

# ASSIGNMENT REPORT

## DISCREATE STRUCTURE IN COMPUTER SCIENCE

*L<sup>A</sup>T<sub>E</sub>X-Beamer* on April 24, 2022

Nhóm 6 - L03  
Faculty of Computer Science and Engineering  
University of Technology - VNUHCM

DISCREATE  
STRUCTURE IN  
COMPUTER SCIENCE

Nhóm 6 - L03



Danh sách thành viên

- i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- iii. Nhóm câu hỏi liên quan thu thập dữ liệu
- iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu
- v. Nhóm câu hỏi về trực quan dữ liệu theo tháng
- vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

- ➊ Danh sách thành viên
- ➋ i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- ➌ ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- ➍ iii. Nhóm câu hỏi liên quan thu thập dữ liệu
- ➎ iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu
- ➏ v. Nhóm câu hỏi về trực quan dữ liệu theo tháng
- ➐ vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- ➑ vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng
- ➒ viii. Nhóm câu hỏi về tất cả quốc gia theo trung bình 7 ngày gần nhất
- ➓ ix. Nhóm câu hỏi về tương quan nhiễm bệnh và tử vong
- ➑ x. Nhóm câu hỏi riêng

# Danh sách thành viên



## Thành viên nhóm 6 - L03

- ① Nguyễn Đức Bình - 2112899
- ② Trần Nguyễn Thái Bình - 2110051
- ③ Vương Trần Khánh - 2111504
- ④ Hoàng Đức Nguyên - 2110393
- ⑤ Nguyễn Phan Hoàng Phúc - 2110457
- ⑥ Nguyễn Hoài Trung - 2110631

### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

### Phần làm chung:

Đầu tiên, ta sẽ tạo dataframe `countries_covid_data` qua hàm `subset` trích từ dataframe `owid_covid_data` bằng cách lọc tên nước qua câu lệnh `subset` và dữ liệu

### 1. Tập mẫu thể hiện thu thập dữ liệu vào các năm nào.

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Từ dataframe `owid_covid_data`, ta dùng hàm `mdy` của thư viện `lubridate` để định dạng cột date trong dataframe.
- Dùng hàm `year` để lấy thông tin về năm.
- Dùng hàm `unique` để lấy ra những năm khác nhau trong dataframe `owid_covid_data`.
- Dùng hàm `cat` để in kết quả ra màn hình.

- Kết quả câu 1:

Tập mẫu thu thập dữ liệu vào các năm 2020, 2021, 2022.



i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



### 2. Số lượng đất nước và định danh của mỗi đất nước (hiển thị 10 đất nước đầu tiên)

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `unique` lấy ra những `iso_code` khác nhau trong tập mẫu và dùng hàm `head` để hiển thị 10 phần tử đầu tiên của `iso_code`, sau đó gán vào `iso_code`.
- Dùng hàm `unique` lấy ra những `location` khác nhau trong tập mẫu và dùng hàm `head` để hiển thị 10 phần tử đầu tiên của `location`, sau đó gán vào `Country`.
- Tạo thêm một dataframe phụ tên là `data_country`, cuối cùng dùng hàm `print` để in kết quả ra màn hình

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

### 2. Số lượng đất nước và định danh của mỗi đất nước (hiển thị 10 đất nước đầu tiên)

- Kết quả câu 2:

iso_code	Country
AFG	Afghanistan
ALB	Albania
DZA	Algeria
AND	Andorra
AGO	Angola
AIA	Anguilla
ATG	Antigua and Barbuda
ARG	Argentina
ARM	Armenia
ABW	Aruba
Count:	225



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 3. Số lượng châu lục trong tập mẫu

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Tạo thêm một dataframe phụ tên là `countries_covid_data` để lọc ra những hàng có nội dung phần châu lục bị trống.
- Dùng hàm `unique` để tìm ra những châu lục khác nhau trong dataframe ban đầu dùng hàm `length` để tìm số lượng châu lục trong tập mẫu:
- Dùng hàm `length` để tìm số lượng châu lục trong tập mẫu:

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

### 3. Số lượng châu lục trong tập mẫu

- Kết quả câu 3:

Continent :	6
Africa:	Châu Phi
Asia:	Châu Á
Europe:	Châu Âu
North America:	Châu Bắc Mỹ
South America:	Châu Nam Mỹ
Oceania:	Châu Đại Dương

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

### 4. Số lượng dữ liệu thể hiện thu thập dữ liệu được trong từng từng châu lục và tổng số

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm **unique** lấy ra những **continent** khác nhau trong tập mẫu và gán vào đối tượng **Continent\_data**.
- Tiếp theo kết hợp vòng lặp **for**, hàm **length** và **sum** để tìm số liệu thu thập được của từng châu lục.
- Dùng hàm **sum** để tính tổng số liệu và hàm gán kết quả cuối cùng vào dataframe mới tên là **data\_continent**.
- Cuối cùng dùng hàm **print** để in kết quả ra màn hình.

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

### 4. Số lượng dữ liệu thể hiện thu thập dữ liệu được trong từng từng châu lục và tổng số

- Kết quả câu 4:

Continent	Observations
Asia	35528
Europe	36375
Africa	38647
North America	24438
South America	9335
Oceania	8993
Total:	153316

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

### 5. Số lượng dữ liệu thể hiện thu thập dữ liệu được trong từng từng đất nước (hiển thị 10 đất nước cuối cùng) và tổng số

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `unique` lấy ra những `iso_code` khác nhau trong tập mẫu và gán vào đối tượng `iso_code.new`.
- Lấy ra 10 phần tử cuối cùng của `iso_code.new` và gán vào `iso_code.new10`.
- Kết hợp vòng lặp `for`, hàm `length` và `subset` để tìm số liệu thu thập được trong 10 nước cuối.
- Dùng hàm `sum` để tính tổng số liệu và hàm gán kết quả cuối cùng vào dataframe mới tên là `data5`.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

### 5. Số lượng dữ liệu thể hiện thu thập dữ liệu được trong từng từng đất nước (hiển thị 10 đất nước cuối cùng) và tổng số

- Kết quả câu 5:

iso_code	Observations
UZB	707
VUT	467
VAT	716
VEN	708
VNM	759
WLF	489
OWID_WRL	760
YEM	681
ZMB	704
ZWE	702
Tổng:	6693



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 6. Cho biết các châu lục nào có lượng dữ liệu thu thập nhỏ nhất và giá trị nhỏ nhất đó?

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `min` và `table` để tìm ra lượng dữ liệu thu thập nhỏ nhất và gán vào `min.continent`.
- Sau đó dùng hàm `subset` để tìm ra các châu lục có lượng dữ liệu bằng `min.continent` và gán vào `min.continent.data`.
- Dùng hàm `cat` để xuất kết quả ra màn hình.

- Kết quả câu 6:

The continent with the least data is Oceania which has 8993 data.

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 7. Cho biết các châu lục nào có lượng dữ liệu thu thập lớn nhất và giá trị lớn nhất đó?

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `max` và `table` để tìm ra lượng dữ liệu thu thập lớn nhất và gán vào `max.continent`.
- Dùng hàm `subset` để tìm ra các châu lục có lượng dữ liệu bằng `max.continent` và gán vào `data7`.
- Dùng hàm `cat` để xuất kết quả ra màn hình.

- Kết quả câu 7:

The continent has the most data is Africa, which has 38647 data.

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

### 8. Cho biết các nước nào có lượng dữ liệu thu thập nhỏ nhất và giá trị nhỏ nhất đó?

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm **min** và **table** để tìm ra lượng dữ liệu thu thập nhỏ nhất và gán vào **min.location**.
- Sau đó dùng hàm **subset** để tìm ra các nước có lượng dữ liệu bằng **min.location** và gán vào **min.location.data**.
- Dùng hàm **cat** để xuất kết quả ra màn hình.

- Kết quả câu 8:

The location with the least data is Pitcairn , which has 85 data.

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

### 9. Cho biết các nước nào có lượng dữ liệu thu thập lớn nhất và giá trị lớn nhất đó?

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `max` và `table` để tìm ra lượng dữ liệu thu thập lớn nhất và gán vào `max.location`.
- Dùng hàm `subset` để tìm ra các nước có lượng dữ liệu bằng `max.location` và gán vào `data9`.
- Dùng hàm `cat` để xuất kết quả ra màn hình.

- Kết quả câu 9:

The locations have the most data are Argentina and Mexico, which have 781 data.

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



### 10. Cho biết các date nào có lượng dữ liệu thu thập nhỏ nhất và giá trị nhỏ nhất đó?

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `min` và `table` để tìm ra lượng dữ liệu thu thập nhỏ nhất và gán vào `min.date`.
- Sau đó dùng hàm `subset` để tìm ra các ngày có lượng dữ liệu bằng `min.date` và gán vào `min.date.data`.
- Dùng hàm `cat` để xuất kết quả ra màn hình.

- Kết quả câu 10:

The dates with the least data are 2020-01-01 2020-01-02  
2020-01-03 , which have 2 data.

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



### 11. Cho biết các date nào có lượng dữ liệu thu thập lớn nhất và giá trị lớn nhất đó?

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `max` và `table` để tìm ra lượng dữ liệu thu thập lớn nhất và gán vào `max.date`.
- Dùng hàm `subset` để tìm ra các ngày có lượng dữ liệu bằng `max.date` và gán vào `data11`.
- Dùng hàm `cat` để xuất kết quả ra màn hình.

- Kết quả câu 11:

The dates have the most data are 8/22/2021, 8/23/2021  
8/24/2021, 8/25/2021, 8/26/2021, 8/27/2021, 8/28/2021  
8/29/2021, which have 238 data.

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 12. Cho biết số lượng dữ liệu thu thập được theo date và châu lục.

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `mdy` của thư viện `lubridate` để định dạng cột `date` của dataframe `countries_covid_data` theo định dạng tháng/ngày/năm.
- Dùng hàm `table` để thống kê số lượng dữ liệu thu thập được theo ngày và theo châu lục.
- Lưu kết quả vào `data12` dưới dạng dataframe để dễ theo dõi và xử lý cho những câu tiếp theo.

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 12. Cho biết số lượng dữ liệu thu thập được theo date và châu lục.

- Kết quả câu 12: lấy dữ liệu 10 dòng đầu tiên

	Africa	Asia	Europe	North America	Oceania	South America
2020-01-01	0	0	0	1	0	1
2020-01-02	0	0	0	1	0	1
2020-01-03	0	0	0	1	0	1
2020-01-04	0	1	0	1	0	1
2020-01-05	0	1	0	1	0	1
2020-01-06	0	1	3	1	0	1
2020-01-07	0	1	0	1	0	1
2020-01-08	0	1	0	1	0	1
2020-01-09	0	1	0	1	0	1
2020-01-10	0	1	0	1	0	1

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

### 13-14. Cho biết số lượng dữ liệu thu thập được là lớn nhất theo date và châu lục

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `max` và `min` để tìm ra dữ liệu lớn nhất (nhỏ nhất) trong `data12` đã tạo ở câu 12, gán giá trị tìm được vào `max.date.continent` và `min.date.continent`.
- Dùng hàm `subset` để tìm ra những ngày và những châu lục có lượng dữ liệu bằng với `max.date.continent` và `min.date.continent`, lưu kết quả vào `res13` và `res14`.
- Dùng hàm `sprintf` để in kết quả ra màn hình.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

### 13-14. Cho biết số lượng dữ liệu thu thập được là lớn nhất theo date và châu lục

- Kết quả câu 13:

Số lượng dữ liệu thu thập được lớn nhất theo ngày và châu lục là 55, ở Africa, từ ngày 2020-09-07 đến ngày 2022-02-19.

- Kết quả câu 14:

Số lượng dữ liệu thu thập được nhỏ nhất theo ngày và châu lục là 0. Các dữ liệu có giá trị bằng 0 được lấy từ các châu lục theo ngày như sau:

+Dữ liệu ở Africa từ 2020-01-01 đến 2020-02-06.

+Dữ liệu ở Asia từ 2020-01-01 đến 2020-01-03.

+Dữ liệu ở Europe từ 2020-01-01 đến 2020-01-05, từ 2020-01-07 đến 2020-01-17, từ 2020-01-19 đến 2020-01-23.

+Dữ liệu ở Oceania từ 2020-01-01 đến 2020-01-25.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

15. Với một date là k và châu lục t cho trước, hãy cho biết số lượng dữ liệu thể hiện thu thập dữ liệu được.

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `readline` để nhập dữ liệu date k và châu lục t.
- Sau đó dùng hàm `subset` để lọc dữ liệu theo date và châu lục đã nhập từ `countries_covid_data` lưu kết quả vào `optional_data`.
- Cuối cùng dùng hàm `print` để in kết quả ra màn hình.

- Kết quả câu 15:

Enter date : 2020-03-25

Enter continent : Europe

The amount of data collected on 2020-03-25 in Europe is 49 data.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

16. Có đất nước nào mà số lượng dữ liệu thu thập được là bằng nhau không? Hãy cho biết các `iso_code` của đất nước đó.

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Ta sẽ tạo một dataframe `Observations_data` gồm số lượng thu thập dữ liệu của các quốc gia, trích từ `countries_covid_data` để lọc những `iso_code` trùng nhau.
- Tiếp theo tạo `duplicated_vector` gồm những số lượng thu thập dữ liệu trùng nhau, sau đó dùng hàm `subset` để lọc những nước có cùng số lượng thu thập dữ liệu.
- Cuối cùng dùng hàm `arrange` để sắp xếp và hàm `print` để in kết quả ra màn hình.

- Kết quả câu 16: Bảng gồm 181 nước.

iso_code	Observations
SPM SSD	686
BDI SLE	691
AIA TCA VGB	694
...	...
ARG MEX	781



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 17. Liệt kê iso\_code, tên đất nước mà chiều dài iso\_code lớn hơn 3.

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng `subset` để dữ liệu mà chiều dài iso\_code lớn hơn 3 từ `countries_covid_data` lưu kết quả vào `iso_code_data`.
- Sau đó dùng hàm `filter` để lọc những iso\_code trùng nhau.
- Cuối cùng dùng hàm `print` để in kết quả ra màn hình.

