

Chapter

CẤU TRÚC RỜI RẠC CHO KHMT

L^AT_EX-Beamer on April 24, 2022

Huynh Tuong Nguyen, Nguyen Ngoc Le
Faculty of Computer Science and Engineering
University of Technology - VNUHCM
htnguyen@hcmut.edu.vn

CẤU TRÚC RỜI RẠC
CHO KHMT

Huynh Tuong Nguyen,
Nguyen Ngoc Le



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết



Dòng cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

① Động cơ nghiên cứu

② Mục tiêu

③ Mô tả dữ liệu

④ Nhiệm vụ

⑤ Tổng kết



Nghiên cứu để làm gì?

COVID-19 đã tạo ra những tác động tiêu cực đến nền đời sống của cư dân trên thế giới. Các đợt bùng phát của COVID-19 hay những biến thể virus đã mang đến những thách thức chưa từng có và được dự báo sẽ có tác động đáng kể đến sự phát triển kinh tế. Phân tích & thống kê dữ liệu về COVID-19 giúp cho ta thấy được số ca nhiễm bệnh, tử vong của một quốc gia, so sánh tình trạng của các quốc gia trong khu vực hay diễn biến dịch trên thế giới. Từ số liệu được báo cáo mới chúng ta muốn biết các ca nhiễm bệnh có xu hướng tăng lên hay giảm xuống quy mô các đợt bùng phát ở mỗi quốc gia.

Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết



Cần có mục tiêu thiết thực

Trong bài tập lớn này, chúng ta sẽ bắt đầu với các bài toán thống kê đơn giản từ những dữ liệu được cung cấp. Qua đó, ta tìm ra những con số thú vị, có ý nghĩa đối với các dữ liệu thực tế từ tình hình dịch corona. Những kết quả mà tìm ra sẽ là bước khởi đầu cho việc khai phá nguồn dữ liệu của hệ thống sau này, nhằm đạt tới mục tiêu nâng cao kỹ năng lập trình, kỹ năng giải quyết vấn đề cho người học, kỹ năng làm việc nhóm cũng như hướng tới mục tiêu cao hơn là đam mê trong làm việc, học tập và nghiên cứu.

Mô tả dữ liệu

Dữ liệu gồm các thuộc tính chính “iso_code, continent, location, date, new_cases, new_deaths” được lưu tron file csv.

- ❶ *iso_code*: Định danh đất nước
- ❷ *continent* Tên châu lục
- ❸ *location*: Tên quốc gia
- ❹ *date*: Ngày quan sát với định dạng Month-Day-Year
- ❺ *new_cases*: Số trường hợp COVID-19 mới được xác nhận
- ❻ *new_deaths*: Số tử vong mới do COVID-19





Cơ sở lý thuyết

- Lưu bảng dữ liệu thô vào biến bigTable
- Lưu bảng dữ liệu sau khi lọc bớt các hàng không có châu lục vào biến tableNotNull
- Lưu các biến này vào preTask.RData bằng câu lệnh `save.image("workspace/preTask.RData")` để có thể tái sử dụng
- Cài đặt một số packages: readr, ggplot2, stringr, dplyr
- Một số hàm quan trọng được dùng trong task 1: `str_sub`, `unique`, `write` và `write.csv`, `subset`, `append`, `tail` và `head`, `length`, `data.frame`
- Cuối cùng, lưu lại các biến để có thể tái sử dụng bằng `save.image`

Task 1-Subtask 1



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Tập mẫu thể hiện thu thập dữ liệu vào các năm nào:

- 2020
- 2021
- 2022

Task 1-Subtask 2



Dòng cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Số lượng đất nước và định danh của mỗi đất nước:

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

Task 1-Subtask 3



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Số lượng đất nước:

238



Dòng cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Số lượng châu lục trong tập mẫu:

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

Task 1-Subtask 5



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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

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

Task 1-Subtask 6



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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 s

iso_code	Observations
VEN	708
VGB	694
VNM	759
VUT	467
WLF	489
WSM	459
YEM	681
ZAF	744
ZMB	704
ZWE	702
Tong:	153316

Task 1-Subtask 7



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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

Châu lục có lượng dữ liệu thu thập nhỏ nhất là Châu Đại Dương với giá trị là 8993

Task 1-Subtask 8



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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

Châu lục có lượng dữ liệu thu thập lớn nhất là Châu Phi với giá trị là 38647

Task 1-Subtask 9



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

Nước có ít dữ liệu thu thập nhỏ nhất là Pitcairn với giá trị là 85

Task 1-Subtask 10



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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

Nước có ít dữ liệu thu thập lớn nhất là Argentina và Mexico với giá trị là 781

Task 1-Subtask 11



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

Date:	Minumum
1/1/2020	2
1/2/2020	2
1/3/2020	2

Task 1-Subtask 12



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

Date:	Maximum
8/22/2021	238
8/23/2021	238
8/24/2021	238
8/25/2021	238
8/26/2021	238
8/27/2021	238
8/28/2021	238
8/29/2021	238

Task 1-Subtask 13



Dòng cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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

continent	date	numbers of data
Africa	1/1/2021	53
Africa	1/1/2022	54
Africa	1/10/2021	54
Africa	1/10/2022	54
Africa	1/11/2021	54
Africa	1/11/2022	54
Africa	1/12/2021	54
Africa	1/12/2022	54
Africa	1/13/2021	54
Africa	1/13/2022	54

... with 4,358 more rows

Task 1-Subtask 14



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Cho biết số lượng dữ liệu thu thập được là lớn nhất theo date và châu lục

số lượng dữ liệu thu thập được là lớn nhất theo date và châu lục:
54

Task 1-Subtask 15



Dòng cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Cho biết số lượng dữ liệu thu thập được là nhỏ nhất theo date và châu lục

Số lượng dữ liệu thu thập được là nhỏ nhất theo date và châu lục:
1

Task 1-Subtask 16



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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.

Số lượng dữ liệu thể hiện thu thập dữ liệu được ngày 23/3/2020 với châu lục Asia là: 22

Task 1-Subtask 17



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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.

"FJI" "VNM" "ABW" "GIN" "MLT" "SOM" "ATG" "BLZ" "HTI" "MWI" "BEN"
"BMU" "BRB" "ETH" "LBR" "SLV" "TCA" "GEO" "KWT" "LIE" "TZA" "KGZ"
"LBY" "LVA" "SUR" "BWA" "COG" "OMN" "ZMB" "BHS" "IMN" "SEN" "SVK"
"BLR" "MMR" "NIC" "BRN" "MRT" "AGO" "BOL" "CIV" "HND" "MCO" "MLI"
"NZL" "SYR" "ZAF" "KAZ" "LKA" "QAT" "URY" "CMR" "JOR" "PSE" "TGO"
"UZB" "VEN" "ARM" "KEN" "CYP" "EST" "MUS" "NER" "TTO" "CPV" "LTU"
"SAU" "AFG" "GMB" "MNE" "NGA" "OWID_KOS" "ZWE" "AND" "BIH" "CHL"
"COL" "CZE" "ISL" "MKD" "ROU" "COD" "GHA" "SGP" "SRB" "ARE" "CRI"
"GAB" "ISR" "MEX" "PRY" "TUN" "CUW" "HRV" "JAM" "MDA" "RUS" "BFA"
"BGD" "BRA" "CUB" "CYM" "ECU" "PAK" "TUR" "DOM" "MYS" "NOR" "PRT"
"BIH" "GTM" "HUN" "DNK" "LUX" "SDN" "SVN" "AUT" "AZE" "UKR" "BEL"
"DZA" "GUY" "IDN" "IND" "IRL" "PAN" "POL" "ALB" "BGR" "GRC" "SWE"
"LBN" "MAR" "CAN" "DEU" "ARG" "EGY" "FIN" "GBR" "NLD" "PER" "ESP"
"IRQ" "AUS" "THA" "SMR" "USA" "FRA" "IRN"

Task 1-Subtask 18



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Liệt kê iso_code, tên đất nước mà chiều dài iso_code lớn hơn 3

iso_code	location
OWID_KOS	Kosovo

Task 2

Cơ sở lý thuyết

- Tứ phân vị là đại lượng mô tả sự phân bố và sự phân tán của tập dữ liệu
- Tứ phân vị có 3 giá trị, đó là tứ phân vị thứ nhất, thứ nhì, và thứ ba.
- Giá trị tứ phân vị thứ hai Q2 chính bằng giá trị trung vị. Giá trị tứ phân vị thứ nhất Q1 bằng trung vị phần dưới. Giá trị tứ phân vị thứ ba Q3 bằng trung vị phần trên
- Trong thống kê mô tả, độ lệch chuẩn là thước đo độ phân tán của một tập hợp các giá trị so với giá trị trung bình của chúng. Độ lệch chuẩn của 1 giá trị càng thấp nghĩa là giá trị đó càng gần với giá trị trung bình của tập hợp.
- Biểu đồ hộp (Box plot) là biểu đồ diễn tả 5 vị trí phân bố của dữ liệu, đó là: giá trị nhỏ nhất (min), tứ phân vị thứ nhất (Q1), trung vị (median), tứ phân vị thứ 3 (Q3) và giá trị lớn nhất (max).
- Các hàm thường sử dụng trong task 2: names, summary, rownames, colnames, boxplot



Task 2



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Số lượng ca nhiễm mới

Countries	Min	Q1	Q2	Q3	Max	Avg	Std	outlier
Australia	1	12	35	939	175271	4296.903	16288.12	98
New Caledonia	6	38	79	327	3343	328.4436	571.7878	22
New Zealand	1	2	6	50	2336	56.02559	195.7582	71

Task 2



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Số ca tử vong

Countries	Min	Q1	Q2	Q3	Max	Avg	Std	outlier
Australia	1	3	7	14	155	13.89266	15.78844	43
New Caledonia	1	1	3	8	16	4.783333	4.231006	2
New Zealand	1	1	1	1	16	4.783333	0.6798493	8

Task 2



Động cơ nghiên cứu

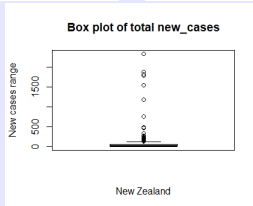
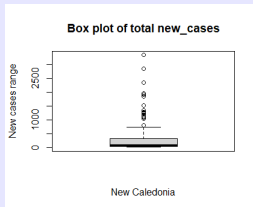
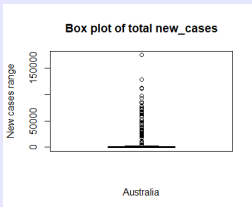
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ boxplot nhiễm coronavirus





Cơ sở lý thuyết

Sử dụng các hàm trong R để xử lý làm sạch dữ liệu ngoài ra có sử dụng vòng lặp để tính toán số lượng dữ liệu cần tìm Các hàm sử dụng:

- `is.na`: Kiểm tra dữ liệu NA
- `na.omit`: Làm sạch dữ liệu

Nhập file và thiết lập một số giá trị ban đầu

- Sử dụng hàm `read.csv` để nhập file "owid-covid-data.csv"
- Tạo dataframe kết hợp các dữ liệu của file lại

Task 3-Subtask 1



Hướng làm

Sử dụng hàm subset để lọc ra các nước cần xử lý trong dataframe đã tạo từ file csv

- Sử dụng hàm is.na và hàm length để đếm dữ liệu không được báo cáo mới

```
19 #description of not report data
20 Countries <- c("Australia","New Caledonia","New Zealand")
21 Infections <- c(
22     sum(is.na(Aus$New_cases))+length(Aus$New_cases[Aus$New_cases==0]),
23     sum(is.na(Donia$New_cases))+length(Donia$New_cases[Donia$New_cases==0]),
24     sum(is.na(Zealand$New_cases))+length(Zealand$New_cases[Zealand$New_cases==0])
25 )
26 Deaths <- c(
27     sum(is.na(Aus$New_deaths))+length(Aus$New_deaths[Aus$New_deaths==0]),
28     sum(is.na(Donia$New_deaths))+length(Donia$New_deaths[Donia$New_deaths==0]),
29     sum(is.na(Zealand$New_deaths))+length(Zealand$New_deaths[Zealand$New_deaths==0])
30 )
31 non_report <- data.frame(Countries,Infections,Deaths)
32 print(non_report)
```

Task 3-Subtask 1



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Có bao nhiêu ngày có số lần dữ liệu không được báo cáo mới.

Countries	Infections	Deaths
Australia	28	437
New Caledonia	507	1183
New Zealand	159	712



Hướng làm

- Khác với subtask 1, ở subtask 2 và 3 này chúng ta sẽ sử dụng hàm `na.omit` để lọc sạch dữ liệu bị NA
 - Ngoài ra chúng ta còn phải quét để lọc ra các dữ liệu bị âm hoặc bằng 0
- Ta sẽ làm như code trong hình sau:

```
ex_3 <- na.omit(ex3)
aus <- subset(ex_3, ex_3$Location=="Australia")
donia <- subset(ex_3, ex_3$Location=="New Caledonia")
zealand <- subset(ex_3, ex_3$Location=="New Zealand")

aus_cases <- aus$New_cases[aus$New_cases>0]
donia_cases <- donia$New_cases[donia$New_cases>0]
zealand_cases <- zealand$New_cases[zealand$New_cases>0]
```


Task 3-Subtask 2+3



Hướng làm

Cuối cùng ở subtask 2 và 3 này ta sẽ dùng hàm `length` có điều kiện để đếm dữ liệu thỏa mãn yêu cầu bài trong đó có sử dụng thêm hàm `min`, `max` để tìm min, max của tập dữ liệu

```
Infections <- c(  
  length(aus_cases[aus_cases==min(aus_cases)]),  
  length(donia_cases[donia_cases==min(donia_cases)]),  
  length(zealand_cases[zealand_cases==min(zealand_cases)])  
)  
Deaths <- c(  
  length(aus_deaths[aus_deaths==min(aus_deaths)]),  
  length(donia_deaths[donia_deaths==min(donia_deaths)]),  
  length(zealand_deaths[zealand_deaths==min(zealand_deaths)])  
)
```

Task 3-Subtask 2



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Có bao nhiêu ngày có số ca nhiễm/ tử vong là thấp nhất được báo cáo mới

Countries	Infections	Deaths
Australia	2	50
New Caledonia	1	20
New Zealand	75	33

Task 3-Subtask 3



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Có bao nhiêu ngày có số ca nhiễm/ tử vong là cao nhất được báo cáo mới

Countries	Infections	Deaths
Australia	1	1
New Caledonia	1	2
New Zealand	1	1

Task 3-Subtask 5+6



Hướng làm

- Ở subtask 5 và 6 này có cách làm như nhau, chúng ta sẽ sử dụng lại hàm `is.na` để xác định dữ liệu có NA hay không
- Sử dụng vòng lặp `for`, `while` và cấu trúc rẽ nhánh `if`, `else` cộng với một chút tư duy lập trình để ta giải quyết được bài toán "dữ liệu dài nhất/ngắn nhất liên tiếp" Có thể minh họa code như hình sau:

```
#Australia-no-New cases
aus_min_no_cases=length(Aus$New_cases)
aus_max_no_cases=0
for(i in 1:length(Aus$New_cases)){
  if(is.na(Aus$New_cases[i]) || Aus$New_cases[i]==0)
  {
    count=0
    while(is.na(Aus$New_cases[i]) || Aus$New_cases[i]==0)
    {
      count= count+1
      i=i+1}
    if(i>length(Aus$New_cases)){break}
  }
  if(count<aus_min_no_cases && count>=1){aus_min_no_cases=count}
  if(count>aus_max_no_cases && count>=1){aus_max_no_cases=count}
  i=i-1
}
if(aus_min_no_cases>aus_max_no_cases){aus_min_no_cases=0}
```

Task 3-Subtask 5+6



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Cho biết số ngày ngắn nhất/dài nhất liên tiếp mà không có dữ liệu được báo cáo

Countries	min_no_cases	max_no_cases	min_no_deaths	max_no_deaths
Australia	1	21	1	105
New Caledonia	1	52	1	540
New Zealand	1	24	1	199

Task 3-Subtask 7+8



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Cho biết số ngày ngắn nhất/dài nhất liên tiếp mà không có người nhiễm bệnh mới

