A procedure for geocoding point addresses using ArcView GIS 3.x[1]

Martin A. Andresen
andresen@sfu.ca
http://www.sfu.ca/~andresen

School of Criminology, Simon Fraser University

and

Institute for Canadian Urban Research Studies,
Simon Fraser University

01 September 2006

---

[1] This procedure is a work in progress, intending to include how to set up your address file, standardising addresses, geocoding addresses, aggregating point data by address and polygon, and table joining.

PREPARING DATA FOR GEOCODING

Though all steps in this geocoding procedure are important, none are more so than setting up your data. Regardless of the steps performed below, if the data are not set up correctly, the geocoding will not be successful.

A table is provided at the end of this document to aid in setting up the addresses to be geocoded.   These street types and their respective abbreviations are from Statistics Canada's roads network (provided by DMTI Spatial Inc.). If this road network is not being used, you should consult the metadata provided with your road network.

All of the address information must be in one field, or column. Some data provide the house/apartment number in one field with the street name in another field.

The easiest way to merge two, or more, columns is within MS Excel. Excel, however, does have limited memory so you may have to perform the following on subsets of your entire data or perform a similar technique with a different software program.

The Excel function to merge the two columns is "concatenate". As with all functions in Excel, the word concatenate is preceded by an equal sign. Each cell to be concatenated is separated by a comma: =concatenate(A1, B1, C1)
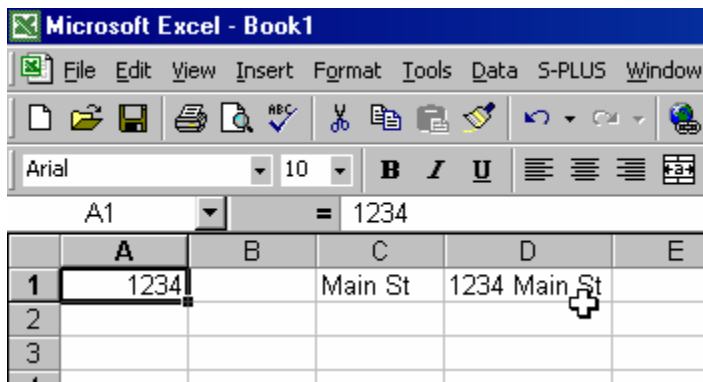
Performing this function on two columns appears as follows:
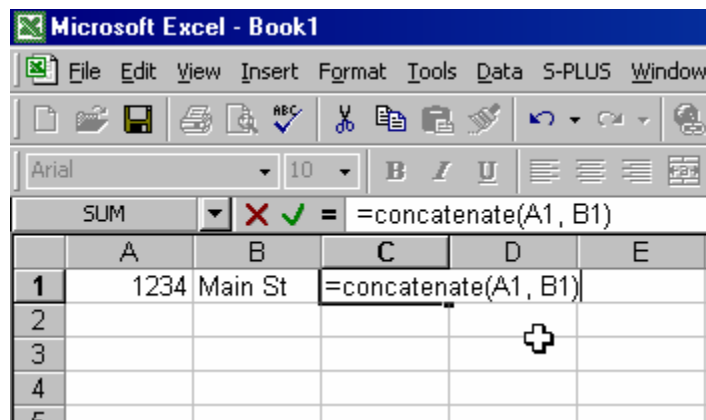
with the following result:
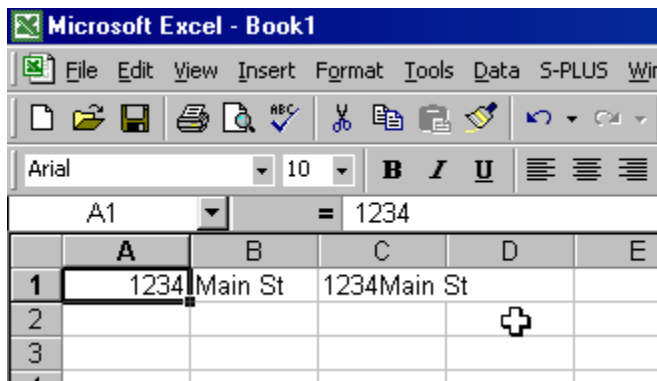
This result shows an obvious shortcoming. There is no space between the street number and the street name. This problem can be solved by inserting a new column in between the street number and name and manually inserting a space.


This problem is now resolved. As with all functions in Excel, concatenate can be "dragged" down so you do not have to type in the function for every observation. Similarly, the space in the middle column can be dragged as well.

Street intersections can also be geocoded using this procedure. As with the above, the streets must both be within the same field. The street names must be separated with a single ampersand "&" with a space on either side. If both streets are in separate columns, the concatenate function can be used again.
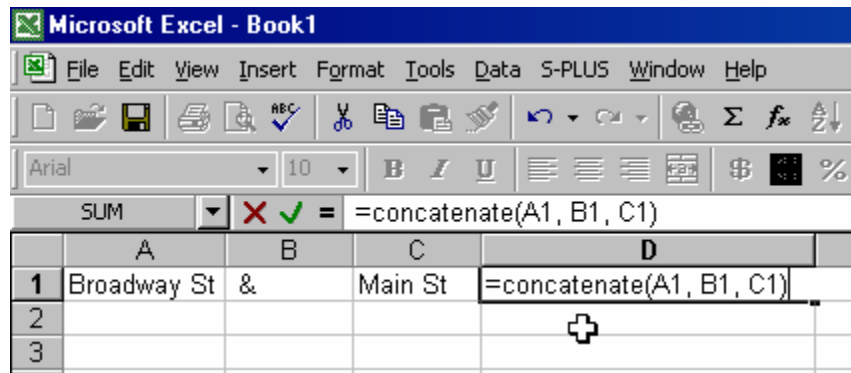

Giving the desired end result:
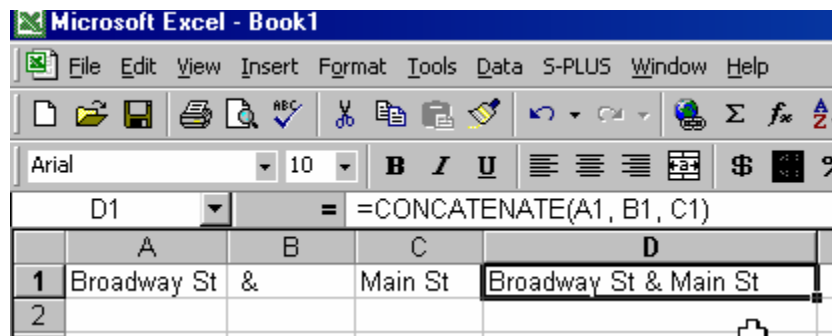
As in the previous example, to ensure there is a space on either side of the ampersand, &, manually insert the space on either side of the ampersand.