- Kết quả câu 17:

iso_code	Location
OWID_KOS	Kosovo
OWID_CYN	Northern Cyprus

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### Chuẩn bị dữ liệu

- Đầu tiên, ta xử lý các dữ liệu số âm thành số dương bằng hàm `abs`.
- Tiếp theo, ta sẽ lọc dữ liệu của ba quốc gia nhóm cần thu thập (Brazil, Chile, Venezuela) bằng hàm `subset`.
- Chuyển các số liệu không xác định thành 0 bằng hàm `is.na`.
- Trước khi thực hiện tính toán, ta sẽ tạo các dataframe chứa thông tin ứng với từng nước trong nhóm bằng hàm `subset`.

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

Câu 1, 2, 3, 4, 5, 6: Tính toán và lập bảng mô tả số liệu thống kê cho từng đất nước trong nhóm

- Ta sẽ tạo một bảng số liệu thống kê ứng với mỗi nước thể hiện các giá trị Min, Q1, Q2, Q3, Max, Avg, Std, Outlier bằng các hàm `min`, `quantile`, `max`, `mean`, `sd` và đếm số Outlier.
- Bảng mô tả số liệu thống kê cho từng đất nước thuộc về nhóm là Brazil, Chile và Venezuela:
- Brazil

Status	New_cases	New_deaths
Min	0	0
Q1	13276	317
Q2	31149	721
Q3	53989	1190
Max	287149	4148
Avg	38747.7	888.54
Std	37015.4	771.6
Outlier	26	40



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Câu 1, 2, 3, 4, 5, 6: Tính toán và lập bảng mô tả số liệu thống kê cho từng đất nước trong nhóm

- Chile

Status	New _cases	New _deaths
Min	0	0
Q1	1176.75	17
Q2	1997.5	40
Q3	4315.25	82
Max	41651	1057
Avg	3872.59	57.17
Std	6093.36	67.16
Outlier	38	28

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Câu 1, 2, 3, 4, 5, 6: Tính toán và lập bảng mô tả số liệu thống kê cho từng đất nước trong nhóm

- Venezuela

Status	New _cases	New _deaths
Min	0	0
Q1	245	3
Q2	590	6
Q3	1100.25	12
Max	4418	35
Avg	721.12	7.93
Std	631.37	7.21
Outlier	15	21

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

### Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus

Phương pháp làm:

- Đầu tiên ta tạo một dataframe chứa thông tin của các quốc gia trong nhóm.
- Tiếp theo, ta sẽ sử dụng thư viện [ggplot2](#) để vẽ biểu đồ.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

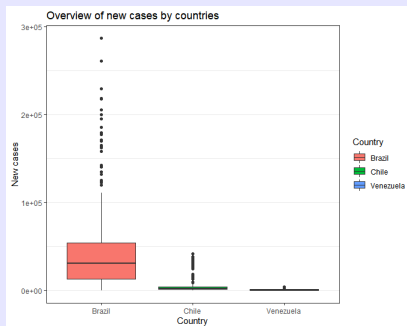
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus

Kết quả:



**Hình:** Biểu đồ boxplot về số ca nhiễm mới của ba quốc gia Brazil, Chile và Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

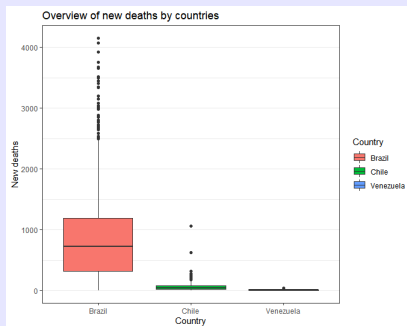
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus

Kết quả:



**Hình:** Biểu đồ boxplot về số ca tử vong mới của ba quốc gia Brazil, Chile và Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

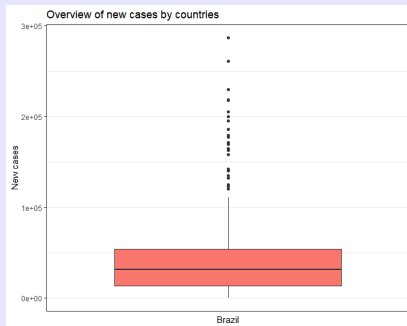
vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus

Kết quả:



**Hình:** Biểu đồ boxplot về số ca nhiễm mới của Brazil



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

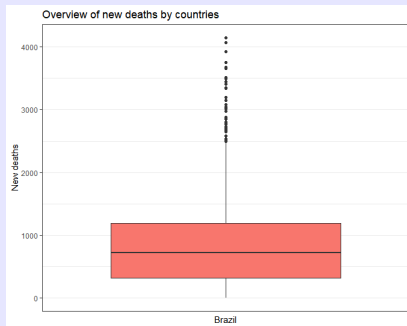
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus

Kết quả:



**Hình:** Biểu đồ boxplot về số ca tử vong mới của Brazil



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

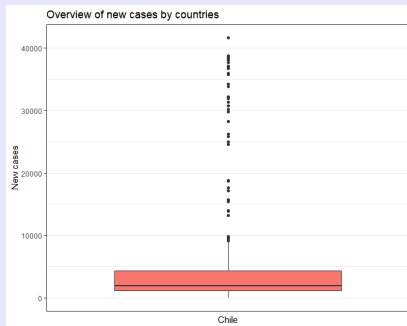
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

**Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus**

**Kết quả:**



**Hình:** Biểu đồ boxplot về số ca nhiễm mới của Chile



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

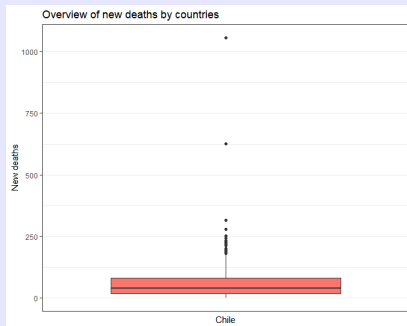
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus

Kết quả:



**Hình:** Biểu đồ boxplot về số ca tử vong mới của Chile



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

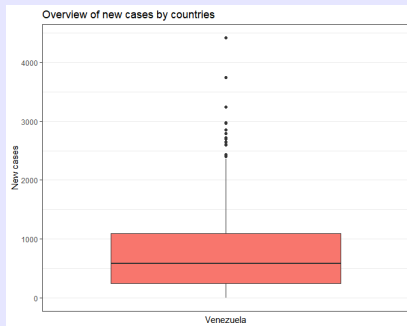
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus

Kết quả:



**Hình:** Biểu đồ boxplot về số ca nhiễm mới của Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

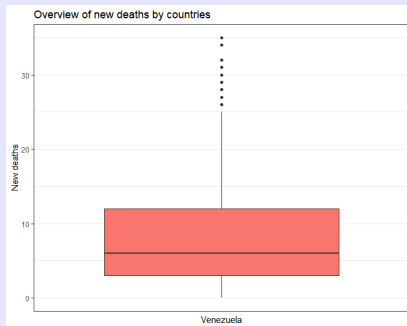
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

Câu 7: Vẽ biểu đồ boxplot hay còn được gọi là box-and-whisker cho nhiễm coronavirus

Kết quả:



**Hình:** Biểu đồ boxplot về số ca tử vong mới của Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### iii. Nhóm câu hỏi liên quan thu thập dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

#### Chuẩn bị dữ liệu

- Với 3 quốc gia được phân công, lọc ra bảng dữ liệu của ba nước thông qua hàm `subset`, sử dụng thêm hàm `group_by`.
- Trích dữ liệu chỉ lấy 3 cột `location`, `new_cases`, `new_deaths` để xử lý.

### iii. Nhóm câu hỏi liên quan thu thập dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

#### 1. Số ngày không được báo cáo mới

- Từ bảng dữ liệu trong phần chuẩn bị, đếm số dòng dữ liệu `new_cases` bằng 0 hoặc NA. Trường hợp `new_deaths` tương tự.
- Kết quả cho `new_cases`:

location	no_report_days
Brazil	9
Chile	14
Venezuela	63

- Kết quả cho `new_deaths`:

location	no_report_days
Brazil	22
Chile	33
Venezuela	129



### iii. Nhóm câu hỏi liên quan thu thập dữ liệu

#### 2-3. Xác định số ngày có số lần thu thập dữ liệu là thấp nhất và cao nhất

Ta trình bày cách làm cho trường hợp `new_cases`. Trường hợp `new_deaths` tương tự.

- Tạo một bảng dữ liệu lưu trữ những dòng mà cột `new_cases` được báo cáo mới.
- Xác định và đếm những dòng mà `new_cases` đạt giá trị nhỏ nhất. Tiến hành tương tự đối với giá trị lớn nhất.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### iii. Nhóm câu hỏi liên quan thu thập dữ liệu

#### Kết quả

Kết quả cho số lần thu thập dữ liệu đạt thấp nhất

- Trường hợp **new\_cases**:

location	min_days
Brazil	2
Chile	1
Venezuela	1

- Trường hợp **new\_deaths**:

location	min_days
Brazil	1
Chile	1
Venezuela	1



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### iii. Nhóm câu hỏi liên quan thu thập dữ liệu

#### Kết quả

Kết quả cho số lần thu thập dữ liệu đạt lớn nhất

- Trường hợp **new\_cases**:

location	max_days
Brazil	1
Chile	1
Venezuela	1

- Trường hợp **new\_deaths**:

location	max_days
Brazil	1
Chile	1
Venezuela	3



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### iii. Nhóm câu hỏi liên quan thu thập dữ liệu

#### 4. Thể hiện bảng số liệu không được báo cáo mới và được báo cáo mới

- Từ bảng dữ liệu trong phần chuẩn bị, đếm các dòng `new_cases` bằng 0 hoặc NA rồi lưu vào Infections, `new_deaths` bằng 0 hoặc NA rồi lưu vào Deaths. Thực hiện tương tự với dữ liệu được báo cáo mới.
- Đổi tên cột đúng với định dạng yêu cầu đề bài.
- Bảng dữ liệu không được báo cáo mới

Countries	Infections	Deaths
Brazil	9	22
Chile	14	33
Venezuela	63	129

- Bảng dữ liệu được báo cáo mới

Countries	Infections	Deaths
Brazil	716	703
Chile	714	695
Venezuela	645	579



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### iii. Nhóm câu hỏi liên quan thu thập dữ liệu

**Tạo hai hàm phụ đếm số ngày ngắn nhất và dài nhất liên tiếp của một điều kiện nào đó**

Ta trình bày ý tưởng tạo hàm đếm số ngày liên tiếp lớn nhất. Hàm đếm số ngày nhỏ nhất làm tương tự

- Tham số đầu vào là một bảng dataframe có cột dữ liệu `date` là chuỗi các ngày.
- Duyệt cột `date` với vòng lặp `for`, nếu có một chuỗi con có số ngày liên tiếp lớn hơn biến `count` lưu độ dài chuỗi dài nhất tới trước thời điểm đang xét, thì gán `count` bằng giá trị lớn hơn đó.
- Hàm trả về giá trị `count`.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### iii. Nhóm câu hỏi liên quan thu thập dữ liệu

#### 5-6. Xác định số ngày ngắn nhất và dài nhất liên tiếp không có dữ liệu được báo cáo mới

- Trích dữ liệu từ `countries_covid_data` những ngày mà dữ liệu không được báo cáo mới (`new_cases` và `new_deaths` mang giá trị 0 hoặc NA)
- Sử dụng hai hàm phụ để tìm số ngày ngắn nhất và lớn nhất của mỗi nước rồi lưu vào một dữ liệu mới.
- Bảng số ngày ngắn nhất liên tiếp không có báo cáo mới

Countries	min_days
Brazil	1
Chile	1
Venezuela	1

- Bảng số ngày dài nhất liên tiếp không có báo cáo mới

Countries	max_days
Brazil	20
Chile	28
Venezuela	14



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### iii. Nhóm câu hỏi liên quan thu thập dữ liệu

#### 7-8. Xác định số ngày ngắn nhất và dài nhất liên tiếp không có người nhiễm bệnh mới

- Trích dữ liệu từ `countries_covid_data` những ngày mà không có người nhiễm bệnh mới (`new_cases = 0`)
- Sử dụng hai hàm phụ để tìm số ngày ngắn nhất và lớn nhất của mỗi nước rồi lưu vào một dữ liệu mới.
- Bảng số ngày ngắn nhất liên tiếp không có người nhiễm bệnh mới

Countries	min_days
Brazil	1
Chile	1
Venezuela	1

- Bảng số ngày dài nhất liên tiếp không có người nhiễm bệnh mới.

Countries	max_days
Brazil	3
Chile	6
Venezuela	1



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 1. Vẽ biểu đồ tần số tích lũy quốc gia cho các châu lục

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm **distinct** để lọc ra những đất nước khác nhau trong dataframe **countries\_covid\_data**, gán vào **unique.location**. Sắp xếp lại dữ liệu trong **unique** bằng hàm **arrange**.
- Vẽ đồ thị: dùng hàm **barplot** để vẽ đồ thị. Vì đề bài yêu cầu tính tần số tích lũy nên ta dùng hàm **cumsum** kết hợp với **table** để tính tần số tích lũy.



## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

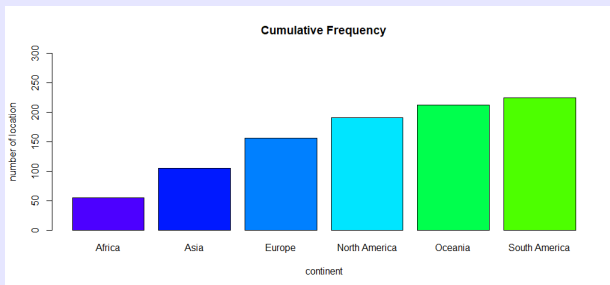
v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 1. Vẽ biểu đồ tần số tích lũy quốc gia cho các châu lục

- Kết quả câu 1:



**Hình:** Biểu đồ tần số tích lũy quốc gia cho các châu lục.

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 2. Vẽ biểu đồ tần số tương đối quốc gia cho các châu lục

- Phương pháp làm: Các bước thực hiện có các ý chính sau:
  - Để tìm tần số tương đối, ta lấy tần số tuyệt đối chia cho tổng của tất cả các tần số trong tập mẫu. Vậy nên ta sẽ tính toán tần số tương đối của các châu lục bằng cách lấy `table(data4$continent)` chia cho `length(data4$continent)`.  
Cách vẽ đồ thị tương tự câu 1.

