POLS/COMS228 – Data Visualisation

HW 4

You will submit your homework as 2 files: 1) a R Markdown (.Rmd) file and 3) an HTML file knitted from your R Markdown file. You must upload both files to AKO|LEARN using Xxx.Rmd and Xxx.html as file names where Xxx is your last name.

I will review both your files. To receive full credit:

- You must submit your files on time. They must be named exactly as specified, and your .Rmd file must knit without errors to produce a .html file. Double check your file for errors before submitting.
- The .html file should read as a well written report with plots and your comments and answers included in the file. Your report should not contain any extra or foreign material, such as leftovers from code or notes.
- The code in your .Rmd files must be clear, readable, and well organized.

1. Plot population changes by ethnic groups as maps

Brian is working with census data and is asked to present to a general audience how the population of regions in New Zealand changes by ethnic groups. He asks you to help him with producing maps for the presentation.

You obtain a shapefile from Stat NZ with the data for population changes by ethnic groups (the zip file supplied with the Problem Set). Upon familiarizing yourself with the dataset and after consulting with Brian, you decide that you will plot the percent change for the ethnic groups from 2018 to 2023 (variables VAR_1_33 through VAR_1_38, see the CSV file included in the zip file to identify which variable corresponds to which group. CSV file is a codebook for all the variables included in the shapefile).

In total, you decide you will plot all these 6 variables and make one data visual with 6 small multiples (1 map per variable). Using either tmap or ggplot2 libraries, produce

1. 6 individual maps plotting changes in the population by ethnic groups.
2. One small multiples data visual containing all 6 maps.

Consider creating one or two common scale/legends for all your graphs. Discuss advantages and limitations of this approach (you should plot histograms to check the variable distributions. What do you notice about the range of values in the variables?

Format your maps professionally (include title, subtitle (if applicable), data source, etc.). Choose an appropriate colour palette for the type of data you are plotting.