The final preparation step is to save the address file in a format readable by ArcView GIS 3.2. The file type needed is Dbase IV, which has the file extension: .dbf
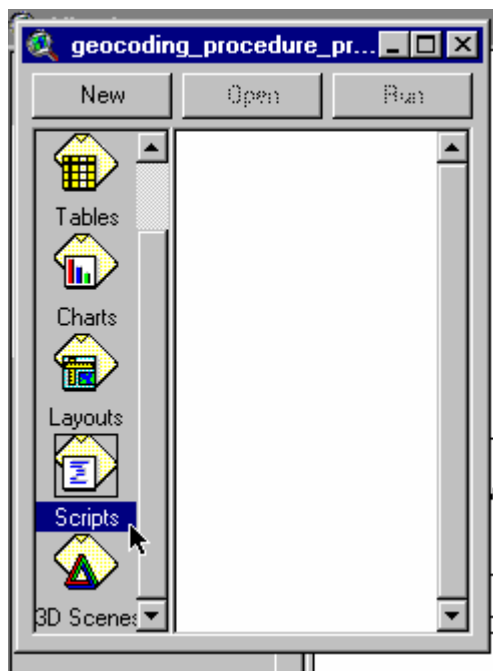
This is done by using the "Save As…" function in the "File" drop down menu. In the steps below, I called the field with my addresses "Address".

STANDARDISING ADDRESSES FOR GEOCODING

Open ArcView project window.
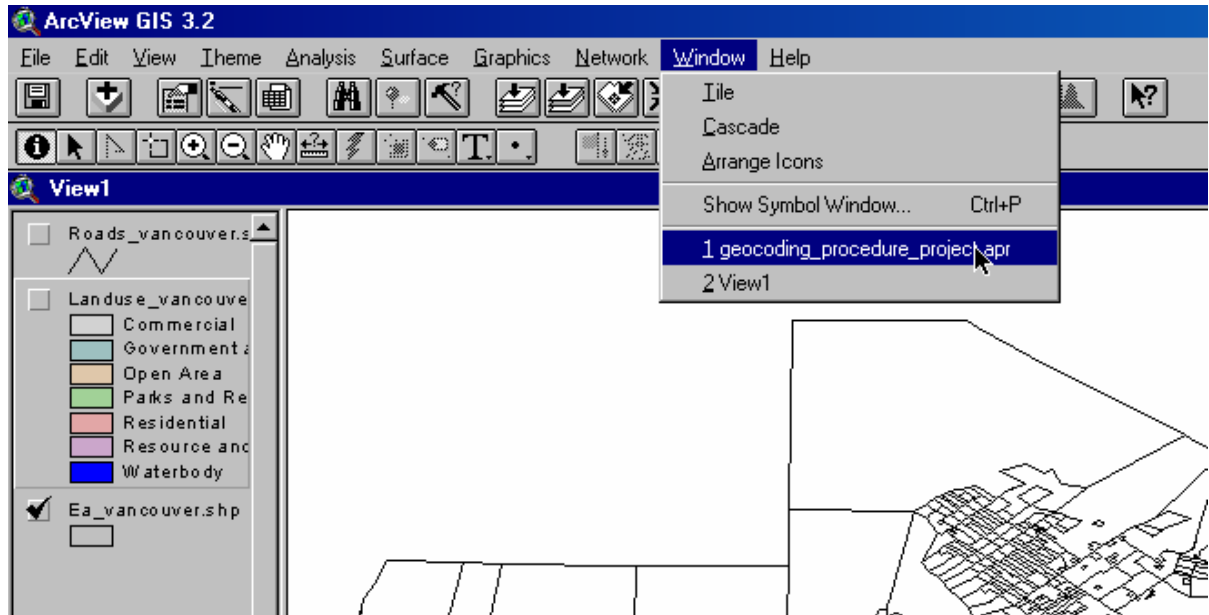
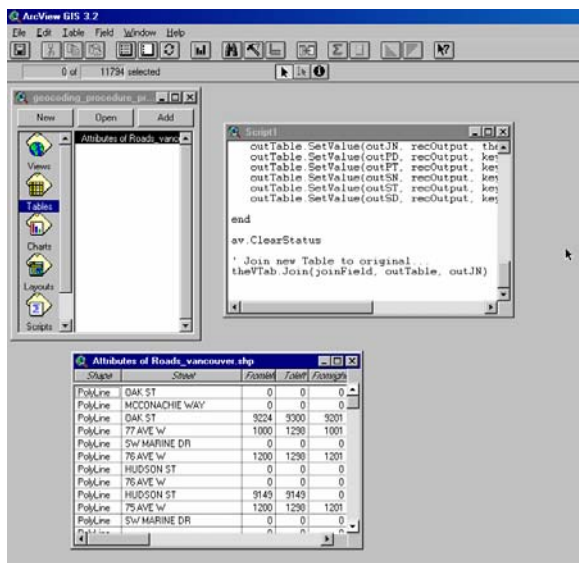

Click on scripts.

Click ``New''.

A window titled ``Script1'' will open.

In the toolbar, click on the ``load text file'' icon (looks like a folder).

Select stanshpe.ave, most likely in its default folder:
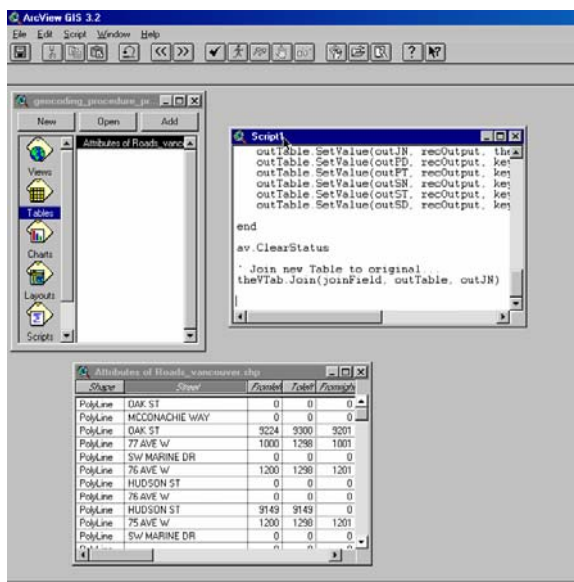C:\ESRI\AV_GIS30\ARCVIEW\Samples\scripts

Click on ``compile'' icon (check mark).