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

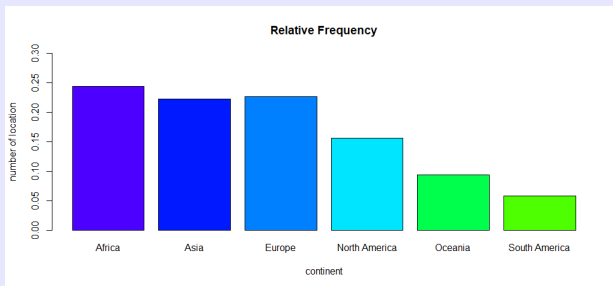
v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 2. Vẽ biểu đồ tần số tương đối quốc gia cho các châu lục

- Kết quả câu 2:



**Hình:** Biểu đồ tần số tương đối quốc gia cho các châu lục.

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 3. Vẽ biểu đồ thể hiện nhiễm bệnh đã báo cáo của các quốc gia mà thuộc về nhóm trong 7 ngày cuối của năm cuối cùng

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Dùng hàm `subset` để lọc ra dữ liệu của 7 ngày cuối của năm cuối cùng.
- Dùng hàm `barplot` để vẽ đồ thị cột thể hiện nhiễm bệnh đã báo cáo của dữ liệu đã lọc ở trên.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

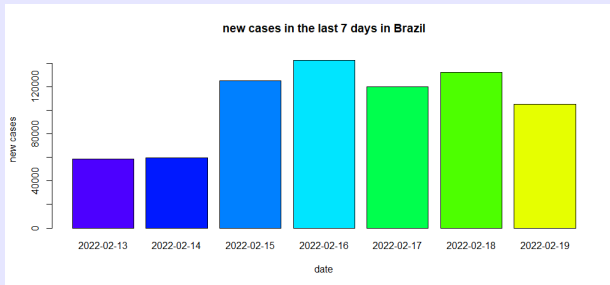
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 3. Vẽ biểu đồ thể hiện nhiễm bệnh đã báo cáo của các quốc gia mà thuộc về nhóm trong 7 ngày cuối của năm cuối cùng

- Kết quả câu 3:



**Hình:** Biểu đồ thể hiện nhiễm bệnh đã báo cáo của Brazil trong 7 ngày cuối cùng của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

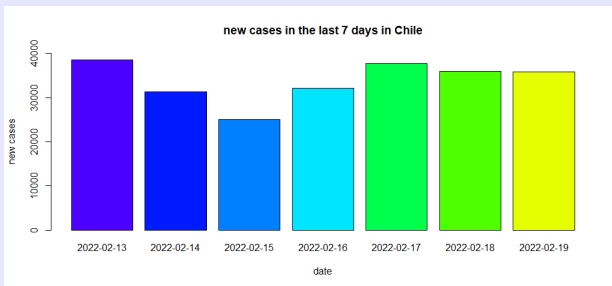
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 3. Vẽ biểu đồ thể hiện nhiễm bệnh đã báo cáo của các quốc gia mà thuộc về nhóm trong 7 ngày cuối của năm cuối cùng

- Kết quả câu 3:



**Hình:** Biểu đồ thể hiện nhiễm bệnh đã báo cáo của Chile trong 7 ngày cuối cùng của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

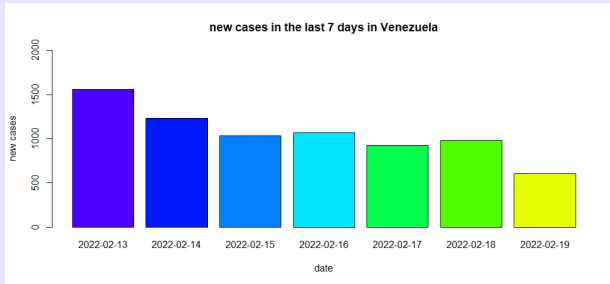
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 3. Vẽ biểu đồ thể hiện nhiễm bệnh đã báo cáo của các quốc gia mà thuộc về nhóm trong 7 ngày cuối của năm cuối cùng

- Kết quả câu 3:



**Hình:** Biểu đồ thể hiện nhiễm bệnh đã báo cáo của Venezuela trong 7 ngày cuối cùng của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 4. Vẽ biểu đồ thể hiện tử vong đã báo cáo của các quốc gia mà thuộc về nhóm trong 7 ngày cuối của năm cuối cùng

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- dùng hàm `subset` để lọc ra dữ liệu của 7 ngày cuối của năm cuối cùng.
- Dùng hàm `barplot` để vẽ đồ thị cột thể hiện tử vong đã báo cáo của dữ liệu đã lọc ở trên.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

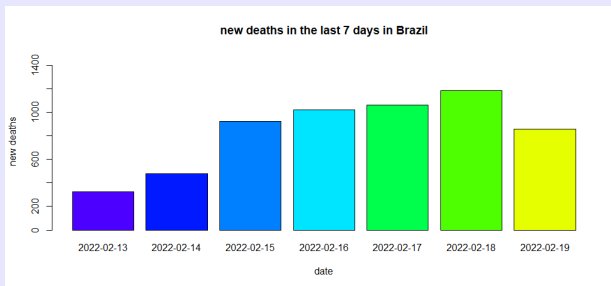
vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 4. Vẽ biểu đồ thể hiện tử vong đã báo cáo của các quốc gia mà thuộc về nhóm trong 7 ngày cuối của năm cuối cùng

- Kết quả câu 4:



**Hình:** Biểu đồ thể hiện tử vong đã báo cáo của Brazil trong 7 ngày cuối cùng của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

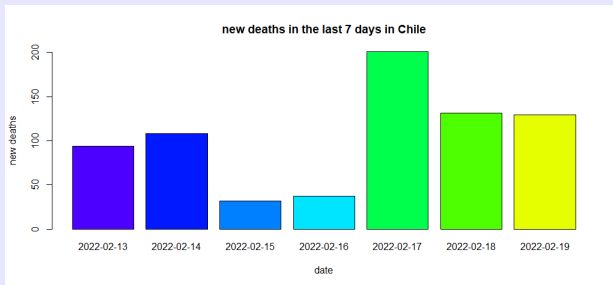
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 4. Vẽ biểu đồ thể hiện tử vong đã báo cáo của các quốc gia mà thuộc về nhóm trong 7 ngày cuối của năm cuối cùng

- Kết quả câu 4:



**Hình:** Biểu đồ thể hiện tử vong đã báo cáo của Chile trong 7 ngày cuối cùng của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

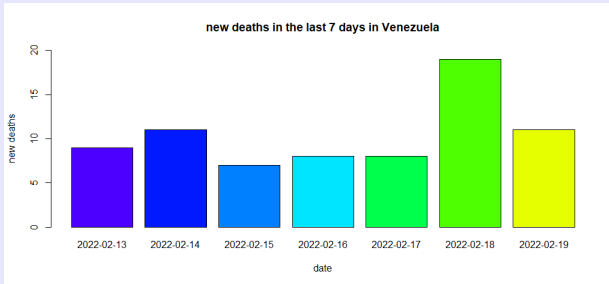
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 4. Vẽ biểu đồ thể hiện tử vong đã báo cáo của các quốc gia mà thuộc về nhóm trong 7 ngày cuối của năm cuối cùng

- Kết quả câu 4:



**Hình:** Biểu đồ thể hiện tử vong đã báo cáo của Venezuela trong 7 ngày cuối cùng của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 5. Vẽ biểu đồ phổ đất nước xuất hiện outliers cho nhiễm bệnh

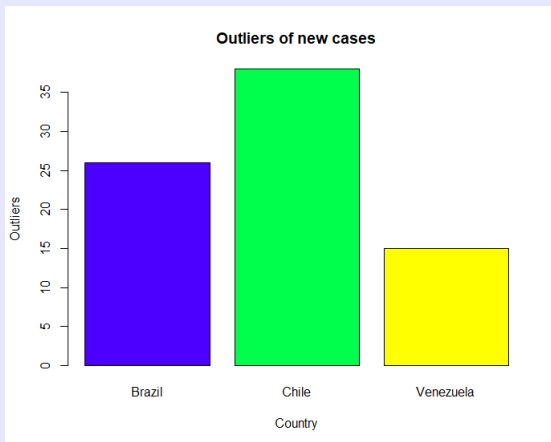
- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Đầu tiên ta dùng hàm `subset` để lọc ra dữ liệu của từng nước từ `countries_covid_group_data` (ở nhóm câu hỏi ii) .
- Tính outliers theo công thức ở nhóm câu hỏi ii.
- Dùng hàm `barplot` để vẽ đồ thị cột thể hiện nhiễm bệnh đã báo cáo của dữ liệu đã lọc ở trên.

#### iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 5. Vẽ biểu đồ phổ đất nước xuất hiện outliers cho nhiễm bệnh

- Kết quả câu 5:



**Hình:** Biểu đồ phổ đất nước xuất hiện outliers cho nhiễm bệnh



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

#### iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

#### 6. Vẽ biểu đồ phổ đất nước xuất hiện outliers cho tử vong

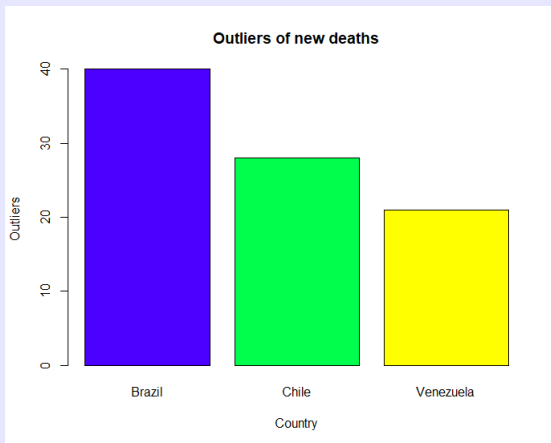
- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Đầu tiên ta dùng hàm `subset` để lọc ra dữ liệu của từng nước từ `countries_covid_group_data` (ở nhóm câu hỏi ii) .
- Tính outliers theo công thức ở nhóm câu hỏi ii.
- Dùng hàm `barplot` để vẽ đồ thị cột thể hiện tử vong đã báo cáo của dữ liệu đã lọc ở trên.

## iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

### 6. Vẽ biểu đồ phổ đất nước xuất hiện outliers cho tử vong

- Kết quả câu 6:



**Hình:** Biểu đồ phổ đất nước xuất hiện outliers cho tử vong



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

viii. Nhóm câu hỏi về...

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### Phương pháp làm câu 1, 2, 3

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

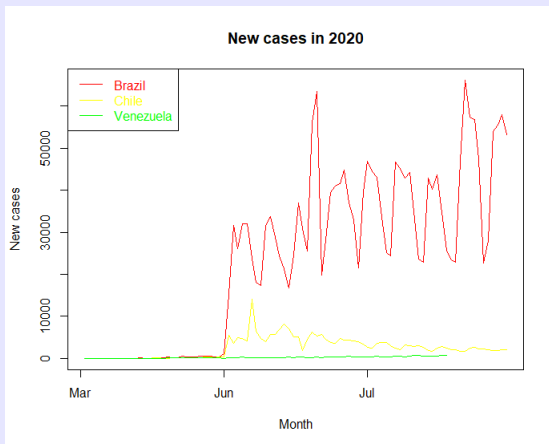
- Đầu tiên, ta sẽ tạo dataframe **brazil** qua hàm **subset** trích từ dataframe **owid\_covid\_data** bằng cách lọc đất nước được giao (ở đây là Brazil).
- Các tháng cần sử dụng (ở đây là tháng 1, 3, 6, 7) sẽ được giữ lại trong dataframe **brazil** bằng lệnh **subset**.
- Sau đó, ta lại chia số liệu ra thành các năm để thuận lợi cho việc vẽ biểu đồ.
- Thực hiện tương tự với 2 nước được giao là Chile và Venezuela.
- Cuối cùng, ta vẽ biểu đồ đường bằng lệnh **plot**



## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh cho từng tháng

- Kết quả câu 1:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh cho từng tháng năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

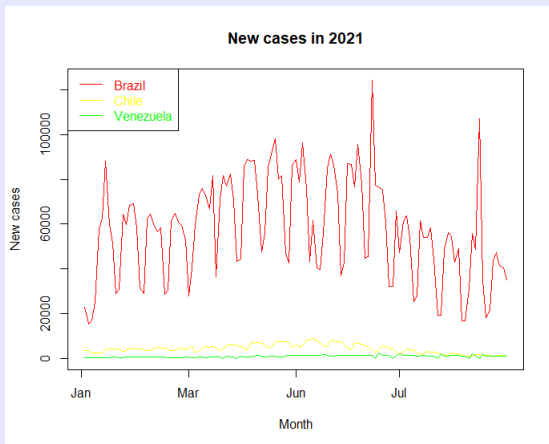
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh cho từng tháng

- Kết quả câu 1:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh cho từng tháng năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

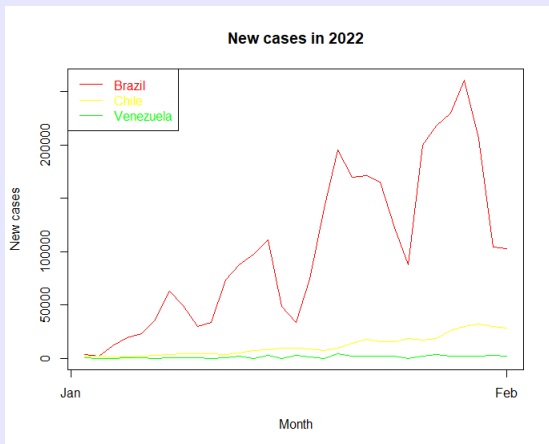
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh cho từng tháng

- Kết quả câu 1:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh cho từng tháng năm 2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu tử vong cho từng tháng

- Kết quả câu 2:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong cho từng tháng năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

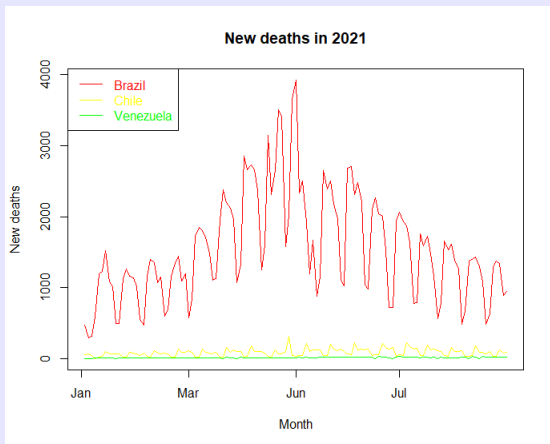
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu tử vong cho từng tháng

- Kết quả câu 2:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong cho từng tháng năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

