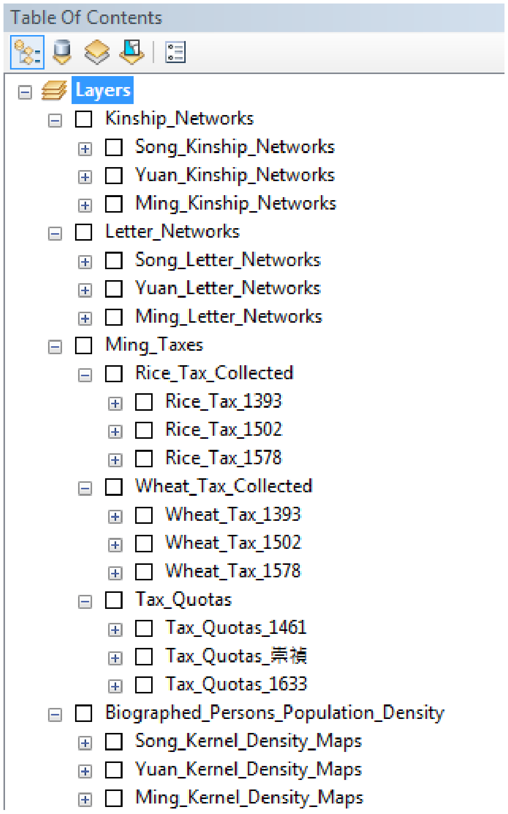
Darius Li and Adam Mitchell
Chinese History 200r
Professor Bol
16 January 2012

Visualizing CBDB Data from the Song to Ming Dynasties: ArcGIS Technique and Methodology

Contents

Layers in CH 200r Project Main.mxd



Kinship_Networks: This is a visual representation of kinship networks for 20-year periods from 960 to 1659. The data was collected from CBDB by running a Social Network Query (for all kinship ties) for all persons with an index year falling within the upper and lower bounds of the Song, Yuan, and Ming dynasties. This representation, created though the XY to Line function includes the full set of data from CBDB, but because of the function, only displays the connections between persons belonging to places with different XY coordinates.

Letter_Networks: This is a visual representation of the networks of literary exchanges for 20-year periods from 960 to 1659. The data was collected from CBDB by running a Social Network Query (all forms of writing checked under the "Select Types of Relations" tab) for all persons with an index year falling within the upper and lower bounds of the Song, Yuan, and Ming dynasties. Like the relationship network data, this representation only displays the written communications between persons belonging to places with different XY coordinates.

For both Kinship_Networks and Letter_Networks, three sub-layers roughly divide the data into the Song, Yuan, and Ming dynasties. We used Excel pivot tables to create a count function to determine the intensity of interaction (kinship and literary exchanges) between any given two locations. This data is however very sparse for the Ming Dynasty and should be updated once more Ming biographies have been added to the CBDB.

Ming_Taxes: The data source is 中國歷代戶口 、田地 、田賦統計 by 梁方仲

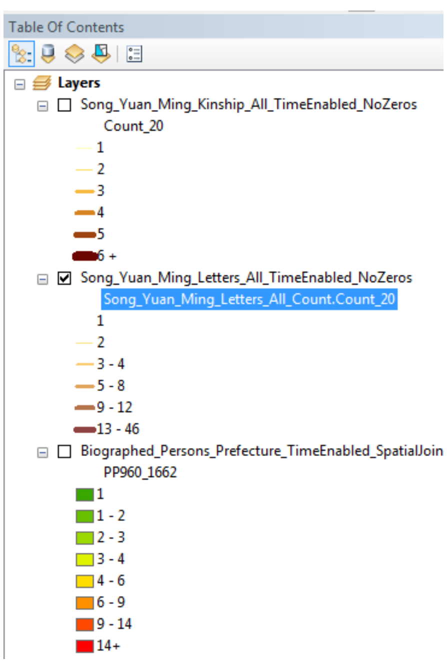
Rice_Tax_Collected:
The data is compiled from pages 332-3 and 344-5 in the book (乙表 29 and 35).
This layer has the rice tax collection data across different locations for the years 1393, 1502, and 1578.

Wheat_Tax_Collected:
The data is compiled from pages 332-3 and 344-5 in the book (乙表 29 and 35).
This layer has the wheat tax collection data across different locations for the years 1393, 1502, and 1578.

Tax_Quotas:
The data is compiled from pages 354-359 in the book (乙表 43-45).
The unit of the data is 糧 (石).
The layer has the tax quotas across different locations for 天順五年 (1461), 崇貞年間 (unknown year), and 崇貞六年 (1633).

Biographed_Persons_Population_Density: This layer has the kernel density maps for the number of biographed persons in CBDB whose index years are in the 20-year periods from 960 to 1659. The layers are named after the first year in the 20-year period. For instance, K0960 is the kernel density for the people whose index years range from 960 to 979.

Layers in CH 200r Project Time Animation.mxd



Song_Yuan_Ming_Kinship_All_Time_Enabled_NoZeros:

This layer has all the kinship connections for people in CBDB whose index years fall between 960 and 1659, and is a time-enabled version of all the 20-year kinship layers in Song_Kinship_Networks.lyr, Yuan_Kinship_Networks.lyr, and Ming_Kinship_Networks.lyr.

This layer has the same data as Song_Yuan_Ming_Kinship_All_TimeEnabled.lyr except that all entries in which the kinship pair lived in the same location are removed for the purposes of not slowing down the time slider with excessive data. Distances of 0 do not appear on the map anyway, so removing these entries do not affect the visual representation

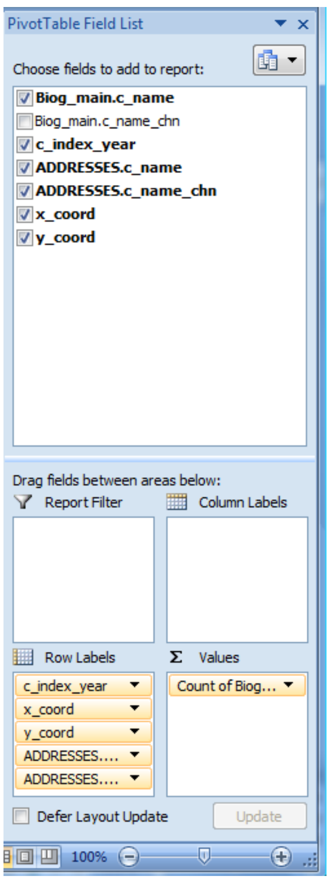
Song_Yuan_Ming_Letters_All_TimeEnabled_NoZeros:

This layer has all the literary connections for people in CBDB whose index years fall between 960 and 1659, and is a time-enabled version of all the 20-year layers in Song_Letters_Networks.lyr, Yuan_Letters_Networks.lyr, and Ming_Letters_Networks.lyr.

This layer has the same data as Song_Yuan_Ming_Letters_All_TimeEnabled.lyr except that all entries in which the sender and recipient lived in the same location are removed for the purposes of faster processing. Distances of 0 do not appear on the map anyway, so removing these entries do not affect the visual representation.

