GIS II: Data Management
Leveraging the *Where* of Geographic Data Exercise

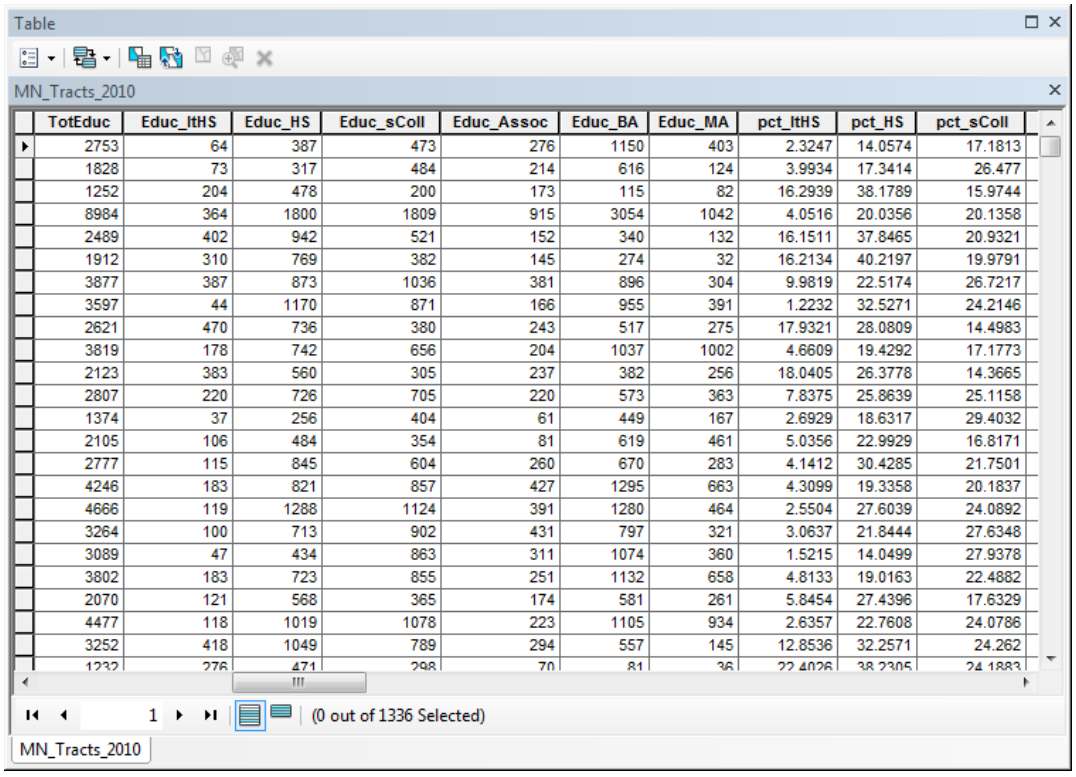
*** Files needed for exercise: MN_*Tracts_2010.shp, and MD_USA_CANADA.shp*

Goals: The goals for this exercise are to gain experience performing spatial joins and selections and to review your understanding of spatial data projections.

Skills: After completing this exercise, you will be able to spatially join a point dataset (businesses of interest) to a polygon dataset (census tracts). This is a useful method for the enumeration of points within polygons: i.e. what is the count for points of interest within geography of interest?

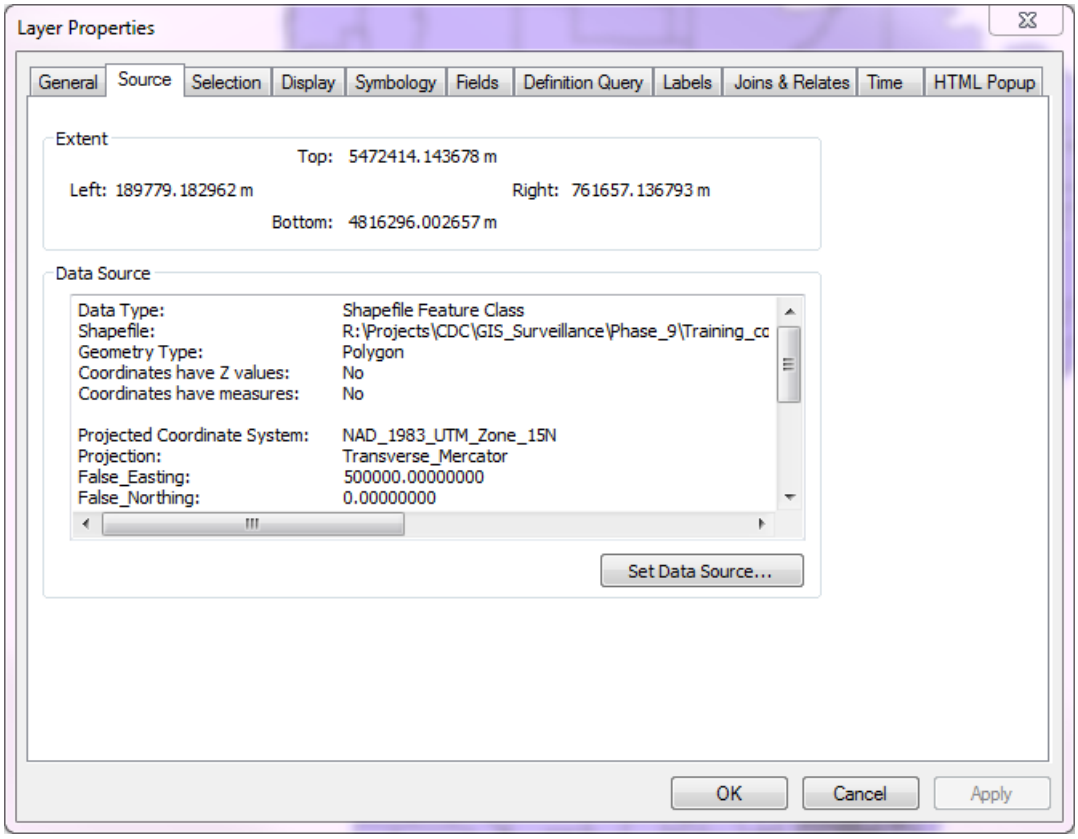
Adding Data and Checking Coordinate Systems

1. Open ArcMap. Choose to start a new blank map.
2. Click the **Add Data** button .
3. Click on the **Connect to Folder** button . Browse to the Leveraging the Where Exercise_Data folder and connect to it. You now have a permanent connection to that folder.
4. Double click on *MN_Tracts_2010.shp* to add it to your project. This shapefile has American Community Survey 2008-2012 socio-economic and demographic data appended to it.
5. Open the table for the shapefile and take a look at the attribute fields.