In the project window, click on ``Tables'' and open your ``roads'' file. Be sure to
make the windows small enough so you can see them both simultaneously—this
makes the following steps easier.

In your "roads" table, click on the tab for the field called ``street''.

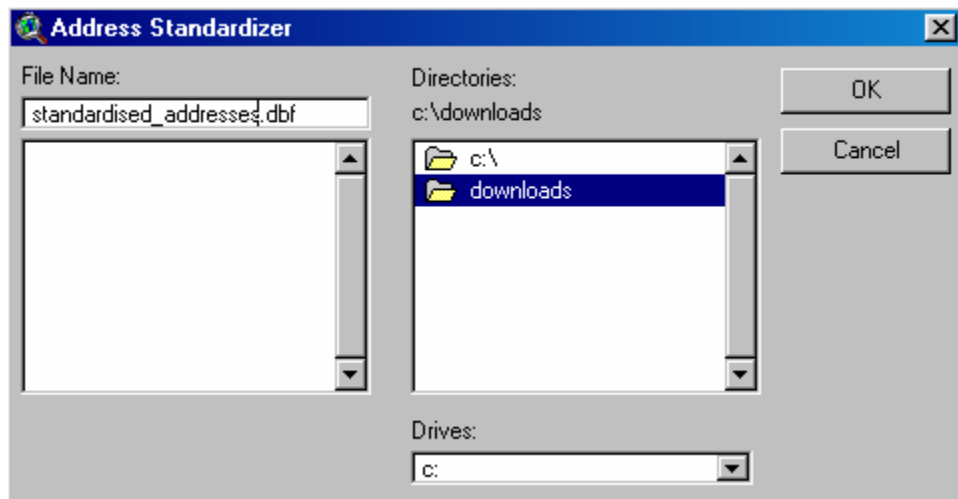Make the ``script1'' file active by clicking on it.

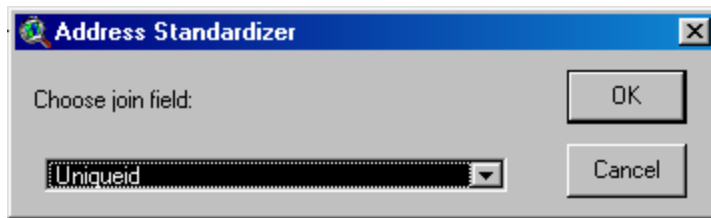Click on the ``run'' icon in toolbar (running man).

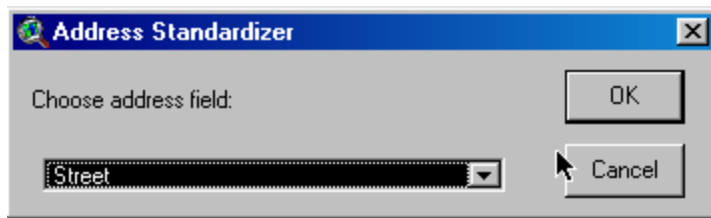An address standardiser window will open.

Save the file so you know it's your standardised addresses.

Another address standardiser window opens to choose join field: select ``uniqueid''.

Another address standardiser window opens to choose address field: select ``street''.

Now you need to edit this ``roads'' table.

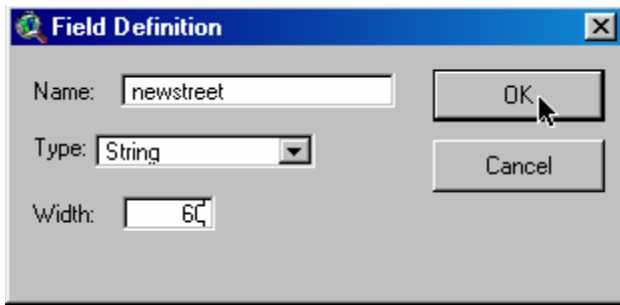In the ``Table'' drop down menu, select ``Start Editing''.

In the ``Edit'' drop down menu, select ``Add Field'':
    name: newstreet
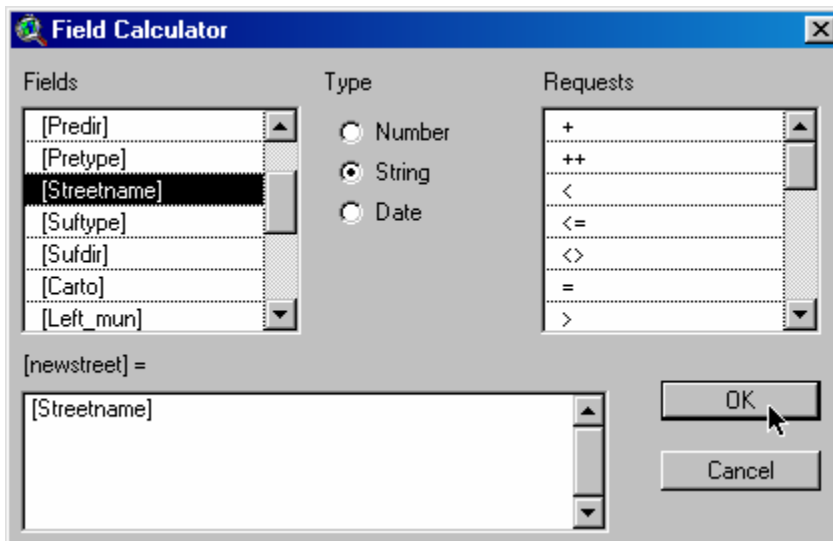   type: string
    width: 60
Click ``OK''

Return to the ``roads'' table.  The new field "newstreet" should be highlighted, but if not select the field ``newstreet'' by clicking on its tab.

In the toolbar, click on the ``Calculator'' icon.

Set the ``Calculator'' to read:
     [newstreet]=
     [streetname]
Click ``OK''.

