must use **ArcCatalog.**

Geocoding in ArcMap

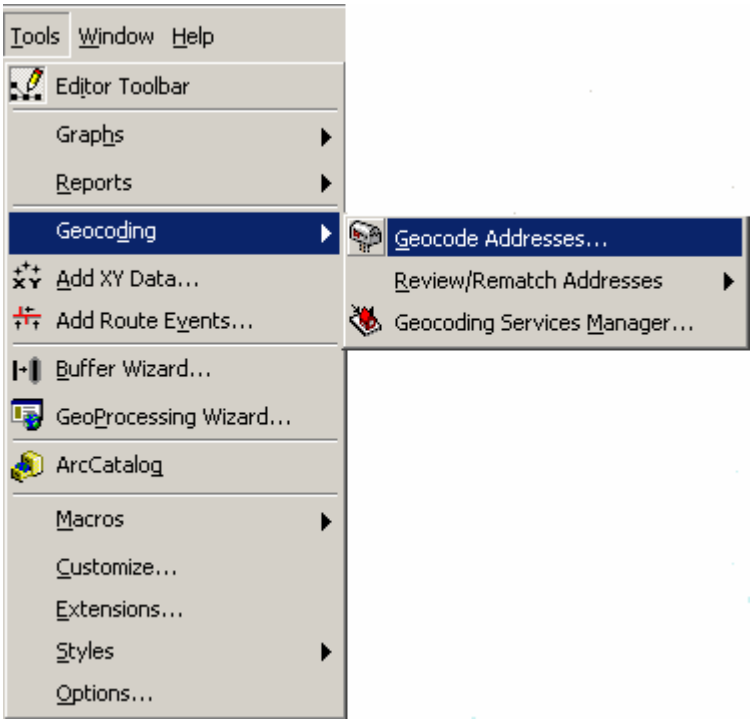
1. Start **ArcMap** .

2. In the welcome window select "Start using ArcMap with:" **A new empty map** and check the **Immediately add data** box. Browse to the **c:\Geocode** or appropriate folder on your hard drive and select **banks.dbf, EAs.shp,** and **Roads.shp.** Click **Open.**

3. Right click on **banks** in the table of contents and select **Open** in order to examine the table.

4. Notice that the table consists of 26 records of various banks in Oakville. The column that you will be using to geocode the points is the **Address** field. **Close** the **attributes of banks** table.

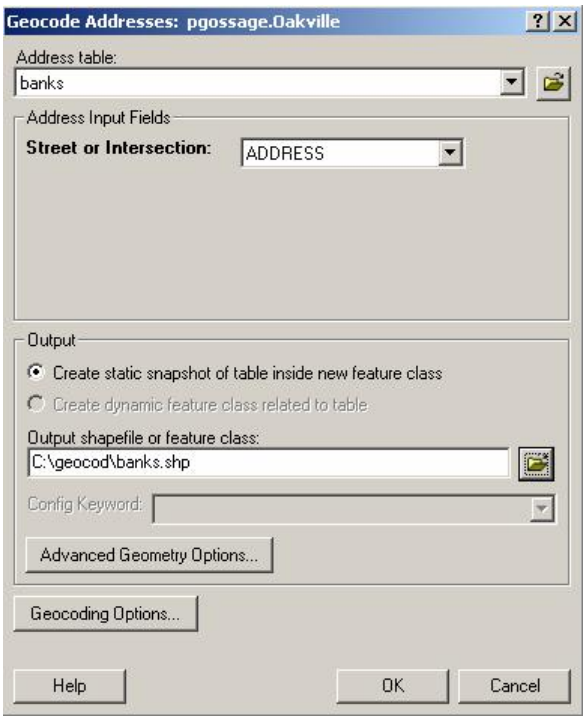
5. Click on the **Tools** menu and select **Geocoding à Geocode Addresses.**



6. You will need to add the **Geocoding service** that you created in **ArcCatalog.** Click the **Add** button and browse to the **Geocoding services** folder. *(Hint: If the geocoding services folder is not initially visible, click the **Connect to Folder** button*  *and browse to it. Click **OK** once you've located it.)* Select the **Oakville** service/locator that you created and click **Add** and then **OK.**

7. In the **Geocode Addresses** dialog box make sure that **banks** is the **Address table** and that **Address** is the **Address Input Field.** In the **Output** box, click the Browse button  and browse to your **Geocode** folder type in **banks.shp** as the name. Click **OK** to execute the geocode.