Countries	min_no_new	max_no_new
Australia	1	1
New Caledonia	1	4
New Zealand	1	24



Cơ sở lý thuyết

- Biểu đồ tần số tích lũy (cumulative frequency plots) biểu thị những thông tin dạng tích lũy.
- Biểu đồ tần số tương đối thể hiện nội dung như biểu đồ tần số tích lũy, chỉ khác một điều đơn vị tính của các cột là tỉ lệ phần trăm.
- Biểu đồ phổ là thể hiện sự phân phối về số lượng của một tập dữ liệu.
- Các hàm sử dụng trong task 4: ggplot, for, if geom_bar, theme_bw, theme_grey, labs, data.frame

Task 4-Subtask 1



Động cơ nghiên cứu

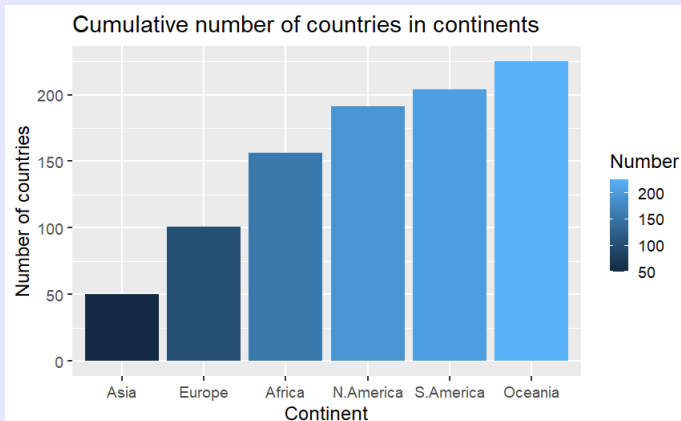
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ tần số tích lũy quốc gia ở các châu lục



Task 4-Subtask 2



Động cơ nghiên cứu

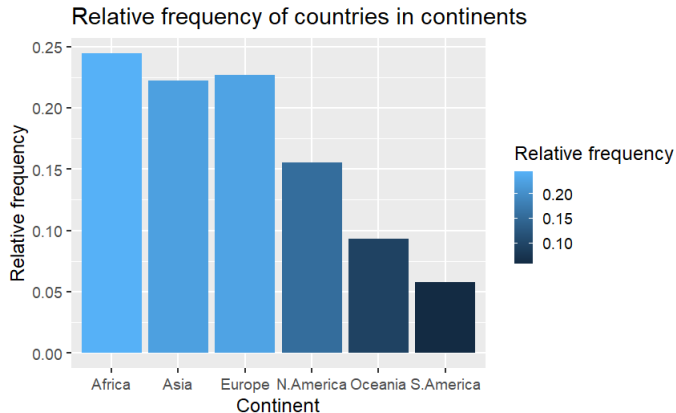
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ tần số tương đối quốc gia ở các châu lục



Task 4-Subtask 3



Động cơ nghiên cứu

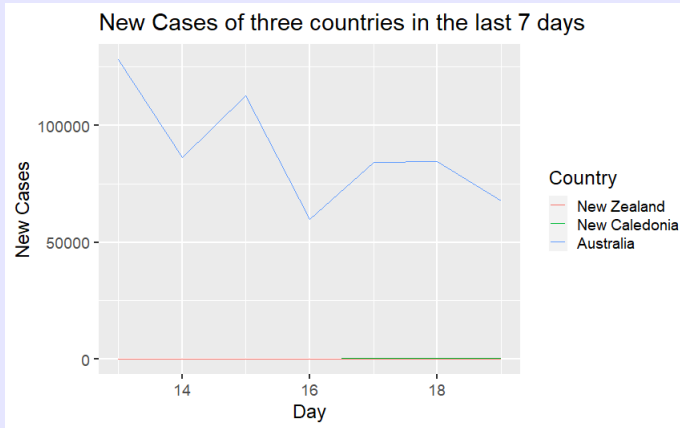
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 4-Subtask 4



Động cơ nghiên cứu

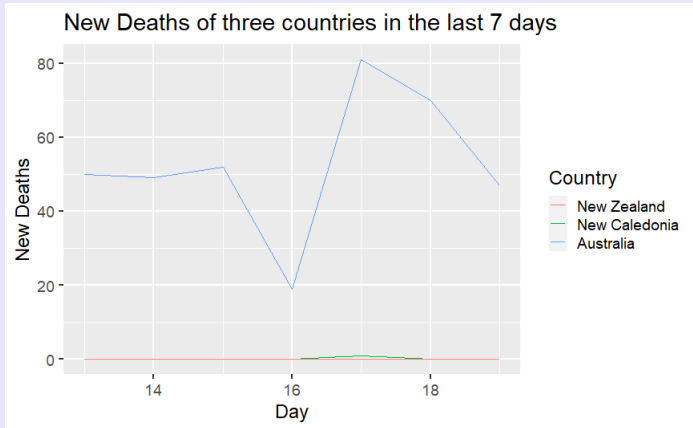
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 4-Subtask 5



Động cơ nghiên cứu

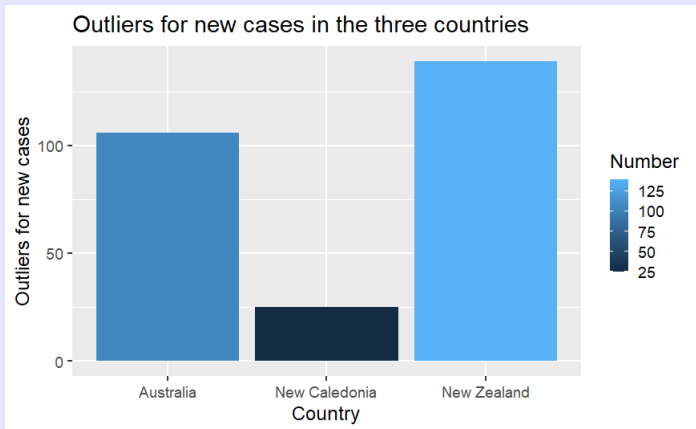
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ phổ đất nước xuất hiện outliers cho nhiễm bệnh



Task 4-Subtask 6



Động cơ nghiên cứu

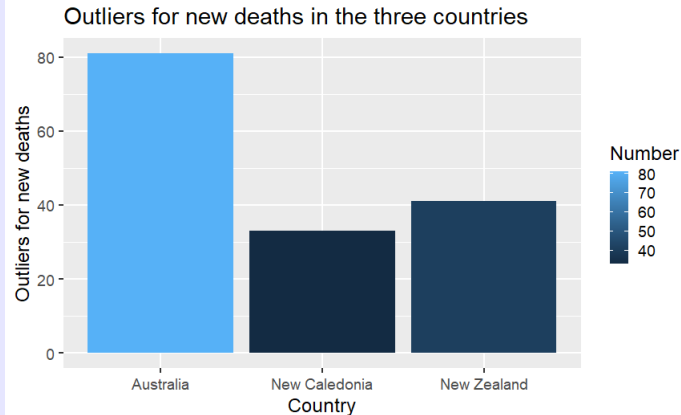
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ phổ đất nước xuất hiện outliers cho tử vong



Task 5



Cơ sở lý thuyết

- Nhiễm bệnh/tử vong tích lũy: Số lượng nhiễm bệnh/tử vong của 1 ngày trong tháng bằng số lượng nhiễm bệnh/tử vong trong ngày hôm đó cộng với tổng số lượng nhiễm bệnh/tử vong của tất cả các ngày trước đó
- Các hàm dùng trong task 5
`subset`, `for`, `mutate`, `as.Date`, `data.frame`, `ggplot`, `goemline`,
`theme_bw`, `labs`, `theme_grey`, `scale_color_discrete`

Task 5

Hướng làm

Đầu tiên ta lấy dữ liệu từ file `owid-covid-data` vào `mydata` và lọc lại các dữ liệu của các nước để chia thành các dữ liệu nhỏ phù hợp theo cú pháp: Nước-tháng-năm

```
{
  mydata<-owid_covid_data
  mydata$date <- as.Date(mydata$date,"%m/%d/%Y")

  australia<-subset(mydata,mydata$location=="Australia")
  new_caledonia<-subset(mydata,mydata$location=="New Caledonia")
  new_zealand<-subset(mydata,mydata$location=="New Zealand")

  australia_1_2020<-australia[australia$date=="2020-01-01"& australia$date<="2020-01-31",]
  australia_7_2020<-australia[australia$date=="2020-07-01"& australia$date<="2020-07-31",]
  australia_8_2020<-australia[australia$date=="2020-08-01"& australia$date<="2020-08-31",]
  australia_10_2020<-australia[australia$date=="2020-10-01"& australia$date<="2020-10-31",]

  australia_1_2021<-australia[australia$date=="2021-01-01"& australia$date<="2021-01-31",]
  australia_7_2021<-australia[australia$date=="2021-07-01"& australia$date<="2021-07-31",]
  australia_8_2021<-australia[australia$date=="2021-08-01"& australia$date<="2021-08-31",]
  australia_10_2021<-australia[australia$date=="2021-10-01"& australia$date<="2021-10-31",]

  australia_1_2022<-australia[australia$date=="2022-01-01"& australia$date<="2022-01-31",]

  new_caledonia_1_2020<-new_caledonia[new_caledonia$date=="2020-01-01"& new_caledonia$date<="2020-01-31",]
  new_caledonia_7_2020<-new_caledonia[new_caledonia$date=="2020-07-01"& new_caledonia$date<="2020-07-31",]
  new_caledonia_8_2020<-new_caledonia[new_caledonia$date=="2020-08-01"& new_caledonia$date<="2020-08-31",]
  new_caledonia_10_2020<-new_caledonia[new_caledonia$date=="2020-10-01"& new_caledonia$date<="2020-10-31",]

  new_caledonia_1_2021<-new_caledonia[new_caledonia$date=="2021-01-01"& new_caledonia$date<="2021-01-31",]
  new_caledonia_7_2021<-new_caledonia[new_caledonia$date=="2021-07-01"& new_caledonia$date<="2021-07-31",]
  new_caledonia_8_2021<-new_caledonia[new_caledonia$date=="2021-08-01"& new_caledonia$date<="2021-08-31",]
  new_caledonia_10_2021<-new_caledonia[new_caledonia$date=="2021-10-01"& new_caledonia$date<="2021-10-31",]

  new_caledonia_1_2022<-new_caledonia[new_caledonia$date=="2022-01-01"& new_caledonia$date<="2022-01-31",]

  new_zealand_1_2020<-new_zealand[new_zealand$date=="2020-01-01"& new_zealand$date<="2020-01-31",]
  new_zealand_7_2020<-new_zealand[new_zealand$date=="2020-07-01"& new_zealand$date<="2020-07-31",]
  new_zealand_8_2020<-new_zealand[new_zealand$date=="2020-08-01"& new_zealand$date<="2020-08-31",]
  new_zealand_10_2020<-new_zealand[new_zealand$date=="2020-10-01"& new_zealand$date<="2020-10-31",]

  new_zealand_1_2021<-new_zealand[new_zealand$date=="2021-01-01"& new_zealand$date<="2021-01-31",]
  new_zealand_7_2021<-new_zealand[new_zealand$date=="2021-07-01"& new_zealand$date<="2021-07-31",]
  new_zealand_8_2021<-new_zealand[new_zealand$date=="2021-08-01"& new_zealand$date<="2021-08-31",]
  new_zealand_10_2021<-new_zealand[new_zealand$date=="2021-10-01"& new_zealand$date<="2021-10-31",]

  new_zealand_1_2022<-new_zealand[new_zealand$date=="2022-01-01"& new_zealand$date<="2022-01-31",]
}#chia thành dữ liệu nhỏ
```



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Hướng làm

Ta sẽ vẽ biểu đồ cho từng nước ở các năm với đường biểu diễn là tháng: January, July, August, October. Tuy nhiên ở câu 4 và câu 5 ta chỉ dùng 2 tháng cuối năm: November, December

Trong bài có một số dữ liệu hiển thị N/A nên trong biểu đồ vẽ thì dữ liệu tại khu vực đó không được thể hiện đồng thời vào năm 2022 chỉ cung cấp dữ liệu đến tháng 2 nên ta chỉ thể hiện được mỗi tháng 1 (January)

VD: Nếu October của New Caledonia dữ liệu là N/A thì chúng ta chỉ còn sử dụng các tháng : January, July, August

Task 5-Subtask 1



Động cơ nghiên cứu

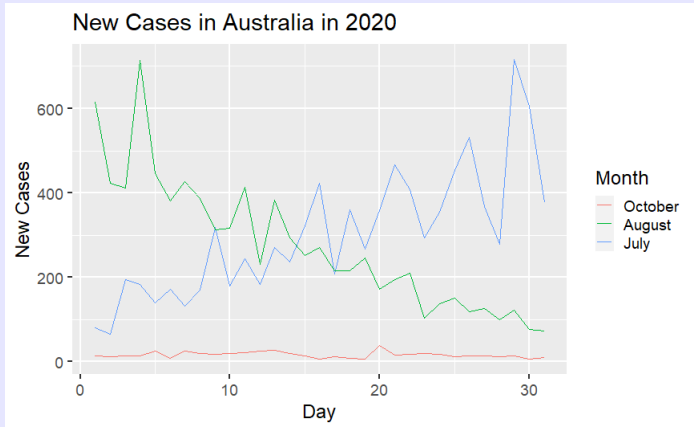
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của Australia vào năm 2020





Động cơ nghiên cứu

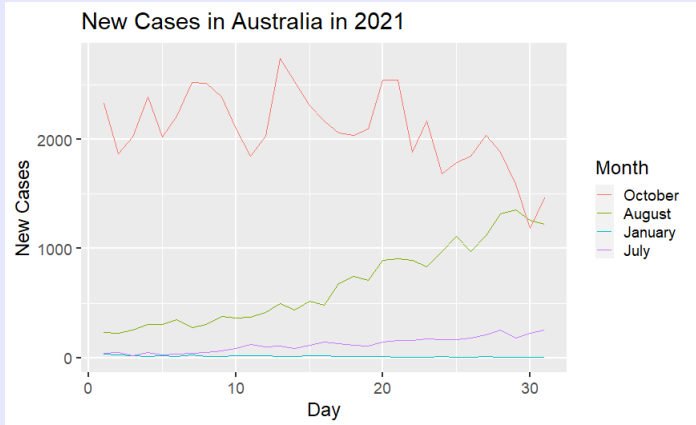
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của Australia vào năm 2021





Động cơ nghiên cứu

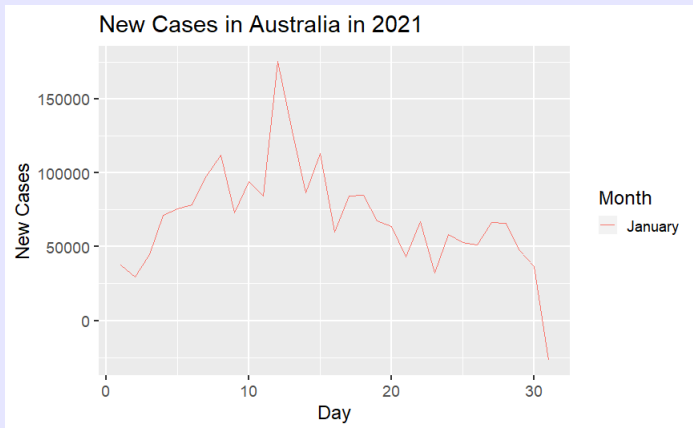
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của Australia vào năm 2022





Động cơ nghiên cứu

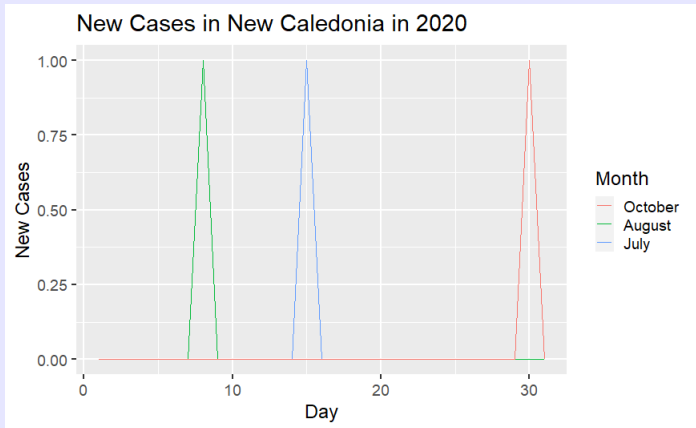
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của New Caledonia vào năm 2020