Select the ``streetname'' tab and click on the ``Sort Ascending'' icon in the toolbar again so the numbered streets are at the top of the table.
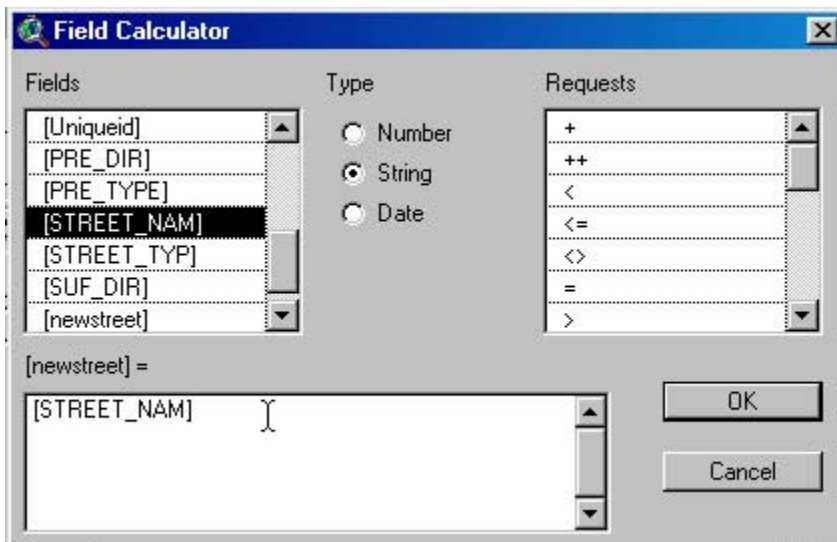
Manually select all the numbered streets by simultaneously holding down the ``shift'' key and selecting/dragging the mouse.

Select the ``newstreet'' field.

Click on the ``Calculator'' icon in the toolbar set the calculator to read:
    [newstreet]=
    [STREET_NAM]

Click ``OK'' now the numbered street should have ST, AVE, etc associated with them again.

Remove the selection and save.

Now you are ready to geocode.

GEOCODING

In project window, click on ``View'' and open your view file that contains the ``roads'' layer.
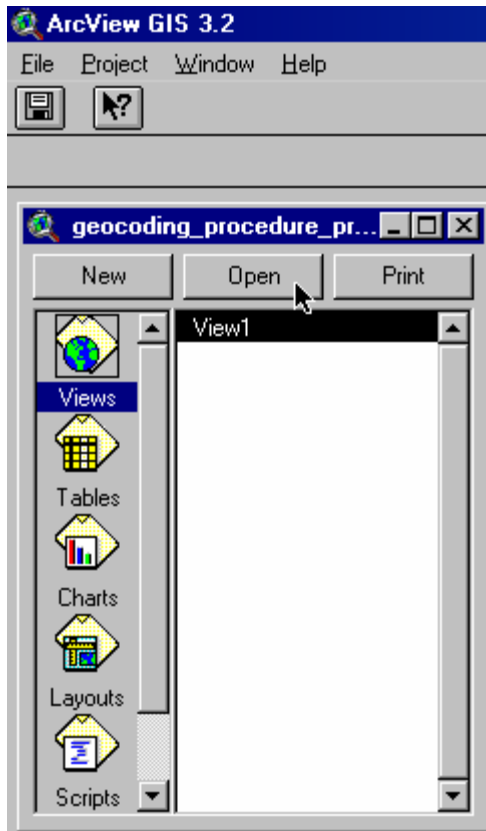
Select ``roads'' by clicking on it in the left hand side of the screen. This makes roads the active layer and will look slightly higher than the other themes.
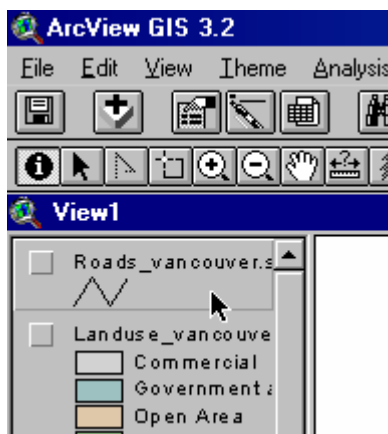
In the ``Theme'' drop down menu, select ``Properties''.
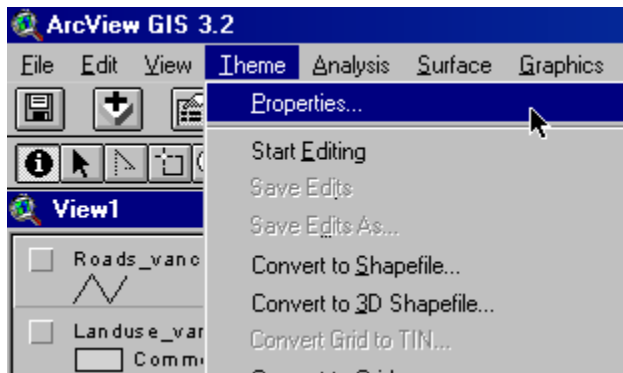
In the left side of this window, select ``Geocoding''
   Theme Name: the name of your roads file
   Address Style: US Streets
   Change: PreDir to SufDir; Sufdir to <None>; Street Type to
   STREET_TYP; and Streetname to newstreet.
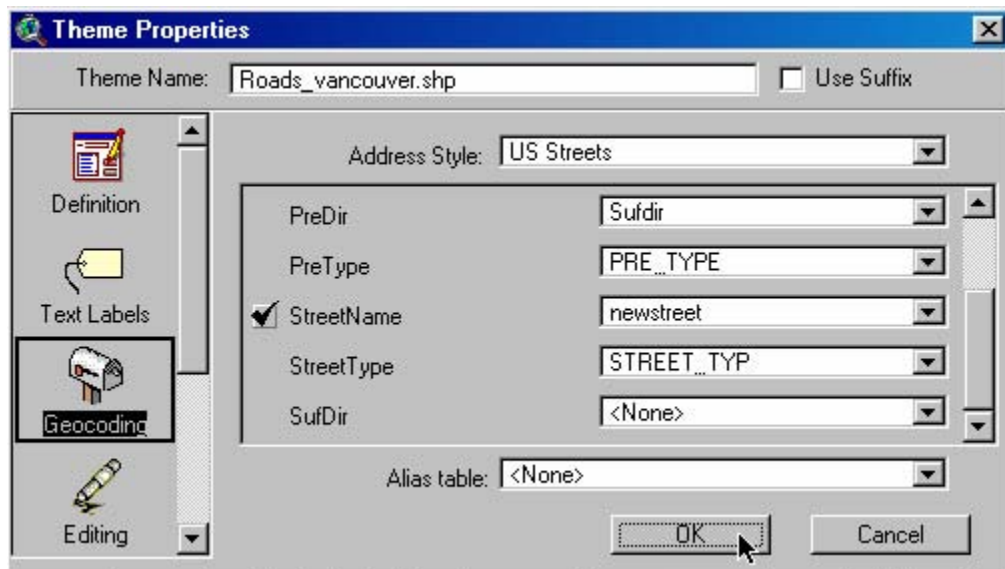
When prompted to build geocoding indexes, click yes.