12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com

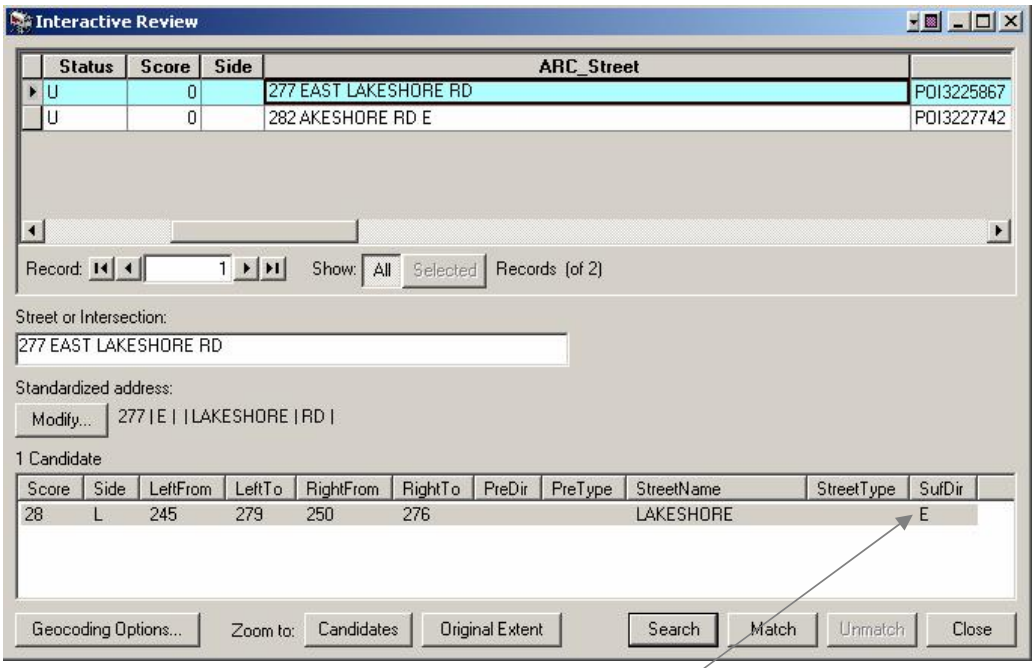


Once the geocode has been processed you will see a **Review/Rematch Addresses** window appear with the results of your geocoding. It should say that 24 addresses (or 92%) have a match score of 80% or greater and 2 addresses (or 8%) are unmatched. This is a very reasonable score however, considering that there are only two outstanding records it is worth trying to rematch them interactively.

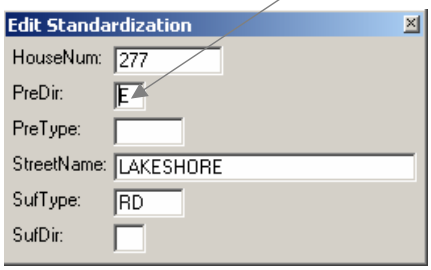
8. Click on the **Match Interactively** button in order to step through each address in an attempt to match them individually. A list of the addresses that have not been matched will appear at the top of the **Interactive Review** window.

9. Highlight the first record which should be 277 EAST LAKESHORE RD. The **candidate** window at the bottom should already be listing a potential match for this address, however the current score for it is only 28% which is quite low. Examine the candidate carefully though. What are the differences between it and the record you're attempting to geocode?

12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com

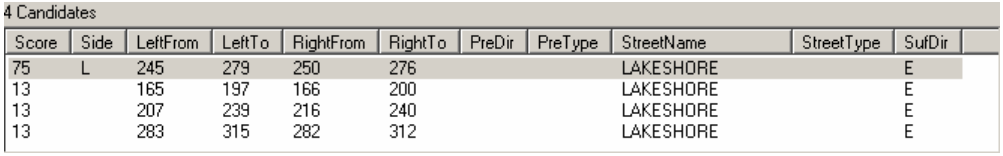


10. Notice that the directional specification *EAST* is listed before *LAKESHORE* in the **Banks** table address spelling at the top of the interactive review window. Click on the **Modify** button. This will show you how the address has been broken down by the geocoding service and allow you to conduct any necessary edits. The geocoding service has placed the letter *"E"* in the PreDir (PreDirectional) window because the directional word *EAST* came at the beginning of the address. However, in the candidate window *EAST* is listed as a **SufDir**.



11. With a little research, you discover the data tale is incorrect and the *East* should actually come after *Lakeshore.* (Hint: Compare this Lakeshore record with other records in the table.) In order to fix this mistake in the data, delete the letter *"E"* in the **PreDir** textbox and enter it into the **SufDir** textbox. Press the **Enter** key.