Aggregating Data in Excel

Because we chose to divide our CBDB biographical population data into 20-year periods, we want the intensity of our relationships (in this case the count of relationship ties that occur between two locations for any given period) to reflect the data for an individual period and not the entire dataset. Since ArcGIS's time series does not have the function to aggregate data in this way, we had to pre-aggregate the data in Excel by assigning each biographical person to a 20-year period associated with his index year. This tricks GIS into thinking that our data moves through time in 20-year increments when in fact we aggregated the data within each 20-year period.



Step 1: Open Song_Yuan_Ming_Elites_960_1660.xlsx (in the Worksheets folder). The Song-Ming Elites 960-1660 tab has the original export from CBDB with all the persons queried from Biog_main whose index years ranged from 960 to 1660.

Note: Make sure to delete duplicated entries. In CBDB, any given biographed individual may have multiple addresses.

Step 2: Create a pivot table. The pivot table options are listed in the diagram on the left. We pivoted the data based on index year, coordinates, address, and then created a count under Values based on the field Biog_main.c_name. This pivot table is shown in the Pivot Table tab in Song_Yuan_Ming_Elites_960_1660.xlsx.

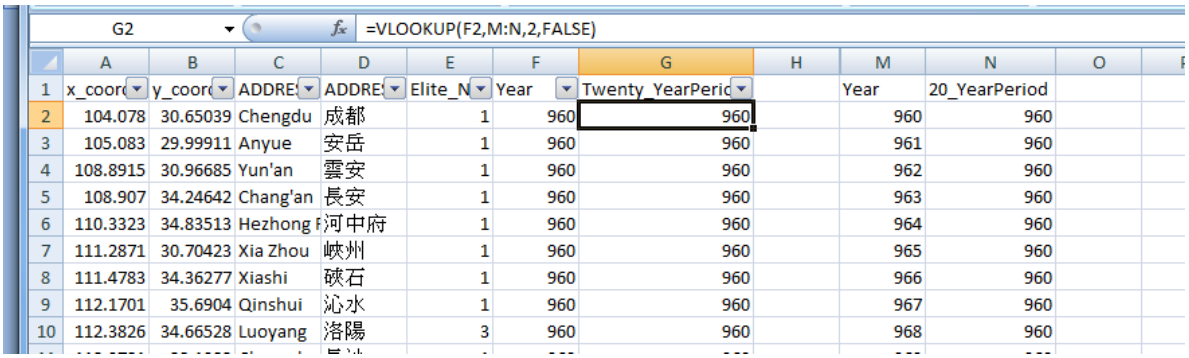
Note: Any given location may have different codes for different time periods. Also, any given location might be assigned to multiple coordinates in the CBDB database. For consistency, consider cleaning the data by assigning a single coordinate and address code to each location.

Step 3: The Formatting Pivot tab adds Excel formulas to repeat row labels for the index year and coordinates, which do not appear automatically in the pivot table as shown below.

	A	B	C	D	E	F
1	c_index_year	x_coord	y_coo	ADDR	ADDRESSES.c_name_chn	Count of Biog_main.c_name
2	960	104.077995	30.65	Cheng	成都	1
3				Huaya	華陽	1
4		105.08304	29.999	Anyue	安岳	1
5		108.891495	30.967	Yun'an	雲安	1
6		108.906976	34.246	Chang'	長安	1
7		110.332306	34.835	Hezhor	河中府	1
8		111.287125	30.704	Xia Zh	峽州	1
9		111.47834	34.363	Xiashi	硤石	1
10		112.170128	35.69	Qinshu	沁水	1
11		112.38263	34.665	Luoyan	洛陽	3
12		112.978134	28.198	Chang	長沙	1
13		113.103485	36.184	Shang	上黨	1
14		113.819748	34.031	Chang	長社	1

Step 4: Open Song_Yuan_Ming_Elites_960_1660_FinalPivot.xlsx (in the Worksheets folder). This Excel sheet further formats the pivot table to assign each index year to the first year of a 20-year period.

Step 5: In the Elites960_1660_Year_Conversion tab, in column M, we listed all the years from 960 to 1660. In cell N2, we typed the formula, =ROUNDDOWN(M2/20,0)*20 and applied it down the column. This assigns each year to the beginning of its corresponding 20-year period. For instance, 968 is assigned to 960.


G2 =VLOOKUP(F2,M:N,2,FALSE)

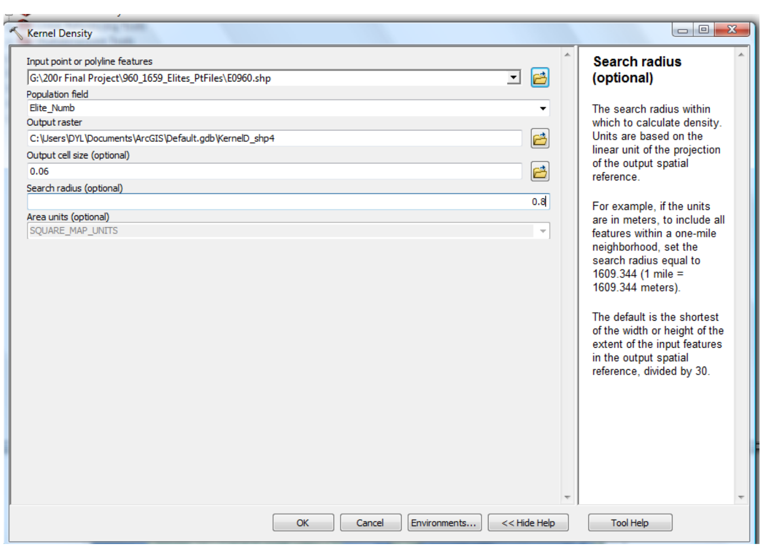
	A	B	C	D	E	F	G	H	M	N	O
1	x_coord	y_coord	ADDRE	ADDRE	Elite_N	Year	Twenty_YearPeric		Year	20_YearPeriod	
2	104.078	30.65039	Chengdu	成都	1	960	960		960	960	
3	105.083	29.99911	Anyue	安岳	1	960	960		961	960	
4	108.8915	30.96685	Yun'an	雲安	1	960	960		962	960	
5	108.907	34.24642	Chang'an	長安	1	960	960		963	960	
6	110.3323	34.83513	Hezhong F	河中府	1	960	960		964	960	
7	111.2871	30.70423	Xia Zhou	峽州	1	960	960		965	960	
8	111.4783	34.36277	Xiashi	硤石	1	960	960		966	960	
9	112.1701	35.6904	Qinshui	沁水	1	960	960		967	960	
10	112.3826	34.66528	Luoyang	洛陽	3	960	960		968	960	

Step 6: In the same tab, we use the vlookup function in Column G to assign all the index years in the data pasted from the original pivot table to a 20-year period.

Step 7: Take Columns A to G in the Elites960_1660_Year_Conversion tab, and recreate the pivot table, but this time, use the 20-year period as the row label for time. This will further aggregate the number of elites in a given location during a given 20-year period. The Elites960_1659Final tab shows the final data that is exported into ArcGIS.

