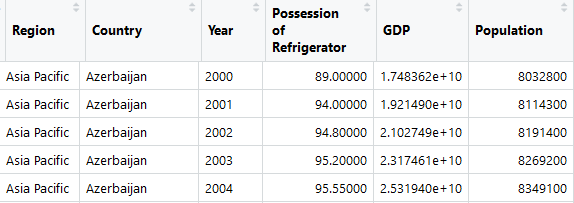
Use comments to explain your code.

1. *ice.cream.data.csv* contains Ice Cream sales (volume and value sold) for 80 countries for years 2000-2019. Using this data, complete the following tasks:
2. For our further analysis, we are only interested in Retail Volume in tonnes and Retail Value in US dollars. Subset the data accordingly.
3. Remove unnecessary columns. Columns to keep: Region, Country, Data Type, Unit, Unit Multiplier and years from 2000 to 2019
4. Unit Multiplier is the number corresponding to Unit. If Unit is millions then it is 1,000,000, if Thousands (000) - 1,000. Convert volume data to kilograms and value data to US dollars. Then remove Unit and Unit Multiplier columns.
5. Convert data to the analysis friendly format shown in the picture below:



1. Rename variables to “Ice Cream Value” and “Ice Cream Volume”
2. As an additional column, calculate Ice Cream Price.
3. For each year, calculate total volume by Region. Which is the second largest region by ice cream volumes in 2018?
4. Try writing a function, which applies step I.7. Make it as general as possible. Using this function calculate:
5. What is the third largest Region by Ice Cream Value?
6. What is the tenth largest country by Ice Cream Volume?
7. Which country had the second largest Ice Cream Price in 2009?
8. *macro.data.csv* contains GDP, Population and Possession of Refrigerator data for 80 countries from 1977 to 2030. Using this data, complete the following tasks:
9. Using similar steps as in Ice cream data convert data in to this format:



1. Combine data:
2. Join both data frames and do all the following tasks on this joint data
3. Subset data to years, which are present in Ice Cream Data. Is this step necessary? Is there something you could have done in previous steps, so that you would not have to do this?
4. Calculate per capita variables where appropriate.
5. Visualize the relationship between Ice Cream Volume per Capita and GDP per Capita. Write your visualization to “Output” folder.
6. Write a function, which calculates percent of missing data for each Country and each Indicator. The result of a function should be easy to read table. Write this table to “Output” folder.