**PROCESSING DOCUMENTATION**

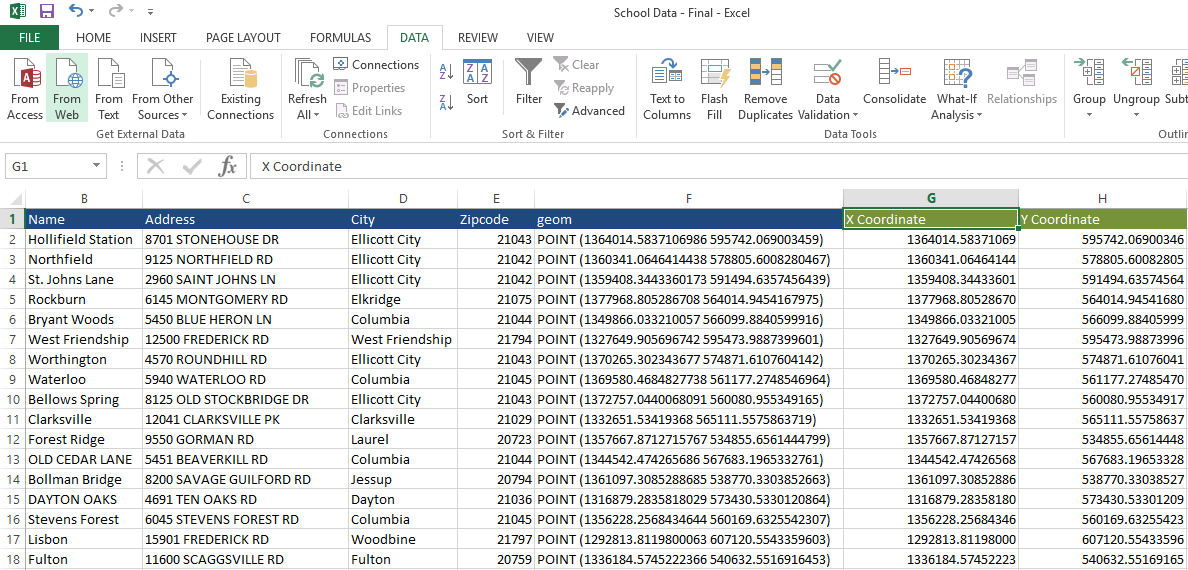
I have combined the two data sets, Schools-Elementary (Focal Data Set) and Sign Stop. The basis on which I was able to join the above two was using the column, ‘geom’ which is present in both the above data sets. The main motive to join these tables was to know exactly how many Stop Signs are present around a School. To perform this analysis, I have assumed the values in the ‘geom’ column to be X and Y coordinates of a School and Stop Sign.

The following are the steps to be followed on Excel while combing the two data sets.

1. For easy readability, the values of the column ‘geom’ for both the tables were separated based on its coordinate values to form two new columns, ‘X Coordinate’ and ‘Y Coordinate’.

Steps:

* Select the Column ‘geom’ by clicking on the particular Column number which would be an Alphabet just above the column name.
* Click on the Data Tab-> Text to Columns->Delimited and click Next ->Check only the Space Textbox and uncheck all the others and click Next->General and click Next->Finish.
* Rename the column values as X-Coordinate and Y-Coordinate
* Next, in order to remove the brackets from the coordinate values. Select the Y Coordinate column and shift it one column to the right, leaving an empty Column in between X and Y Coordinate.
* Select the X Coordinate Column(In the same way as Step 1)->Click on Data Tab-> Text to Columns->Delimited and click Next->Select Other and type ’(’ in the Text box beside it and Click Finish.
* Perform the Similar Step for removing ‘)’ from Y Coordinate Column.
* The Document would then look as below



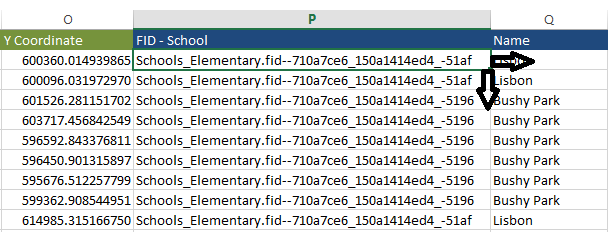
1. Perform all the above steps to separate the coordinate values of the data above Stop Signs as well.
2. Open the Above two Excel sheets on a common Workbook for the join process.
3. On the same workbook, create a new excel sheet and copy the Sign Stop data into it.
4. Click on Cell P2, and paste the Array Formula: “**=INDEX(A2:A26,MATCH(MIN(ABS(A2:A26-C1)),ABS(A2:A26-C1),0))**” on the formula bar. This formula will have to be modified as per the column and table names under consideration.

The Exact formula for Cell P2 would be the following:

**“=INDEX(Schools\_Elementary!A$2:A$42,MATCH(MIN(ABS('Joined Table\_With Formula'!$N2-Schools\_Elementary!$G$2:$G$42)+ABS('Joined Table\_With Formula'!$O2-Schools\_Elementary!$H$2:$H$42)),(ABS('Joined Table\_With Formula'!$N2-Schools\_Elementary!$G$2:$G$42)+ABS('Joined Table\_With Formula'!$O2-Schools\_0Elementary! $H$2:$H$42)),0))”**

Where Schools\_Elementary is my table Schools-Elementary sheet and Joined Table\_With Formula is my present sheet.

1. The values for the rest of the rows and columns can be found by dragging along the row or dragging along the column from Cell P2.

****

1. The above process gives us a combined data set which all the columns.
2. The combination of the two data sets was done based on the distance between the coordinate values using the distance formula. (X2-X1)2-(Y2-Y1)2 which could be approximated as taking the absolute values of the difference i.e. X2-X1 and Y2-Y1.
3. Another Column that could be added to another sheet with School name and another column could be ‘Count’ which calculates the number of stop signs associated with a particular School.

* Select Cell B2 and paste “=COUNTIF('Joined Table\_School data first'!$B$2:$B$5629,'No of Stop Signs'!$A2)” on the formula bar.
* Drag along the column to get the rest of the values as shown below.