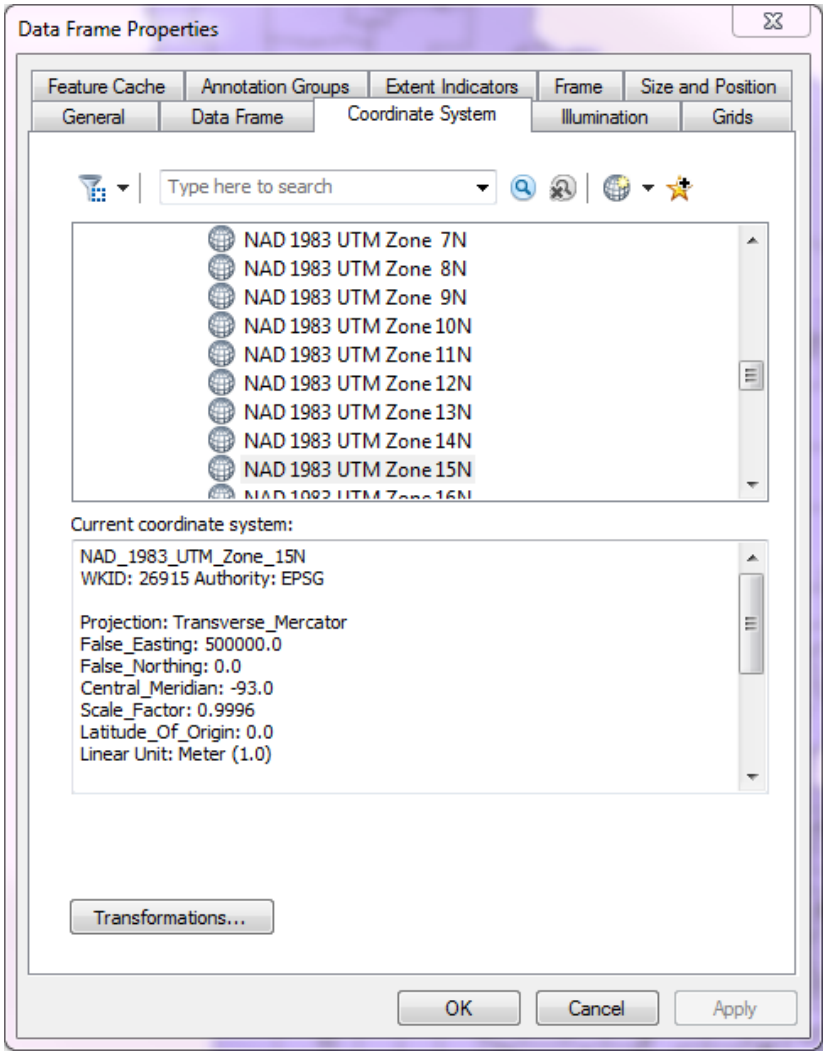
TotEduc	Educ_ltHS	Educ_HS	Educ_sColl	Educ_Assoc	Educ_BA	Educ_MA	pct_ltHS	pct_HS	pct_sColl
2753	64	387	473	276	1150	403	2.3247	14.0574	17.1813
1828	73	317	484	214	616	124	3.9934	17.3414	26.477
1252	204	478	200	173	115	82	16.2939	38.1789	15.9744
8984	364	1800	1809	915	3054	1042	4.0516	20.0356	20.1358
2489	402	942	521	152	340	132	16.1511	37.8465	20.9321
1912	310	769	382	145	274	32	16.2134	40.2197	19.9791
3877	387	873	1036	381	896	304	9.9819	22.5174	26.7217
3597	44	1170	871	166	955	391	1.2232	32.5271	24.2146
2621	470	736	380	243	517	275	17.9321	28.0809	14.4983
3819	178	742	656	204	1037	1002	4.6609	19.4292	17.1773
2123	383	560	305	237	382	256	18.0405	26.3778	14.3665
2807	220	726	705	220	573	363	7.8375	25.8639	25.1158
1374	37	256	404	61	449	167	2.6929	18.6317	29.4032
2105	106	484	354	81	619	461	5.0356	22.9929	16.8171
2777	115	845	604	260	670	283	4.1412	30.4285	21.7501
4246	183	821	857	427	1295	663	4.3099	19.3358	20.1837
4666	119	1288	1124	391	1280	464	2.5504	27.6039	24.0892
3264	100	713	902	431	797	321	3.0637	21.8444	27.6348
3089	47	434	863	311	1074	360	1.5215	14.0499	27.9378
3802	183	723	855	251	1132	658	4.8133	19.0163	22.4882
2070	121	568	365	174	581	261	5.8454	27.4396	17.6329
4477	118	1019	1078	223	1105	934	2.6357	22.7608	24.0786
3252	418	1049	789	294	557	145	12.8536	32.2571	24.262
1232	276	471	298	70	81	36	22.4026	38.2305	24.1883

6. This shapefile is projected properly to conform to the standard projected coordinate system for Minnesota. Confirm this by right clicking on the shapefile in your table of contents (TOC), clicking **Properties**, and then selecting the **Source** tab. Once you've verified the projected coordinate system, close the Layer Properties.



7. Since this is the first projected layer you have added to your map data frame, the data frame projection now matches *MN_Tracts_2010.shp.* Check this by right clicking **Layers** and going to **Properties**. Take a look at the **Coordinate System tab**. Once you've verified that it is the same, close the Data Frame Properties.

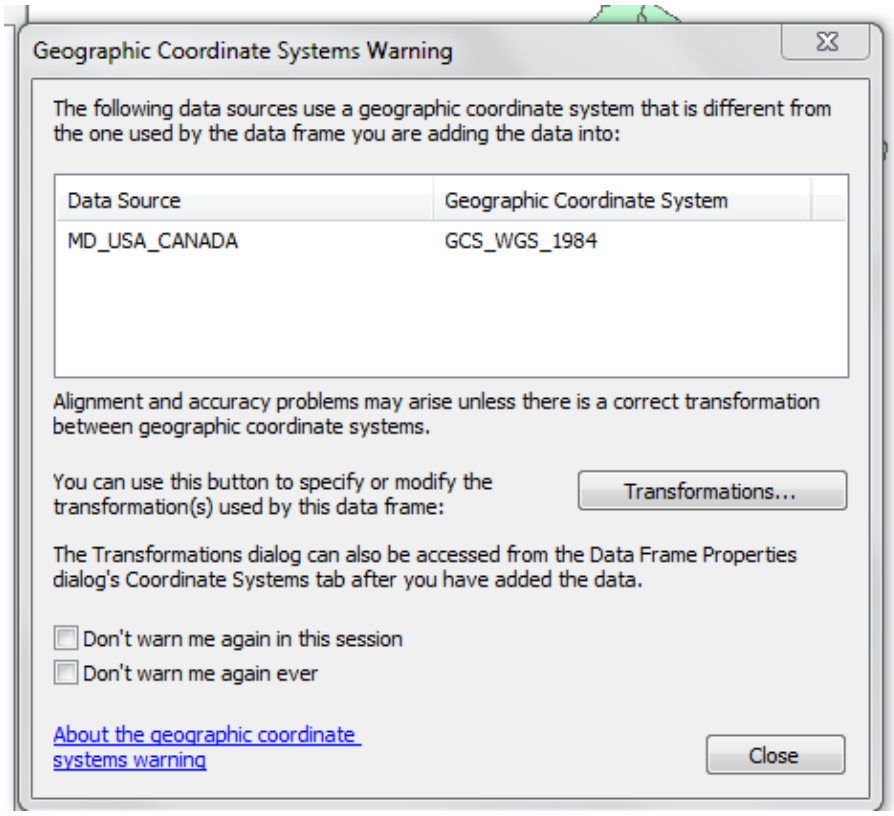
GIS II: Data Management
Leveraging the *Where* of Geographic Data Exercise