Động cơ nghiên cứu

Mục tiêu

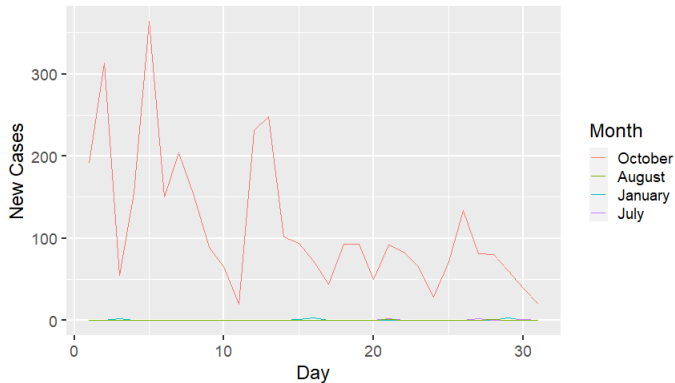
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của New Caledonia vào năm 2021

New Cases in New Caledonia in 2021





Động cơ nghiên cứu

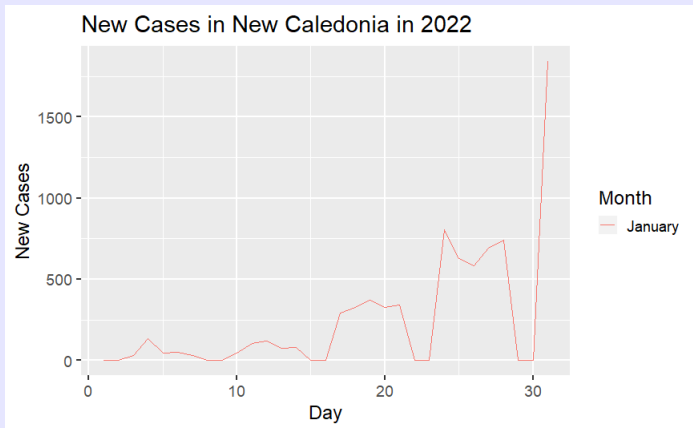
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của New Caledonia vào năm 2022





Động cơ nghiên cứu

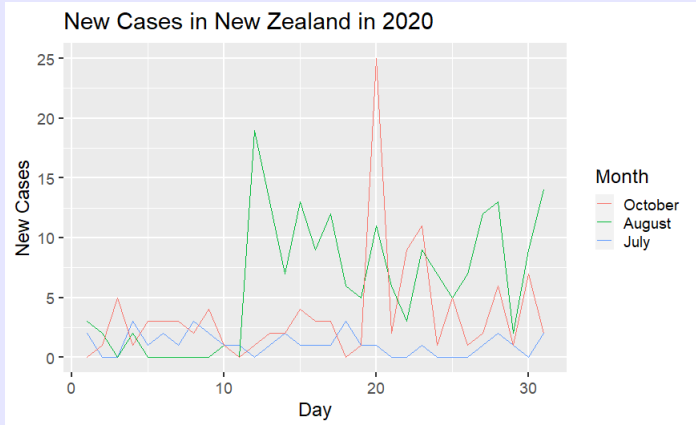
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của New Zealand vào năm 2020





Động cơ nghiên cứu

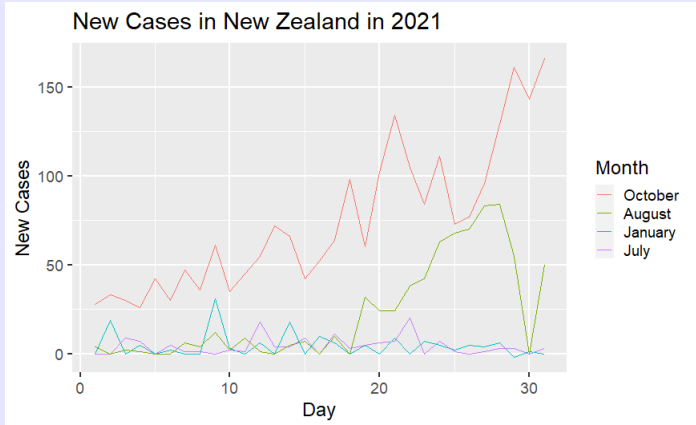
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của New Zealand vào năm 2021





Động cơ nghiên cứu

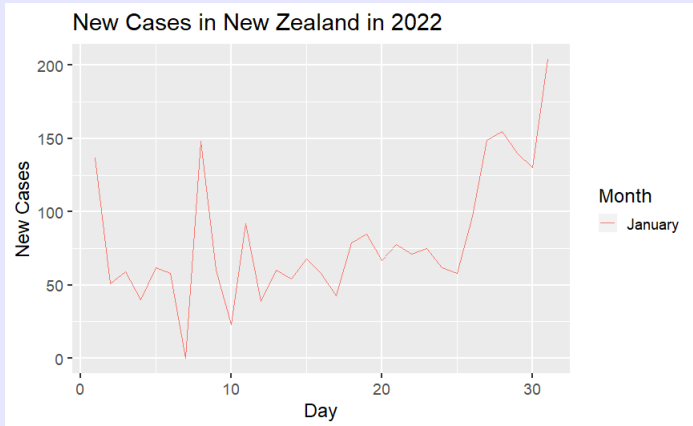
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của New Zealand vào năm 2022



Task 5-Subtask 2



Động cơ nghiên cứu

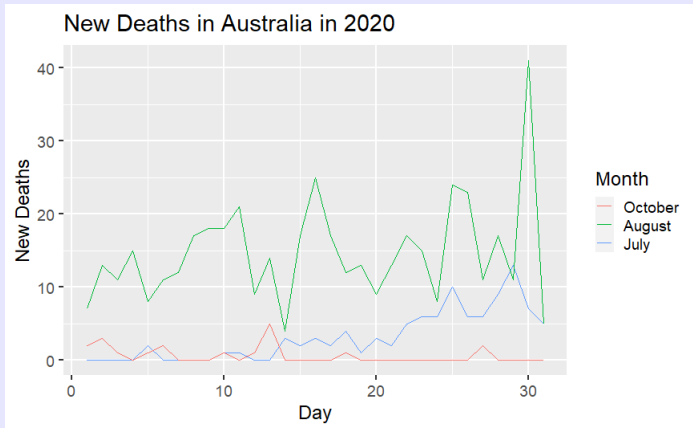
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của Australia vào năm 2020





Động cơ nghiên cứu

Mục tiêu

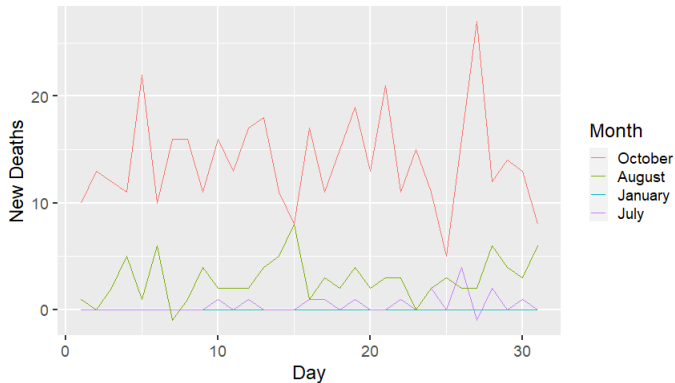
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của Australia vào năm 2021

New Deaths in Australia in 2021





Động cơ nghiên cứu

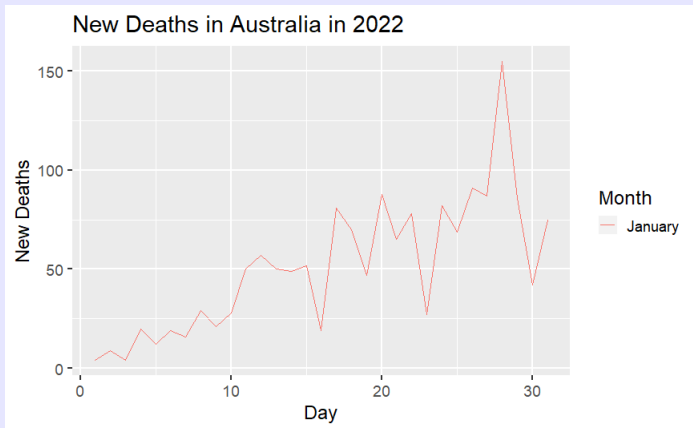
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của Australia vào năm 2022





Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

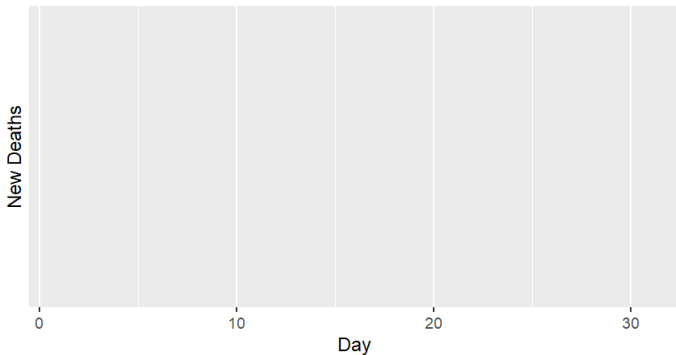
Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của New Caledonia vào năm 2020

New Deaths in New Caledonia in 2020

[NO DATA WAS GIVEN]





Động cơ nghiên cứu

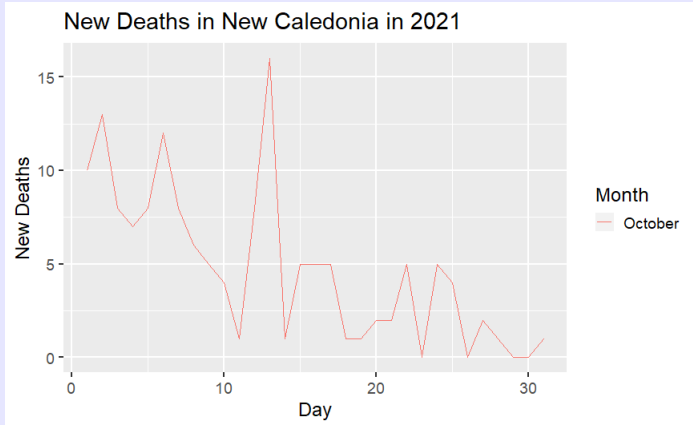
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của New Caledonia vào năm 2021





Động cơ nghiên cứu

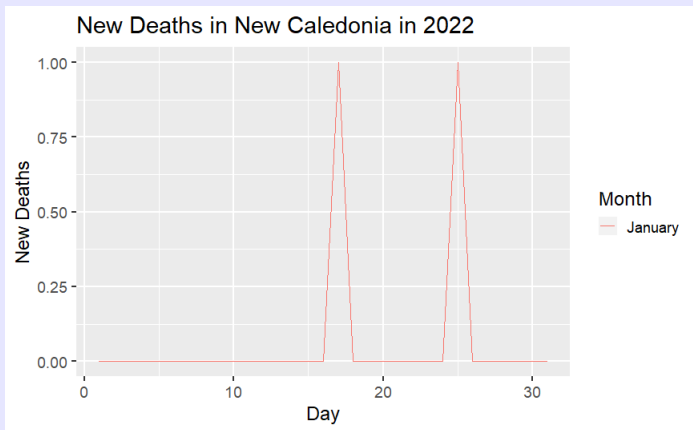
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của New Caledonia vào năm 2022





Động cơ nghiên cứu

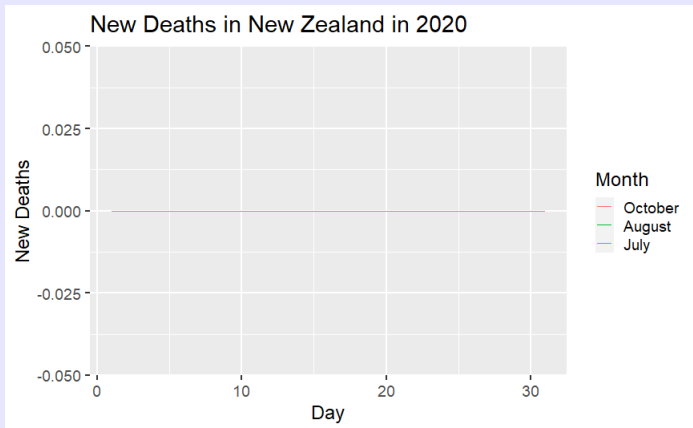
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của New Zealand vào năm 2020





Động cơ nghiên cứu

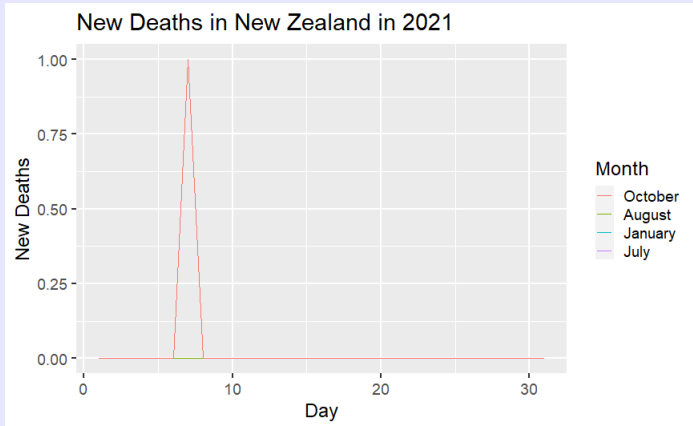
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của New Zealand vào năm 2021





Động cơ nghiên cứu

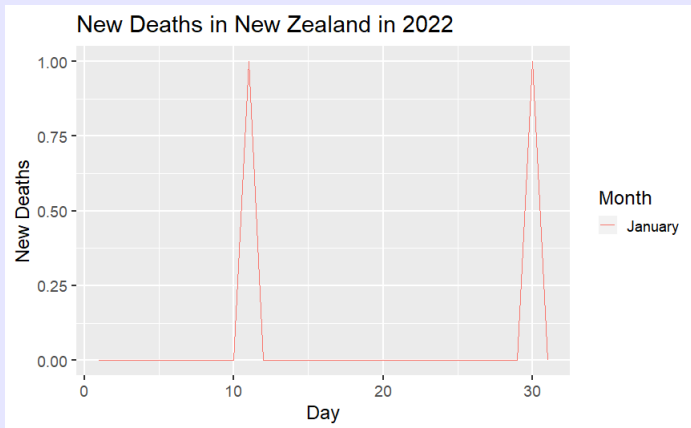
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của New Zealand vào năm 2022



Task 5-Subtask 3



Động cơ nghiên cứu

Mục tiêu

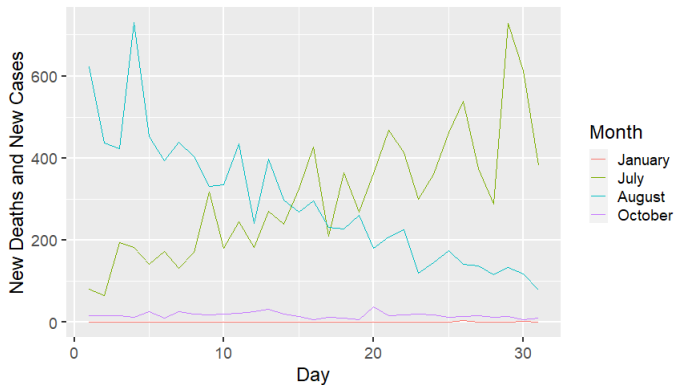
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của Australia vào năm 2020

New Deaths and New Cases in Australia in 2020





Động cơ nghiên cứu

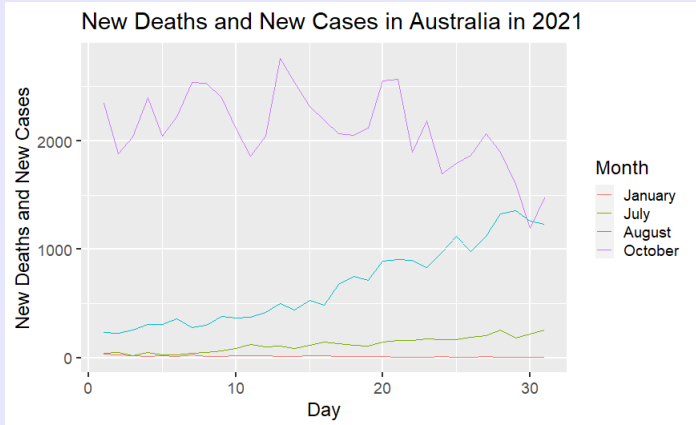
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của Australia vào năm 2021





