

## **Week Five – Assignment**

1. Load and merge the population and education datasets.
2. Show the distribution of age across regions using boxplot. What are your observations?
3. Adjust the x-axis labels by rotating the text to a 90-degree angle and enhance the visual clarity by removing both the horizontal and vertical gridlines within the chart using theme ().
4. Now let's make use of what we have learnt in the past weeks. In one chunk of code,
  - a. Filter age greater or equal to 3 (convert age to numeric).
  - b. Mutate the educational level into Basic (Primary, JSS/JHS, Middle), Secondary (Secondary, SSS/SHS, Vocational/technical/commercial), Post-Secondary (Post middle/secondary Diploma, Post middle/secondary Certificate), tertiary (HND, Bachelor's degree, Master's, Tertiary - Post graduate Certificate/Diploma and PhD), other and assign the rest to no education.

Try to transform the educational level into a factor and arrange it in the appropriate order, ranging from No education to the highest level of education.

5. Using the transformed data, calculate the percentage distribution of individuals with different levels of education in the Savannah region. visualize the findings using an appropriate graph.