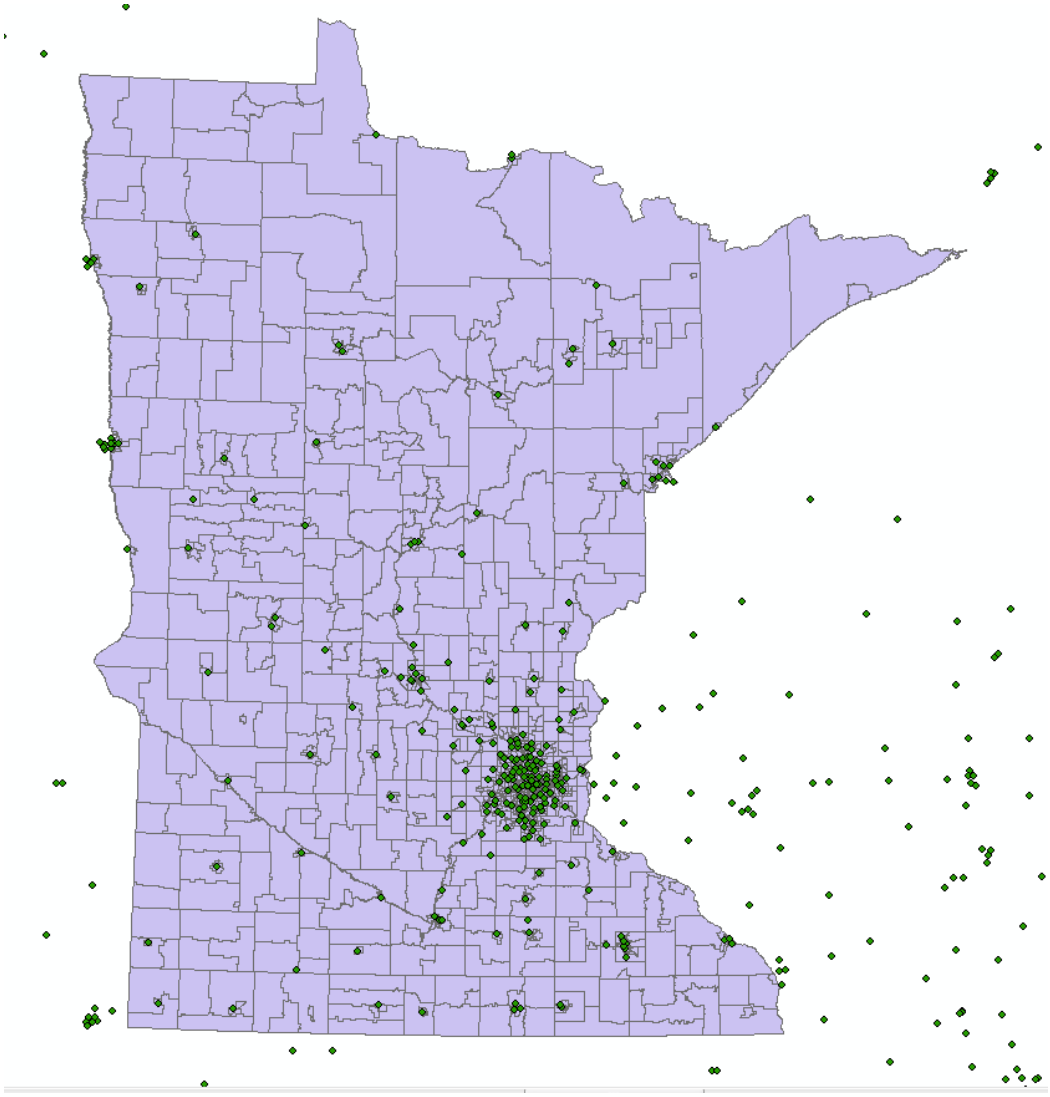
GIS II: Data Management
Leveraging the *Where* of Geographic Data Exercise

Projecting On the Fly

1. Add *MD_USA_CANADA.shp* to your project from the Exercise_Data folder. A Geographic Coordinate Systems Warning dialogue box will appear.