Creating Kernel Densities of the CBDB Biographical Population

Step 1: Under ArcToolbox, click Spatial Analyst Tools → Density → Kernel Density



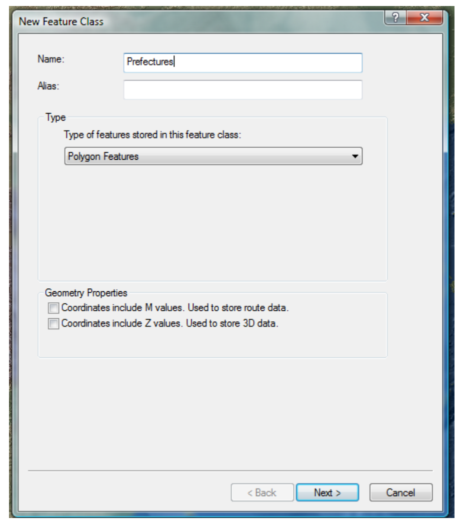
Step 2: Input a point file under the 960_1659_Elites_PtFiles folder. The files in this folder have data on the number of elites (CBDB persons in biog_main) in each given location for a given 20-year period from 960 to 1659.

Under the Population field, select Elite_Number.

We selected the output cell size as 0.06 (this determines the pixelation of the kernel density map). The map is more pixelated as the cell size increased. We selected the search radius as 0.8, so that the program counts the number of elites in circles of radius 0.8. The unit in our case is degrees. Click OK. A kernel density map will automatically appear in the Table of Contents pane.

Time Animation for the CBDB Biographical Population Density: Spatial Join

In this part of the project, we aim to produce a spatial-join time series in which the map aggregates the number of biographical persons whose coordinates fall under particular prefectural boundaries. One tedious approach would be to spatial-join the prefectural polygons to the point files. However, this approach will only join the polygon to the first point that falls under each polygon. In order to create a time series, we would have to spatially join each polygon to all points in each 20-year period.

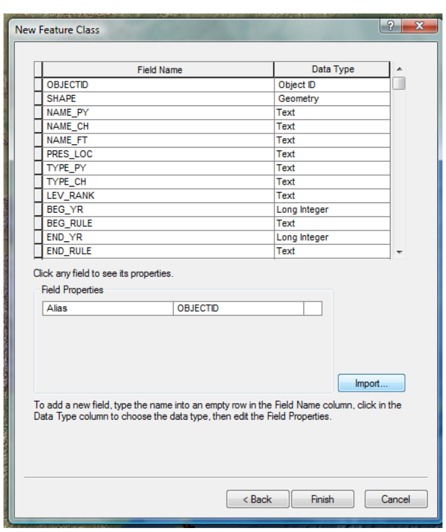


Please be aware that prefectural boundaries change throughout different historical periods. Since we aggregated our data into 20-year periods, the effect of boundary changes is reduced. Also, since we only have the shape file for the 1820 prefectural boundaries, using one set of prefectural boundaries provides some stability in the spatial-join technique.

Here is the method:

Step 1: Create geodatabase Song_Ming_Animation.gdb.

In the Catalog pane, right click on the folder that you want to store the geodatabase. Click New → File Geodatabase. We named the geodatabase Song_Ming_Animation.gdb.

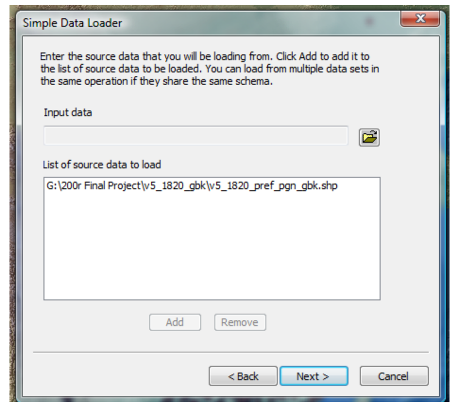


Step 2: In the Catalog pane, right click Song_Ming_Animation.gdb and click New → Feature Class.

In the New Feature Class dialog box, give the feature class a name (Prefectures). Select the correct coordinate system. Then continue to click Next until you reach the following step:

Click Import and select v5_1820_pref_pgn_gbk.shp. The fields from this file will appear. Click Finish to create the Prefecture feature class.

This only imports the fields of the shape file. In order to input the actual data, we must load the data.

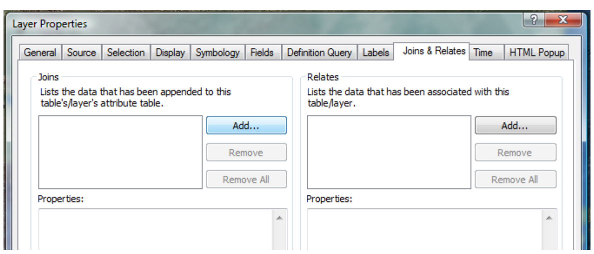


In the Catalog pane, right click the newly created Prefecture feature class. Click Load → Load Data.

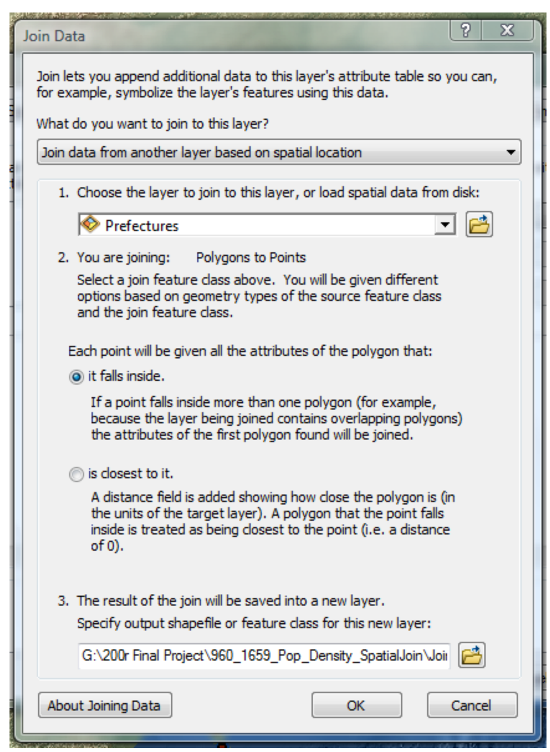
In the Simple Data Loader dialog box, input v5_1820_pref_pgn_gbk.shp and then click Add. Then continue to click Next, retaining the default options, and then click Finish. This adds all the data from v5_1820_pref_pgn_gbk.shp into the Prefecture feature class.

Step 3: Add Export960_1659.shp into Table of Contents. This is the shape file created by exporting Song_Yuan_Ming_Elites_960_1660_FinalPivot.xlsx.

Step 4: In the Table of Contents pane, right click on Export960_1659 and click Properties.



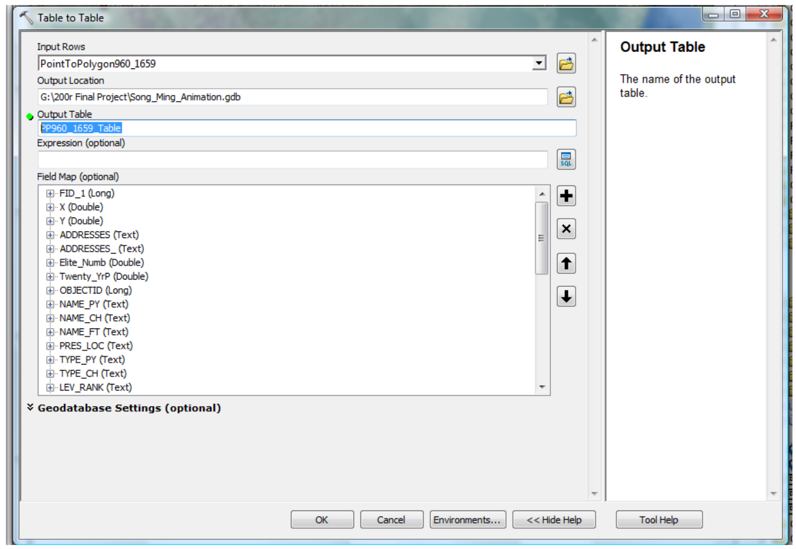
Under the Joins & Relates tab, under the Joins Section, click Add. This will take you to the Join Data dialog box.



