DATA VIZ - HOMEWORK IV

NATHANIEL ASIEDU SAKYI

2023-02-27

Loading Complete Data Set into R

```
## [1] "data.frame"
```

```
## [1] 259146     67
```

```
##  [1] "iso_code"
##  [2] "continent"
##  [3] "location"
##  [4] "date"
##  [5] "total_cases"
##  [6] "new_cases"
##  [7] "new_cases_smoothed"
##  [8] "total_deaths"
##  [9] "new_deaths"
## [10] "new_deaths_smoothed"
## [11] "total_cases_per_million"
## [12] "new_cases_per_million"
## [13] "new_cases_smoothed_per_million"
## [14] "total_deaths_per_million"
## [15] "new_deaths_per_million"
## [16] "new_deaths_smoothed_per_million"
## [17] "reproduction_rate"
## [18] "icu_patients"
## [19] "icu_patients_per_million"
## [20] "hosp_patients"
## [21] "hosp_patients_per_million"
## [22] "weekly_icu_admissions"
## [23] "weekly_icu_admissions_per_million"
## [24] "weekly_hosp_admissions"
## [25] "weekly_hosp_admissions_per_million"
## [26] "total_tests"
## [27] "new_tests"
## [28] "total_tests_per_thousand"
## [29] "new_tests_per_thousand"
## [30] "new_tests_smoothed"
## [31] "new_tests_smoothed_per_thousand"
## [32] "positive_rate"
## [33] "tests_per_case"
## [34] "tests_units"
```

```
## [35] "total_vaccinations"
## [36] "people_vaccinated"
## [37] "people_fully_vaccinated"
## [38] "total_boosters"
## [39] "new_vaccinations"
## [40] "new_vaccinations_smoothed"
## [41] "total_vaccinations_per_hundred"
## [42] "people_vaccinated_per_hundred"
## [43] "people_fully_vaccinated_per_hundred"
## [44] "total_boosters_per_hundred"
## [45] "new_vaccinations_smoothed_per_million"
## [46] "new_people_vaccinated_smoothed"
## [47] "new_people_vaccinated_smoothed_per_hundred"
## [48] "stringency_index"
## [49] "population_density"
## [50] "median_age"
## [51] "aged_65_older"
## [52] "aged_70_older"
## [53] "gdp_per_capita"
## [54] "extreme_poverty"
## [55] "cardiovasc_death_rate"
## [56] "diabetes_prevalence"
## [57] "female_smokers"
## [58] "male_smokers"
## [59] "handwashing_facilities"
## [60] "hospital_beds_per_thousand"
## [61] "life_expectancy"
## [62] "human_development_index"
## [63] "population"
## [64] "excess_mortality_cumulative_absolute"
## [65] "excess_mortality_cumulative"
## [66] "excess_mortality"
## [67] "excess_mortality_cumulative_per_million"

## [1] "Asia"          ""              "Europe"        "Africa"
## [5] "North America" "South America" "Oceania"
```

Comments: The data set is a dataframe of 259146 observations and 67 variables collected on several nations falling in one of 7 continents.

Subsetting the South American Continent

```
## 'data.frame':    14106 obs. of  6 variables:
##  $ iso_code          : chr  "ARG" "ARG" "ARG" "ARG" ...
##  $ date              : POSIXct, format: "2020-01-01" "2020-01-02" ...
##  $ hosp_patients     : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ total_vaccinations: num  0 0 0 0 0 0 0 0 0 0 ...
##  $ month             : Ord.factor w/ 12 levels "Jan"<"Feb"<"Mar"<..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ year              : Factor w/ 4 levels "2020","2021",..: 1 1 1 1 1 1 1 1 1 1 ...

##  [1] "ARG" "BOL" "BRA" "CHL" "COL" "ECU" "FLK" "GUY" "PRY" "PER" "SUR" "URY"
## [13] "VEN"
```