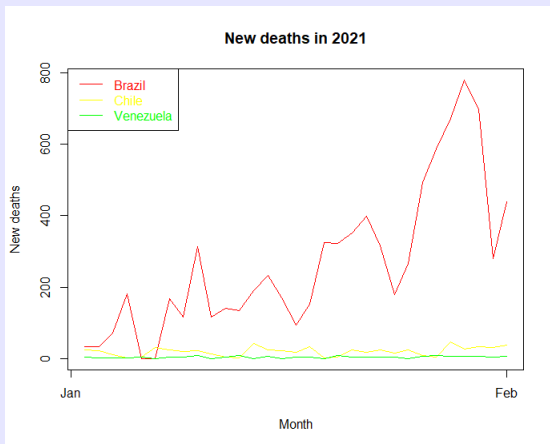
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu tử vong cho từng tháng

- Kết quả câu 2:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong cho từng tháng năm 2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

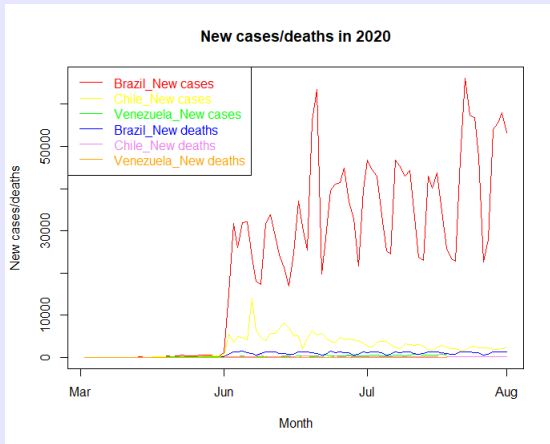
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong cho từng tháng

- Kết quả câu 3:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong cho từng tháng năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

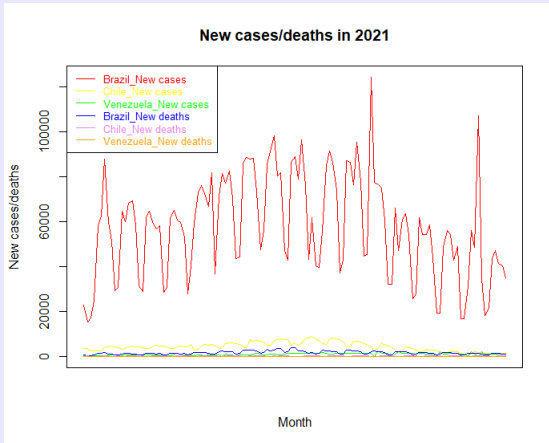
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong cho từng tháng

- Kết quả câu 3:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong cho từng tháng năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

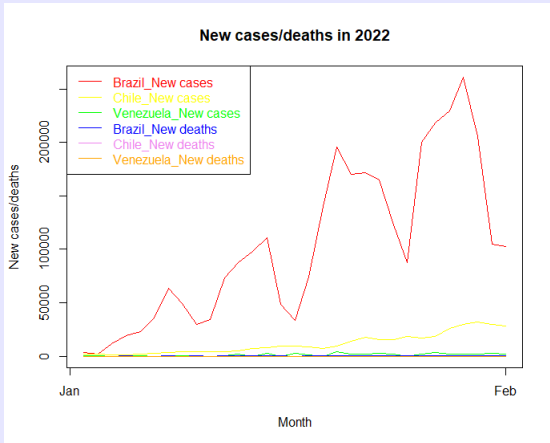
vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong cho từng tháng

- Kết quả câu 3:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong cho từng tháng năm 2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng



### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### Phương pháp làm câu 4, 5, 6

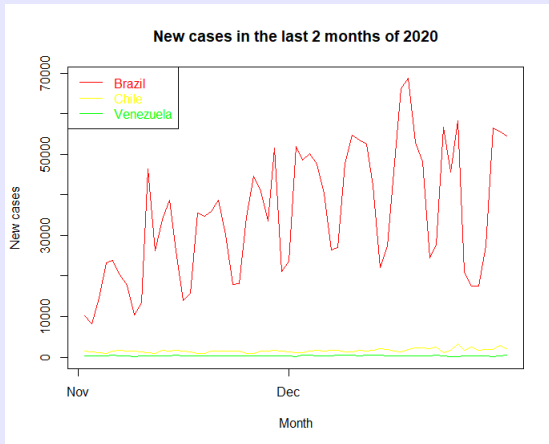
- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Đầu tiên, ta sẽ tạo dataframe **brazil** qua hàm **subset** trích từ dataframe **owid\_covid\_data** bằng cách lọc đất nước được giao (ở đây là Brazil).
- Các tháng cần sử dụng (ở đây là tháng 11, 12) sẽ được giữ lại trong dataframe **brazil** bằng lệnh **subset**.
- Sau đó, ta lại chia số liệu ra thành các năm để thuận lợi cho việc vẽ biểu đồ.
- Thực hiện tương tự với 2 nước được giao là Chile và Venezuela.
- Cuối cùng, ta vẽ biểu đồ đường bằng lệnh **plot**

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh gồm 2 tháng cuối của năm

- Kết quả câu 4:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh gồm 2 tháng cuối của năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

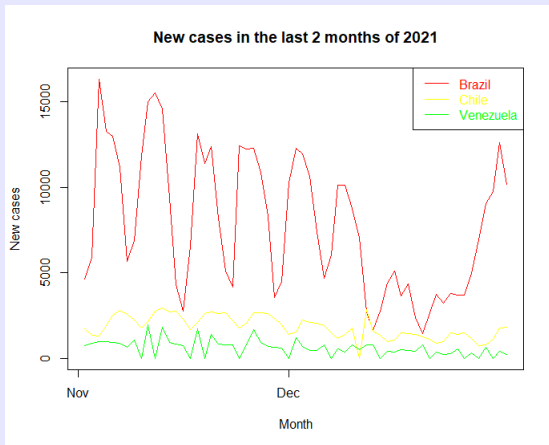
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh gồm 2 tháng cuối của năm

- Kết quả câu 4:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh gồm 2 tháng cuối của năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

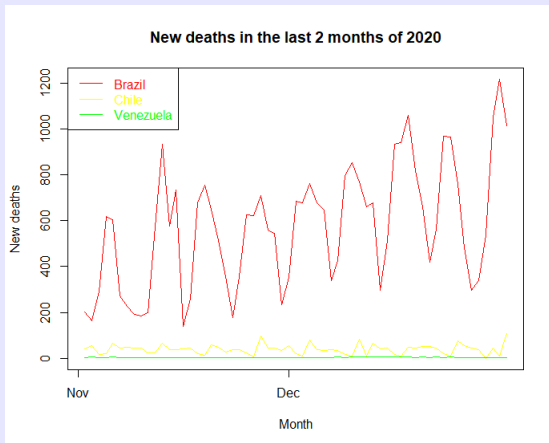
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong gồm 2 tháng cuối của năm

- Kết quả câu 5:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong gồm 2 tháng cuối của năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

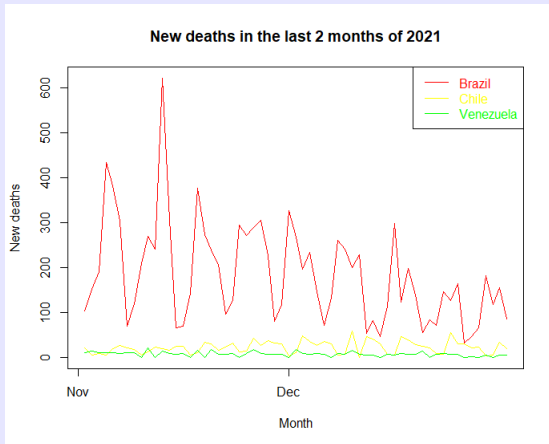
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong gồm 2 tháng cuối của năm

- Kết quả câu 5:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong gồm 2 tháng cuối của năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

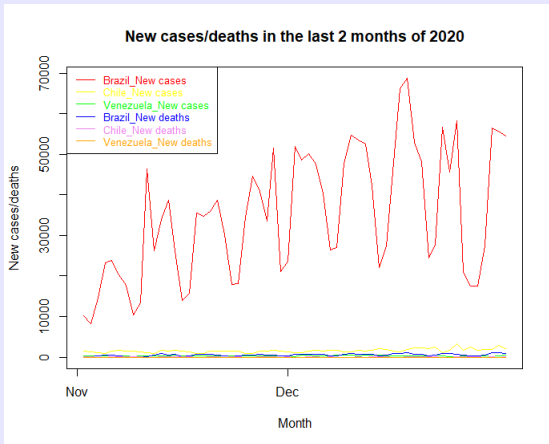
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong gồm 2 tháng cuối của năm

- Kết quả câu 6:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong gồm 2 tháng cuối của năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

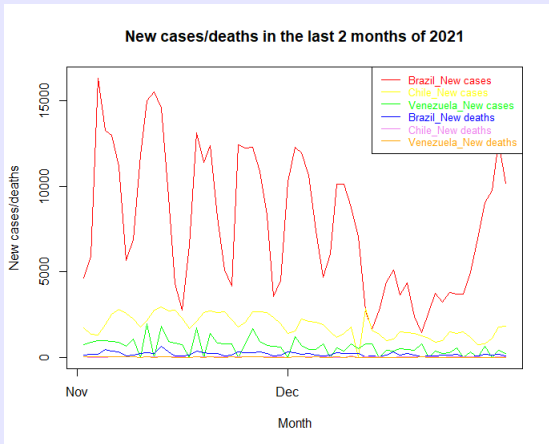
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong gồm 2 tháng cuối của năm

- Kết quả câu 6:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong gồm 2 tháng cuối của năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Phương pháp làm câu 7, 8

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Đầu tiên, ta sẽ tạo dataframe **brazil** qua hàm **subset** trích từ dataframe **owid\_covid\_data** bằng cách lọc đất nước được giao (ở đây là Brazil).
- Ở đây ta dùng lệnh **cumsum** để có thể tính tích lũy số lượng bệnh nhân nhiễm bệnh/tử vong sau khi đã lọc ra nước được giao.
- Các tháng cần sử dụng (ở đây là tháng 11, 12) sẽ được giữ lại trong dataframe **brazil** bằng lệnh **subset**.
- Sau đó, ta lại chia số liệu ra thành các năm để thuận lợi cho việc vẽ biểu đồ.
- Thực hiện tương tự với 2 nước được giao là Chile và Venezuela.
- Cuối cùng, ta vẽ biểu đồ đường bằng lệnh **plot**



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

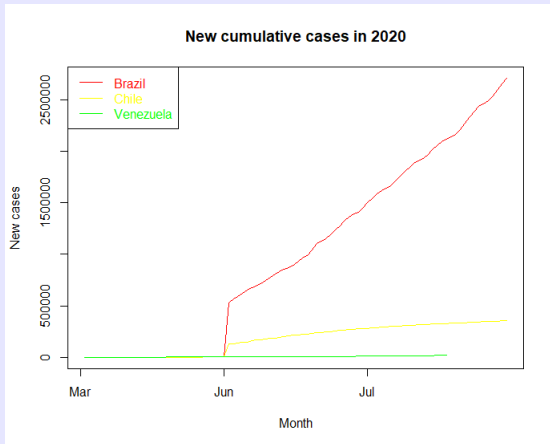
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy cho từng tháng

- Kết quả câu 7:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy cho từng tháng của năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

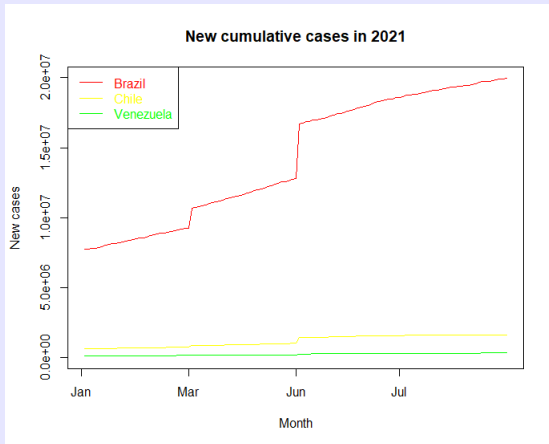
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy cho từng tháng

- Kết quả câu 7:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy cho từng tháng của năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

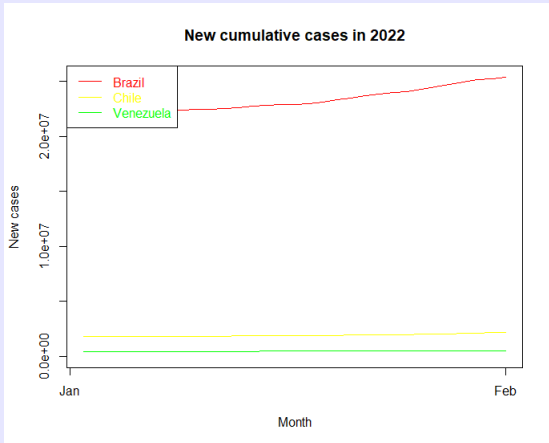
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy cho từng tháng

- Kết quả câu 7:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy cho từng tháng năm 2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

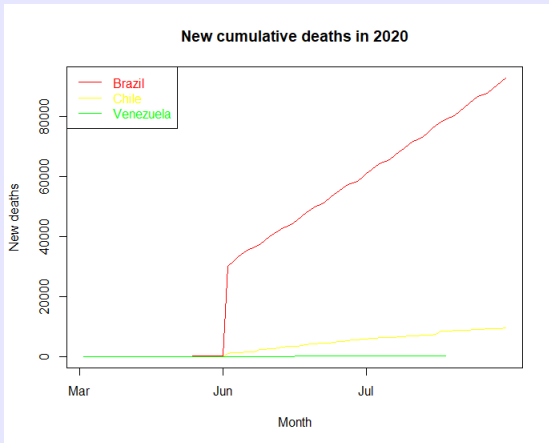
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy cho từng tháng

- Kết quả câu 8:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy cho từng tháng của năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

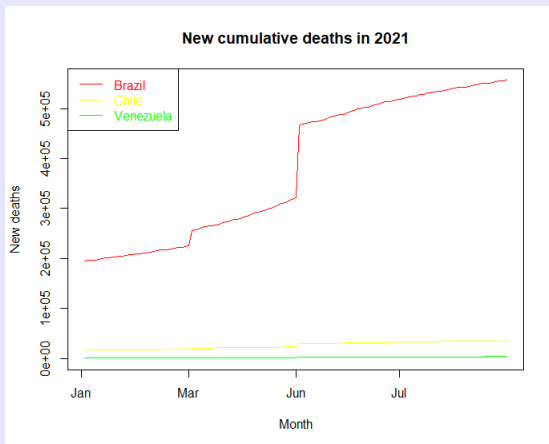
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy cho từng tháng