Step 5: In the Join Data dialog box, select Join Data from another layer based on spatial location. Choose the Prefecture feature class for #1. Retain the default selection “it falls inside” under #2. In #3, save this spatially joined shape file as PointToPolygon960_1659.shp. This shape file is a point-to-polygon shape file, since we associated each point with a polygon into which each point falls.

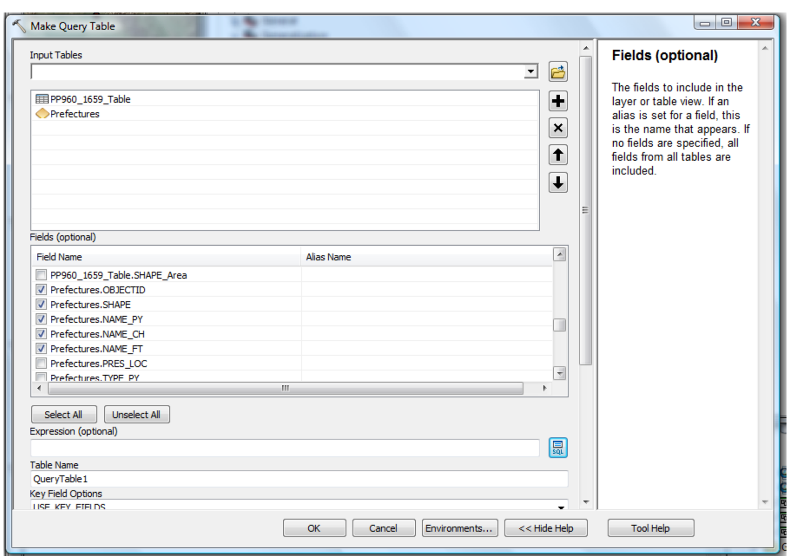
Note: Do not spatially join the points to polygons by selecting the original v5_1820_pref_pgn_gbk.shp. For some reason, the entries in the shape file will associate the polygons with an ID number that does not match the ID number in the Prefectures feature class. In fact, the ID numbers will be off by one, but this makes it very difficult to make a query in a later step. Theoretically, you could add a field in the attribute table, using the field calculator to add one to the ID numbers. However, the query function will not allow SQL syntax to apply to fields created from the field calculator.

Step 6: In the Catalog pane, right click Song_Ming_Animation.gdb. Click Import → Table (single). The Table to Table dialog box will appear. Import PointToPolygon960_1659. Under Output Table, name the table PP960_1659_Table. Click OK. We now have the point-to-polygon shape file and prefecture shape file data in one geodatabase, so we can now create a query table.



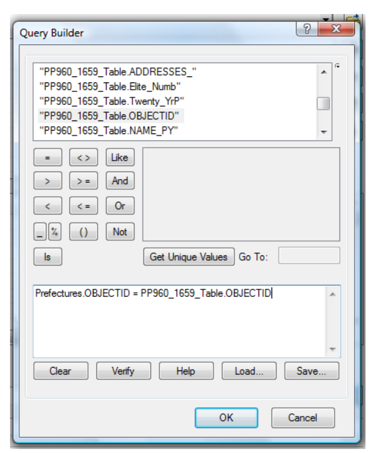
Step 7: Make a query table.

Under Arc Toolbox, select Data Management Tools → Layers and Table Views → Make Query Table.



Input both PP960_1659_Table and the Prefecture feature class. Under Fields, select all the fields that you would want to appear in the new query table.

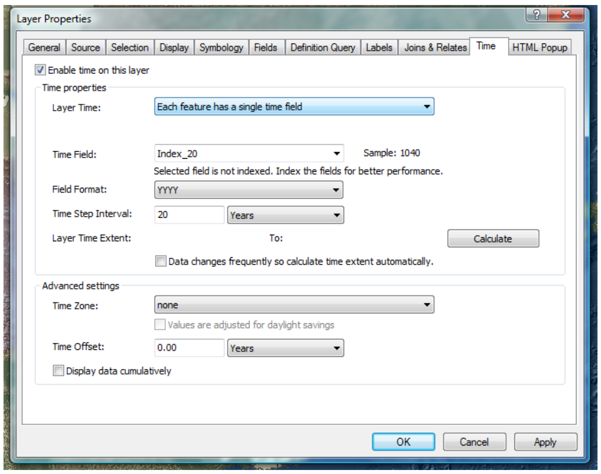
Note: you must include Prefectures.SHAPE in order to associate all the points into polygons.



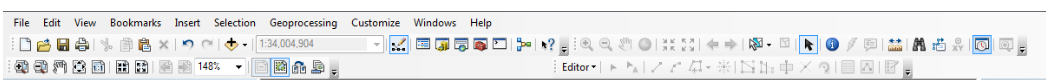
Click the SQL button under the Expression (optional) field. A Query Builder dialog box will appear. Create the SQL query Prefectures.OBJECTID = PP960_1659_Table.OBJECTID. This will join all the entries in PP960_1659_Table with the corresponding polygon shape. PP960_1659_Table.OBJECTID is the ID obtained from the original point-to-polygon spatial join that created PointToPolygon960_1659.shp. This query simply associates the prefectural ID with the corresponding polygon shape.

Now, the shape file QueryTable will automatically appear in the Table of Contents pane. Right click on QueryTable and click Data → Export Data. Save the exported data as TimeEnabled960_1659.shp.

Step 8: The exported data will change the field names of the shape file. Open the attribute table for TimeEnabled960_1659.shp and right click on all the fields to change the field names accordingly.

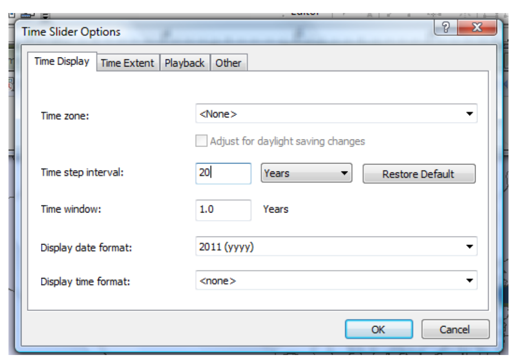


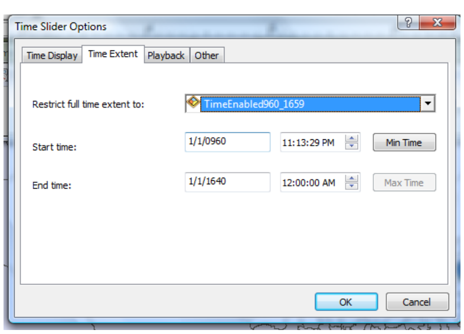
Step 9: Right click on TimeEnable960_1659.shp in the Table of Contents pane and click Properties. After adjusting the symbology, click on the Time tab. Click “Enable time on this layer.” Then choose the time field and field format (YYYY). The time step interval in this case is 20 years.



