**Instructions for Clustering**

1. Download data from [Maryland Election Data](https://elections.maryland.gov/elections/2016/index.html).
2. Separate the different tables into different worksheets: one for Republican, one for Democrat, one for Unaffiliated, and one for Statewide data.
3. For the Statewide Election Data worksheet, calculate the means and standard deviations for each of the variables (Polls, Early Voting, Absentee, Provisional, Eligible Voters, Turnout).
   1. Calculate the z values for each of the variables using the STANDARDIZE function.
   2. Create a table using the Statewide data downloaded from the website and the z values just calculated. Name it “statewide.”
   3. Create another table with the columns “anchor,” “county #,” “county name,” and the z values for each of the variables. Pull the county name and z variables using a VLOOKUP function and the “statewide” table.
   4. Add five columns to the original “statewide” table. Name these “dist 2\_1,” “dist 2\_2,” and so on up to “dist 2\_5.”
   5. Use the SUMXMY2 function to pull values for each of the five “dist 2\_x” columns.
   6. Add another column and call it “min dist2” (minimum distance squared). Input numbers in this column by using the MIN function, with the values coming from one row of the five “dist 2\_x” columns.
   7. Create a new column at the end called “anchor” and use the MATCH function to see which anchor it corresponds with.
   8. Calculate the sum of the minimum distance squared by using the SUM function for all the values in the “min dist2” column.
4. Run a cluster for the Statewide Election Data worksheet by using the Solver add-in.
   1. Make sure the solving method is “Evolutionary.”
5. Repeat steps 3-4 for the Republican, Democrat, and Unaffiliated worksheets. When naming tables, name them according to the name of each worksheet.
6. Interpret the new cluster data and z values for each anchor.
7. Create a new worksheet that compiles all the raw data for all variables across the four categories (Statewide, Democrat, Republican, and Unaffiliated). Categorize by color and then reorder the data/group them into variables (ex: compile all “Polls” data together, etc.).
8. Create individual clustered column graphs as a data visualization. There should be a total of 6 graphs, one per variable.