Save.

From the ``View'' drop down menu, select ``Geocode Addresses...''
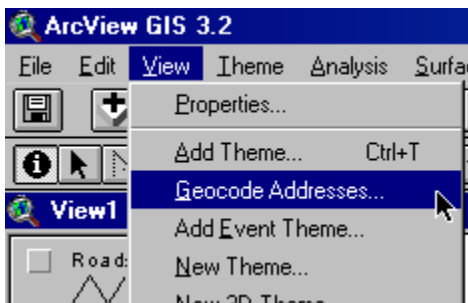
Reference Theme: Roads file
Join Field: <None>
Using Address : US Streets
Address Table: file containing addresses to geocode
Address Field: address variable name in above file
Display Field: <None>
Offset Distance: 0
Alias Table: <None>
Geocoded Theme: Name your output file to what you want
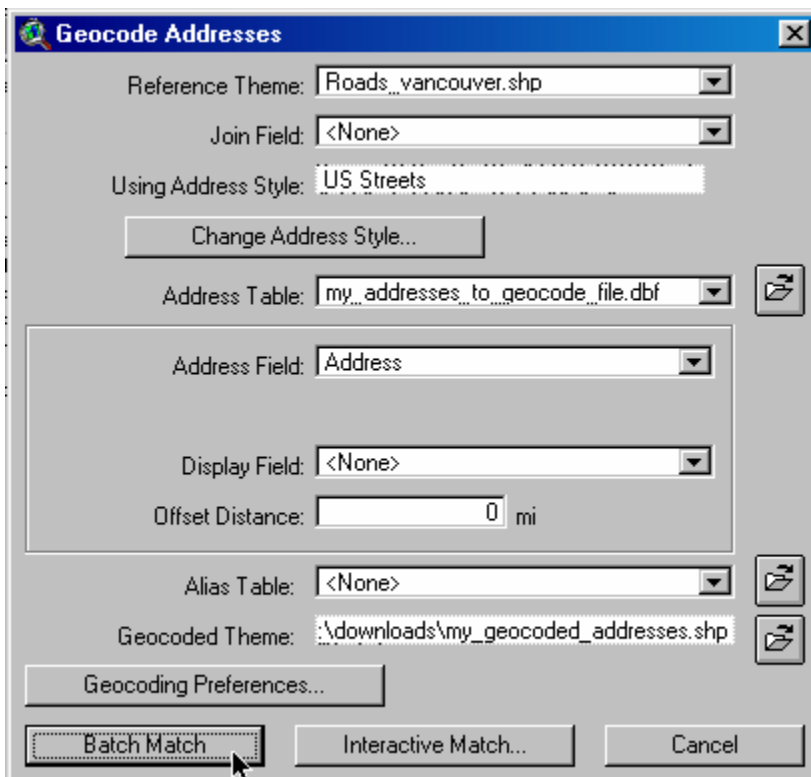
Click on ``Batch Match'', wait and cross your fingers.

SPATIAL JOINING OF POINTS WITH POLYGONS

If all you desired from geocoding was to make maps, you do not need to read on. You now have a new layer of data to display the geocoded points. However, if any analysis is to be performed, these points will most likely need to be associated with polygons that provide socio-demographic and socio-economic information. This procedure is performed using a function in ArcView GIS 3.2.

From the "View" drop down menu, select "GeoProcesing Wizard…"
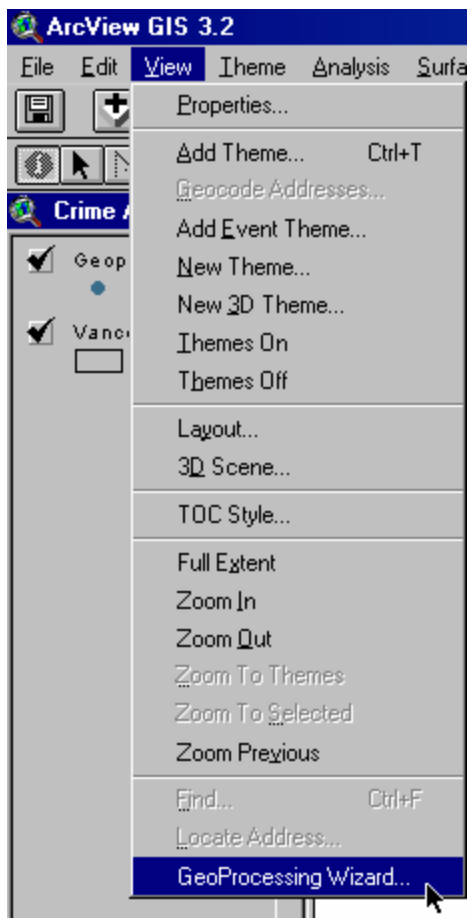
The following window will appear:
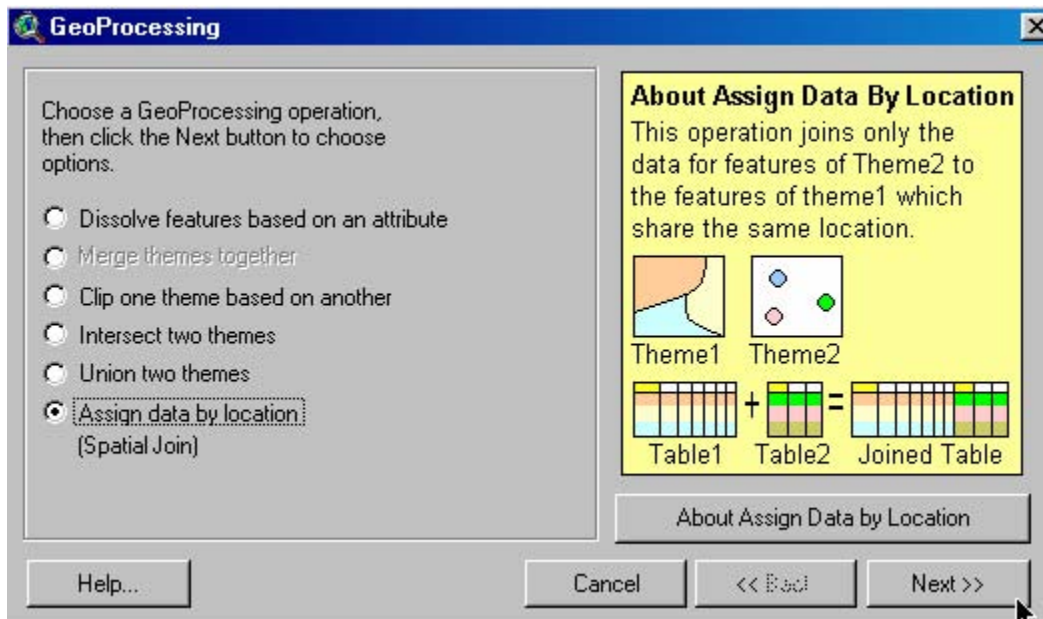
Select the option of "Assign data by location (Spatial Join)" and click Next. Read the "About…" section if you wish to learn more about this procedure.