Động cơ nghiên cứu

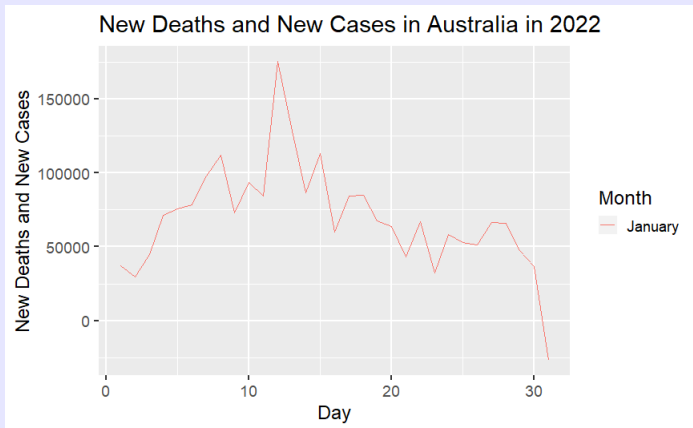
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của Australia vào năm 2022





Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

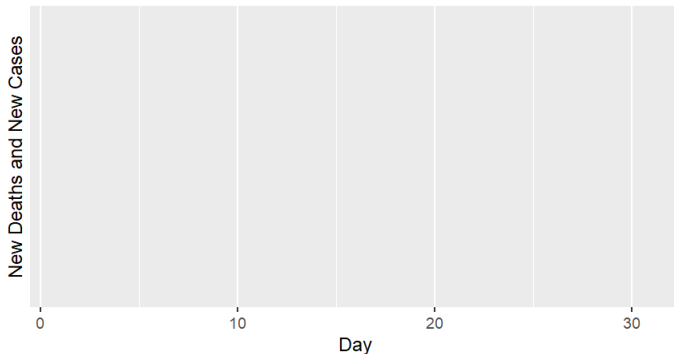
Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của New Caledonia vào năm 2020

New Deaths and New Cases in New Caledonia in 2020

[NO DATA WAS GIVEN]





Động cơ nghiên cứu

Mục tiêu

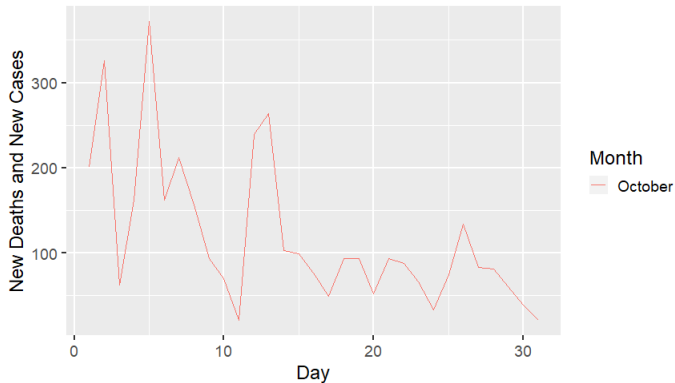
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của New Caledonia vào năm 2021

New Deaths and New Cases in New Caledonia in 2021





Động cơ nghiên cứu

Mục tiêu

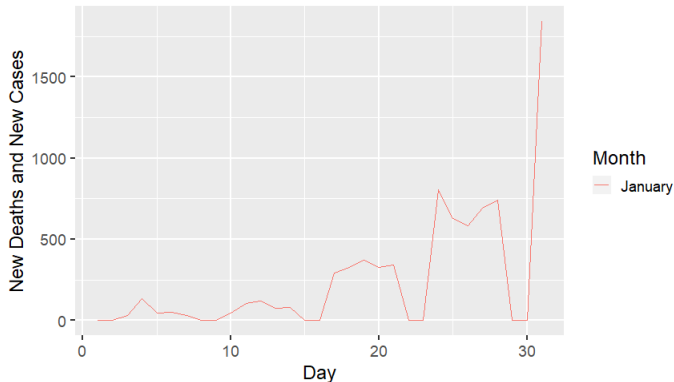
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của New Caledonia vào năm 2022

New Deaths and New Cases in New Caledonia in 2022





Động cơ nghiên cứu

Mục tiêu

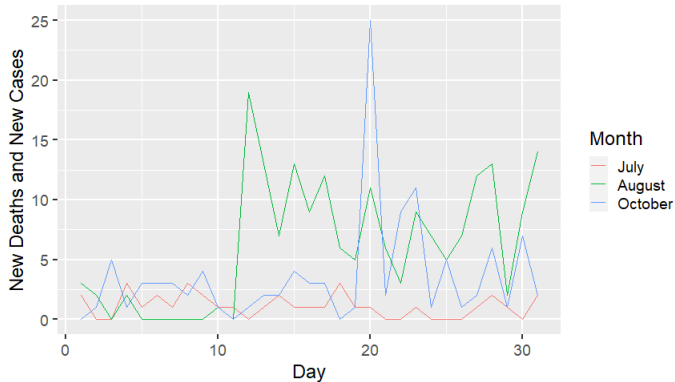
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của New Zealand vào năm 2020

New Deaths and New Cases in New Zealand in 2020 (NO DA





Động cơ nghiên cứu

Mục tiêu

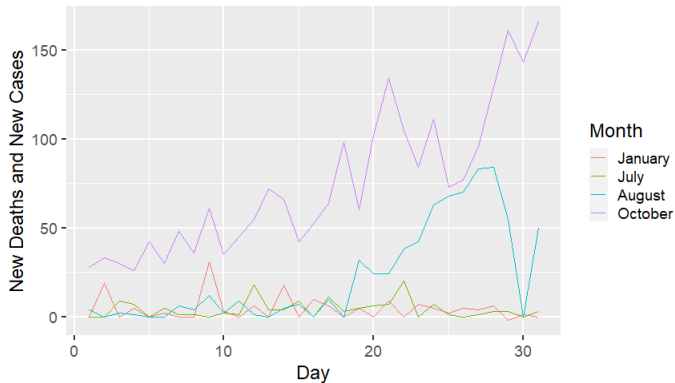
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của New Zealand vào năm 2021

New Deaths and New Cases in New Zealand in 2021





Động cơ nghiên cứu

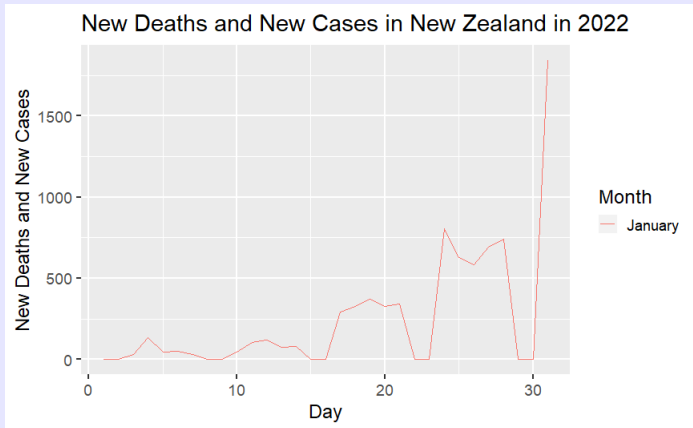
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của New Zealand vào năm 2022



Task 5-Subtask 4



Động cơ nghiên cứu

Mục tiêu

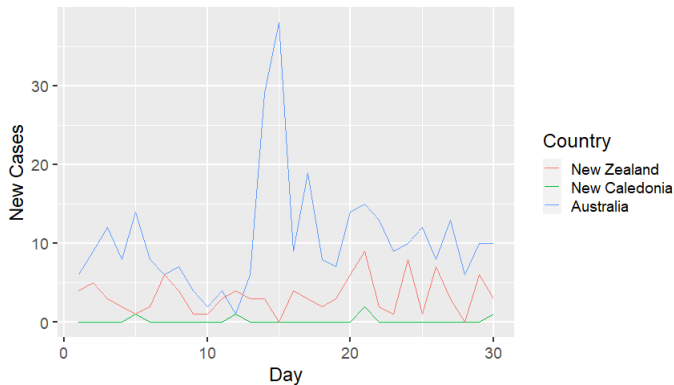
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của tháng 11 năm 2020

New Cases on November 2020 in the three countries



Task 5-Subtask 4



Động cơ nghiên cứu

Mục tiêu

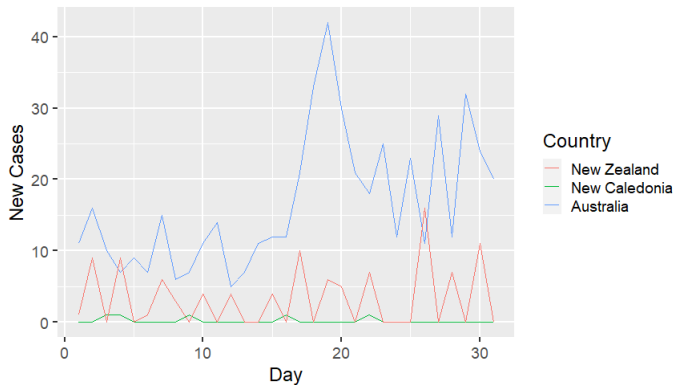
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của tháng 12 năm 2020

New Cases on December 2020 in the three countries





Động cơ nghiên cứu

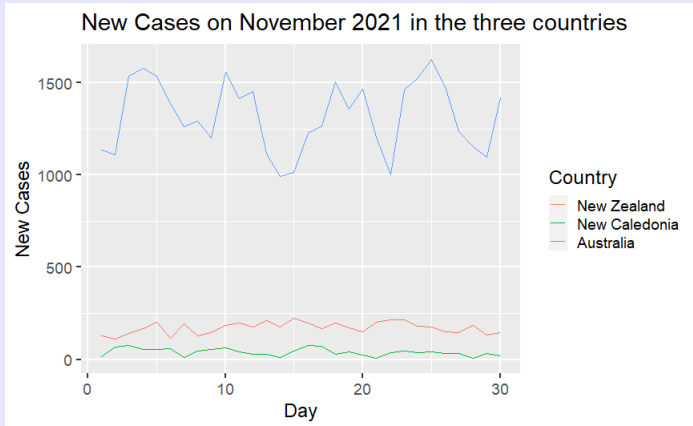
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của tháng 11 năm 2021





Động cơ nghiên cứu

Mục tiêu

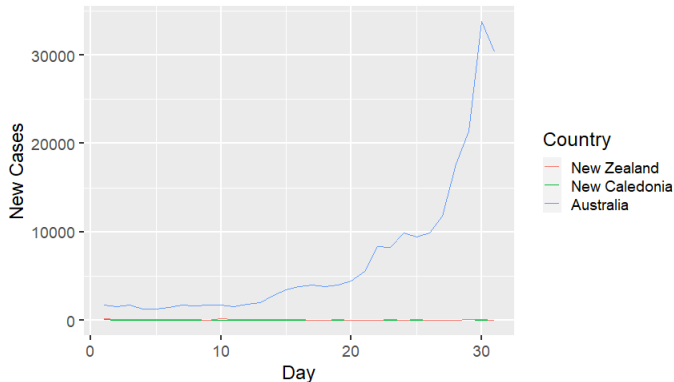
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh của tháng 12 năm 2021

New Cases on December 2021 in the three countries



Task 5-Subtask 5



Động cơ nghiên cứu

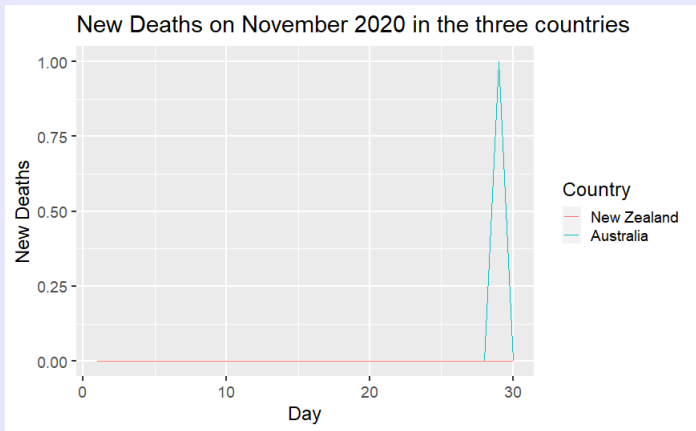
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của tháng 11 năm 2020





Động cơ nghiên cứu

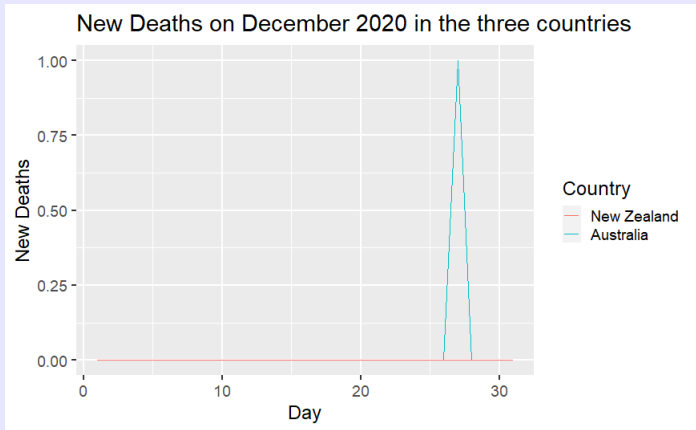
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của tháng 12 năm 2020





Động cơ nghiên cứu

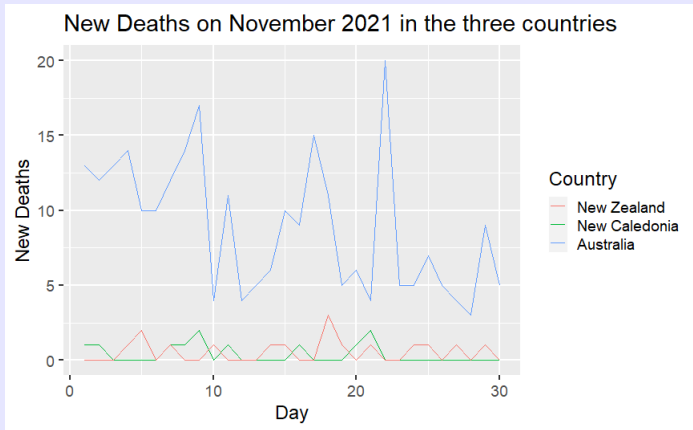
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của tháng 11 năm 2021





Động cơ nghiên cứu

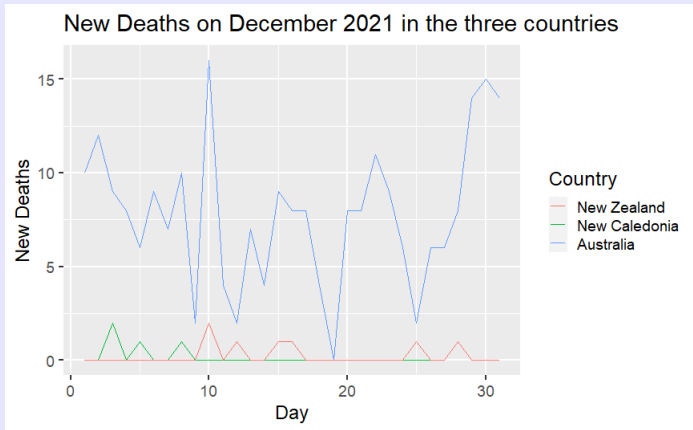
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong của tháng 12 năm 2021



Task 5-Subtask 6



Động cơ nghiên cứu

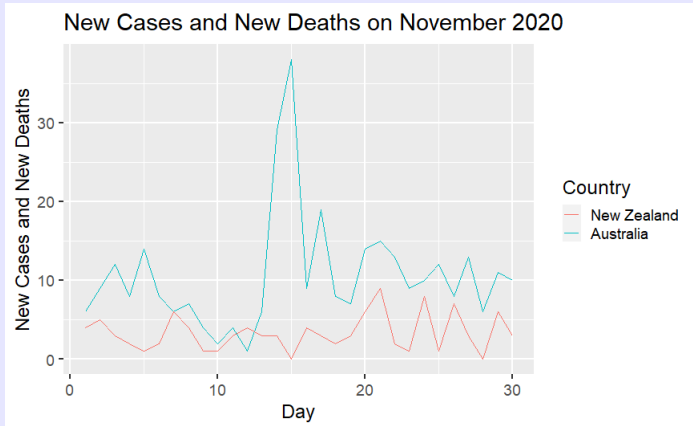
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của tháng 11 năm 2020