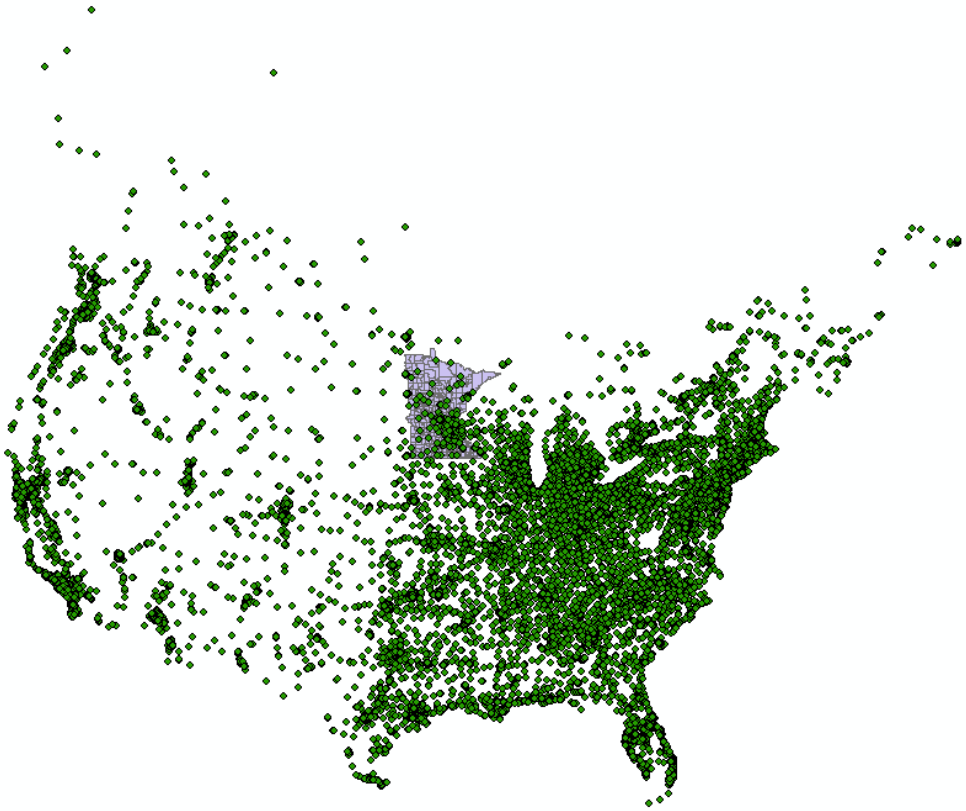


2. Click **Close**. You should now see *MD_USA_CANADA.shp* in the TOC on the left hand side.
3. *MD_USA_CANADA.shp* is a continental (and also Hawaii!) dataset of McDonald's restaurants, so it probably uses a coordinate system that differs from *MN_Tracts_2010.shp;* confirm this. What coordinate system is defined? **GCS_WGS_1984** is the de facto projected coordinate system for Global Position System (GPS) gathered data (these came from ***Points of Interest Factory*** – check it out: http://www.poi-factory.com/node/11154). Note these data are current as of 12/11/2016
4. Since the projected and geographic coordinate systems are both defined, ArcMap can project on the fly and make the McDonald's data work in the same data frame as the Minnesota tract data.
5. The McDonald's eating establishments should be displayed above your tracts.



6. For kicks, right click on *MD_USA_CANADA.shp* in your TOC and select the **Zoom To Layer** option. You should see the many McDonald's restaurants across the USA & Canada that have been projected on the fly to NAD_1983_UTM_Zone_15N.

GIS II: Data Management
Leveraging the *Where* of Geographic Data Exercise

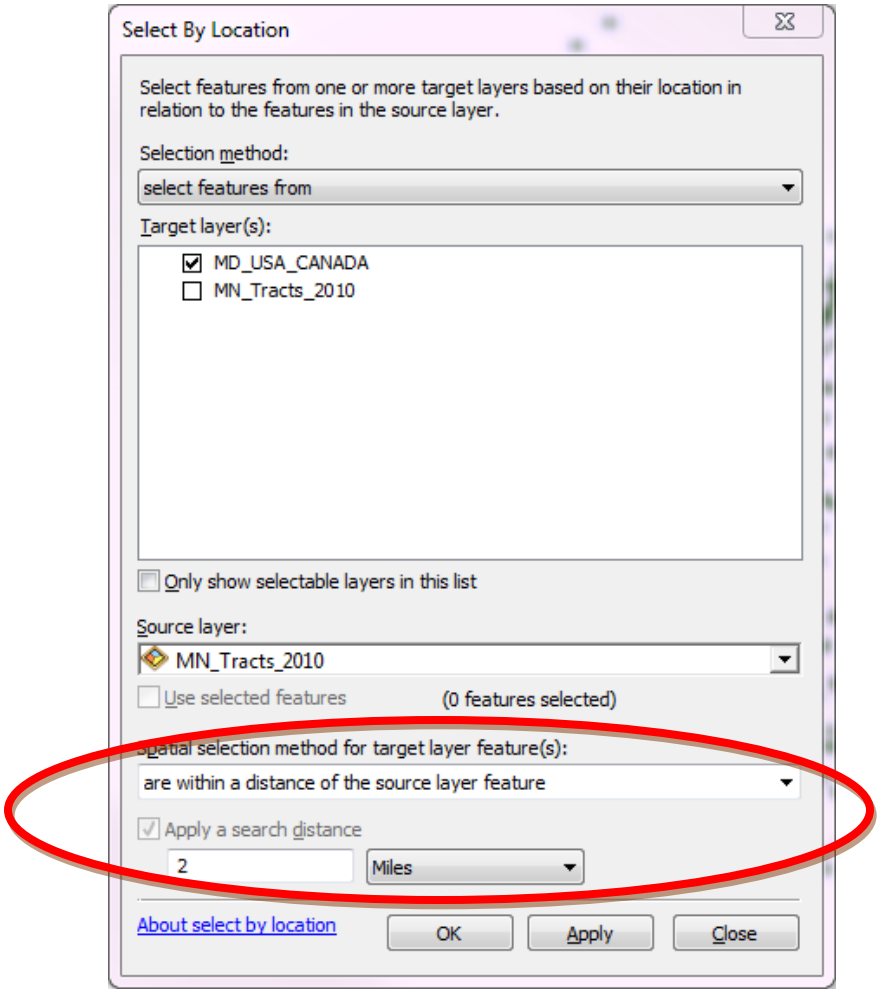


7. Open the attribute table to find out how many McDonald's are in the table; you should see 15,669 records - each represents one McDonald's restaurant.
8. Once you have done this zoom back to the Minnesota tracts by right clicking on the *MN_Tracts_2010.shp* and selecting **Zoom to Layer**.