Step 10: Click on the “Open Time Slider Window” button which has a picture of a clock.

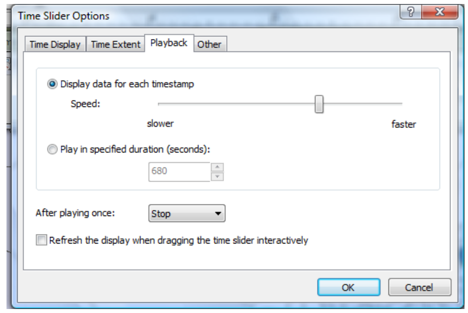
In the Time Slider Options dialog box, under the Time Display tab, you can adjust the time step interval, which we set to 20 years. You can also change the display data format to YYYY.





Under the Time Extent tab, you can select that time extent of the time slider by manually adjusting the start and end time or selecting the shape file for the full time extent.

Under the Playback tab, you can adjust the speed of the time slider presentation.



Social Network Representation in GIS

This tutorial is for the visual representation of spatial data with interactions between nodes and their edges using ArcGIS. Lines will be drawn between coordinates representing the locations of interacting nodes, and the boldness and color of the lines can be adjusted to represent the intensity of the relationships. This data can be viewed as static datasets or as a moving time series. In order to use this method of visual representation, each relationship must occur between two actors, and each node must have XY coordinates. As examples, we demonstrate how to visually represent kinship network interactions and letter correspondences for the Song, Yuan, and Ming dynasties.

Organization of Data in Excel

Step 1: After we have generated an edgelist and attribute list, import them into a single Excel workbook as two separate spreadsheets.[i] After these spreadsheets have been created, open a new blank spreadsheet and copy and paste in columns A and B the names of your nodes and edges. In the column C, create a vlookup for the name of the node1 in column one and the location of the node from the attribute table. In columns D and E, do vlookups for the x and y coordinates for the location of the node.[ii] In columns F through H repeat this process for the attributes for the node2.

Step 2: Next we have to determine a year for when the relationship took place. If the kind of relationship that you are analyzing happens to have occurred in a static period of time (i.e. the sending of a letter, a marriage, etc.), and you happen to know the date, then create a column that lists all of the corresponding dates for the relationships. If the relationship occurs over a period of time, and you do not know the exact period, but you do have time data for your nodes, then construct a way of averaging the time data using node attributes.[iii]

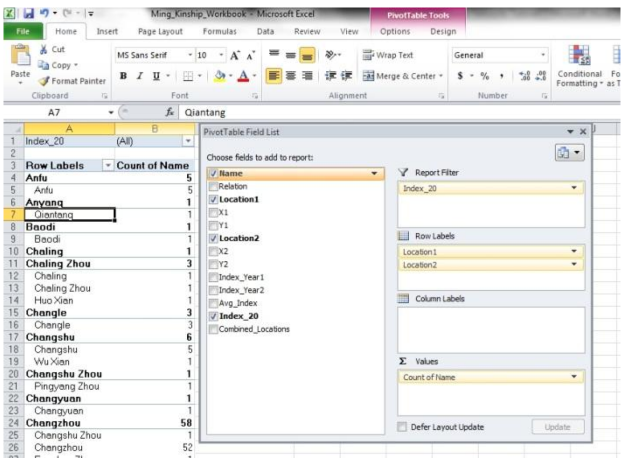
Step 3: Now that we have our time data for each of our realtionships, unless we want to show the data year-by-year, we should group the data into periods. This can be accomplished using a number of methods, but in our case we choose to use nested If statements and a Rounding function (see the **Aggregating Data in Excel** section). Using one of these methods, create a new column that lists the period in which each relationship occurred. For using the If expression, see the below example:

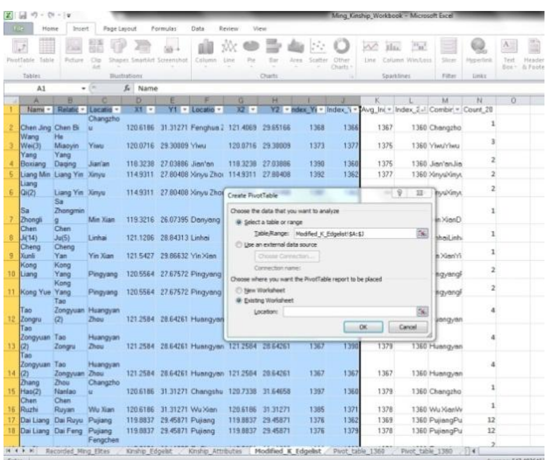
=IF(K2<1380,1360,IF(J2<1400,1380,IF(K2<1420,1400,IF(K2<1440,1420,IF(K2<1460,1440,IF(K2<1480,1460,IF(K2<1500,1480,IF(K2<1520,1500,IF(K2<1540,1520,IF(K2<1560,1540,IF(K2<1580,1560,IF(K2<1600,1580,IF(K2<1620,1600,IF(K2<1640,1620,IF(K2<1660,1640,IF(K2<1680,1660,FALSE))))))))))))))))

This equation was used to designate 20-year time periods within the Ming dynasty. It begins by checking to see if the value in the cell with our relationship date (i.e. K2) is less than my starting period (i.e. 1380). If it is smaller, then the equation outputs 1360 (here the period I designated as 1360 represents the years 1360-1379). If the date was larger, then the equation moves to the second If statement and checks to see if the date is smaller that our second earliest period. This process is continued to the final expression, at which point, if the date is still larger than our final period (i.e. 1660) then the statement outputs FALSE.[iv]

Creating Relationship Intensity

Step 4: Once we have our column that organizes the relationships into periods, we can assign an intensity to each relationship. Intensity can be defined however best suits your particular dataset, but the standard method would be to generate a count of the number of interactions that occur between two unique locations during a given period of time. Because ArcGIS does not have the capability to aggregate data, we will have to do this using Excel. First, we will need to create a series of PivotTables, one for each time period. So select your columns as the pivot array, making sure to include the nodes as well as the locations. Click the Insert tab at the top of the screen, and select the PivotTable button at the far left. A Create PivotTable menu will pop up in the center of the screen. Select the Existing Worsheet button at the bottom, move to a blank sheet, and click any empty cell. After hitting OK, a new PivotTable will be generated at the cell you selected. Click the blank PivotTable in the spreadsheet to pull up the PivotTable Field List.