This should prompt some more potential *candidates* to appear in the lower window. Notice the **score** in the column on the left. The *candidates* are automatically sorted by **score** with the highest appearing at the top. The initial candidate should now have a score of 75. Notice that this is the same candidate as before but the score has been improved since you corrected the data.



12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com

12. Since it is clear that this candidate is the best match, make sure it is highlighted and click on the **Match** button. This will add this record and its score to the table of geocoded/matched addresses and the status will change from **U** to **M**.

13. Close the **Edit Standardization** window and highlight the remaining unmatched address (282 AKESHORE RD E). What does it look like the problem is with this record?

14. Hopefully you noticed that the spelling of this address is incorrect. The street should in fact be spelt like the previous record, *LAKESHORE RD E*. Click on the **Modify** button, change the *StreetName* to **LAKESHORE,** and click the **Enter** key.

15. This will prompt 4 potential candidates to appear in the candidate window. The candidate with the high score of 75 is the most appropriate match. Select that record and click **Match.**

16. All 26 points should now be matched. Close the **Edit Standardization** window and the **Interactive Review** window and click **Done** on the **Review/Rematch Addresses.**

17. The new geocoded layer will be added to your **Data Frame.** Double click on its symbol in the table of contents to bring up the **Symbol Selector.** Click on the **More Symbols** button to show the other symbol themes and select **Business.** This will add the business symbols to your selection list. Choose one of the **Bank** symbols and click **OK.**

18. Double click on the new layer name to bring up the **Layer Properties.** Click on the **General** tab and change the layer name to **Banks.**

Now you can conduct some analysis about the location of the banks. What areas might you expect the banks to be located?

Thematic Map

In this section you are going to make a graduated color map of average income to visually analyze whether there is a correlation between the location of banks and the wealthiest areas in Oakville.

1. Double click on the **EAs** theme in your table of contents to bring up the **Layer Properties.**

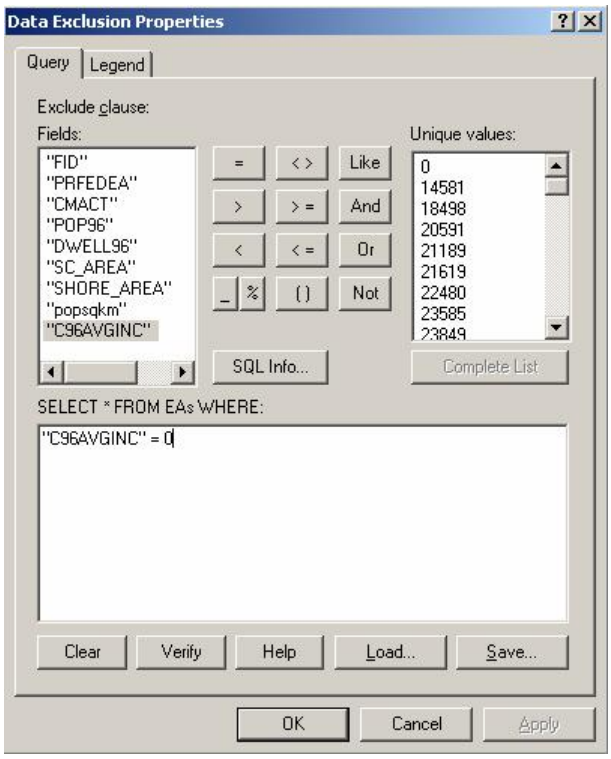
2. Click on the **Symbology** tab and click on **Quantities à Graduated Color.**

3. Click on the **Value** drop down menu and choose **C96AVGINC.**

4. Click on the **Classify** button. Choose **Quantile** for the **Method** and **4** for the number of **Classes.** This way you are roughly dividing up Oakville's population into quarters based on wealth.

5. You will notice from the graph that there are a few **0** or **Null** values that are skewing the intervals. Click on the **Exclusion** button. Double click on the **C96AVGINC** in the **Field** window. Click on the **=** sign and click on or type **0.** Click **OK** three times when this is complete to apply the changes.

12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com



You will now have a graduated color map of average income in Oakville. Where are most of Oakville's wealthiest populations located?