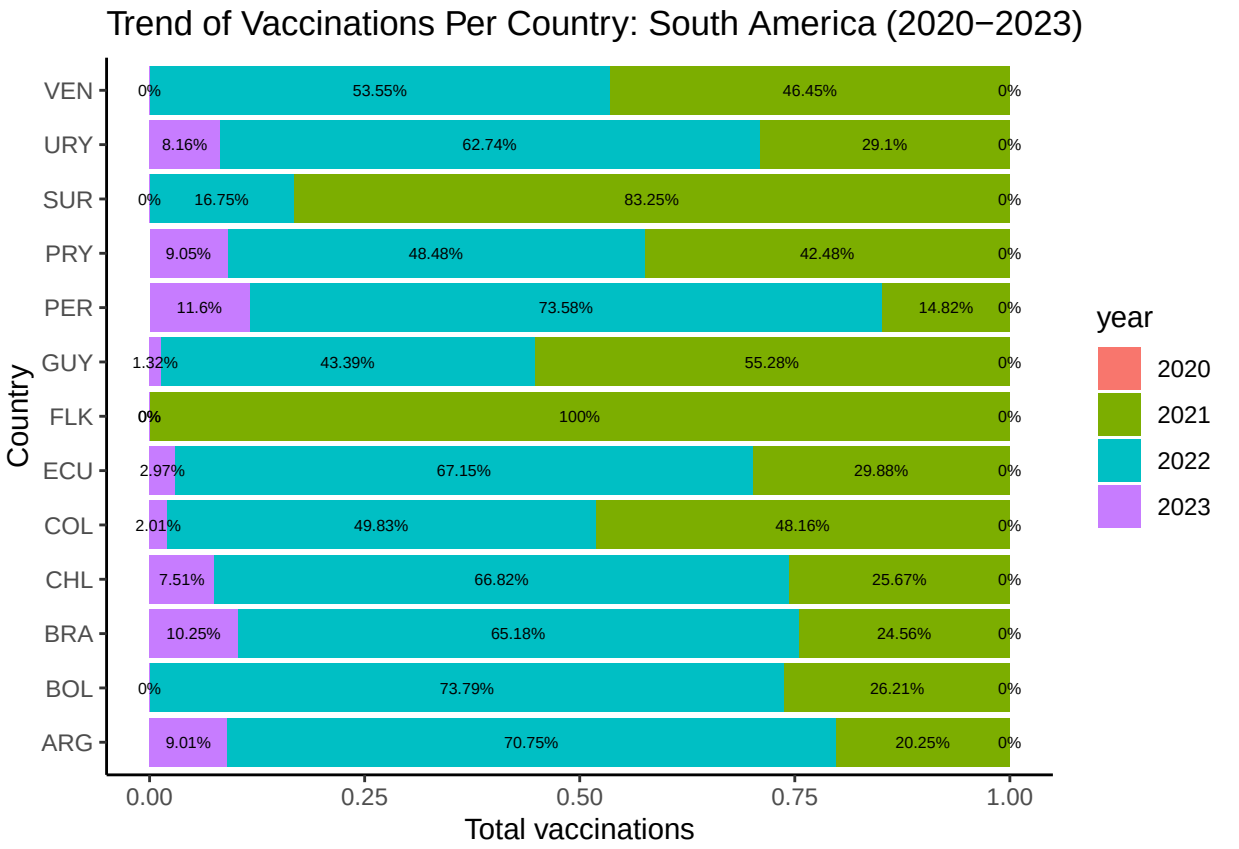
Comment: For the South American continent, there were a total of 13 continents, named as can be seen above.

Fetching The Total Vaccinations of Respective Nations

```
## # A tibble: 15 x 4
## # Groups:   iso_code [4]
##    iso_code year  total_vaccinations labs
##    <chr>    <fct>              <dbl> <chr>
##  1 ARG      2020              104484 0%
##  2 ARG      2021         10755209034 20.25%
##  3 ARG      2022         37581832225 70.75%
##  4 ARG      2023          4785496718 9.01%
##  5 BOL      2020                   0 0%
##  6 BOL      2021          1330732443 26.21%
##  7 BOL      2022          3746377098 73.79%
##  8 BOL      2023                   0 0%
##  9 BRA      2020                   0 0%
## 10 BRA      2021         47441604401 24.56%
## 11 BRA      2022        125910962899 65.18%
## 12 BRA      2023         19807550399 10.25%
## 13 CHL      2020               22606 0%
## 14 CHL      2021          7926071394 25.67%
## 15 CHL      2022         20633039238 66.82%
```