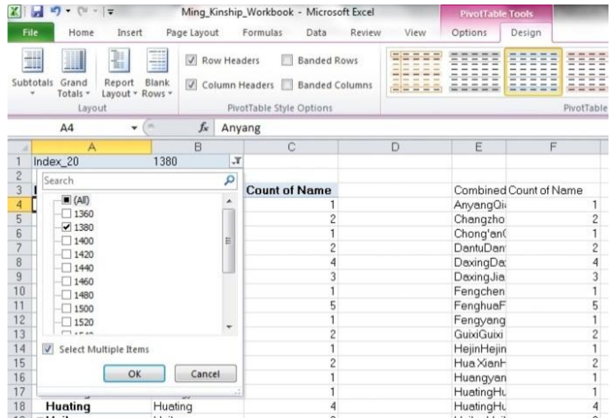


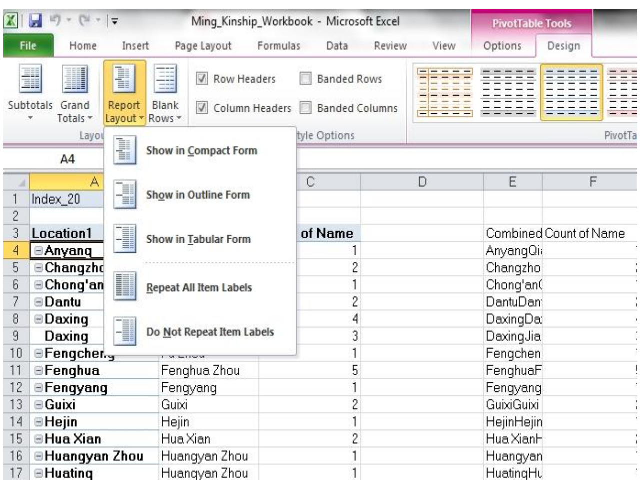


In the PivotTable Field List, you will see a main box with all of your column headers, and four sub boxes. To generate counts for relationships between places, select the location for each node and drag them from the list into the Row Labels box. Next we have to designate what column is being counted. In many cases, our nodes are people, and persons are listed for each unique realtionship. Therefore, we should drag the Name label, or whatever you designated as your identifyer for the nodes, into the Values box. Finally, we want to filter the data by periods; so drag the period designator into the ReportFilter box. After inputting these settings, select OK to generate the PivotTable.

Note: The relationship intensities shown in GIS are the number of connections between any two given locations aggregated from people associated with these places. The connections between people associated with locations are still perserved in the Excel documents.

Step 5: Now we have an unfiltered PivotTable, but the data is not yet in the proper format. Depending on your default setting, the location data is most likely being displayed as stacked cells, with locations associated with node2 being displayed in a row below the locations associated with node1. To fix this click on a cell within the PivotTable, then click the Design tab. Under the design tab, on the left side of the screen will be a dropdown button named Report Layout. Click on the button to bring up the drop down menu, and select the Show in Tabular Form option. Click on the button once more and select Repeat All Item Labels. Now the data will be displayed in the proper format with each pair of locations being listed in adjacent cells and an associted count given to a column to their right.





Now that the data is being displayed in the proper format, the final step in the creation of the PivotTable is to select the appropriate filter. At the very top left of the PivotTable, you will find a bar named as the category you selected as the filter. Click the button to on the right side of the bar to pull up a dropdown menu. Deselect the All filter and choose a single period. After selecting OK, you will have generated intensity counts for all related locations in your selected period of time. Instead of modifying this PivotTable to generate the counts for each period, copy this table into separate sheets, creating a unique PivotTable for each given period of time.

Finalizing Data in Excel

Step 6: Since we have generated a series of PivotTables, one for each of our designated periods of time, the last step in the creation of our spreadsheet is to input the intesensity counts. The easiest way to get the data from the PivotTables back into you main spreadsheet is to do a vlookup. But before we can perform a vlookup, we will need to create a new column that expresses the locations as a single field because vlookups cannot handle multiple expressions (i.e. because our intensity data is based on the number of relationships between two places, we need to perform a vllokup for location1 and its corresponding location2). To do this use the Concatenate function in excel, selecting the the location for node1 and location for node2 as your two fields. This will generate a single column that combines the names of the two locations as a single entry.

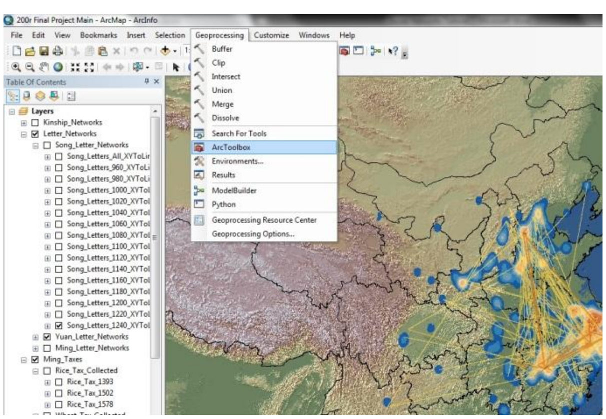
Now we must perform this same step for each of our PivotTables, concatenating the two location columns into a third column outside of the PivotTable. Also because PivotTables are dynamic and updatable, the vlookup function has difficulty searching within them. Therefore, we should also copy the column of counts from the PivotTable and paste the values to the left of our concatenated locations column. Go back to your main spreadsheet and create a new column with a nested If statement that performs a vlookup for each period. This formula is slightly complex so I will provide an example below:

=IF(Modified_K_Edgelist!L2=1360,VLOOKUP(Modified_K_Edgelist!M2,Pivot_table_1360!E:F,2,FALSE),
IF(Modified_K_Edgelist!L2=1380,VLOOKUP(Modified_K_Edgelist!M2,Pivot_table_1380!E:F,2,FALSE),
IF(Modified_K_Edgelist!L2=1400,VLOOKUP(Modified_K_Edgelist!M2,Pivot_table_1400!E:F,2,FALSE),
IF(Modified_K_Edgelist!L2=1420,VLOOKUP(Modified_K_Edgelist!M2,Pivot_table_1420!E:F,2,FALSE),
IF(Modified_K_Edgelist!L2=1440,VLOOKUP(Modified_K_Edgelist!M2,Pivot_table_1440!E:F,2,FALSE))))))

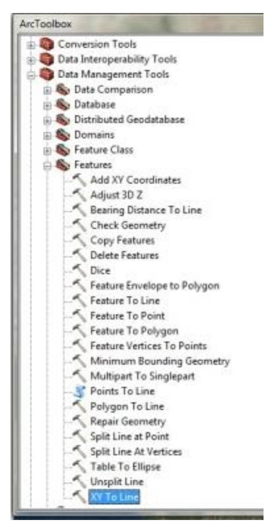
This formula is first matching an entry to its corresponding PivotTable; we do this by selecting the cell with our period data (i.e. Modified_K_Edgelist!L2) and setting it equal to the year of the earliest period. [v] After the date has been matched to its corresponding period, the equation uses a vlookup to search the appropriate PivotTable. Notice that in the vlookups, our concatenated location expression from our main spreadsheet (i.e. Modified_K_Edgelist!M2) is being used as the reference to find the matching term in our PivotTables (i.e. Pivot_table_xxxx!E:F). Then the vlookup is feeding back the information from the second column in the array, in this case, the column of counts that we copied from the PivotTable.

