1. Dataset was downloaded from FAO data dissemination page from year 2000 to 2022
2. Load library in R
3. Import data set
4. Data cleaning – Remove unneeded columns
5. Removed aggregations – I removed Saint Kitts and Nevis because only one year data is included, likewise Africa- First create a list of the strings using paste(c()) then remove them using filter(!grgl(list, column))
6. Data exploration – str(), summary(), colnames(), dim()
7. Find unique values in columns unique()
8. Find min and max percentages
9. Check for duplicates using sum(duplicated())
10. Drop duplicates using unique(dataframe) or distinct(dataframe) and store in new df
11. Segment data