Động cơ nghiên cứu

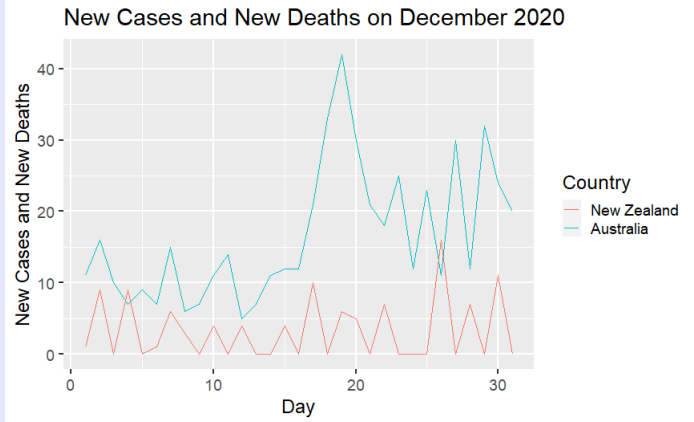
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của tháng 12 năm 2020





Động cơ nghiên cứu

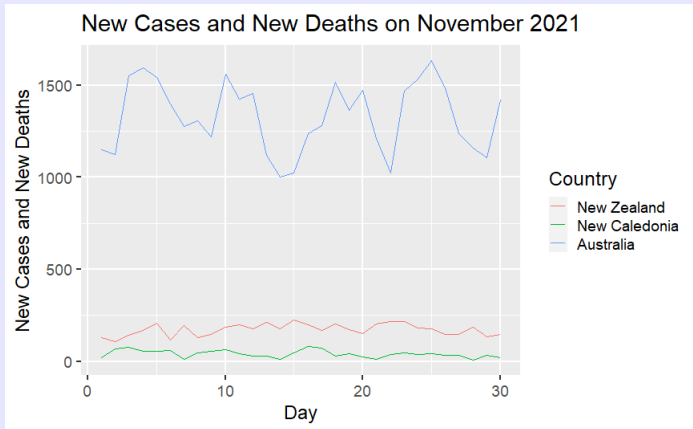
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của tháng 11 năm 2021





Động cơ nghiên cứu

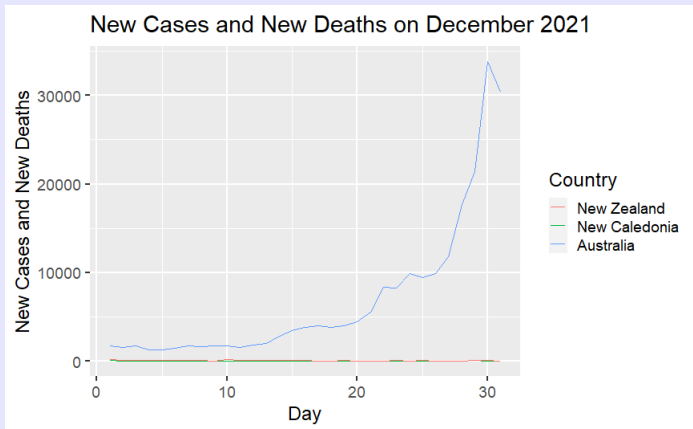
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu gồm nhiễm bệnh và tử vong của tháng 12 năm 2021





Động cơ nghiên cứu

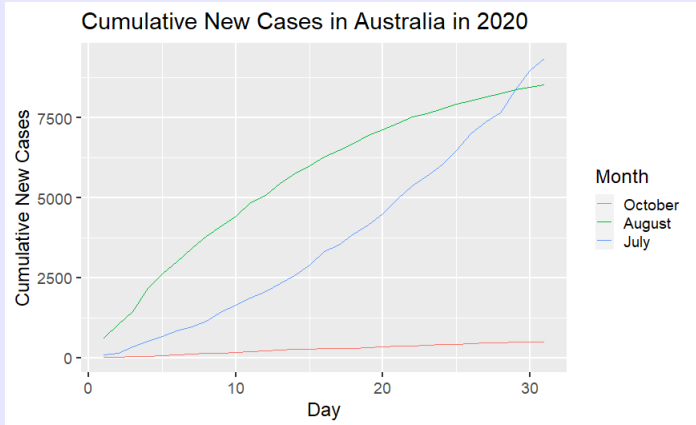
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của Australia năm 2020





Động cơ nghiên cứu

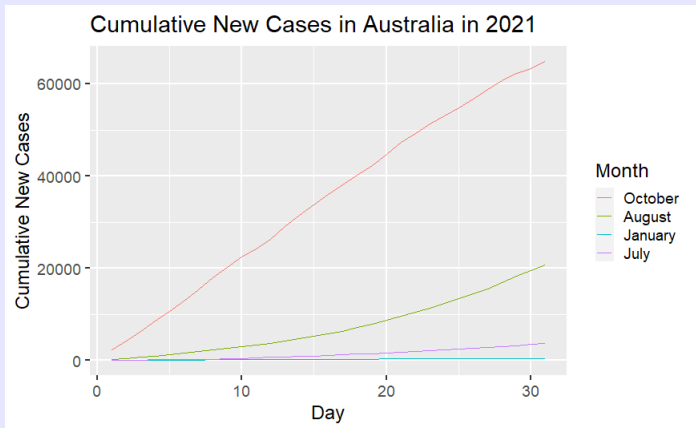
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của Australia năm 2021





Động cơ nghiên cứu

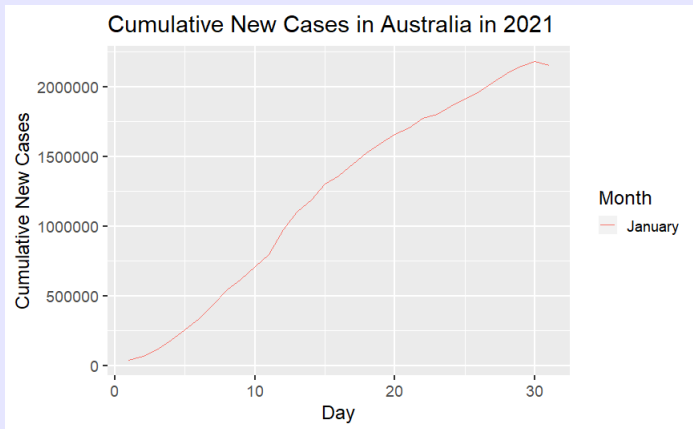
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của Australia năm 2022





Động cơ nghiên cứu

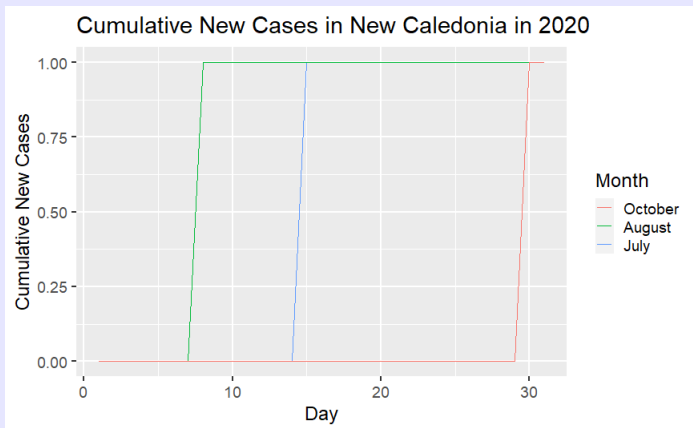
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của New Caledonia năm 2020





Động cơ nghiên cứu

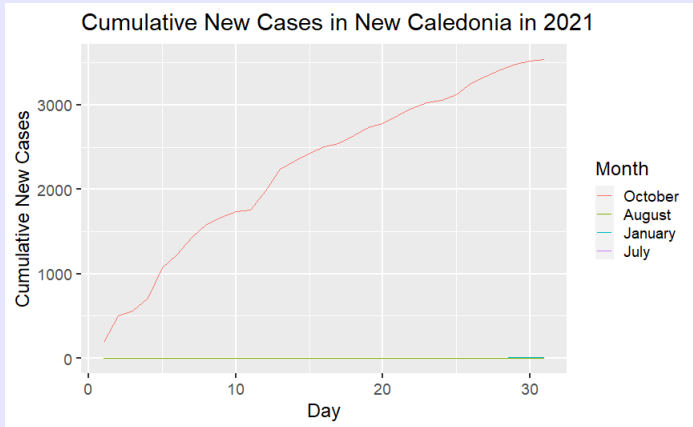
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của New Caledonia năm 2022





Động cơ nghiên cứu

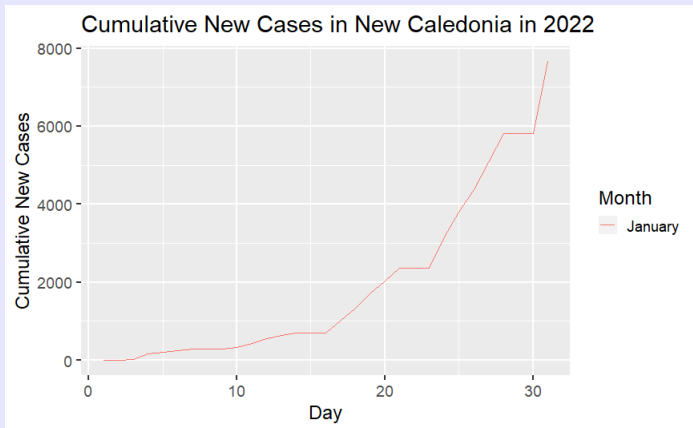
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của New Caledonia năm 2022





Động cơ nghiên cứu

Mục tiêu

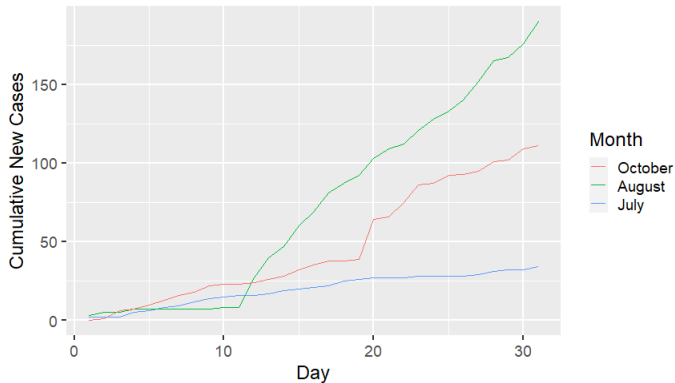
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của New Zealand năm 2020

Cumulative New Cases in New Zealand in 2020





Động cơ nghiên cứu

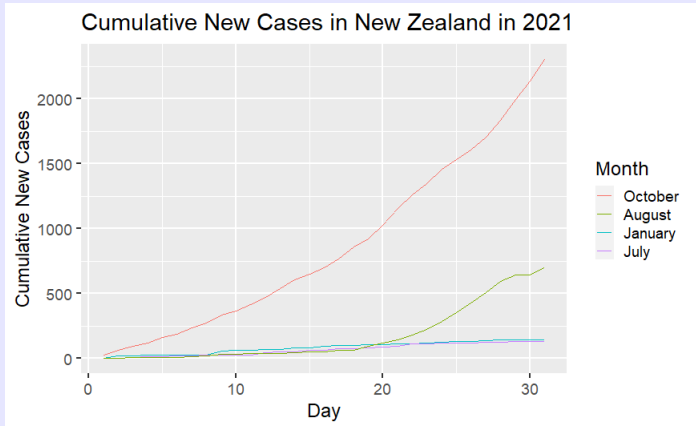
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của New Zealand năm 2021





Động cơ nghiên cứu

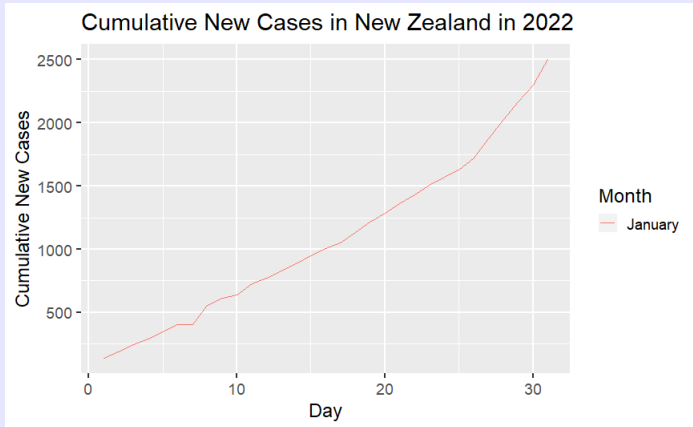
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy của New Zealand năm 2022



Task 5-Subtask 8



Động cơ nghiên cứu

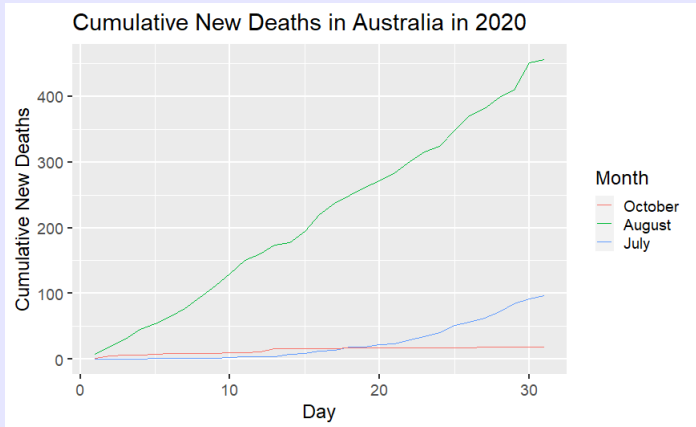
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của Australia năm 2020





Động cơ nghiên cứu

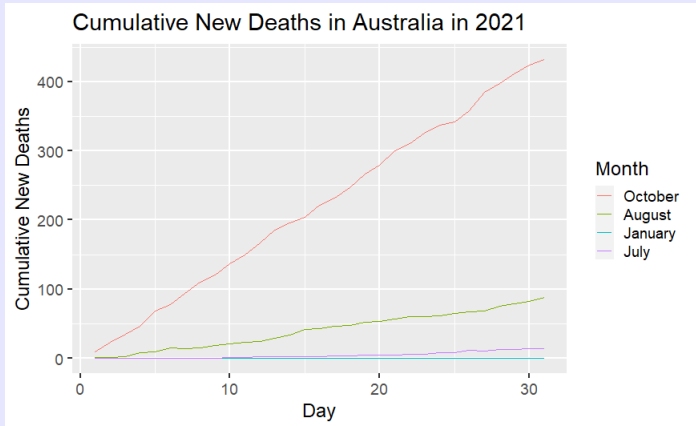
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của Australia năm 2021





Động cơ nghiên cứu

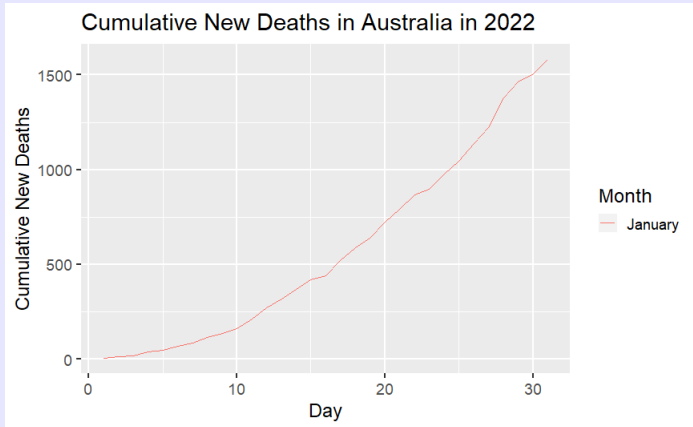
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của Australia năm 2022





Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

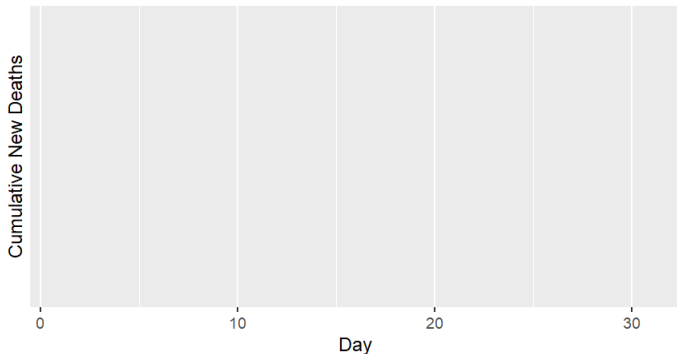
Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của New Caledonia năm 2020