Step 7: Now that we have all of the pertinent information into a single Excel spreadsheet, there is one final thing that we must do before importing our data into ArcGIS. The function that we will be using in ArcGIS, XY to Line, does not accept data in the form of .xsl. Therefore we need to copy this data into a new workbook, and save it as a series of .csv files. If you wish to represent all of the data as one picture, simply save you main worksheet as a single .csv. However, to see the individual periods in ArcGIS, you will need to break up the data into periods, and save each period as a separate csv.[vi]

Importing the Data into ArcGIS

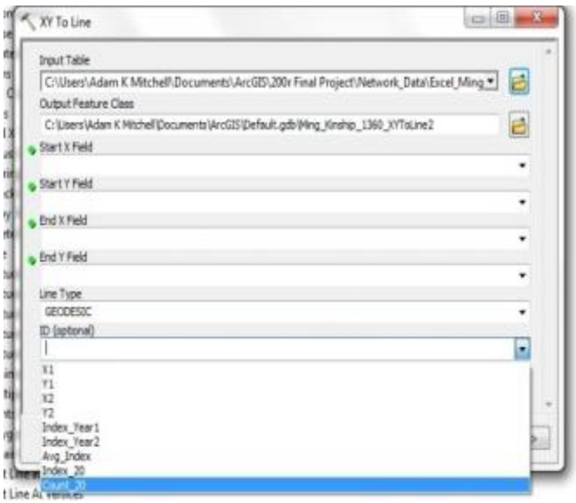


Step 8: After opening ArcMap and selecting your project folder, select the Geoprocessing tab at the top of the sceen and choose ArcToolbox. This should bring up a new window with various tools and functions. Click the plus sign next to Data Management Tools to



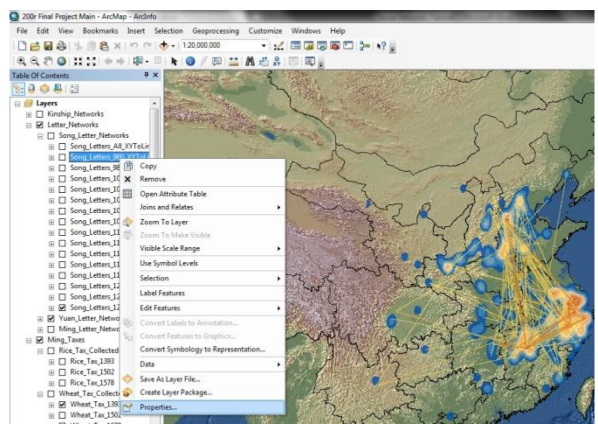
expand its subfields, and then click the plus sign next to Features. Under Features you should see a function at the very bottom called XY To Line; select this function to open a separate window.

Under the Input Table field, click the folder button and select the destination of your first .csv file. Once you have selected the file, the first two rows in the box (Input Table and Output Feature Class) should automatically be filled in.[vii] Next, hit the down arrows next to the Start and End XY fields (there are 4 in total) and select the appropriate column headers. For Start X Field and Start Y feild, select the XY coordintes for the location associated with node1; then use the location for node2 for the End fields. In the final field, be sure to select the geographic coordinate system that cooresponds with the rest of your data (it is important that you are using only one coordinate system). Finally, in order to include the intensity feature that we generated in Excel, under the ID field select the appropriate column name. Select OK and ArcGIS will create a new layer in your project with the .csv file data. Repeat this process for each of your .csv files.

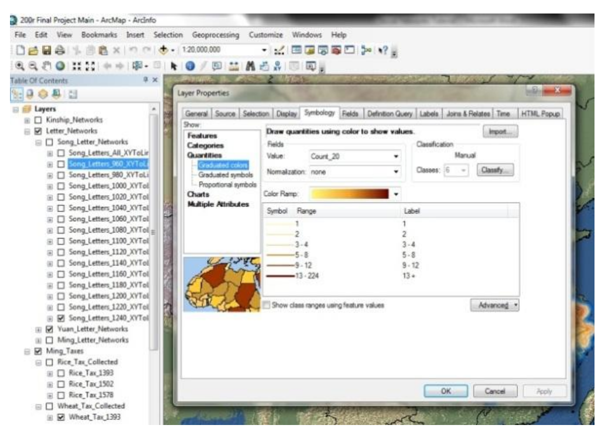


Changing Symbology in ArcGIS

Step 9: Now that the data has been imported into ArcGIS, our final step is to figure out how to best visually represent our realtionships. To change the appearance of the data, right click one of your layers and select Properties. This will open a separate Layer Properties window. Select the Symbology tab, and under Quantities you will see three options: Graduated Colors, Graduated Symbols, and Proportional Symbols.

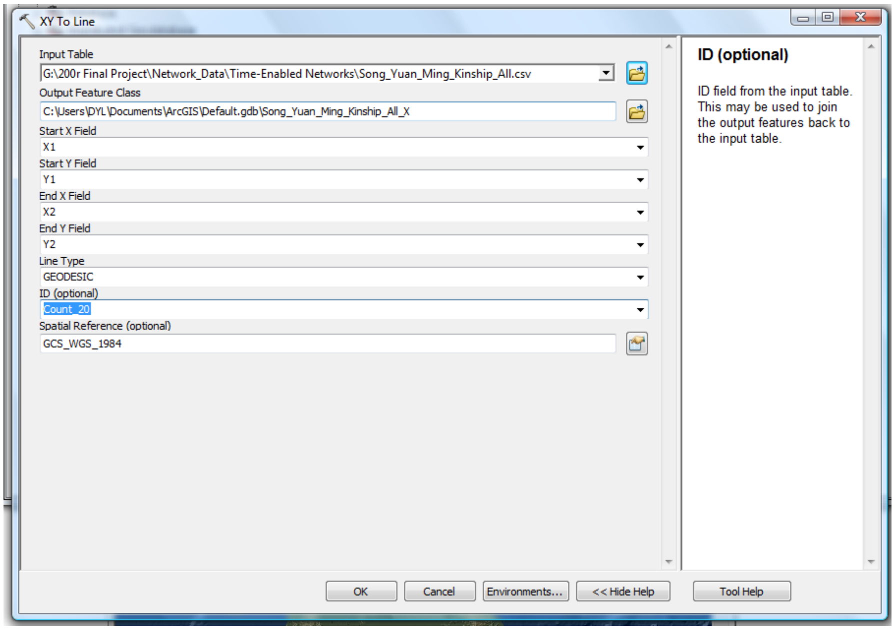


For the purpose of relational data, one of the first two will likely provide the best representation. By clicking on either, and designating as the Value the column name for your intensity counts, you will get a display that shows your relationship network with intensity. If you want to make your relational lines vary by both color and thinkness, choose the Graduated Color option, select intensity counts as your value, and manually click on each line and change its thinkness.



XY to Line: Time-Enabled

Step 1:   Go to ArcToolbox → Data Management Tools → Features → XY to Line.



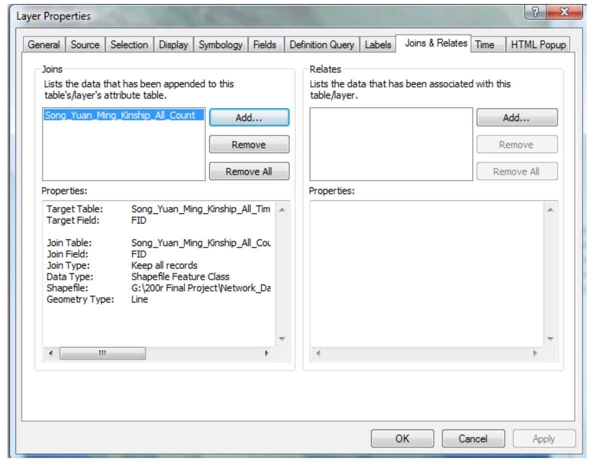
Step 2: Select a file under the Network_Data folder. All the .csv files for kinship and letter networks are organized under this folder. The Song_Yuan_Ming_Kinship_All.csv and Song_Yuan_Ming_Letters_All.csv (saved under the Time-Enabled Networks subfolder under the Network_Data folder) have the entire network data from 960 to 1659 necessary for time-enabled the network visualization. If you only want to create a static visualization, you can select the .csv file for a particular 20-year period.