Performing a Spatial Selection of Points

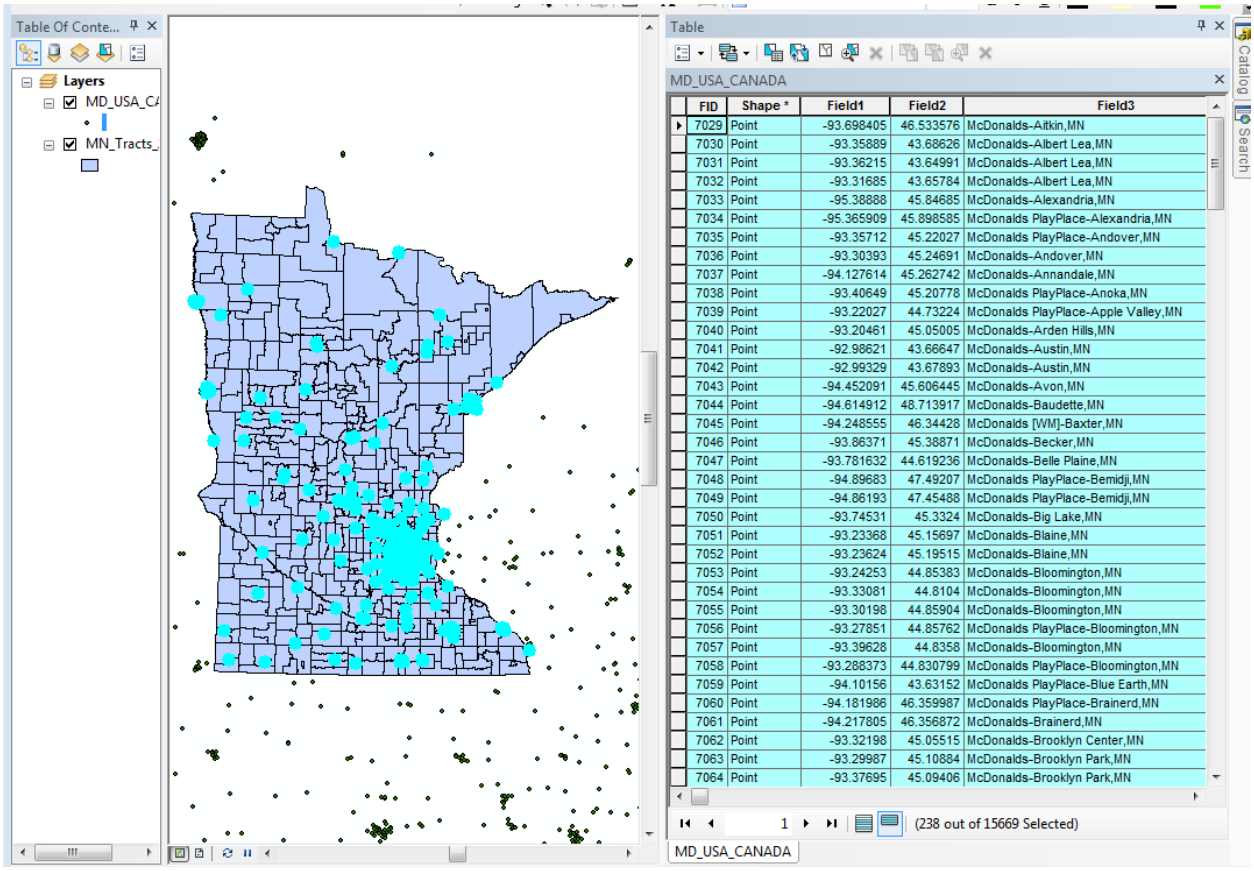
1. Under the selection menu select choose **Select by Location** to bring up the Select by Location dialogue box. You want to use the location information of both the McDonald's points and tracts (Minnesota) to **select features from the target layer (McDonald's points)** that**: are within a distance of… 2 miles from the source layer (Minnesota tracts).** *Plainly put: which McDonald's restaurant points are in located inside of, or within a distance of 2 miles from the border of Minnesota?*

GIS II: Data Management
Leveraging the *Where* of Geographic Data Exercise

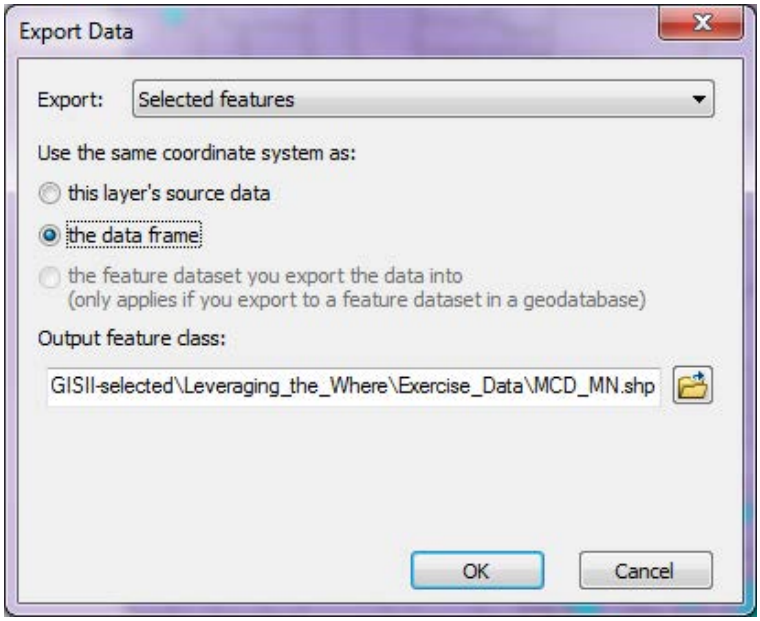


2. **Apply** your spatial selection and then click **OK**.
3. Right click on *MD_USA_CANADA.shp* in the TOC and open the table; how many records have been selected based on your spatial query?

GIS II: Data Management
Leveraging the *Where* of Geographic Data Exercise

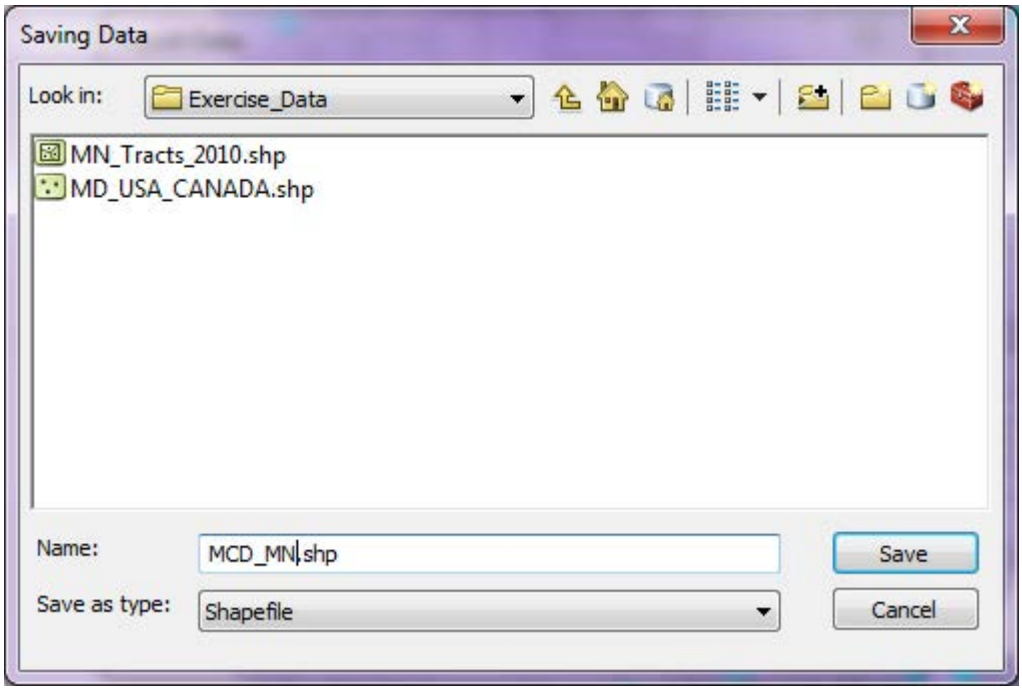


4. Once you have confirmed that **238 records** have been selected, close the table.
5. How could you have created this selection with a tabular selection?
6. Now you will export your selected McDonald's to a new shapefile. In the process you will set a new coordinate system, making sure it matches your Minnesota tracts. Right click on *MD_USA_CANADA.shp* in the TOC. Select **Data** > **Export Data**.
7. Choose the radio button to indicate that you want to use the coordinate system from "the data frame". Recall that you have defined the projection for the data frame by adding *MN_Tracts_2010.shp* as the first layer. With this information ArcMap can project the data to Minnesota's projected coordinate system.