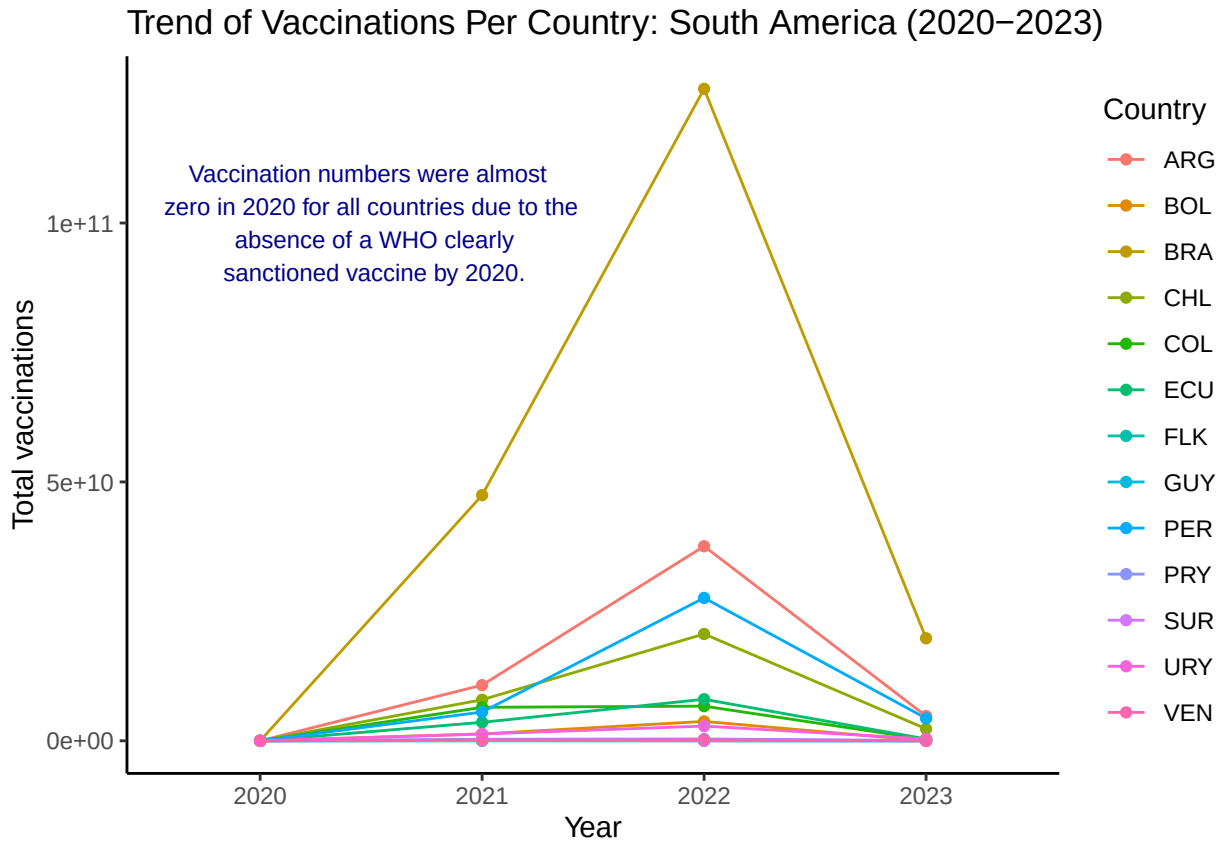
Data Visualization of Nations Total Vaccinations Across Time

Cookbook Attempt: Bar Plot of Total Vaccinations



Drawbacks of Barplot: The barplot above illustrates the total vaccinations per country from 2020 to 2023. As it can be seen, although the visualization is clear, the information of the vaccinations trend across time does not come out so obviously. Much effort is required to make comparisons for a given year across countries; and this undermines Tufte's 2nd Design Principle of Maximizing Every Graphical Element. Secondly, although it tries to apply Tufte's Principle of Utilizing Layering, the plot area seems to have an overly fancy appeal to color, and that makes it a bit clumsy.

Improved Final Attempt: Line Plot of Total Vaccinations



Remarks: Similarly, the time series plot is the preferred plot for revealing the information that must be conveyed in our Data Visualization. This is because, besides making trend comparisons of total vaccinations both across countries and within a country easy to do, it raises a genuine interest to further investigate the clear fluctuations in the trend for any country across time. It Secondly, it is preferred because it applies Tufte's 1st Design Principle of Maximizing Data-to-ink ratio. There is not too much ink used on the background of the visualization which makes it simple and easy to grasp. Finally, it does well to apply Tufte's 5th Design Principle as it Provides the User with Details on Demand.

Source

https://ourworldindata.org/coronavirus