Querying based on Attributes

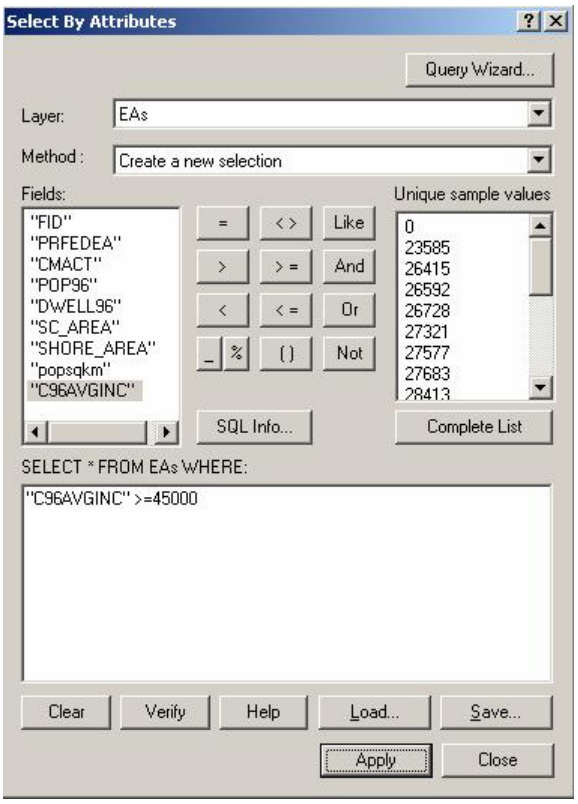
You are now going to select the wealthiest areas in Oakville; you want to select all EAs with an average income of $45,000 or more.

1. Click the **Selection** drop down menu and choose **Select by Attributes.**

2. Choose **EAs** as the layer and **Create new selection** as the **Method.**

3. Double click on **C96AVGINC** in the **fields** window, click on the **">="** sign, and type **"45000".** Click **Apply** and **Close** when this is complete.

12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com



You should now see a number of enumeration areas highlighted in blue. Right click on the **EAs** theme in your table of contents and select **Open attribute table.** You should see **12** of **71** records selected.

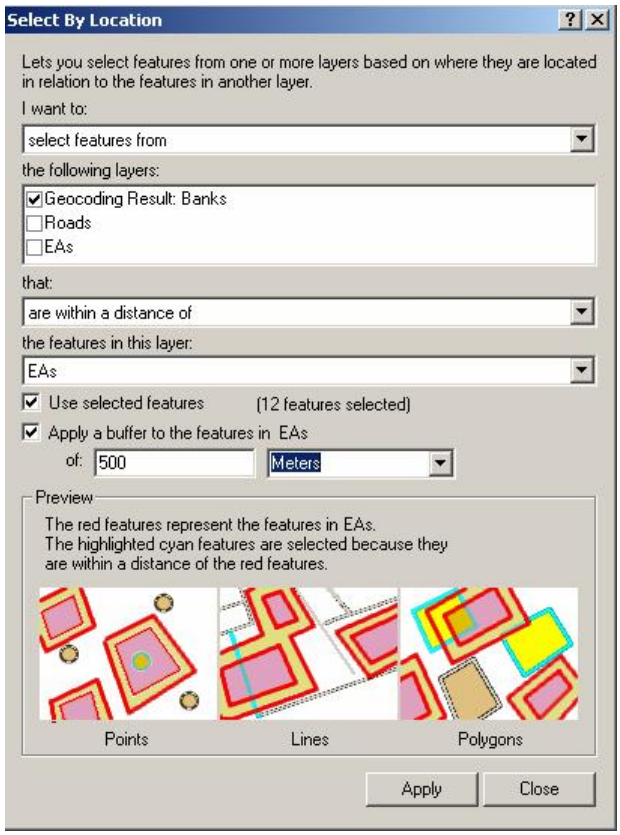
Querying by Location

*You are now going to see how many **Banks** are contained by or within 500 metres of these 12 enumeration areas.*

1. Click on the **Selection** menu and choose **Select by Location.**

2. In the resulting **Select By Location** dialog box, choose **select features from** the following layers: **Banks** that: **are within a distance of** the features in this layer: **EAs.** Check the **Use selected features** box, so that only the 12 selected features will be incorporated into the query. Also check the **Apply a buffer to the features in EAs** of **500 Meters.** Click **Apply** and **Close.**

12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com



3. Right click on the **Banks** theme in the table of contents and select **Open Attribute Table.** Notice that **15** out of **26 Banks** are located within 500 metres of one of the wealthiest census tracts.

More than half the banks are located within close proximity to the wealthiest quarter of the population. What other conclusions can you draw about the location of banks? Conduct some analysis on your own! Pretend you are the owner of a chain of banks, where would be an ideal location to build a new one?

12 Concorde Place, Suite 900 ● Toronto, Ontario, M3C 3R8● Tel: (416) 441-6035 ● Fax: (416) 441-6838
Email: k-12@esricanada.com ● World Wide Web: http://k12.esricanada.com