8. Name your file *MCD_MN* and save it in your Exercise_Data folder.

   Make sure you save as type: ***Shapefile.*** Click **Save** and add your new layer to the map.

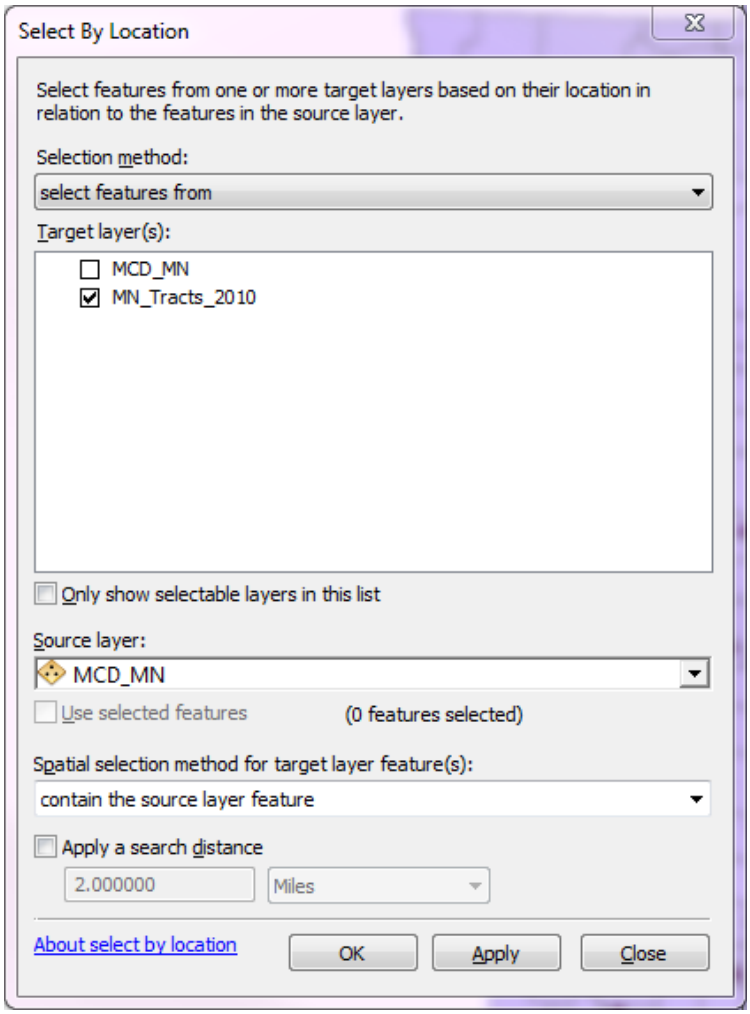


9. Add *the MCD_MN.shp* to your TOC. You can now remove *MD_USA_CANADA.shp* from your TOC.

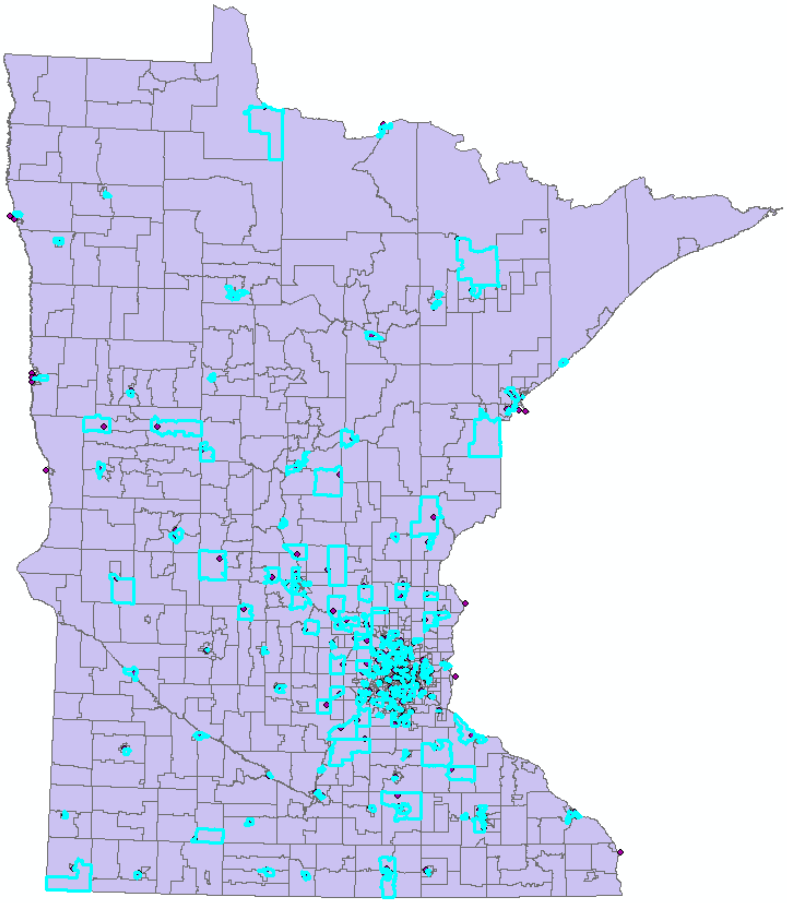
GIS II: Data Management
Leveraging the *Where* of Geographic Data Exercise

Performing a Spatial Selection of Polygons

1. You now have a shapefile that contains only McDonald's within/or within a distance of 2 miles of the border of Minnesota: *MCD_MN*. Let's say you are interested in knowing all census tracts that contain at least one McDonald's restaurant. You could not do this using a tabular selection and so you must use a spatial selection.
2. Click on the **Selection** tab and choose **Select by Location** to answer the question: which census tracts contain a McDonald's restaurant?
3. You will select features from *MN_Tracts_2010.shp* (Target layer) that contain features from the *MCD_MN* layer (Source layer). To do this your spatial selection method will be: **Target layer features <u>contain</u> the Source layer feature**. Click **Apply**, and then click **OK**.