Cumulative New Deaths in New Caledonia in 2020

[NO DATA GIVEN]





Động cơ nghiên cứu

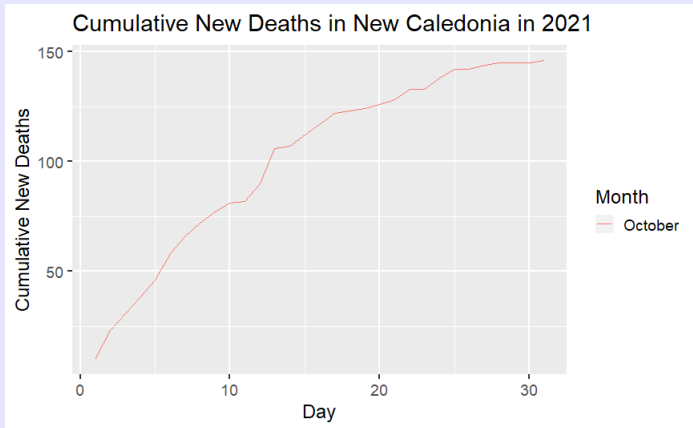
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của New Caledonia năm 2021





Động cơ nghiên cứu

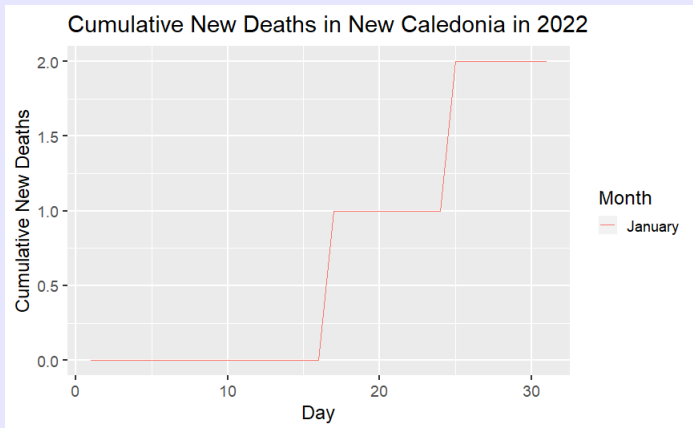
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của New Caledonia năm 2022





Dòng cơ nghiên cứu

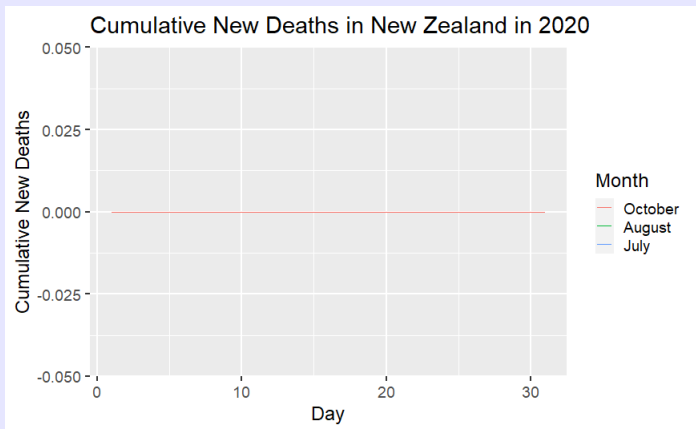
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của New Zealand năm 2020





Động cơ nghiên cứu

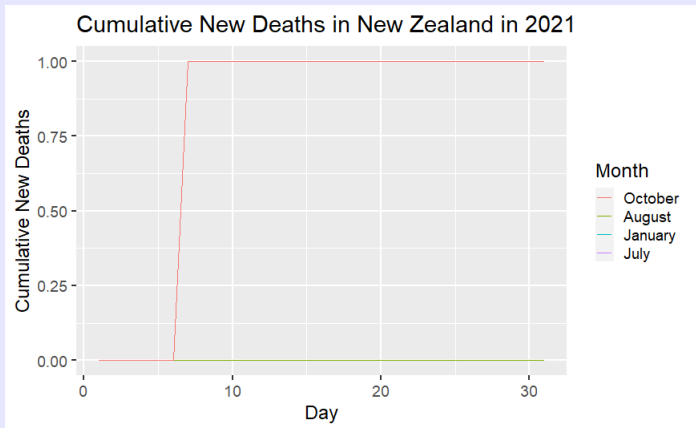
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của New Zealand năm 2021





Động cơ nghiên cứu

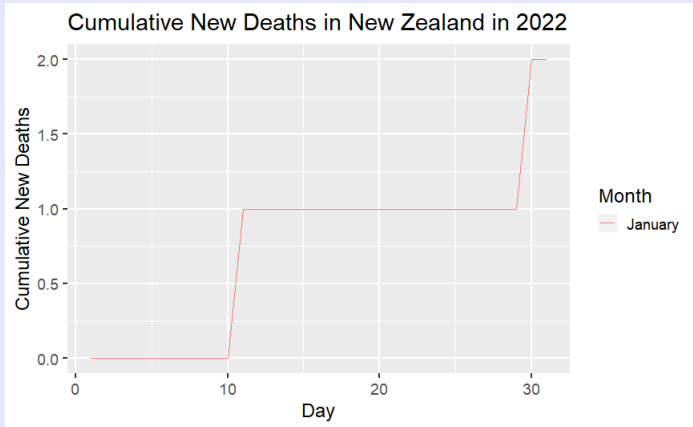
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy của New Zealand năm 2022



Task 6

Cơ sở lý thuyết

- Yêu cầu làm biểu đồ theo thời gian là tháng, những tháng thuộc về nhóm là 1, 7, 8, 10.
- Các quốc gia thuộc về nhóm là Australia, New Caledonia, New Zealand
- Sử dụng ngôn ngữ lập trình R để lọc dữ liệu theo từng tháng của nhóm.
- Tính giá trị trung bình 7 ngày gần nhất theo quy luật sau:
 - + Ngày 1 của tháng giữ nguyên giá trị
 - + Ngày 2 của tháng tính các giá trị trung bình giữa ngày 1 và ngày 2
 - ...
 - + Từ ngày thứ 7 trở đi lấy trung bình của 7 ngày gần nhất kể từ đó
- Sử dụng package ggplot2 để vẽ các biểu đồ theo yêu cầu sau khi đã xử lý số liệu



Task 6 - Subtask 1



Động cơ nghiên cứu

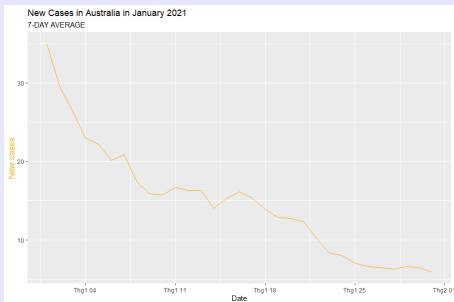
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 1



Động cơ nghiên cứu

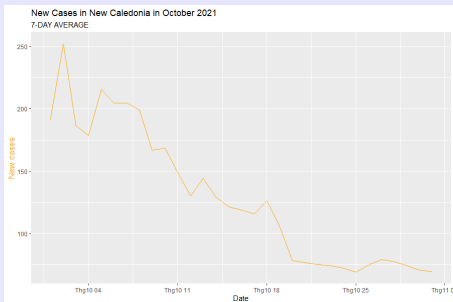
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 1



Động cơ nghiên cứu

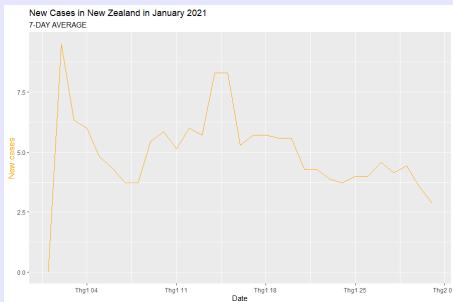
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 2



Động cơ nghiên cứu

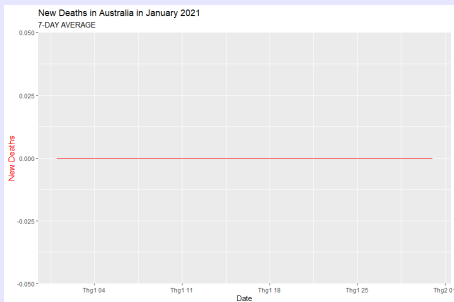
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 2



Động cơ nghiên cứu

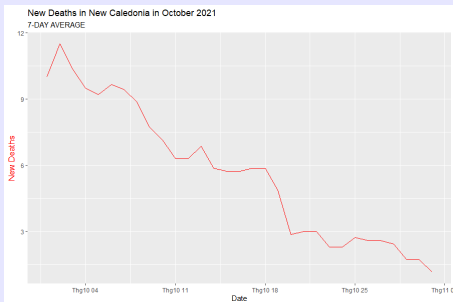
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 2



Động cơ nghiên cứu

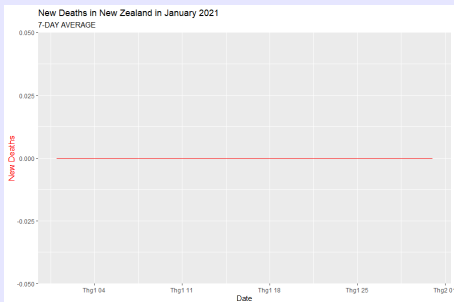
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 3

CẤU TRÚC RỜI RẠC
CHO KHMT

Huỳnh Tường Nguyễn,
Nguyễn Ngọc Lê



Động cơ nghiên cứu

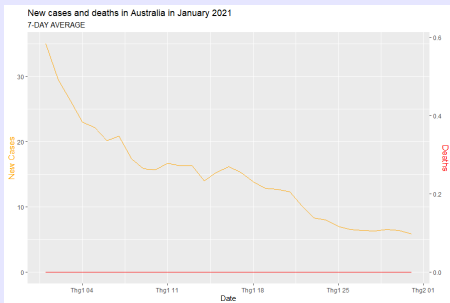
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 3



Động cơ nghiên cứu

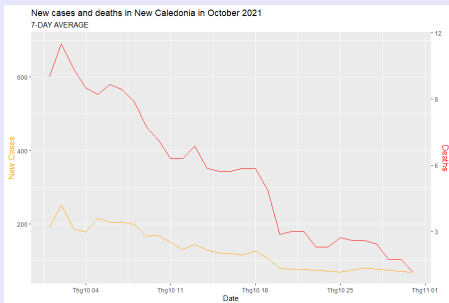
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 3

CẤU TRÚC RỜI RẠC
CHO KHMT

Huynh Tuong Nguyen,
Nguyen Ngoc Le



Động cơ nghiên cứu

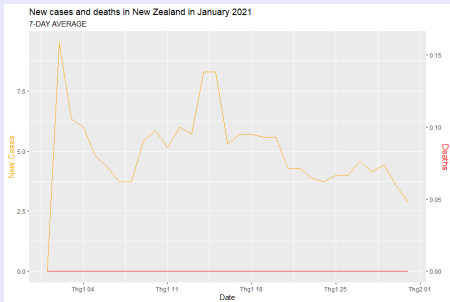
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 4



Động cơ nghiên cứu

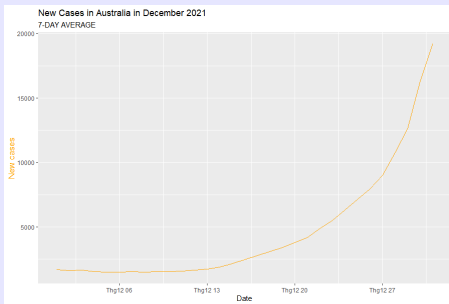
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 4



Động cơ nghiên cứu

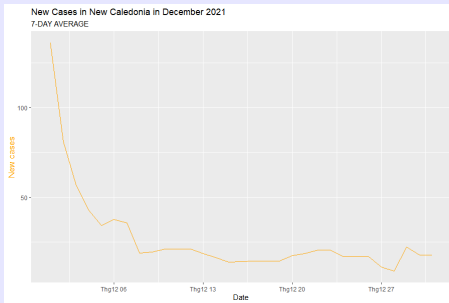
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 4



Động cơ nghiên cứu

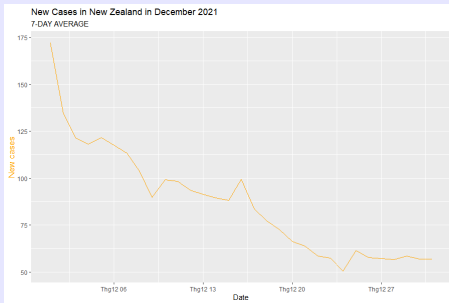
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 5

CẤU TRÚC RỜI RẠC
CHO KHMT

Huynh Tuong Nguyen,
Nguyen Ngoc Le



Động cơ nghiên cứu

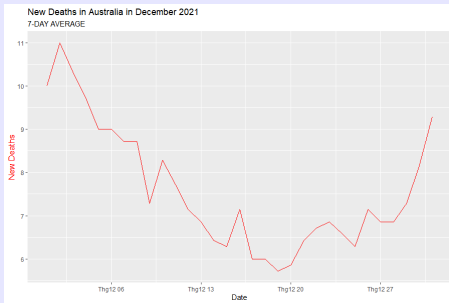
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 5



Động cơ nghiên cứu

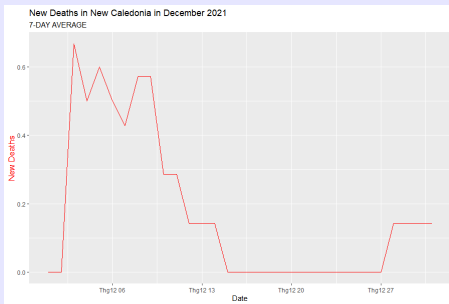
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 5



Động cơ nghiên cứu

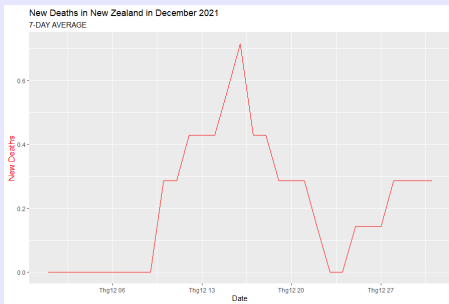
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 6

CẤU TRÚC RỜI RẠC
CHO KHMT

Huỳnh Tường Nguyễn,
Nguyễn Ngọc Lê



Động cơ nghiên cứu

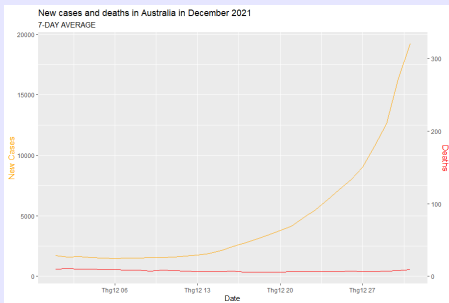
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 6

CẤU TRÚC RỜI RẠC
CHO KHMT

Huỳnh Tường Nguyễn,
Nguyễn Ngọc Lê



Động cơ nghiên cứu

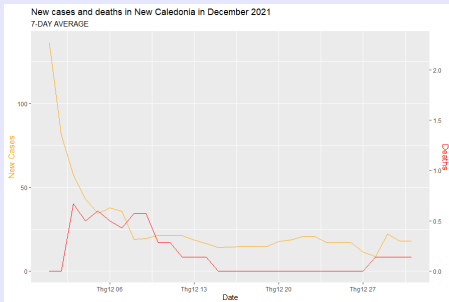
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 6

CẤU TRÚC RỜI RẠC
CHO KHMT

Huỳnh Tường Nguyễn,
Nguyễn Ngọc Lê



Động cơ nghiên cứu

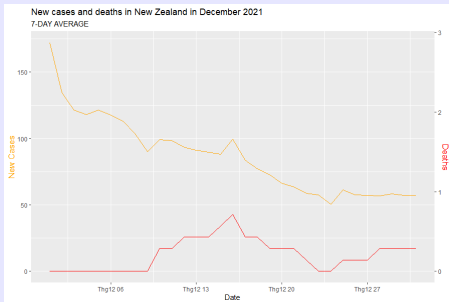
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 7