- Kết quả câu 8:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy cho từng tháng của năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

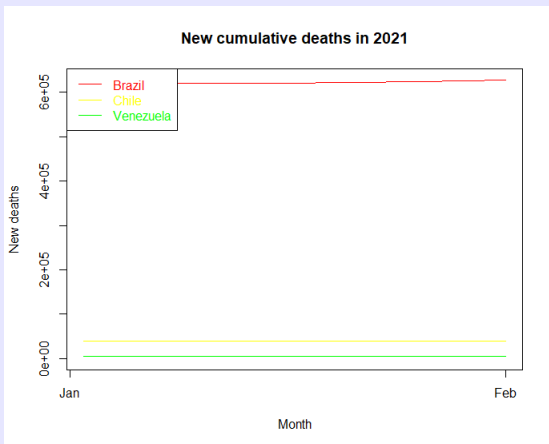
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

### Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy cho từng tháng

- Kết quả câu 8:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy cho từng tháng của năm 2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

### 1-3 Vẽ biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh/ tử vong/ nhiễm bệnh và tử vong cho từng tháng

- Đầu tiên, ta sẽ tạo dataframe `countries_covid_data` qua hàm `subset` trích từ dataframe `owid_covid_data` bằng cách lọc tên nước qua câu lệnh `subset` và dữ liệu
- Tính toán lại giá trị tử vong và nhiễm bệnh các số liệu theo trung bình thời gian 7 ngày
- Lọc ra các số liệu theo từng nước, theo từng tháng các tháng 1, 3, 6, 7 theo *MADE*
- Cuối cùng, ta vẽ biểu đồ đường qua lệnh `ggplot()` trong thư viện `ggplot2`



#### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu về

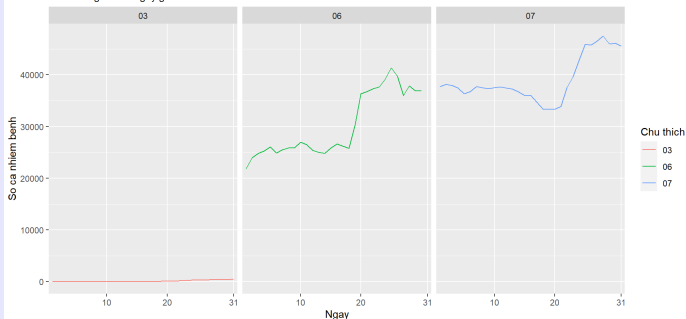
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 1-3 Vẽ biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh/ tử vong/ nhiễm bệnh và tử vong cho từng tháng

Biểu đồ minh họa một phần output câu 1, hình output đầy đủ của câu 1 và các câu 2, 3 ở trong phần báo cáo:

Biểu đồ thể hiện thu thập nhiễm bệnh vào các tháng 03,06,07-2020 của Brazil  
Theo trung bình 7 ngày gần nhất



## vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

### 4-6 Vẽ biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh/ tử vong/ nhiễm bệnh và tử vong cho 2 tháng cuối của năm

- Đầu tiên, ta sẽ tạo dataframe `countries_covid_data` qua hàm `subset` trích từ dataframe `owid_covid_data` bằng cách lọc tên nước qua câu lệnh `subset` và dữ liệu
- Tính toán lại giá trị tử vong và nhiễm bệnh các số liệu theo trung bình thời gian 7 ngày
- Lọc ra các số liệu theo từng nước, theo từng tháng 11, 12.
- Cuối cùng, ta vẽ biểu đồ đường qua lệnh `ggplot()` trong thư viện `ggplot2`



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất



### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

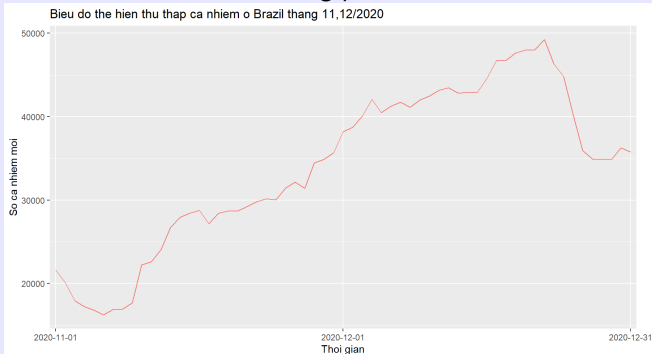
v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 4-6 Vẽ biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh/ tử vong/ nhiễm bệnh và tử vong cho 2 tháng cuối của năm

Biểu đồ minh họa một phần output câu 4, hình output đầy đủ của câu 4 và các câu 5, 6 ở trong phần báo cáo:



## vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

### 7-8 Vẽ biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh/ tử vong/ nhiễm bệnh và tử vong cho từng tháng

- Đầu tiên, ta sẽ tạo dataframe `countries_covid_data` qua hàm `subset` trích từ dataframe `owid_covid_data` bằng cách lọc tên nước qua câu lệnh `subset` và dữ liệu
- Tính toán lại giá trị tử vong và nhiễm bệnh các số liệu theo trung bình thời gian 7 ngày.
- Lọc ra các số liệu theo từng nước, theo các tháng 1, 3, 6, 7 theo *MADE*. Tính toán số ca tử vong tích lũy và nhiễm bệnh tích lũy qua lệnh `cumsum`.
- Cuối cùng, ta vẽ biểu đồ đường qua lệnh `ggplot()` trong thư viện `ggplot2`



#### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

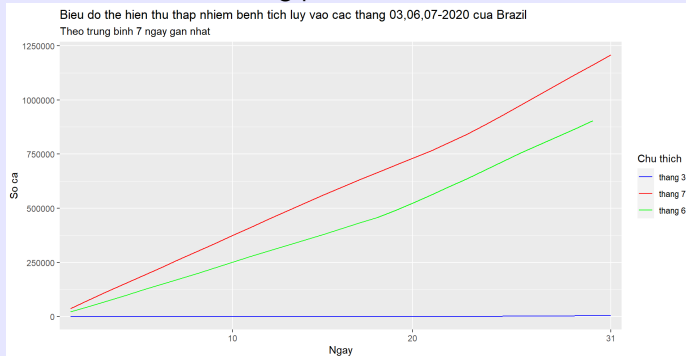
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

### 7-8 Vẽ biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh/ tử vong/ nhiễm bệnh và tử vong cho từng tháng

Biểu đồ minh họa một phần output câu 7, hình output đầy đủ của câu 7 các câu 7,8 ở trong phần báo cáo:



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### Phương pháp làm câu 1, 2

Đầu tiên, ta sẽ tạo dataframe **vii** qua hàm **subset** trích từ dataframe **owid\_covid\_data** bằng cách lọc các tháng cần sử dụng (ở đây là tháng 1, 3, 6, 7).

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- Sử dụng lệnh **summarise** để lấy ra những cột cần thiết.
- Để tính toán được số liệu nhiễm bệnh/ tử vong của tất cả quốc gia trong từng ngày, ta sử dụng câu lệnh **aggregate**.
- Lệnh **names** giúp ta đặt tên cho phù hợp các cột được trích ra.
- Sau đó, ta lại chia số liệu ra thành các năm để thuận lợi cho việc vẽ biểu đồ.
- Cuối cùng, ta vẽ biểu đồ đường bằng lệnh **plot**



#### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

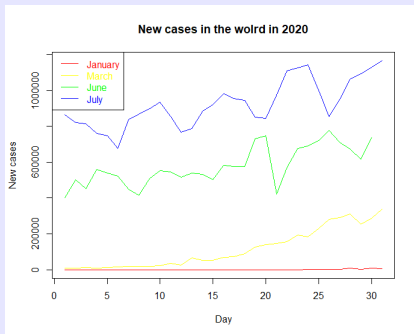
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia

- Kết quả câu 1:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

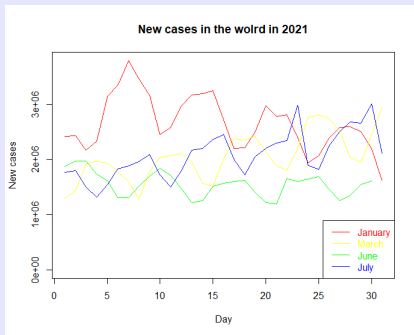
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia

- Kết quả câu 1:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia năm 2021





Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

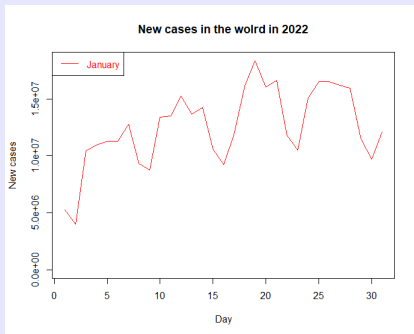
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia

- Kết quả câu 1:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia năm 2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

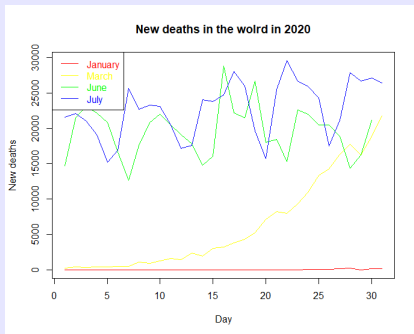
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia

- Kết quả câu 2:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

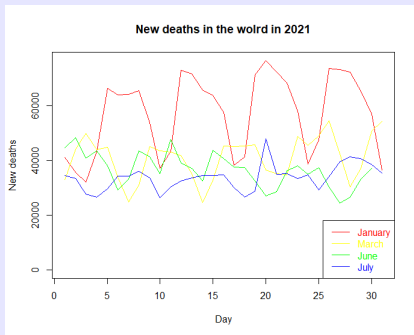
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia

- Kết quả câu 2:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

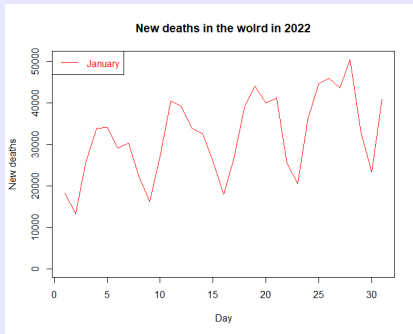
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia

- Kết quả câu 2:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia năm 2022

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### Phương pháp làm câu 3, 4

Tương tự như câu 1, 2, ta sẽ tạo dataframe **vii** qua hàm **subset** trích từ dataframe **owid\_covid\_data** bằng cách lọc các tháng cần sử dụng (ở đây là tháng 11, 12).

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- a) Sử dụng lệnh **summarise** để lấy ra những cột cần thiết.
- b) Để tính toán được số liệu nhiễm bệnh/ tử vong của tất cả quốc gia trong từng ngày, ta sử dụng câu lệnh **aggregate**.
- c) Lệnh **names** giúp ta đặt tên cho phù hợp các cột được trích ra.
- d) Sau đó, ta lại chia số liệu ra thành các năm để thuận lợi cho việc vẽ biểu đồ.
- e) Cuối cùng, ta vẽ biểu đồ đường bằng lệnh **plot**



#### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

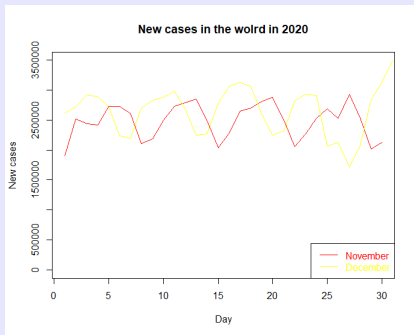
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là 2 tháng cuối của năm của tất cả quốc gia

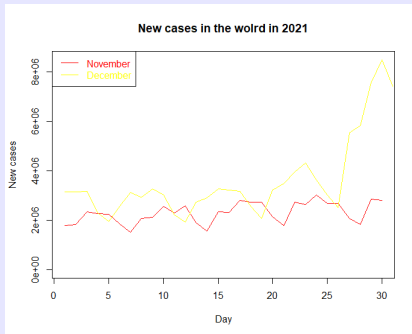
- Kết quả câu 3:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là 2 tháng cuối của năm của tất cả quốc gia năm 2020

## Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là 2 tháng cuối của năm của tất cả quốc gia

- Kết quả câu 3:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là 2 tháng cuối của năm của tất cả quốc gia năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



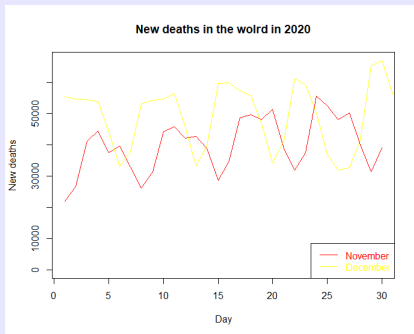
## Danh sách thành viên

- i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- iii. Nhóm câu hỏi liên quan thu thập dữ liệu
- iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu
- v. Nhóm câu hỏi về trực quan dữ liệu theo tháng
- vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là 2 tháng cuối của năm của tất cả quốc gia

- Kết quả câu 4:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là 2 tháng cuối của năm của tất cả quốc gia năm 2020





Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

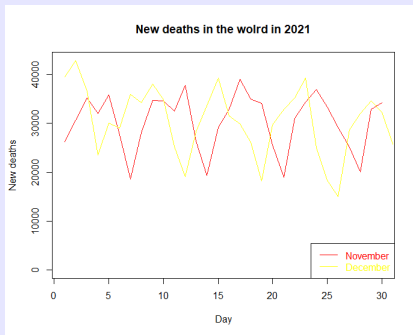
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là 2 tháng cuối của năm của tất cả quốc gia

- Kết quả câu 4:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là 2 tháng cuối của năm của tất cả quốc gia năm 2021

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### Phương pháp làm câu 5, 6

- Phương pháp làm: Các bước thực hiện có các ý chính sau:

- a) Sử dụng lệnh **summarise** để lấy ra những cột cần thiết.
- b) Để tính toán được số liệu nhiễm bệnh/ tử vong của tất cả quốc gia trong từng ngày, ta sử dụng câu lệnh **aggregate**.
- c) Lệnh **names** giúp ta đặt tên cho phù hợp các cột được trích ra.
- d) Ở đây ta dùng lệnh **cumsum** để có thể tính tích lũy số lượng bệnh nhân nhiễm bệnh/tử vong sau khi đã sắp xếp thời gian cho phù hợp.
- e) Sau đó, ta lại chia số liệu ra thành các năm để thuận lợi cho việc vẽ biểu đồ.
- f) Cuối cùng, ta vẽ biểu đồ đường bằng lệnh **plot**



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



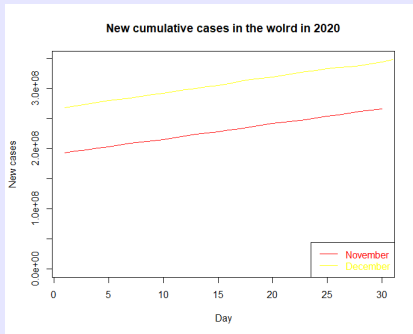
## Danh sách thành viên

- i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- iii. Nhóm câu hỏi liên quan thu thập dữ liệu
- iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu
- v. Nhóm câu hỏi về trực quan dữ liệu theo tháng
- vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tương đối tích lũy theo thời gian là 2 tháng cuối của năm của tất cả quốc gia

- Kết quả câu 5:

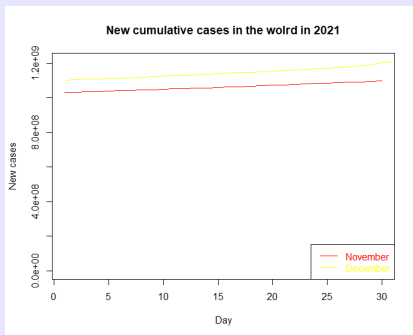


**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tương đối tích lũy theo thời gian là 2 tháng cuối của năm của tất cả quốc gia năm 2020

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tương đối tích lũy theo thời gian là 2 tháng cuối của năm của tất cả quốc gia

- Kết quả câu 5:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tương đối tích lũy theo thời gian là 2 tháng cuối của năm của tất cả quốc gia năm 2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

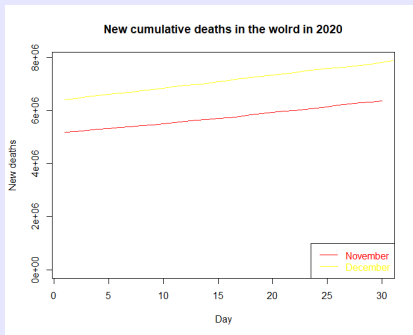
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong tương đối tích lũy theo thời gian là 2 tháng cuối của năm của tất cả quốc gia

- Kết quả câu 6:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong tương đối tích lũy theo thời gian là 2 tháng cuối của năm của tất cả quốc gia năm 2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

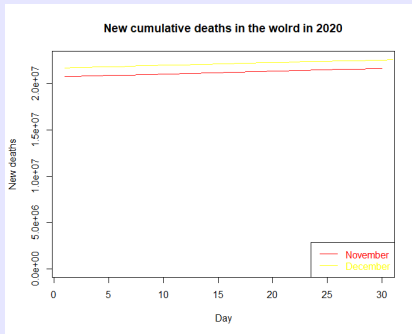
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

Biểu đồ thể hiện thu thập dữ liệu tử vong tương đối tích lũy theo thời gian là 2 tháng cuối của năm của tất cả quốc gia

- Kết quả câu 6:



**Hình:** Biểu đồ thể hiện thu thập dữ liệu tử vong tương đối tích lũy theo thời gian là 2 tháng cuối của năm của tất cả quốc gia năm 2021

## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### Chuẩn bị dữ liệu

Để giải quyết câu hỏi nhóm này, ta sẽ trích dữ liệu những hàng có `location = World` vào một dataframe mới

### 1-2. Biểu đồ thu thập nhiễm bệnh/tử vong theo 7 ngày gần nhất

- Sử dụng vòng lặp for tính toán thu thập nhiễm bệnh/tử vong theo trung bình 7 ngày gần nhất rồi lưu vào cột mới. Sau đó, trích xuất dữ liệu 3 tháng 1, 3, 6, 7 vào một dataframe mới.
- Vẽ đồ thị từ dataframe vừa tạo bằng hàm ggplot từ thư viện ggplot2.



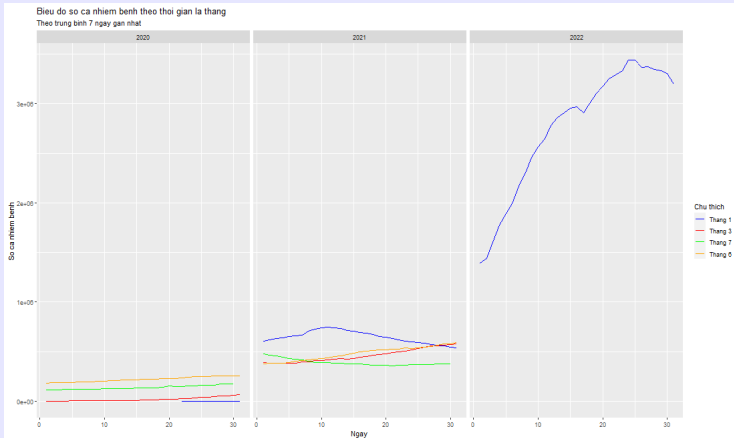
#### Danh sách thành viên

- Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- Nhóm câu hỏi liên quan thu thập dữ liệu
- Nhóm câu hỏi liên quan đến trực quan dữ liệu
- Nhóm câu hỏi về trực quan dữ liệu theo tháng
- Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### 1-2. Biểu đồ thu thập nhiễm bệnh/tử vong theo trung bình 7 ngày gần nhất

- Kết quả câu 1:



**Hình:** Biểu đồ nhiễm bệnh thể giới theo trung bình 7 ngày gần nhất



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

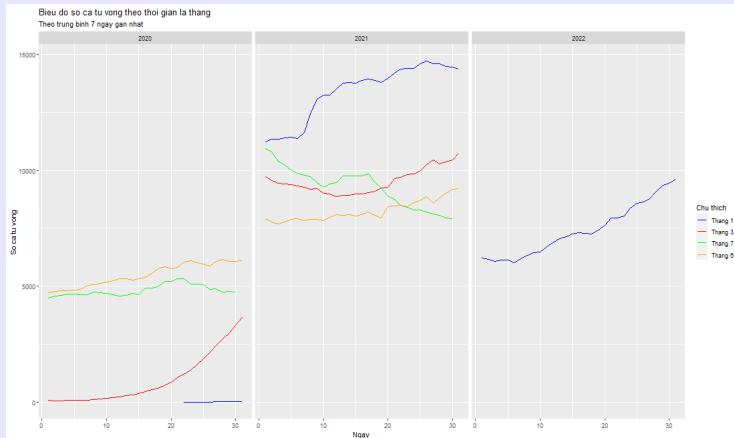
vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### 1-2. Biểu đồ thu thập nhiễm bệnh/tử vong theo trung bình 7 ngày gần nhất

- Kết quả câu 2:



**Hình:** Biểu đồ nhiễm bệnh thể giới theo trung bình 7 ngày gần nhất



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### 3-4. Biểu đồ thu thập nhiễm bệnh/tử vong theo thời gian là 2 tháng cuối năm

- Trích dữ liệu tháng 11, 12 của dataframe trong phần chuẩn bị dữ liệu nhóm viii.
- Sử dụng vòng lặp for tính toán thu thập nhiễm bệnh/tử vong theo trung bình 7 ngày gần nhất rồi lưu vào cột mới. Tuy nhiên trong hai câu này, ta phải chú ý thêm điều kiện nếu thời điểm đang xét là ngày 1 tháng 11. Sau đó, trích xuất dữ liệu 3 tháng 1, 3, 6, 7 vào một dataframe mới.
- Vẽ đồ thị từ dataframe vừa tạo bằng hàm ggplot từ thư viện ggplot2.



#### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

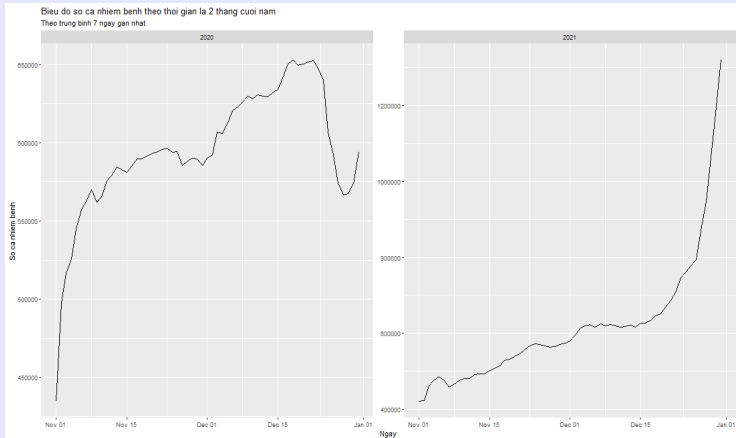
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### 3-4. Biểu đồ thu thập nhiễm bệnh/tử vong theo thời gian là 2 tháng cuối năm

- Kết quả câu 3:



**Hình:** Biểu đồ nhiễm bệnh thể giới theo trung bình 7 ngày gần nhất hai tháng cuối năm



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

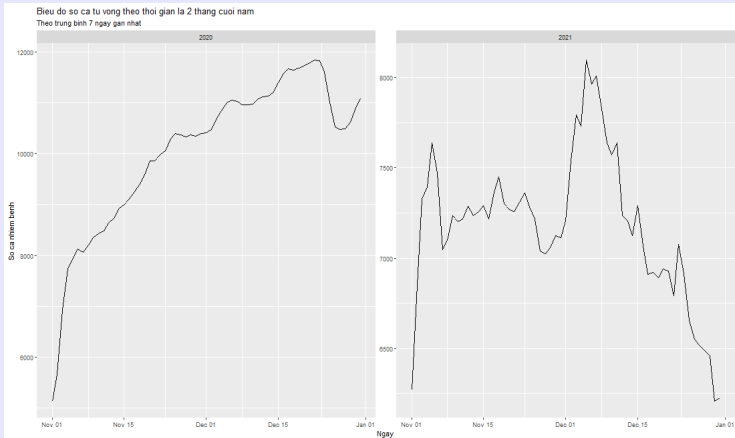
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### 3-4. Biểu đồ thu thập nhiễm bệnh/tử vong theo thời gian là 2 tháng cuối năm

- Kết quả câu 4:



**Hình:** Biểu đồ tử vong thể giới theo trung bình 7 ngày gần nhất hai tháng cuối năm



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### 5-6. Biểu đồ nhiễm bệnh/tử vong tích lũy theo thời gian 2 tháng cuối năm của thể giới

- Trích dữ liệu tháng 11, 12 của thể giới từ dataframe ở phần chuẩn bị dữ liệu
- Sử dụng vòng lặp for để tính nhiễm bệnh/tử vong tích lũy trong 2 tháng cuối năm, lưu vào một dataframe mới.
- Sử dụng vòng lặp for để tính trung bình theo 7 ngày gần nhất của các tích lũy lưu vào cột mới của dataframe vừa tạo, sau đó sử dụng hàm ggplot để vẽ đồ thị.



#### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

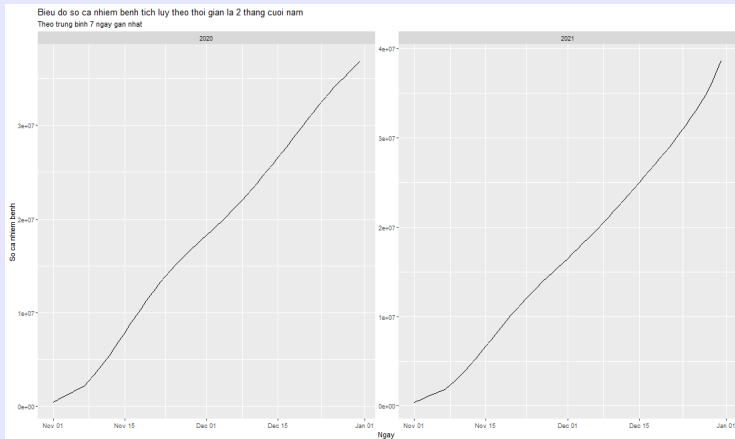
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### 5-6. Biểu đồ thu thập nhiễm bệnh/tử vong tích theo thời gian là 2 tháng cuối năm

- Kết quả câu 5:



**Hình:** Biểu đồ nhiễm bệnh tích thể giới theo trung bình 7 ngày gần nhất hai tháng cuối năm



#### Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

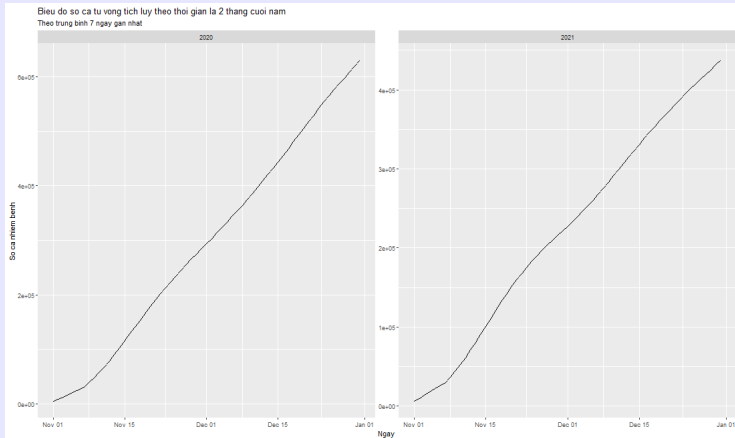
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## viii. Nhóm câu hỏi về thể giới theo trung bình 7 ngày gần nhất

### 5-6. Biểu đồ thu thập nhiễm bệnh/tử vong theo thời gian là 2 tháng cuối năm

- Kết quả câu 6:



**Hình:** Biểu đồ tử vong tích thể giới theo trung bình 7 ngày gần nhất hai tháng cuối năm



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ix. Nhóm câu hỏi liên quan đến sự tương quan giữa nhiễm bệnh và tử vong

### Phần làm chung

Đầu tiên, ta sẽ tạo dataframe `countries_covid_data` qua hàm `subset` trích từ dataframe `owid_covid_data` bằng cách lọc tên nước qua câu lệnh `subset` và dữ liệu

### 1. Biểu đồ thể hiện phần trăm giữa nhiễm bệnh tích lũy trên tổng nhiễm bệnh và phần trăm tử vong tích lũy trên tổng số tử vong của từng quốc gia