4. Take a look at your map; these are the census tracts that contain McDonald's.



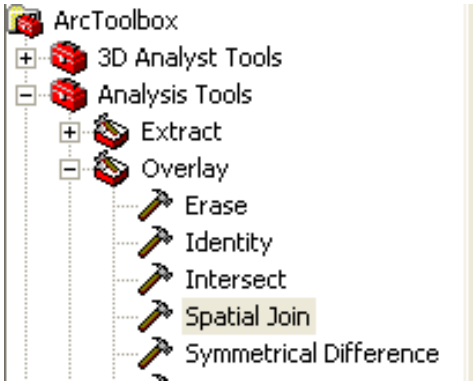
5. Now open the attribute table associated with *MN_Tracts_2010.shp* and toggle the buttons on the bottom of the table to **Show only selected records**. How many tracts contain McDonald's? You should note that there are 218 tracts that contain a McDonald's for the state of Minnesota.
6. Take a look at median household income (MedHHinc) for the tracts that contain McDonald's as opposed to those that do not. To do this, right click on the MedHHinc field. Select **Statistics**.
7. Switch the selection and check the descriptive statistics on those tracts that do not contain McDonald's.
8. This is great information and a good way to understand your data with a query, but like a tabular selection it is not permanent. What about making a more formal connection between

the data? How about determining the count of the McDonald's restaurants by census tract? You can accomplish this with a spatial join.

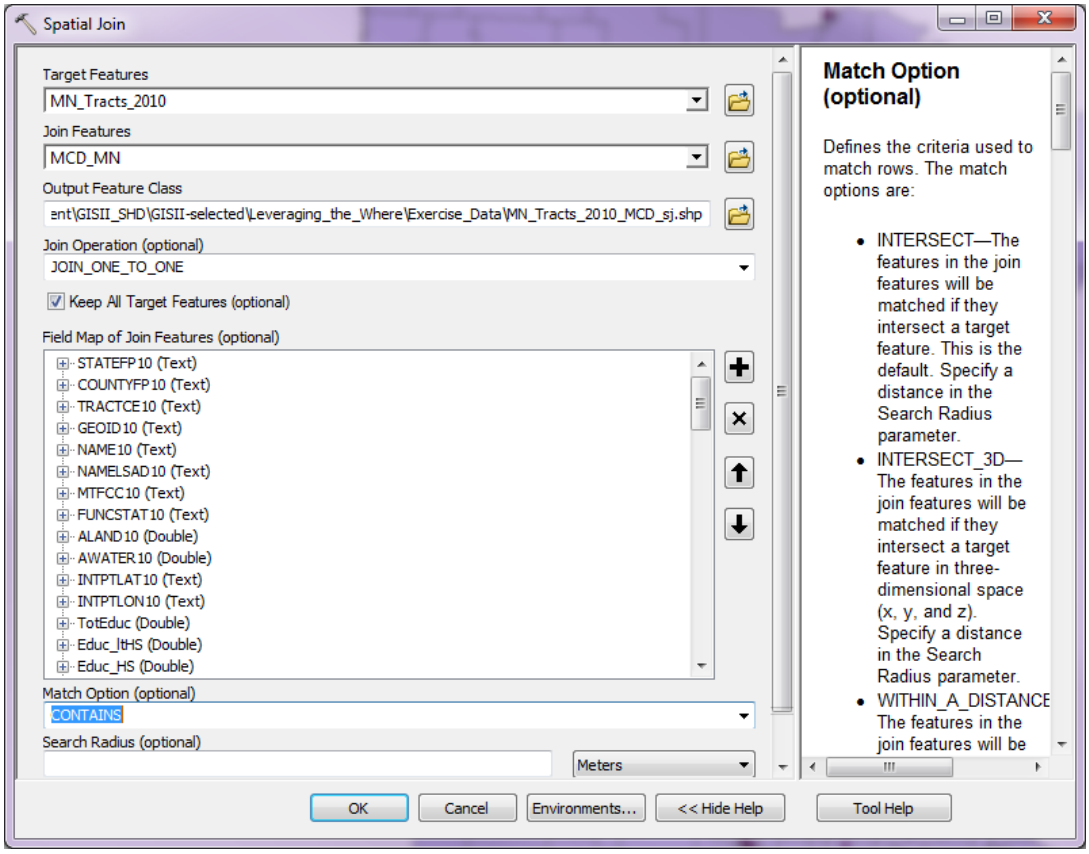
9. Clear your selected features using the  button.

Performing a Spatial Join

1. You are now ready to perform a spatial join between two related layers of data with the goal of answering a question that was initiated by a spatial selection: How many restaurants of interest are present with each host tract?
2. Open **ArcToolbox.** Go to **Analysis Tools > Overlay > Spatial Join**.



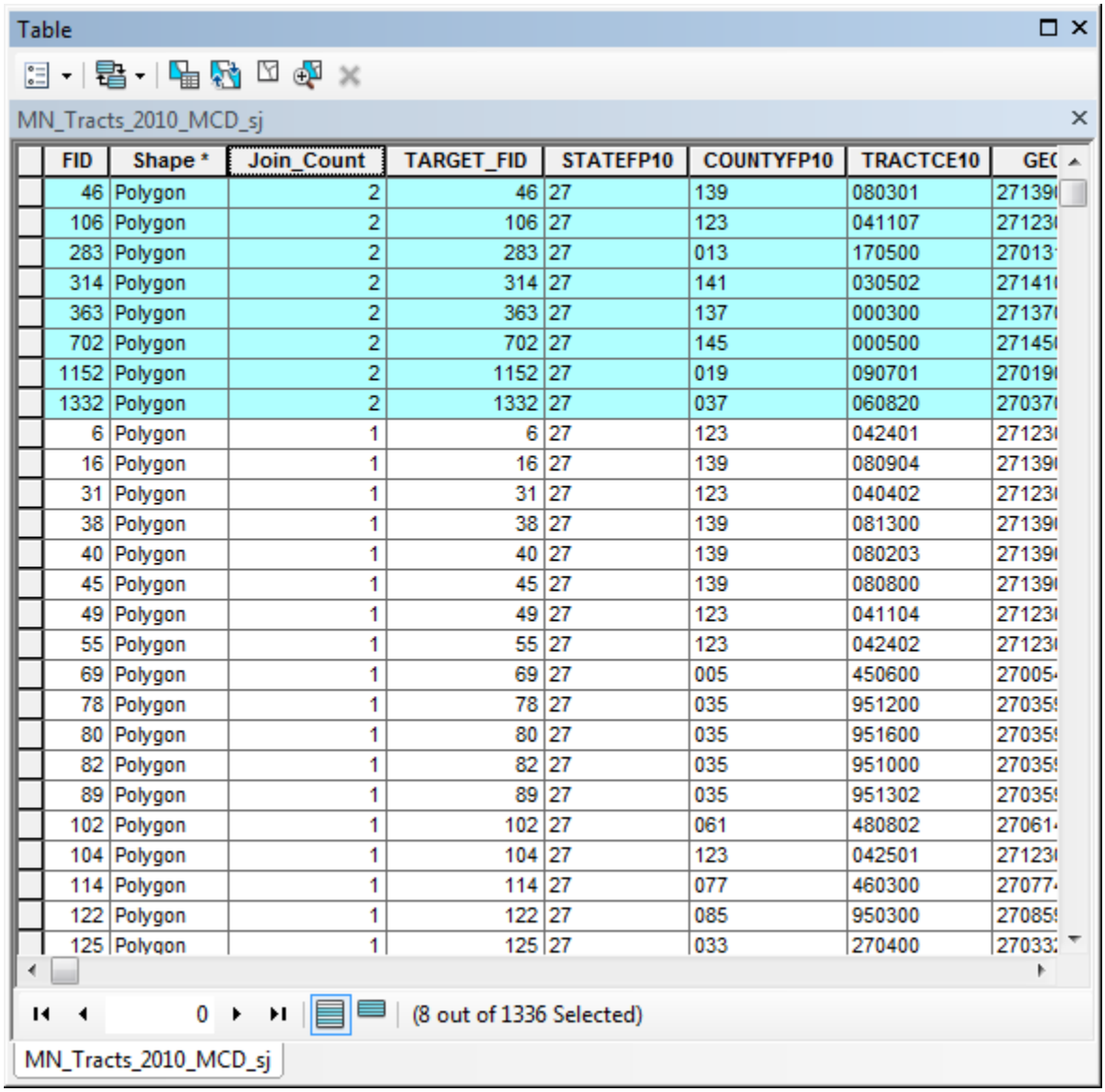
3. Our Target Features (the data to which you are appending) will be: *MN_Tracts_2010.shp* . Our Join Features (the data you are appending) will be: *MCD_MN* . Join Operation: you want to JOIN_ONE_TO_ONE. Keep the default: keep all target features checked, and Match option: CONTAINS.



