**BASIC STEPS FOR DATA VISUALIZATION**

Objectives: Create figures summarizing the population, births, and deaths, using SRS data. Display cause-specific mortality fraction figures for neonates, children 1-59 months, children 5-14 years, adults 15-49 years, and adults 50 years or older, stratified by sex.

Required: R Studio and R packages: *ggplot2, lubridate, plyr, tidyverse, ggpubr, grid, gridExtra, cowplot, RColorBrewer, ggpubr*

**GENERAL STEPS:**

1. Prior to visualizing data, demographic indicators will need to be calculated
	1. For a population pyramid summary figure, use weighted population by age category and sex, and separately by province
	2. For birth summary figures, use weighted births by province and sex, and by province and place of birth
	3. For death summary figures, use weighted deaths by province and sex, by province and age, and by province and place of death
	4. Cause-specific mortality fraction figures require cause by assignment method (InterVA5, InSilicoVA, or EAVA) and by age and sex
2. Consider consistency across figures. For example, province names, color palettes and the order of categorical variables such as causes of death.
	1. Colors can be distinguished using Hex Codes such as #0000FF for the color Blue which could be assigned to a specific cause of death across different age groups
3. Generate source subtitles to identify data sources and the date figures were last updated. Note variables which require manual update (eg. current year)

**WEIGHTED DEMOGRAPHIC SUMMARY GRAPHS**

STEPS FOR POPULATION PYRAMID:

1. Read in data which has weighted population counts by age category and sex
2. Make proportions for males and females by stratifiers (eg. age categories)
3. Decide what to display in graph and ensure those elements have been created as objects which can be called in ggplot (population sums, labels, titles, etc.)
4. Graph using *ggplot*
5. Check plot
6. Save ggplot image as downloadable image

STEPS FOR BAR GRAPH:

1. Read in data stratified by province
2. Decide what elements to display in graph and render subtitles, color pallets, etc.
3. Graph using *ggplot*
4. Check plot
5. Save ggplot image as downloadable image

**CAUSE-SPECIFIC MORTALITY FRACTION GRAPHS**

STEPS FOR STACKED BAR CHART:

1. It is helpful to create R functions, for repeated actions, such as creating graphs which will have the same layout and cohesive titles/color schemes. Here are some suggested functions:
	1. Data cleaning function to round mortality to nearest percentage point and to order causes of death.
	2. Graphs by age group, and for each age group male/female stratification by cause: 0 to 28 days, 1 to 59 months, 5 to 14 years, 15 to 49 years, and 50+.
2. NOTE: it is possible to combine multiple plots from *ggplot* into one image, creating a function, using *grid.arrange* from the gridExtra package
3. Graph using *ggplot*
4. Save graph image for each age group as downloadable image

EXAMPLE OF CAUSE-SPECIFIC MORTALITY FRACTION GRAPH: Neonates

1. Load datasets with neonate data using *read.csv*
2. Undergo data cleaning
	1. Add sample sizes and convert NA sample sizes to 0; make data frame names and column names are generic if planning to use functions; reformat data as needed
3. Filter needed data by sex. Create variables to graph using generic names which can be used across age groups (example: total neonates and neonates by sex)
4. Start building your graphs
5. Add subtitles, check color scheme, and labels and refine graph aesthetics
6. Combine different graphs created (example: total neonates and neonates by sex) into one graph
7. Allow the graph to be saved as an image