In this procedure, you want to assign each point to a polygon using the following options:
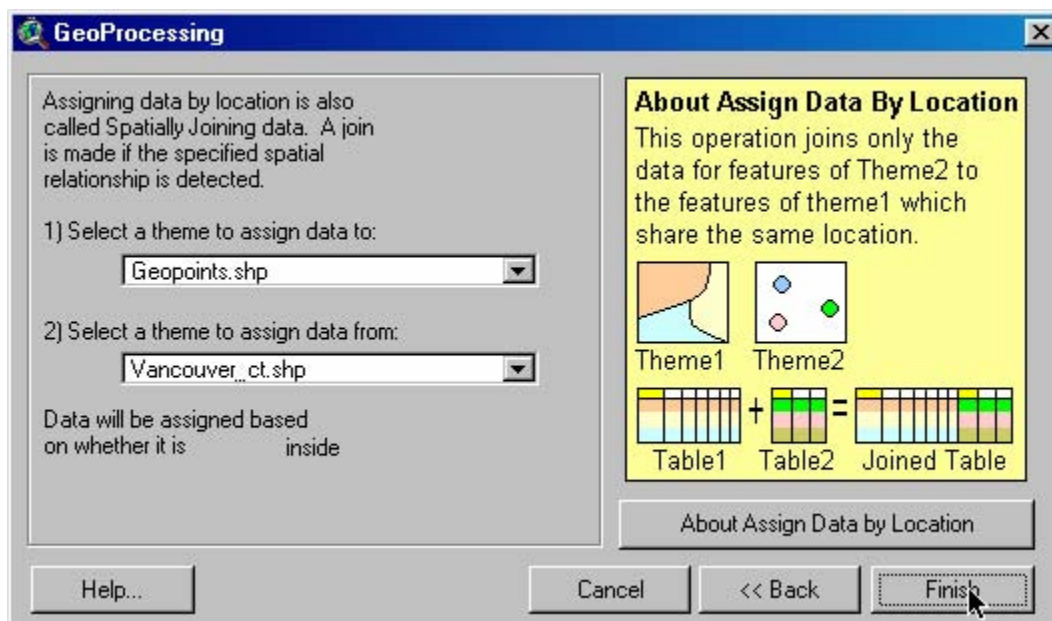
This will assign a polygon number to a point if it falls inside that polygon. Click Finish.

If you check the table in you Geopoints shape file, you will now see that the spatial and attribute information from your polygon file. In this example, I have joined some points to Vancouver Census Tracts from Statistics Canada. Now any information you have associated with your polygons is now associated with each point.

However, if you wish to do a count of all the points in each polygon to calculate rates or other measures, the dissolve feature must be used:
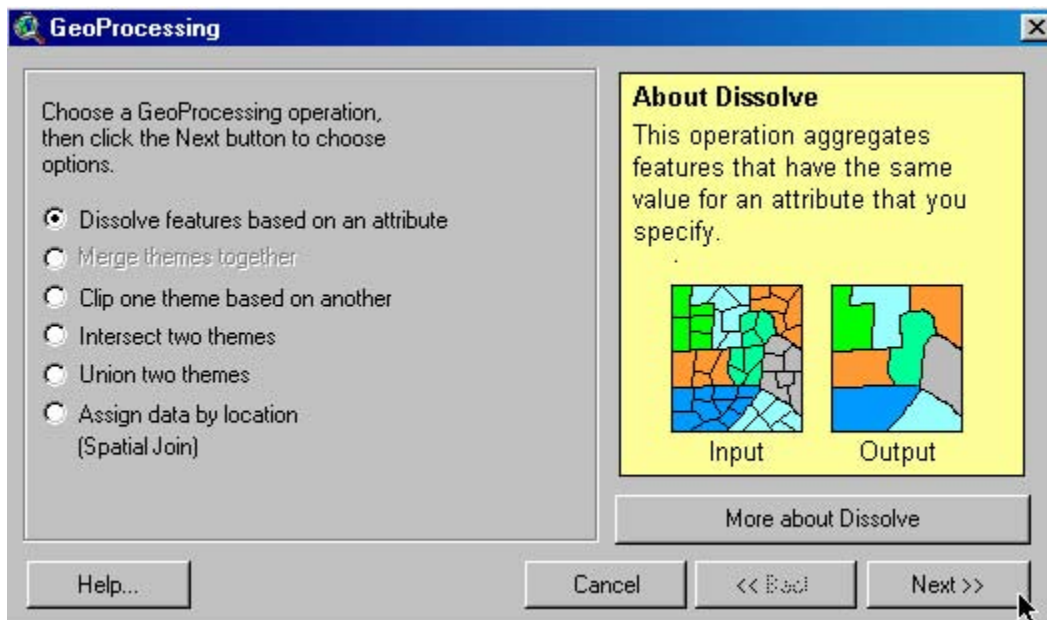
This feature can do a number of functions, but we will focus on counting the number of points in each polygon. Basically, the dissolve feature, selected in the above figure, joins together features based on parameters you choose. Click on "More about Dissolve" to get more information. Then click on Next.