Select the start and end X/Y fields. This connects the coordinates for each network pair. Under ID (optional), select Count_20, which is the intensity of the kinship or letter connection between two given locations. Click OK.

If you simply want a static network of a given 20-year period, this process is complete, since a shape file will automatically be generated in the Table of Contents pane. You can then edit the symbology using Count_20. If you want a time-enabled network, move onto Step 3.

Step 3: Repeat Step 2. However, for ID (optional), instead of Count_20, select Index_20, which is the 20-year period under which the average index years of the network pair falls. We will now join the two generated shape files based on the FID field in both shape files. This is the workaround method for the XYtoLine limitation of one attribute export under ID (optional)

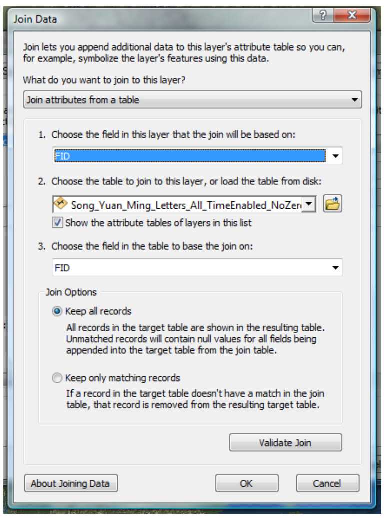
Step 4: In the Table of Contents, right click on the shape file with the Index_20 attribute. Then click on Properties.



Step 5: In the Joins & Relates tab, click Add under the Joins section.

The Join Data dialog box appears.

Step 6: Select "Join attributes from a table" in the first drop-down menu in the Join Data dialog box. Under #1, select FID. For #2, select the other shape file with the Count_20 field. Under #3, select FID. Click OK. The shape file with the Index_20 field will now also have Count_20 as a field. This will allow us to use Count_20 for geospatial symbology and Index_20 as the time field.

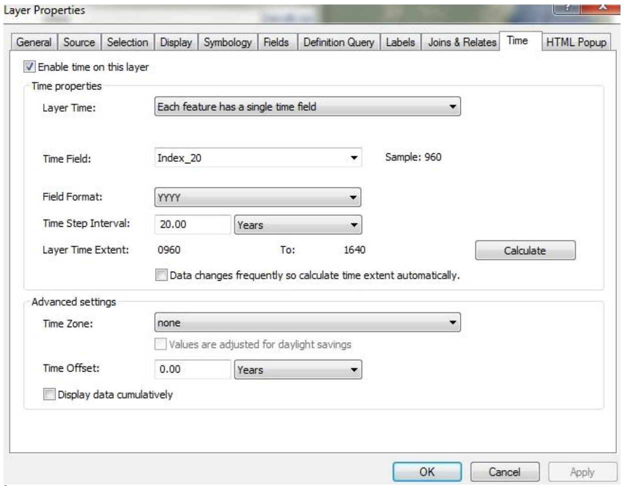


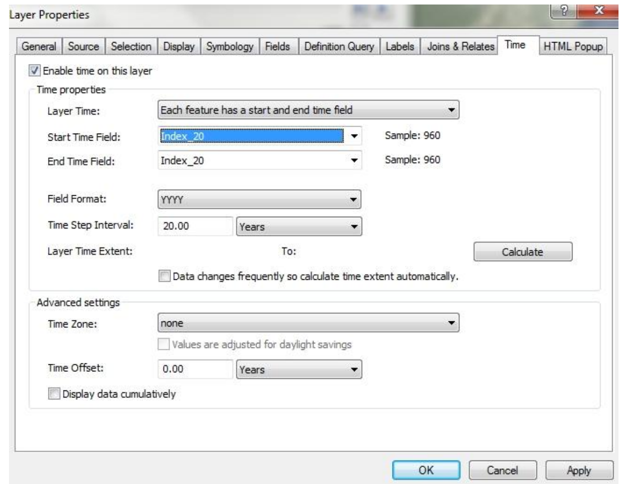
---

[i] Make sure that when you name your columns, you do not use any spaces or non-standard characters. It is okay to use Chinese characters, but if you use a space instead of an underscore, then you will get an error message when importing your data into ArcGIS, and none of the functions will run properly.

[ii] If your data does not come from CBDB, you may have the locations of nodes but not their corresponding XY coordinates. If this is the case, create an additional spreadsheet within the workbook and copy the table from CBDB that lists all known locations and with their XY coordinates. Follow the same procedure as above, but when you are importing the XY coordinates, simply do a vlookup of location from the CBDB table of addresses, instead of in your attribute table.

[iii] If your relationship data does occur over a specific range, and you happen to know that range, you can represent this using ArcGIS's time function. First create two columns in your Excel sheet that list your start times and end times. Then after importing the data into ArcGIS, right click the layer, select Properties, then select the Time tab. Under the Time tab, in the Time properties field, there is an option to change layer time. By default this field will be set to "Each feature has a single time field," but if you click the dropdown arrow you can select "Each feature has a start and end time field." After selecting this option you will be prompted to import you start and time data. Simply designate the two columns you created in Excel, and your data will now be displayed over varying periods.





[iv] It is important to note that because of the nature of the nested If statements, the functions must be written in one direction, from smallest to largest. If written in the other direction, or if an If statement is disordered, then the date could be misclassified into a later period.

[v] If your data has not been sorted into periods, i.e. has specific years that represent ranges in time, then you can perform this if statement with > or < operators (much in the same way as the period function example I provided earlier). Also, be sure to keep track of your number of If expressions and assign the appropriate number of closing parentheses. For me it was easier to first write the expression in Microsoft Word, and then copy it into the formula box in Excel.

[vi] Instead of doing this manually, a more elegant way of creating these spreadsheets would be to write an equation in each sheet that will become an individual .csv file that searches for a given year within the x_All sheet, and copies all of the corresponding data. This could easily be accomplished with a simple vlookup that reports multiple columns for each term (in this case a year) that it identifies within the table array. If the spreadsheets were created in this way, they would be more easily updatable; all you would have to do is input your new data on the first sheet and the sheets for individual periods would automatically update.

[vii] While opening the .csv files in ArcGIS using the XY to Line tool, you may encounter an error that will prevent you from importing some of your data files. This is due to an internal bug in the XY to Line tool function which sometimes prevents it from properly validating the .csv files. If this error occurs, red X's will appear to the left of your file destination, line type, and spatial reference, and you will be unable to choose any categories in the four XY coordinate input fields. In order to fix this problem reopen one of your .csv files in excel. At the bottom left of the page, click on a tab and open a blank spreadsheet (do not create a new workbook). Copy any of your datasets into this spreadsheet and save it as a new .csv file. Then recreate your spreadsheet within this file. By saving a new .csv file under a separate tab of a file that was already designated as a .csv, the validation error in ArcGIS will be resolved. Now you can return to the XY to Line tool, and you should be able to input all of your .csv files.