Động cơ nghiên cứu

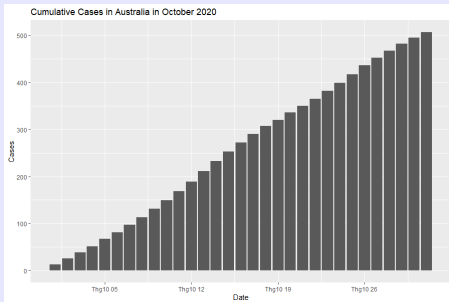
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 7

CẤU TRÚC RỜI RẠC
CHO KHMT

Huynh Tuong Nguyen,
Nguyen Ngoc Le



Động cơ nghiên cứu

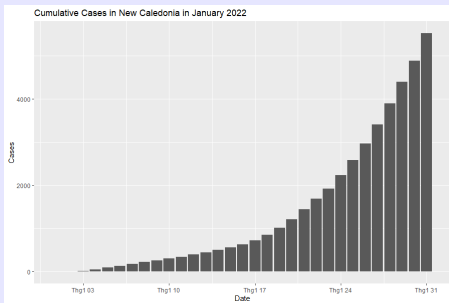
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 7



Động cơ nghiên cứu

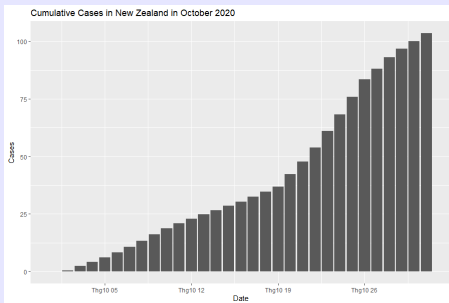
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 8



Động cơ nghiên cứu

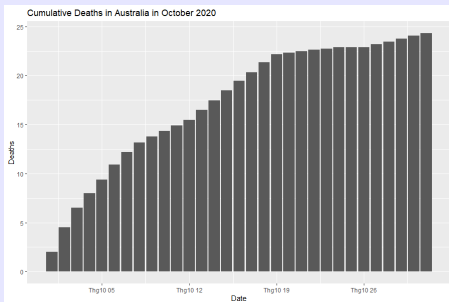
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 8



Động cơ nghiên cứu

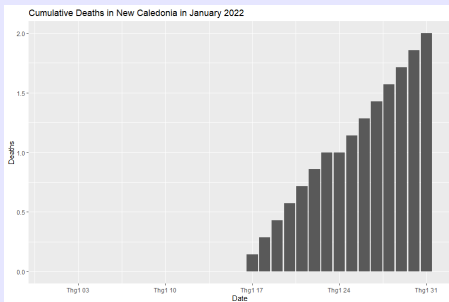
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 6 - Subtask 8



Động cơ nghiên cứu

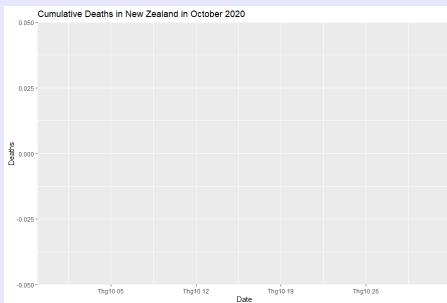
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 7



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Hướng làm

- Yêu cầu làm biểu đồ theo thời gian là tháng, những tháng thuộc về nhóm là 1, 7, 8, 10.
- Lọc dữ liệu dựa trên các dữ liệu có iso_code là OWID_WRL
- Sử dụng package ggplot2 để vẽ các biểu đồ theo yêu cầu của đề

Task 7 - Subtask 1



Động cơ nghiên cứu

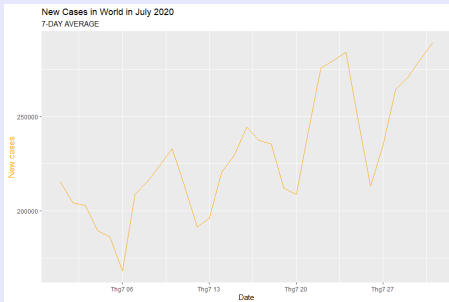
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia



Task 7 - Subtask 2



Động cơ nghiên cứu

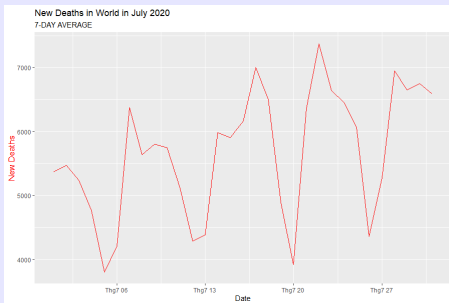
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia



Task 7 - Subtask 3

CẤU TRÚC RỜI RẠC
CHO KHMT

Huỳnh Tường Nguyễn,
Nguyễn Ngọc Lê



Động cơ nghiên cứu

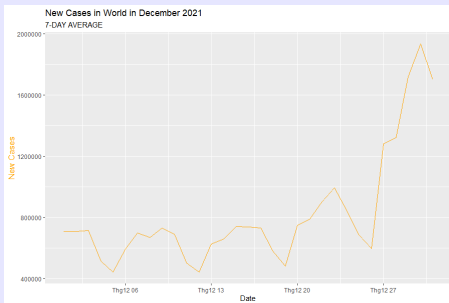
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là 2 tháng cuối của năm của tất cả quốc gia



Task 7 - Subtask 4



Động cơ nghiên cứu

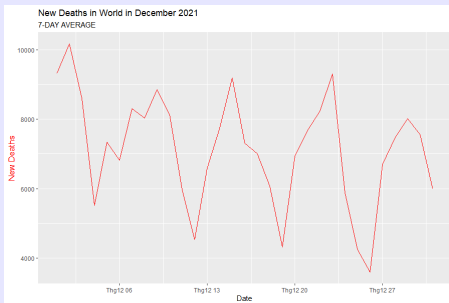
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là 2 tháng cuối của năm của tất cả quốc gia



Task 7 - Subtask 5



Động cơ nghiên cứu

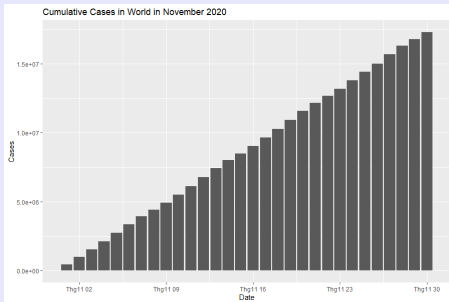
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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



Task 7 - Subtask 6



Động cơ nghiên cứu

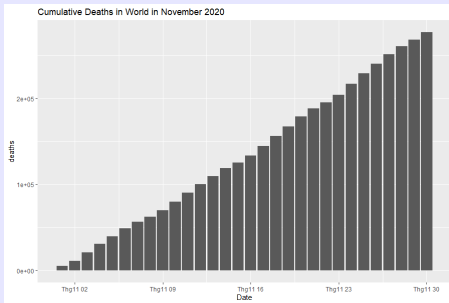
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

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





Cơ sở lý thuyết

-Tiến hành lọc lấy số liệu về số ca nhiễm mới và số ca tử vong của thế giới (dữ liệu có isocode là OWIDWRL) của từng năm trong các tháng 1,7,8,10, đồng thời tạo bảng cho từng mốc thời gian đó. Tiếp theo, tính toán số liệu theo 7 ngày gần nhất của từng ngày và tính các số liệu tích lũy (đối với câu 5 và câu 6).

-Bảng giá trị trung bình 7 ngày gần nhất (TB) của một dãy số A được tính như sau:

- TB[1]: nhận A[1]
- TB[2]: lấy giá trị trung bình của A[1], A[2]
- TB[3]: lấy giá trị trung bình của A[1],A[2],A[3]
- .
- .
- .
- TB[7]: lấy giá trị trung bình của A[1], A[2], ..., A[7]
- TB[i] ($i > 7$): lấy giá trị trung bình của A[i-6], A[i-5],..., A[i]



Cơ sở lý thuyết

-Sau đó, thêm các giá trị vừa tính vào bảng số liệu. Lúc này, bảng số liệu sẽ bao gồm ngày, số ca nhiễm/ tử vong, số ca nhiễm/ tử vong theo trung bình 7 ngày, số ca nhiễm/tử vong tích lũy. Việc lưu tất cả các số liệu vào cùng 1 bảng sẽ thuận tiện cho việc vẽ biểu đồ.

-Dùng hàm ggplot trong thư viện ggplot2 để vẽ biểu đồ cho từng nhóm câu hỏi.

-ggplot(data, aes(x = <ngày>, y = <số ca>, group = <tháng>, color = <tháng>)): khởi tạo hệ trục tọa độ Descartes với data là bảng số liệu tương ứng, trục Ox là các ngày, trục Oy là số ca nhiễm (tùy theo câu hỏi mà giá trị tung độ sẽ khác nhau). Ta phân biệt đường biểu diễn của mỗi tháng qua đối số group và color.

-geomline() : thể hiện đường biểu diễn đi qua các điểm.

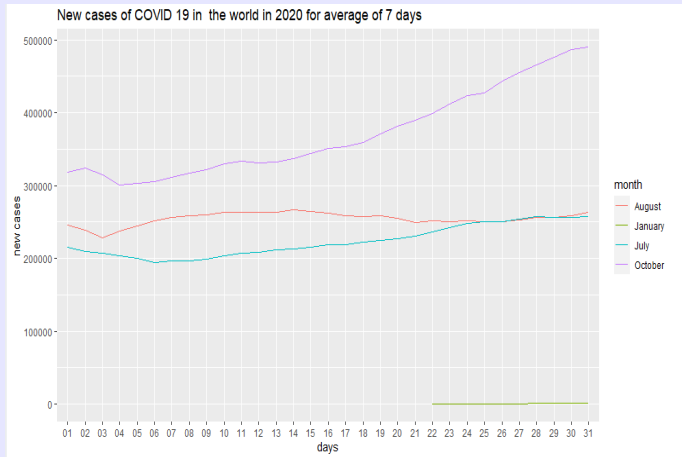
-labs(title="<tên biểu đồ>",x="<tên trục Ox>,y="<tên trục Oy>"): đặt tên biểu đồ và tên các trục Ox, Oy.

- Cuối cùng, ta dùng toán tử + để gắn kết các lệnh.

- ggplot(...)+ geomline() + labs(...) + ...

Task 8 - Subtask 1

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

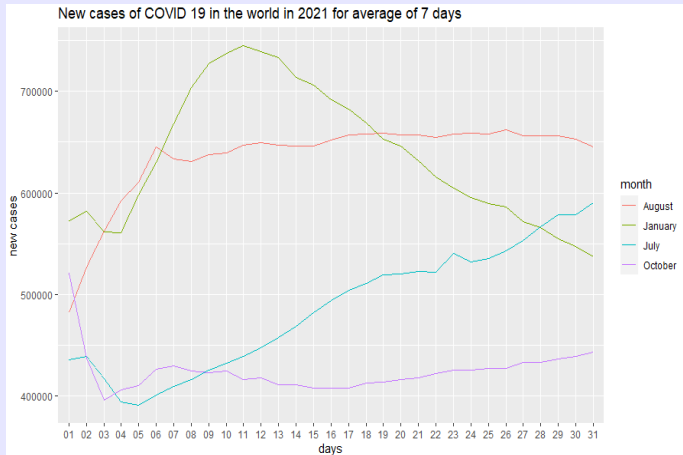
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 1

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

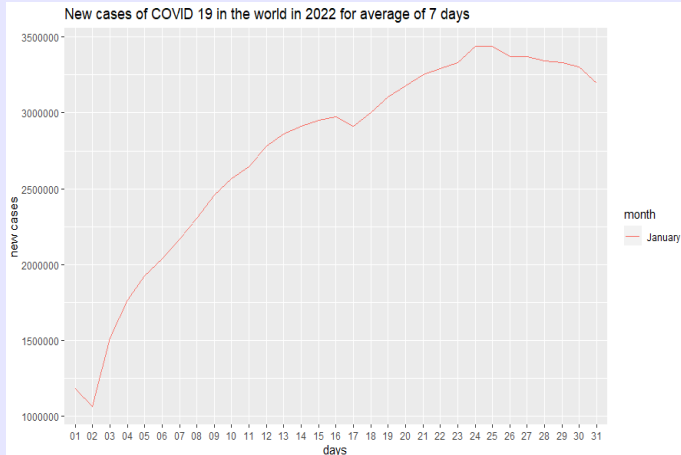
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 1

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là tháng của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

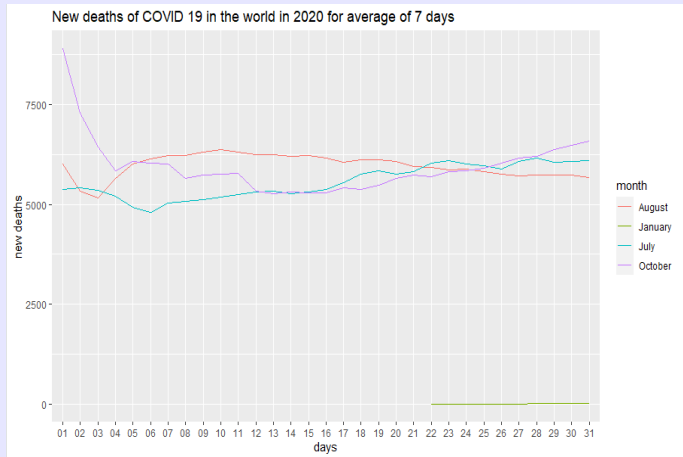
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 2

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

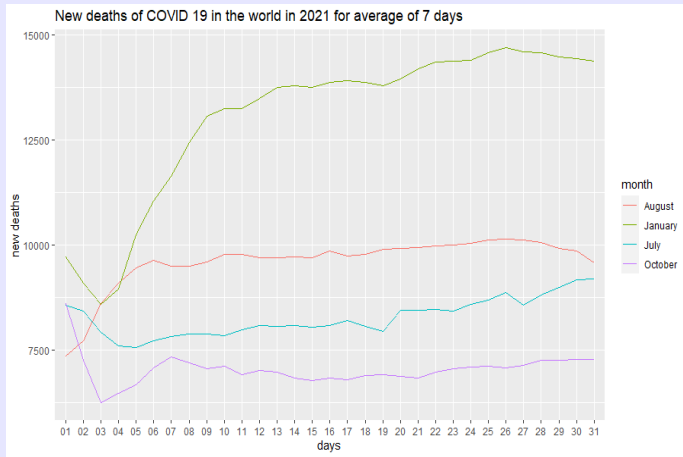
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 2

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

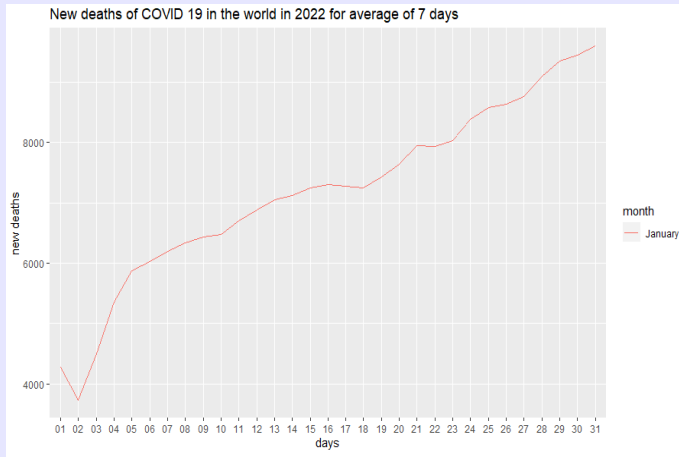
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 2

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là tháng của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

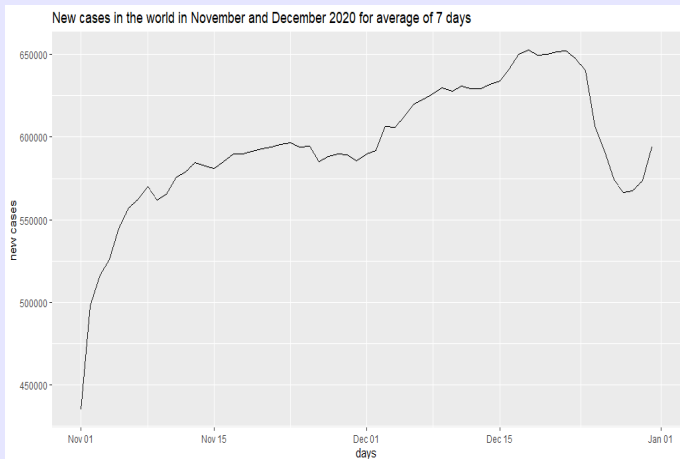
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 3