Since each point is associated with a polygon (Census Tract in this case) we will want to dissolve based on that attribute.
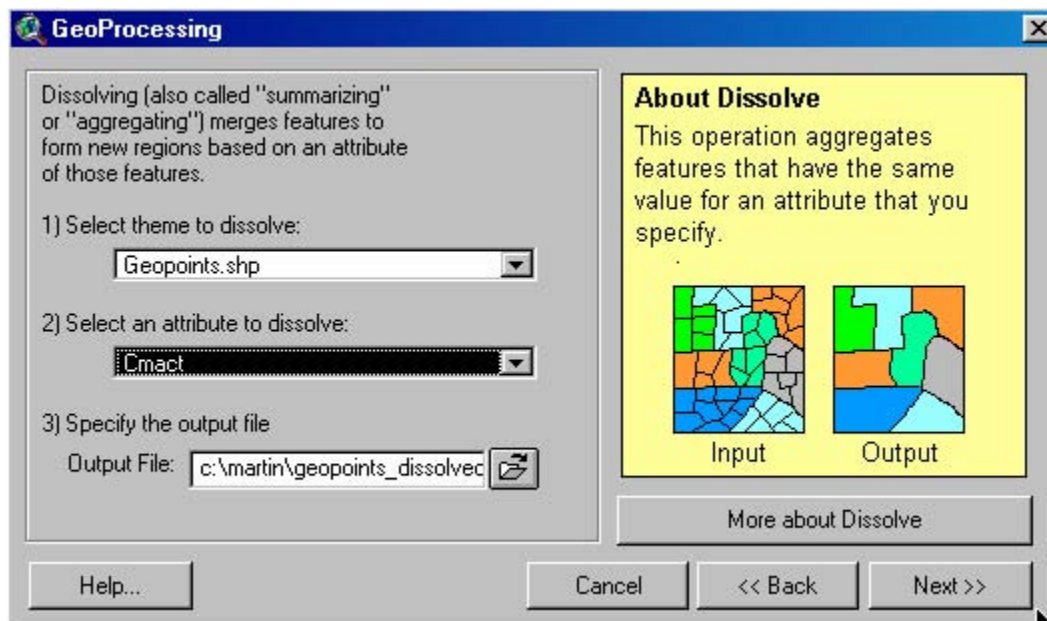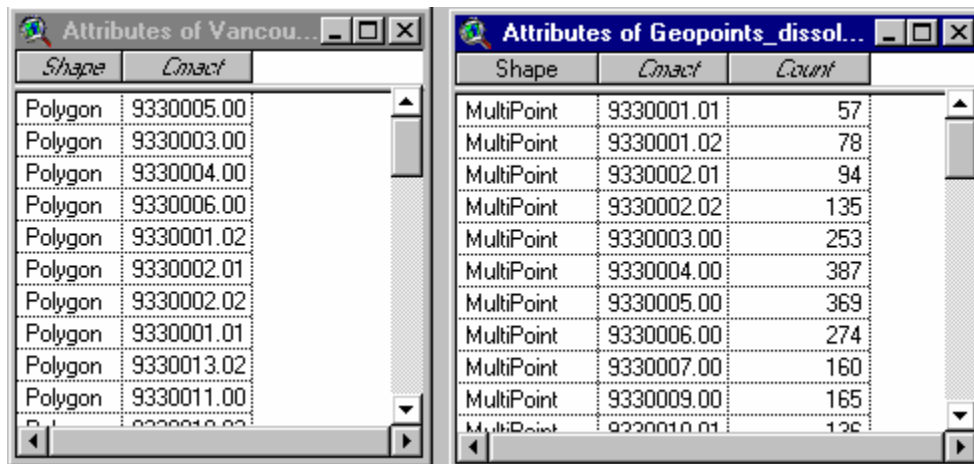
The above image shows that we want to dissolve the Geopoints shape file base on Cmact (the name of the Census Tract field). The resulting output file must also be given a name. Click Next.

The dissolve feature allows you to choose how to dissolve the field variables. If none of the options are selected a variable called "Count" will be created. Do not select any of the options displayed—your options will include different variable names than shown. Click Finish.
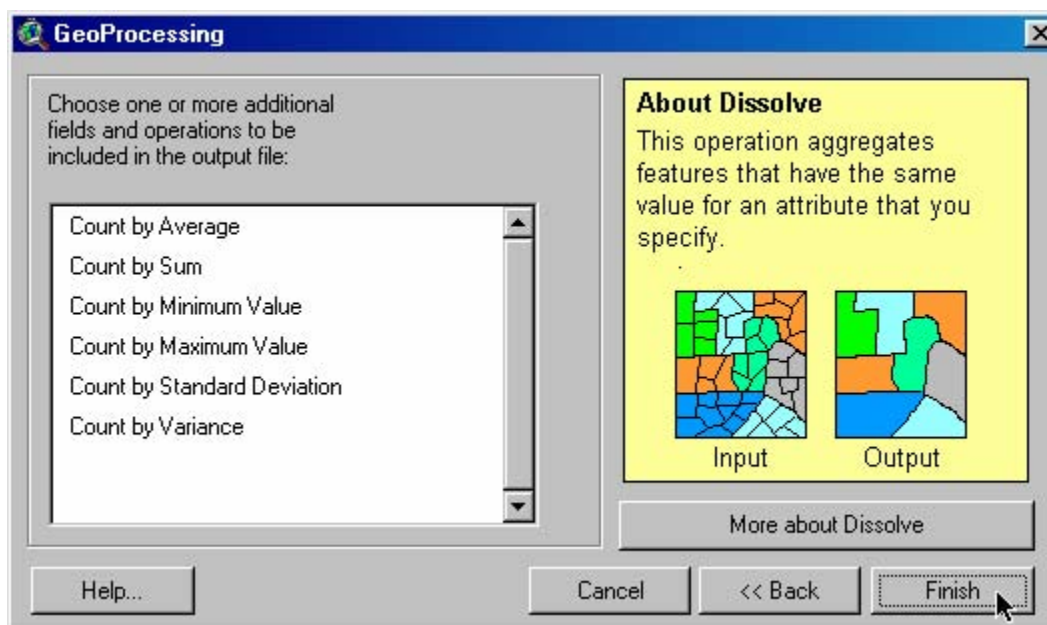
Once the procedure has finished there will be a new layer added to your View. This new layer will contain an attribute table that has a field identifying the polygon (Census Tract in this case) and another field that has the number of points within that polygon.   Of course, the tables shown would normally have more variables:

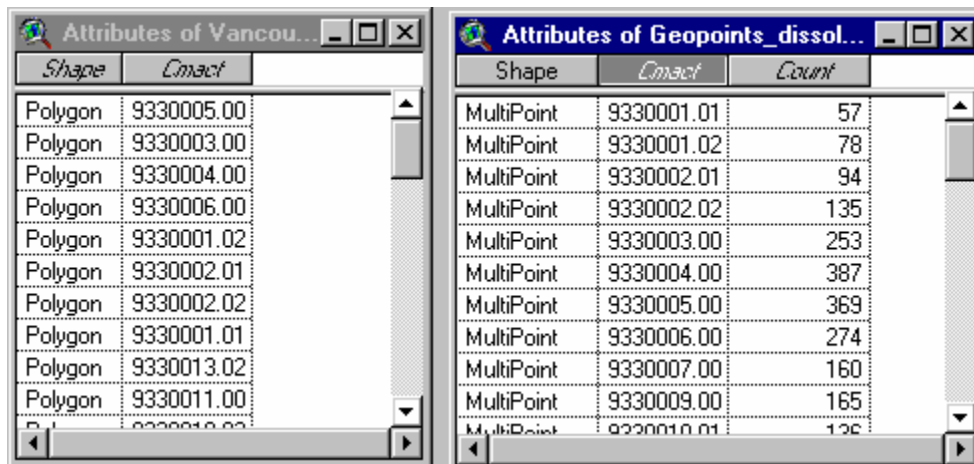| Attributes of Vancou...			Attributes of Geopoints_dissol...		
Shape	Cract		Shape	Cract	Count
Polygon	9330005.00		MultiPoint	9330001.01	57
Polygon	9330003.00		MultiPoint	9330001.02	78
Polygon	9330004.00		MultiPoint	9330002.01	94
Polygon	9330006.00		MultiPoint	9330002.02	135
Polygon	9330001.02		MultiPoint	9330003.00	253
Polygon	9330002.01		MultiPoint	9330004.00	387
Polygon	9330002.02		MultiPoint	9330005.00	369
Polygon	9330001.01		MultiPoint	9330006.00	274
Polygon	9330013.02		MultiPoint	9330007.00	160
Polygon	9330011.00		MultiPoint	9330009.00	165
			MultiPoint	9330010.01	136

Now this information needs to be joined.

Open the tables you wish to join. Make the source (Geopoints) table active and click the name of the common field you wish to use to join the table—this will be the polygon identification (Census Tract) number.

Attributes of Vancou...			Attributes of Geopoints_dissol...		
Shape	Cract		Shape	Cract	Count
Polygon	9330005.00		MultiPoint	9330001.01	57
Polygon	9330003.00		MultiPoint	9330001.02	78
Polygon	9330004.00		MultiPoint	9330002.01	94
Polygon	9330006.00		MultiPoint	9330002.02	135
Polygon	9330001.02		MultiPoint	9330003.00	253
Polygon	9330002.01		MultiPoint	9330004.00	387
Polygon	9330002.02		MultiPoint	9330005.00	369
Polygon	9330001.01		MultiPoint	9330006.00	274
Polygon	9330013.02		MultiPoint	9330007.00	160
Polygon	9330011.00		MultiPoint	9330009.00	165
			MultiPoint	9330010.01	136

Make the destination (Vancouver Census Tract) table active and click the same name of the common field you wish to use to join the table.

From the Table menu choose Join, or click the Join button .

If you have done the steps correctly, the source table will disappear and the destination table will remain with the new variable.

The count of points in each polygon is now associated with the polygon file. These counts can now be displayed for each polygon, rather than points, and can be used to calculate rates or other numbers by adding new fields and using the calculating function, described above.

Abbey = ABBEY	Crossing = CROSS	Knoll = KNOLL	Rang = RANG
Acres = ACRES	Cul-de-sac = CDS	Landing = LANDNG	Range = RG
Allee = ALLEE	Dale = DALE	Lane = LANE	Ridge = RIDGE
Alley = ALLEY	Dell = DELL	Limits = LMTS	Rise = RISE
Autoroute = AUT	Diversion = DIVERS	Line = LINE	Road = RD
Avenue = AVE (E)	Downs = DOWNS	Link = LINK	Rond-point = PDPT
Avenue = AV (F)	Drive = DR	Lookout = LKOUT	Route = RTE
Bay = BAY	Exchangeur = ECH	Loop = LOOP	Row = ROW
Beach = BEACH	End = END	Mall = MALL	Rue = RUE
Bend = BEND	Esplanade = ESPL	Manor = MANOR	Ruelle = RLE
Boulevard = BLVD (E)	Estates = ESTATE	Maze = MAZE	Run = RUN
Boulevard = BOUL (F)	Expressway = EXPY	Meadow = MEADOW	Sentier = SENT
By-Pass = BYPASS	Extension = EXTEN	Mews = MEWS	Square = SQ
Byway = BYWAY	Farm = FARM	Montee = MONTEE	Sideroad = SR
Campus = CAMPUS	Field = FIELD	Moor = MOOR	Street = ST
Cape = CAPE	Forest = FOREST	Mount = MOUNT	Subdivision = SUBDIV
Carre = CAR	Freeway = FWY	Mountain = MTN	Terrace = TERR
Carrefour = CARREF	Front = FRONT	Orchard = ORCH	Terrasse = TSSE
Centre = CTR (E)	Gardens = GDNS	Parade = PARADE	Thicket = THICK
Centre = C (F)	Gate = GATE	Parc = PARC	Towers = TOWERS
Cercle = CERCLE	Glade = GLADE	Park = PK	Townline = TLINE
Chase = CHASE	Glen = GLEN	Parkway = PKY	Trail = TRAIL
Chemin = CH	Green = GREEN	Passage = PASS	Turnabout = TRNABT
Circle = CIRCLE	Grounds = GRNDS	Path = PATH	Vale = VALE
Circuit = CIRCT	Grove = GROVE	Pathway = PTWAY	Via = VIA
Close = CLOSE	Harbour = HARBR	Pines = PINES	View = VIEW
Common = COMMON	Heights = HTS	Place = PL (E)	Village = VILLGE
Concession = CONC	Highlands = HGHLDS	Place = PLACE (F)	Vista = VISTA
Corners = CRNRS	Highway = HWY	Plateau = PLAT	Voie = VOIE
Cote = COTE	Hill = HILL	Plaza = PLAZA	Walk = WALK
Cour = COUR	Hollow = HOLLOW	Point = PT	Way = WAY
Court = CRT	Ile = ILE	Port = PORT	Wharf = WHARF
Cove = COVE	Impasse = IMP	Private = PVT	Wood = WOOD
Crescent = CRES	Island = ISLAND	Promenade = PROM	Wynd = WYND
Croissant = CROIS	Key = KEY	Quay = QUAY	