- Đầu tiên, ứng với mỗi dataframe riêng cho mỗi nước, ta sẽ tính toán dữ liệu ca nhiễm tích lũy `cases_tichluy` và ca tử vong tích lũy `death_tichluy` qua hàm `cumsum()`
- Tiếp theo, ứng với mỗi ngày ta tính toán phần trăm giữa nhiễm bệnh tích lũy trên tổng nhiễm bệnh và phần trăm tử vong tích lũy trên tổng số tử vong của từng quốc gia
- Cuối cùng, ta vẽ biểu đồ đường qua lệnh `ggplot()` trong thư viện `ggplot2`



#### Danh sách thành viên

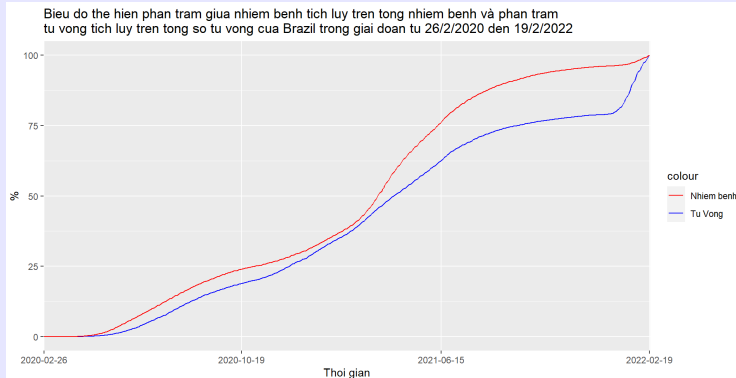
- Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- Nhóm câu hỏi liên quan thu thập dữ liệu
- Nhóm câu hỏi liên quan đến trực quan dữ liệu
- Nhóm câu hỏi về trực quan dữ liệu theo tháng
- Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## ix. Nhóm câu hỏi liên quan đến sự tương quan giữa nhiễm bệnh và tử vong

### 1. Biểu đồ thể hiện phần trăm giữa nhiễm bệnh tích lũy trên tổng nhiễm bệnh và phần trăm tử vong tích lũy trên tổng số tử vong của từng quốc gia

Biểu đồ minh họa một phần output câu 1, các hình còn lại ở trong phần báo cáo:



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## ix. Nhóm câu hỏi liên quan đến sự tương quan giữa nhiễm bệnh và tử vong



Danh sách thành viên

- i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- iii. Nhóm câu hỏi liên quan thu thập dữ liệu
- iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu
- v. Nhóm câu hỏi về trực quan dữ liệu theo tháng
- vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

### 2-3. Xét tương quan mỗi tháng / tương quan mỗi tháng theo trung bình 7 ngày gần nhất

- Phương pháp làm: Vẽ biểu đồ tương quan giữa số ca nhiễm bệnh và số ca tử vong theo ngày bằng lệnh `ggplot()` trong thư viện `ggplot2`, tính hệ số tương quan qua hàm `cor()` từ đó rút ra nhận xét về sự tương quan (hướng tương quan, tính tương quan mạnh, yếu,...).
- Lưu ý: Ở câu 3, ta cần tính lại số liệu theo trung bình 7 ngày trước khi vẽ biểu đồ

## ix. Nhóm câu hỏi liên quan đến sự tương quan giữa nhiễm bệnh và tử vong

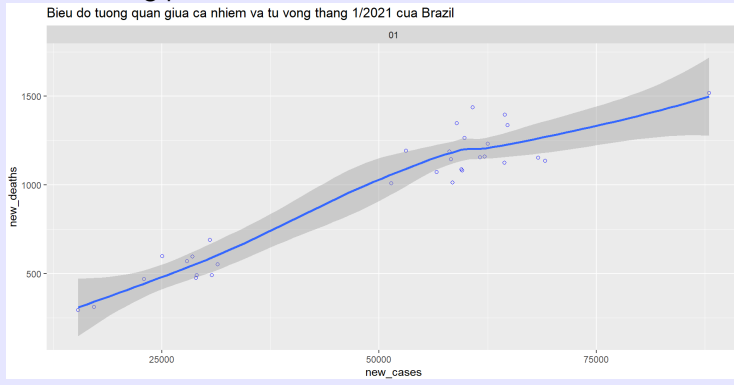


Danh sách thành viên

- i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- iii. Nhóm câu hỏi liên quan thu thập dữ liệu
- iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu
- v. Nhóm câu hỏi về trực quan dữ liệu theo tháng
- vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng
- viii. Nhóm câu hỏi về...

### 2-3. Xét tương quan mỗi tháng / tương quan mỗi tháng theo trung bình 7 ngày gần nhất

Biểu đồ minh họa 1 phần output câu 2, các hình còn lại và output câu 3 ở trong phần báo cáo:



## x. Nhóm câu hỏi riêng

### Chuẩn bị dữ liệu

- Đầu tiên, ta lọc dữ liệu ba quốc gia trong nhóm, sau đó xử lý số liệu không xác định thành giá trị 0 bằng hàm `is.na`.



#### Danh sách thành viên

- Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- Nhóm câu hỏi liên quan thu thập dữ liệu
- Nhóm câu hỏi liên quan đến trực quan dữ liệu
- Nhóm câu hỏi về trực quan dữ liệu theo tháng
- Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

### Câu 1: So sánh tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng

Phương pháp làm:

- Ta tiến hành lọc dữ liệu 7 ngày cuối của năm cuối cùng của ba quốc gia trong nhóm bằng hàm `subset`.
- Sau đó, ta vẽ biểu đồ tình trạng nhiễm bệnh của các quốc gia bằng thư viện `ggplot2`.
- Do quy mô dân số của ba quốc gia chênh lệch quá lớn, nên để dễ dàng so sánh tình trạng nhiễm bệnh của các quốc gia, ta vẽ biểu đồ tình trạng nhiễm bệnh của từng quốc gia trong nhóm.
- Cuối cùng ta so sánh và nhận xét tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

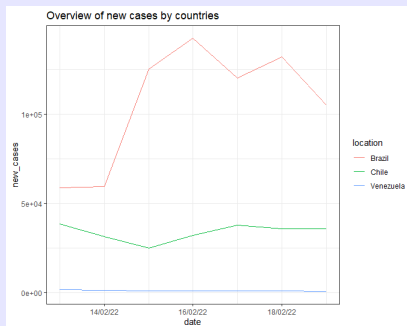
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

Câu 1: So sánh tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng

Kết quả:



**Hình:** Biểu đồ tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 1: So sánh tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng**

**Kết quả:**



**Hình:** Biểu đồ tình trạng nhiễm bệnh của Brazil trong 7 ngày cuối của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

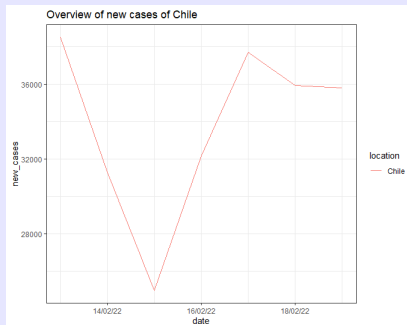
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 1: So sánh tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng**

**Kết quả:**



**Hình:** Biểu đồ tình trạng nhiễm bệnh của Chile trong 7 ngày cuối của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

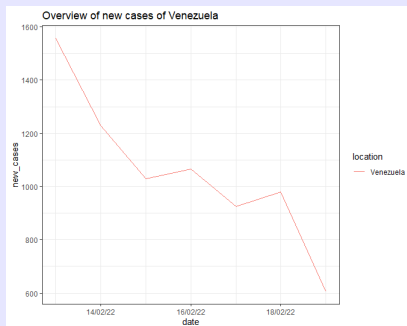
vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## x. Nhóm câu hỏi riêng

Câu 1: So sánh tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng

Kết quả:



**Hình:** Biểu đồ tình trạng nhiễm bệnh của Venezuela trong 7 ngày cuối của năm cuối cùng



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

### Câu 1: So sánh tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng

Nhận xét:

- Brazil có số ca nhiễm mới tăng mạnh vào những ngày 14-16/02/2022, sau đó có xu hướng giảm nhẹ.
- Chile có số ca nhiễm mới giảm mạnh vào những ngày 13-15/02/2022, nhưng sau đó lại tăng mạnh vào những ngày 16-17/02/2022, những ngày cuối có xu hướng ổn định.
- Venezuela có số ca nhiễm mới giảm liên tục trong 7 ngày cuối cùng.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

Phương pháp làm:

- Ta xử lý dữ liệu thời gian thành từng tháng trong năm bằng hàm `substr`.
- Ta tính tỉ lệ nhiễm bệnh tích lũy và tỉ lệ tử vong tích lũy theo từng tháng của từng quốc gia trong nhóm. - Ta cho chạy vòng lặp `for` để tính số ca nhiễm bệnh và số ca tử vong tích lũy của từng quốc gia trong nhóm.
  - Trong vòng lặp, ta tính tỉ lệ nhiễm bệnh tích lũy và tỉ lệ tử vong tích lũy của từng quốc gia trong nhóm.
- Tính độ biến thiên tỉ lệ tích lũy qua các tháng bằng cách cho chạy vòng lặp `for`, lấy tỉ lệ tháng sau trừ đi tỉ lệ tháng trước.
- Lọc ra các khoảng thời gian mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại của từng quốc gia (ta lấy ngưỡng tăng mạnh là  $\frac{2}{3}$  độ tăng lớn nhất và giảm mạnh là  $\frac{2}{3}$  độ giảm lớn nhất).
- Vẽ biểu đồ trực quan bằng thư viện `ggplot2`.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

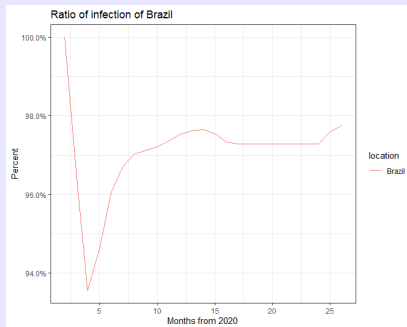
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ nhiễm bệnh tích lũy của Brazil



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

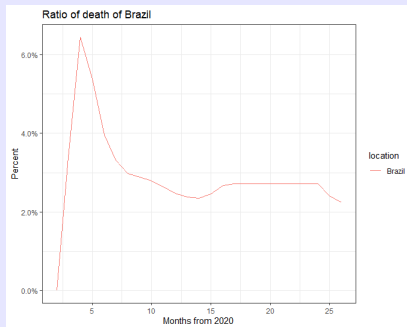
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ tử vong tích lũy của Brazil



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

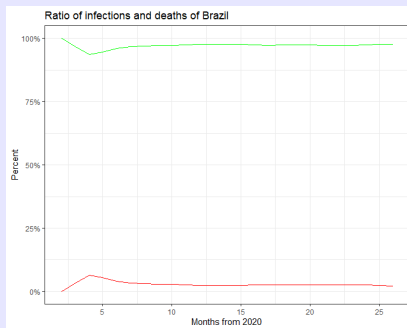
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ nhiễm bệnh tích lũy và tử vong tích lũy của Brazil



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

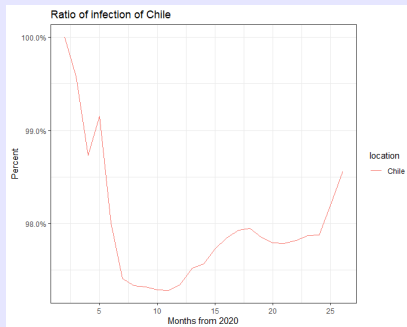
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ nhiễm bệnh tích lũy của Chile



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

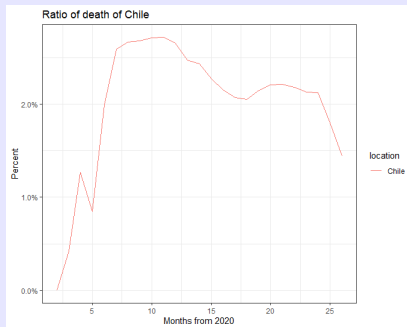
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ tử vong tích lũy của Chile



Danh sách thành viên

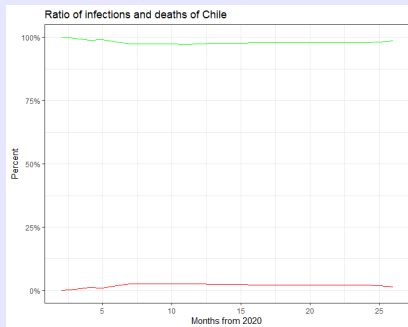
- Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- Nhóm câu hỏi liên quan thu thập dữ liệu
- Nhóm câu hỏi liên quan đến trực quan dữ liệu
- Nhóm câu hỏi về trực quan dữ liệu theo tháng
- Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ nhiễm bệnh tích lũy và tử vong tích lũy của Chile



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

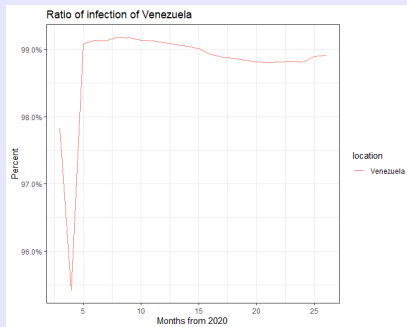
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ nhiễm bệnh tích lũy của Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

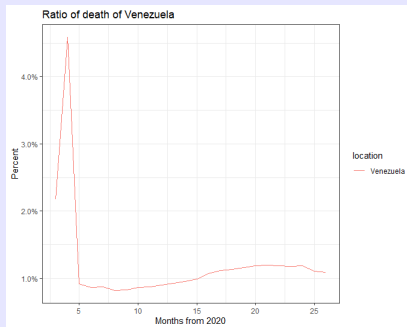
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ tử vong tích lũy của Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

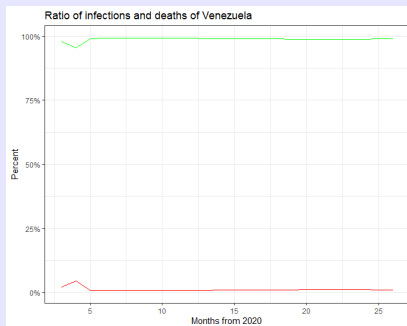
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3:** Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.

