

Excel Step by Step

- 1) We used the Baltimore Neighborhood Indicator Alliance open data
- 2) Did 2 single linear regression analyses for 2 independent variables: free/reduced meals and median household income
 - a) Used 1 different dependent variables, one for each time (life expectancy)
 - b) Created formula for regressions using coefficients from single linear regression data
- 3) Created data visuals in scatterplot form for the single linear regression analyses
 - a) 2 scatterplots total: life expectancy vs. free/reduced meals, life expectancy vs. median household income
- 4) Imported data on poverty rate and life expectancy rate for both Baltimore and Chicago using VLOOKUP
- 5) Created simple regression summary for the poverty rate and life expectancy rate in both Baltimore and Chicago by utilizing data analysis tab
- 6) Did a cluster analysis with 3 anchors.
- 7) Copied and pasted data for the 3 different variables: free/reduced meals, median household income, and life expectancy
- 8) Numbered the Baltimore neighborhoods for easy use in future analysis.
- 9) Found the mean and standard deviation for all the data points of each of my 5 variables using the average() and stdev().
- 10) Used the Standardize() function to find the z scores for all the variables.
- 11) Made 3 tentative anchors, with associated z scores for each variable. Used VLOOKUP to match the associated z score with the anchor department.
- 12) Used the Sumxmy2() function to calculate the squared distance between the different anchors and the department/center.
- 13) Found the minimum squared distance between the anchors and the department/center.
- 14) Used the Match() function to find between which anchor the most minimum squared distance was.
- 15) Found the sum of the minimum squared distances using sum().
- 16) Used Solver to find the 3 anchors that would give the minimum sum of the minimum squared distances.
 - a. Set objective cell to the sum of the minimum squared distances.
 - b. Checked min.
 - c. Changed variable cells to the department/center number.
 - d. Added 3 constraints: must be an integer, greater than 1, and less than number of data points (43). Selected Evolutionary as "solving method."

- 17) Used scatter plot function for the data visualization for three anchors. Added axes and chart titles.
- 18) Used countif() to find the number of neighborhoods in clusters with highest mortality rate and lowest life expectancy.