4. Notice that this spatial operation creates a new dataset that you will need to name and find a home for on your computer. Make sure you save it in your Exercise_Data folder as a **shapefile** (a good name: MN_Tracts_2010_MCD_sj.shp).
    a. Think about this join: McDonald's to tract. It is possible that one tract may contain more than one McDonald's, right? So how will the attribute fields be handled for your join features (McDonald's) when multiple restaurants are present in a tract? **There are no fields that you are interested in capturing in from our Minnesota McDonald's table other than a count of restaurants by tract, but if you had numeric values that you would like to keep tabs on by tract, like sales volume or number of employees, you would adjust the merge rules for your fields in the Field Map of the Join Features window.** This controls what happens to fields when there is more than one record that fulfills your match option (i.e. tracts that contain more than one McDonald's).
5. Click **OK** to run the spatial join tool. The new dataset will be automatically added to your table of contents. Open the attribute table and take a look the fields. For every record in your tract

table (each record represents a closed polygon geometry) where a McDonald's is present, information for the restaurant has been appended to the tract.

6. You can also check this by looking at the *Join_Count* field; it will be the first field in your new table. A count greater than 0 means that the tract contains a McDonald's.
7. Are there any tracts with multiple McDonald's? Sort the *Join_Count* field in descending order by right clicking on the field and choosing **Sort descending**.



FID	Shape *	Join_Count	TARGET_FID	STATEFP10	COUNTYFP10	TRACTCE10	GEO
46	Polygon	2	46	27	139	080301	27139
106	Polygon	2	106	27	123	041107	27123
283	Polygon	2	283	27	013	170500	27013
314	Polygon	2	314	27	141	030502	27141
363	Polygon	2	363	27	137	000300	27137
702	Polygon	2	702	27	145	000500	27145
1152	Polygon	2	1152	27	019	090701	27019
1332	Polygon	2	1332	27	037	060820	27037
6	Polygon	1	6	27	123	042401	27123
16	Polygon	1	16	27	139	080904	27139
31	Polygon	1	31	27	123	040402	27123
38	Polygon	1	38	27	139	081300	27139
40	Polygon	1	40	27	139	080203	27139
45	Polygon	1	45	27	139	080800	27139
49	Polygon	1	49	27	123	041104	27123
55	Polygon	1	55	27	123	042402	27123
69	Polygon	1	69	27	005	450600	27005
78	Polygon	1	78	27	035	951200	27035
80	Polygon	1	80	27	035	951600	27035
82	Polygon	1	82	27	035	951000	27035
89	Polygon	1	89	27	035	951302	27035
102	Polygon	1	102	27	061	480802	27061
104	Polygon	1	104	27	123	042501	27123
114	Polygon	1	114	27	077	460300	27077
122	Polygon	1	122	27	085	950300	27085
125	Polygon	1	125	27	033	270400	27033

Take a look at the resulting data; can you think of useful applications in your own work?

GIS II: Data Management
Leveraging the *Where* of Geographic Data Exercise

This type of examination is particularly useful in assessing the food environment for your state or community. More information on assessing the food environment may be found here:

http://www.cdc.gov/obesity/downloads/HFRassessment.pdf

Modified Retail Food Environment Index Across Census Tracts within State:

- Measures the number of healthy and unhealthy food retailers across census tracts in a state
- Lower scores indicate:
    - Less healthy food retailers
    - More convenience stores and fast food restaurants relative to healthy food retailers

Modified Retail Food Environment Index Across Impoverished Census Tracts within State:

- Measures only census tracts where 20% or more of the residents are below the poverty line

Community Food Environment

Strategies to improve the community food environment include increased access and availability to healthier food retailers. The modified Retail Food Environment Index (mRFEI) measures the number of healthy and less healthy food retailers within a census tract using the formula shown.

$$mRFEI = 100 \times \frac{\text{\# Healthy Food Retailers}}{\text{\# Healthy Food Retailers} + \text{\# Less Healthy Food Retailers}}$$

For this indicator, healthy food retailers include supermarkets, supercenters, and produce stores.† Less healthy food retailers include convenience stores, fast food restaurants, and small grocery stores with 3 or fewer employees.† Higher mRFEI scores indicate more access to healthy food retailers and fewer less healthy food retailers.

Indicator	National Data*
• Modified Retail Food Environment Index across census tracts within state	10
• Modified Retail Food Environment Index across impoverished census tracts within state	7

*The numbers shown for national data are the medians of all state scores.
† Areas without these types of healthy food retailers may still provide adequate access if smaller stores and fast food restaurants provide quality and affordable healthy foods and beverages.

More information on mRFE can be found here:

http://www.cdc.gov/obesity/downloads/NationalActionGuide.pdf