Kết quả:



**Hình:** Biểu đồ tỉ lệ nhiễm bệnh tích lũy và tử vong tích lũy của Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

Kết quả:

- Khoảng thời gian tỉ lệ nhiễm bệnh tích lũy tăng mạnh ở Brazil:

Location	month	year
Brazil	05	2020
Brazil	06	2020

- Khoảng thời gian tỉ lệ nhiễm bệnh tích lũy giảm mạnh ở Brazil:

Location	month	year
Brazil	03	2020
Brazil	04	2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

Kết quả:

- Khoảng thời gian tỉ lệ tử vong tích lũy tăng mạnh ở Brazil:

Location	month	year
Brazil	03	2020
Brazil	04	2020

- Khoảng thời gian tỉ lệ tử vong tích lũy giảm mạnh ở Brazil:

Location	month	year
Brazil	05	2020
Brazil	06	2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

**Kết quả:**

- Khoảng thời gian tỉ lệ nhiễm bệnh tích lũy tăng mạnh ở Chile:

Location	month	year
Chile	05	2020
Chile	01	2022
Chile	02	2022

- Khoảng thời gian tỉ lệ nhiễm bệnh tích lũy giảm mạnh ở Chile:

Location	month	year
Chile	04	2020
Chile	06	2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

**Kết quả:**

- Khoảng thời gian tỉ lệ tử vong tích lũy tăng mạnh ở Chile:

Location	month	year
Chile	04	2020
Chile	06	2020

- Khoảng thời gian tỉ lệ tử vong tích lũy giảm mạnh ở Chile:

Location	month	year
Chile	05	2020
Chile	01	2022
Chile	02	2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

Kết quả:

- Khoảng thời gian tỉ lệ nhiễm bệnh tích lũy tăng mạnh ở Venezuela:

Location	month	year
Venezuela	05	2020

- Khoảng thời gian tỉ lệ nhiễm bệnh tích lũy giảm mạnh ở Venezuela:

Location	month	year
Venezuela	04	2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

Kết quả:

- Khoảng thời gian tỉ lệ tử vong tích lũy tăng mạnh ở Venezuela:

Location	month	year
Venezuela	04	2020

- Khoảng thời gian tỉ lệ tử vong tích lũy giảm mạnh ở Venezuela:

Location	month	year
Venezuela	05	2020



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

**Kết luận:**

- Các khoảng thời gian mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh ở Brazil là tháng 05 và tháng 06 năm 2020.
- Các khoảng thời gian mà tỉ lệ tử vong tích lũy tăng mạnh nhưng tỉ lệ nhiễm bệnh tích lũy giảm mạnh ở Brazil là tháng 03 và tháng 04 năm 2020.
- Các khoảng thời gian mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh ở Chile là tháng 05 năm 2020 và tháng 01, 02 năm 2022.
- Các khoảng thời gian mà tỉ lệ tử vong tích lũy tăng mạnh nhưng tỉ lệ nhiễm bệnh tích lũy giảm mạnh ở Chile là tháng 04 và tháng 06 năm 2020.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 3: Cho biết các khoảng thời gian nào mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh hoặc ngược lại cho các quốc gia.**

**Kết luận:**

- Các khoảng thời gian mà tỉ lệ tử vong tích lũy giảm mạnh nhưng tỉ lệ nhiễm bệnh tích lũy tăng mạnh ở Venezuela là tháng 05 năm 2020.
- Các khoảng thời gian mà tỉ lệ tử vong tích lũy tăng mạnh nhưng tỉ lệ nhiễm bệnh tích lũy giảm mạnh ở Venezuela là tháng 04 năm 2020.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6: Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Phương pháp làm:

- Ta xử lý dữ liệu thời gian thành từng tháng trong năm bằng hàm `substr`.
- Vẽ biểu đồ số ca nhiễm mới của 3 quốc gia trong nhóm trên cùng 1 biểu đồ bằng thư viện `ggplot2`
- Tính ngưỡng  $k$  là ngưỡng bùng phát nhiễm bệnh của từng quốc gia (ở đây ta lấy  $k$  là giá trị Q3 của số ca nhiễm mới) và lọc ra những tháng, năm có số ca nhiễm mới cao hơn ngưỡng bùng phát dịch.
- Vẽ biểu đồ tình trạng bùng phát nhiễm bệnh ở từng quốc gia bằng thư viện `ggplot2`
- Xác định khoảng thời gian bùng phát nhiễm bệnh lớn nhất của từng quốc gia và nhận xét những khoảng thời gian này có chồng lên nhau hay không, nếu có thì nêu ra những khoảng thời gian đó.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

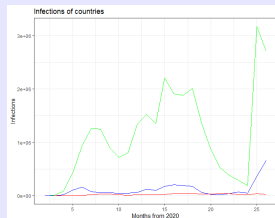
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6:** Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?

Kết quả:



**Hình:** Biểu đồ tình trạng nhiễm bệnh ở 3 quốc gia Brazil, Chile và Venezuela

Chú thích:

- Đường màu xanh: Brazil.
- Đường màu lam: Chile.
- Đường màu đỏ: Venezuela.



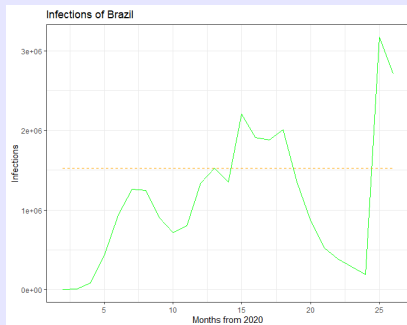
Danh sách thành viên

- Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- Nhóm câu hỏi liên quan thu thập dữ liệu
- Nhóm câu hỏi liên quan đến trực quan dữ liệu
- Nhóm câu hỏi về trực quan dữ liệu theo tháng
- Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6:** Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?

Kết quả:



**Hình:** Biểu đồ tình trạng bùng phát nhiễm bệnh của Brazil



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

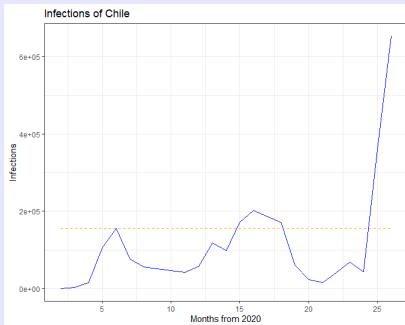
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6:** Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?

Kết quả:



**Hình:** Biểu đồ tình trạng bùng phát nhiễm bệnh của Chile



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

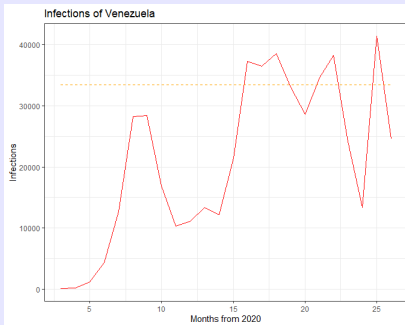
vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## x. Nhóm câu hỏi riêng

**Câu 6:** Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?

Kết quả:



**Hình:** Biểu đồ tình trạng bùng phát nhiễm bệnh của Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6: Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Kết quả:

- Những khoảng thời gian bùng phát nhiễm bệnh ở Brazil:

Location	month	year
Brazil	01	2021
Brazil	03	2021
Brazil	04	2021
Brazil	05	2021
Brazil	06	2021
Brazil	01	2022
Brazil	02	2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6: Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Kết quả:

- Những khoảng thời gian bùng phát nhiễm bệnh ở Chile:

Location	month	year
Chile	06	2020
Chile	03	2021
Chile	04	2021
Chile	05	2021
Chile	06	2021
Chile	01	2022
Chile	02	2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6: Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Kết quả:

- Những khoảng thời gian bùng phát nhiễm bệnh ở Venezuela:

Location	month	year
Venezuela	04	2021
Venezuela	05	2021
Venezuela	06	2021
Venezuela	09	2021
Venezuela	10	2021
Venezuela	01	2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6: Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Nhận xét:

- Khoảng thời gian bùng phát nhiễm bệnh lớn nhất của Brazil là từ tháng 03 đến tháng 06 năm 2021.
- Khoảng thời gian bùng phát nhiễm bệnh lớn nhất của Chile là từ tháng 03 đến tháng 06 năm 2021.
- Khoảng thời gian bùng phát nhiễm bệnh lớn nhất của Venezuela là từ tháng 04 đến tháng 06 năm 2021.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 6: Khoảng thời gian bùng phát nhiễm bệnh lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Kết luận:

- Khoảng thời gian bùng phát nhiễm bệnh lớn nhất của 3 quốc gia trùng nhau từ tháng 04 đến tháng 06 năm 2021.
- Khoảng thời gian bùng phát nhiễm bệnh lớn nhất của Brazil và Chile trùng nhau hoàn toàn (từ tháng 03 đến tháng 06 năm 2021).
- hoảng thời gian bùng phát nhiễm bệnh lớn nhất của Brazil và Venezuela trùng nhau từ tháng 04 đến tháng 06 năm 2021.
- Khoảng thời gian bùng phát nhiễm bệnh lớn nhất của Chile và Venezuela trùng nhau từ tháng 04 đến tháng 06 năm 2021.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Phương pháp làm tương tự như câu 6, ta thay dữ liệu nhiễm bệnh thành dữ liệu tử vong.



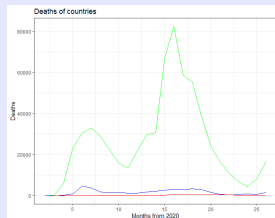
Danh sách thành viên

- i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- iii. Nhóm câu hỏi liên quan thu thập dữ liệu
- iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu
- v. Nhóm câu hỏi về trực quan dữ liệu theo tháng
- vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7:** Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?

Kết quả:



**Hình:** Biểu đồ tình trạng tử vong ở 3 quốc gia Brazil, Chile và Venezuela

Chú thích:

- Đường màu xanh: Brazil.
- Đường màu lam: Chile.
- Đường màu đỏ: Venezuela.



Danh sách thành viên

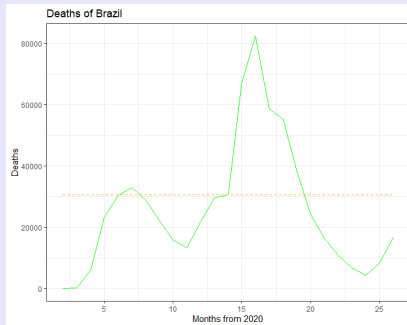
- Nhóm câu hỏi liên quan đến tổng quát dữ liệu
- Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu
- Nhóm câu hỏi liên quan thu thập dữ liệu
- Nhóm câu hỏi liên quan đến trực quan dữ liệu
- Nhóm câu hỏi về trực quan dữ liệu theo tháng
- Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất
- Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng



## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

**Kết quả:**



**Hình:** Biểu đồ tình trạng bùng phát tử vong của Brazil



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

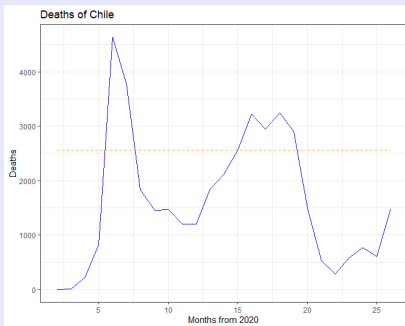
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

**Kết quả:**



**Hình:** Biểu đồ tình trạng bùng phát tử vong của Chile



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

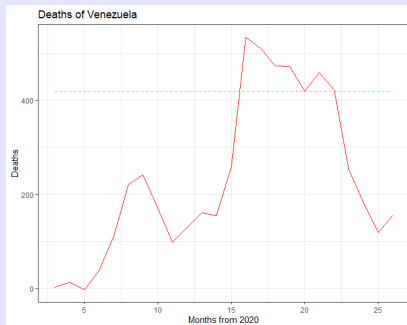
vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

**Kết quả:**



**Hình:** Biểu đồ tình trạng bùng phát tử vong của Venezuela



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Kết quả:

- Những khoảng thời gian bùng phát tử vong ở Brazil:

Location	month	year
Brazil	07	2020
Brazil	02	2021
Brazil	03	2021
Brazil	04	2021
Brazil	05	2021
Brazil	06	2022
Brazil	07	2022



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Kết quả:

- Những khoảng thời gian bùng phát tử vong ở Chile:

Location	month	year
Chile	06	2020
Chile	07	2020
Chile	03	2021
Chile	04	2021
Chile	05	2021
Chile	06	2021
Chile	07	2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Kết quả:

- Những khoảng thời gian bùng phát tử vong ở Venezuela:

Location	month	year
Venezuela	04	2021
Venezuela	05	2021
Venezuela	06	2021
Venezuela	07	2021
Venezuela	09	2021
Venezuela	10	2021



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Nhận xét:

- Khoảng thời gian bùng phát tử vong lớn nhất của Brazil là từ tháng 02 đến tháng 07 năm 2021.
- Khoảng thời gian bùng phát tử vong lớn nhất của Chile là từ tháng 03 đến tháng 07 năm 2021.
- Khoảng thời gian bùng phát tử vong lớn nhất của Venezuela là từ tháng 04 đến tháng 07 năm 2021.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng

## x. Nhóm câu hỏi riêng

**Câu 7: Khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia có chồng lên nhau không, Cho biết khoảng thời gian giao nhau đó?**

Kết luận:

- Khoảng thời gian bùng phát tử vong lớn nhất của 3 quốc gia trùng nhau từ tháng 04 đến tháng 07 năm 2021.
- Khoảng thời gian bùng phát tử vong lớn nhất của Brazil và Chile trùng nhau từ tháng 03 đến tháng 07 năm 2021.
- Khoảng thời gian bùng phát tử vong lớn nhất của Brazil và Venezuela trùng nhau từ tháng 04 đến tháng 07 năm 2021.
- Khoảng thời gian bùng phát tử vong lớn nhất của Chile và Venezuela trùng nhau từ tháng 04 đến tháng 07 năm 2021.



Danh sách thành viên

i. Nhóm câu hỏi liên quan đến tổng quát dữ liệu

ii. Nhóm câu hỏi liên quan đến mô tả thống kê cơ bản dữ liệu

iii. Nhóm câu hỏi liên quan thu thập dữ liệu

iv. Nhóm câu hỏi liên quan đến trực quan dữ liệu

v. Nhóm câu hỏi về trực quan dữ liệu theo tháng

vi. Nhóm câu hỏi về trực quan dữ liệu theo trung bình 7 ngày gần nhất

vii. Nhóm câu hỏi liên quan đến tất cả quốc gia theo thời gian là tháng