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

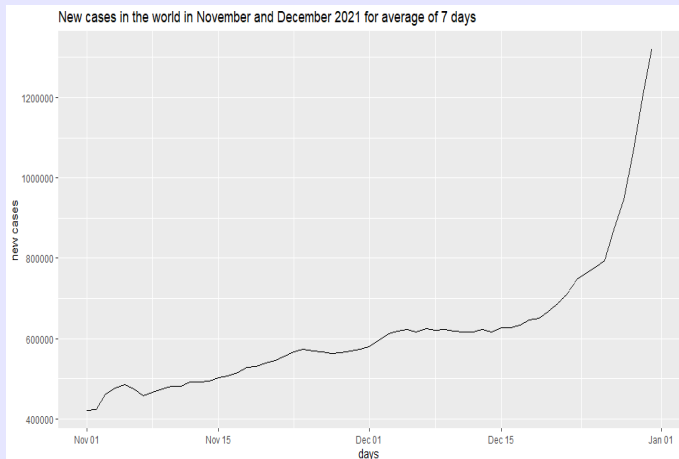
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 3

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

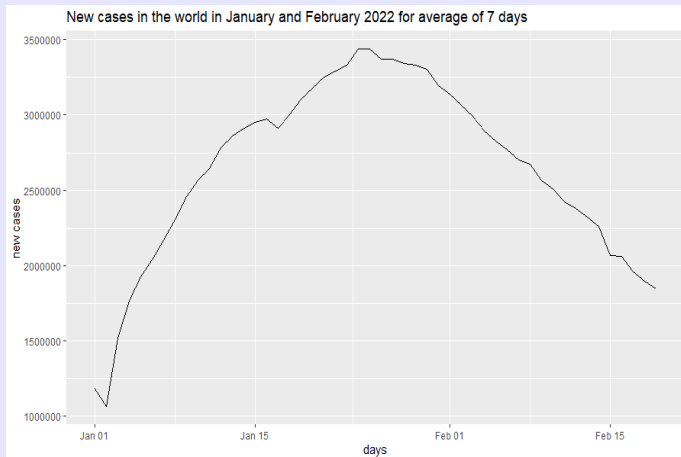
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 3

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

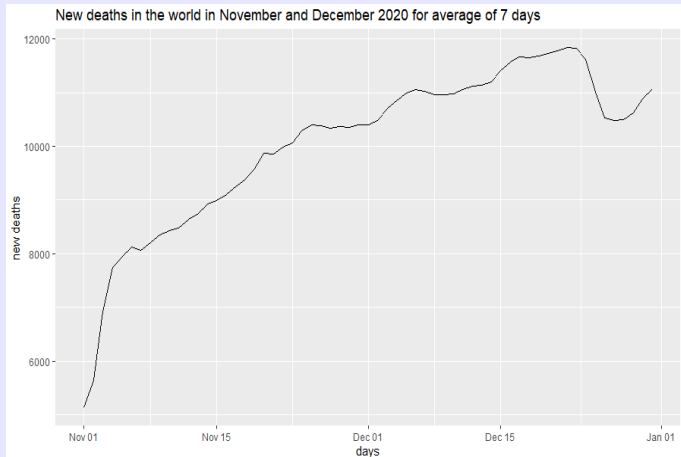
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 4

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

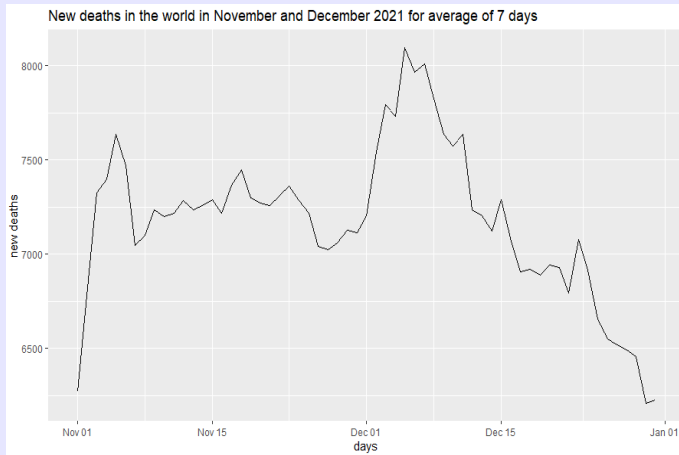
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 4

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

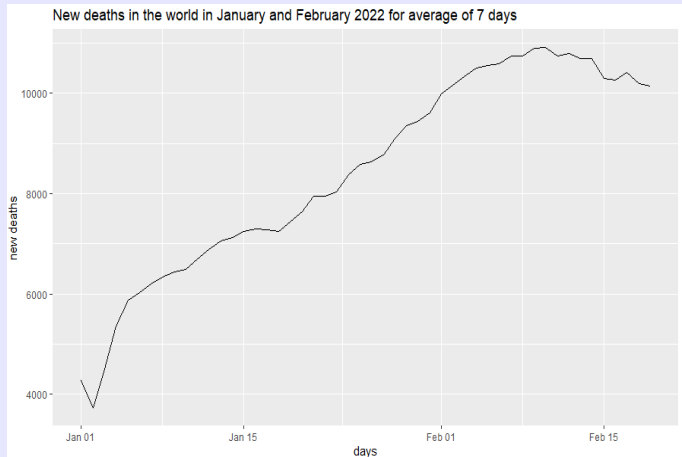
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 4

Biểu đồ thể hiện thu thập dữ liệu tử vong theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

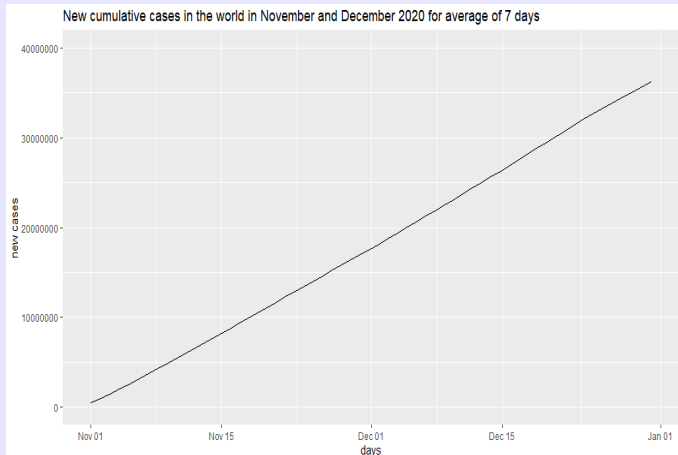
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 5

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

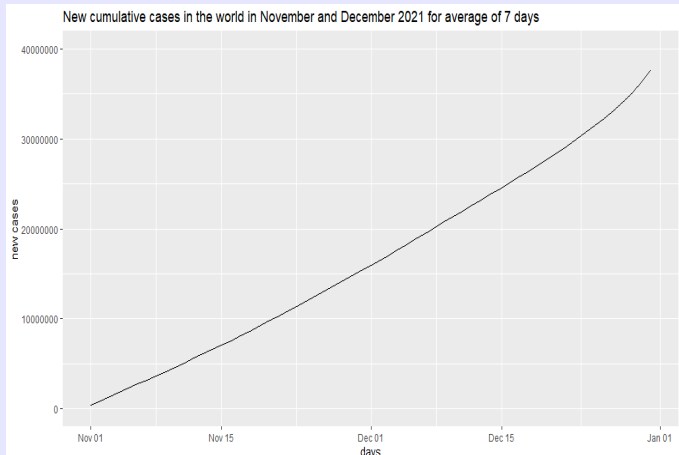
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 5

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

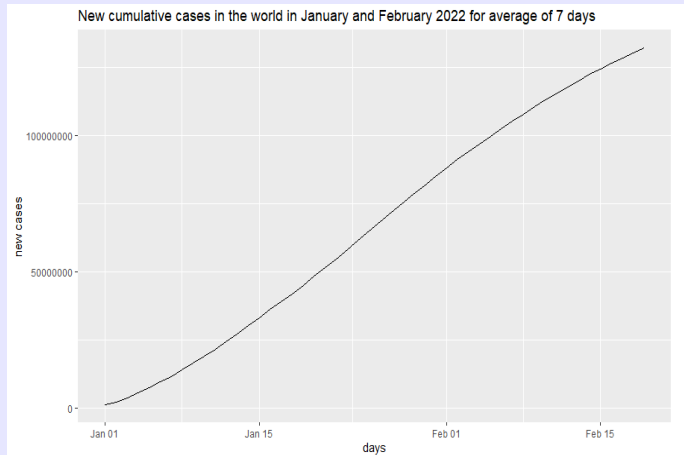
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 5

Biểu đồ thể hiện thu thập dữ liệu nhiễm bệnh tích lũy theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

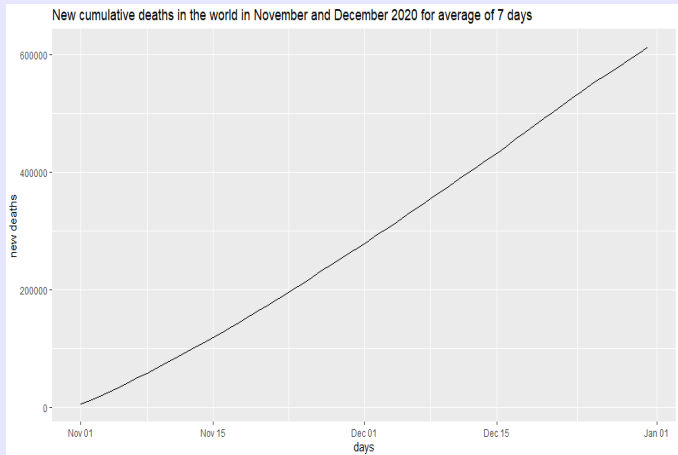
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 6

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

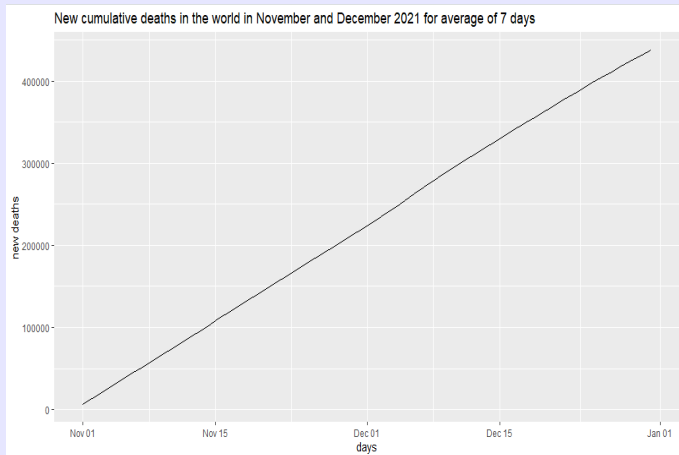
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 6

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

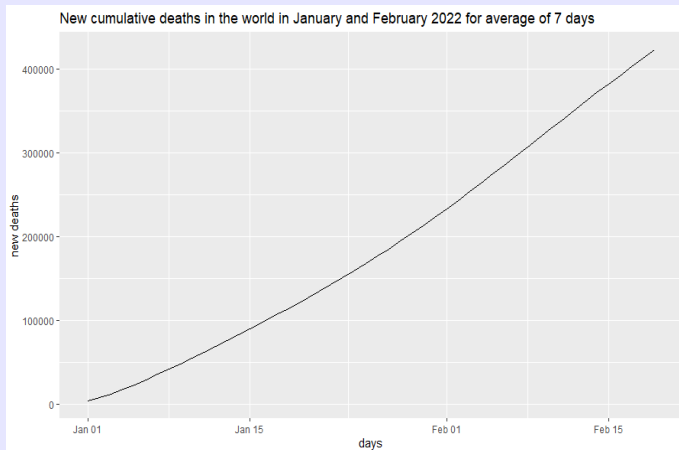
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 8 - Subtask 6

Biểu đồ thể hiện thu thập dữ liệu tử vong tích lũy theo thời gian là 2 tháng của năm của tất cả quốc gia theo trung bình 7 ngày gần nhất



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết



Cơ sở lý thuyết

Hệ số tương quan và các khái niệm liên quan

- Hệ số tương quan (Correlation Coefficient) là một chỉ số đo lường của một số loại tương quan, nghĩa là mối liên hệ thống kê giữa hai biến số. Các biến có thể là hai cột của một bộ dữ liệu quan sát đã cho, thường được gọi là mẫu hoặc hai phần của một biến ngẫu nhiên đa biến số có phân phối đã biết trước. Hệ số tương quan nằm trong phạm vi chạy từ -1 đến $+1$, trong đó $+1$ và -1 biểu thị hai biến số có mối tương quan tuyệt đối và 0 chỉ hai biến số không có liên hệ gì với nhau.
- Có một số loại hệ số tương quan, nhưng phổ biến nhất là hệ số tương quan Pearson (R), hệ số này chỉ ra độ mạnh và hướng của mối quan hệ tuyến tính giữa hai biến. Đây cũng là hệ số sẽ được sử dụng để tính toán trong phần trình bày này.

Cơ sở lý thuyết

- Độ mạnh của mối quan hệ thay đổi theo mức độ dựa trên giá trị của hệ số tương quan. Ví dụ, giá trị 0,2 cho thấy có mối tương quan dương giữa hai biến. Các nhà phân tích trong một số lĩnh vực nghiên cứu không coi các mối tương quan là quan trọng cho đến khi giá trị vượt qua ít nhất 0,8. Tuy nhiên, hệ số tương quan có giá trị tuyệt đối từ 0,9 trở lên sẽ thể hiện một mối quan hệ rất chặt chẽ.
- Trong bài báo cáo này, một mối quan hệ giữa hai biến được coi là:
 - Mạnh (Strong) nếu trị tuyệt đối hệ số tương quan giữa chúng lớn hơn 0,7,
 - Trung bình (Moderate) nếu trị tuyệt đối hệ số tương quan giữa chúng lớn hơn 0,5 và bé hơn bằng 0,7,
 - Yếu (Weak) nếu trị tuyệt đối hệ số tương quan giữa chúng lớn hơn 0,3 và bé hơn bằng 0,5,
 - Rất yếu (Very Weak) nếu trị tuyệt đối hệ số tương quan giữa chúng bé hơn bằng 0,3.





Cơ sở lý thuyết

- Lọc và xử lý dữ liệu có isocode là AUS,NCL,NZL ta được các thông tin cần thiết về số ca mắc mới và số ca tử vong do COVID 19 tại các nước Australia, New Caledonia và New Zealand trong các tháng 1,7,8,10.

- Sử dụng ngôn ngữ lập trình R để tính toán:

Tiến hành tính tổng số ca nhiễm/ tử vong và số ca nhiễm/tử vong tích lũy của mỗi ngày, sau đó tính tỉ lệ phần trăm tích lũy so với tổng. (câu 1)

Sử dụng hàm `cor(x,y,method = "pearson")` để tính Pearson coefficient correlation giữa 2 biến, sau đó kết hợp kết quả với biểu đồ để đối chiếu về độ mạnh yếu của hai đại lượng. (ở đây là số ca nhiễm mới và số ca tử vong) (câu 2,3)



Cơ sở lý thuyết

- Sử dụng các lệnh của ngôn ngữ R để vẽ biểu đồ:
- Đầu tiên, tính toán các số liệu cần thiết (số ca theo trung bình 7 ngày, số ca tích lũy...)
 - + `ggplot(data,aes(x = <các ngày>, y = <số ca mới>))`: khởi tạo hệ trục Oxy với các miền giá trị tương ứng Ox,Oy
 - + `geompoint()` thể hiện các điểm trong hệ trục Oxy.
 - + `geomsMOOTH(method = "lm")` thể hiện đường biểu diễn hồi quy tuyến tính về mối tương quan giữa 2 biến.
 - + `labs(title="<tên biểu đồ>",x="<tên trục Ox>",y="<tên trục Oy>")`: thêm nhãn cho biểu đồ
- Tiếp theo, kết hợp các biến vào một để tạo biểu đồ `ggplot()+geompoint()+geomsMOOTH(method = "lm")+labs()`: tạo biểu đồ theo tháng đang tính toán (4 tháng tất cả)



Động cơ nghiên cứu

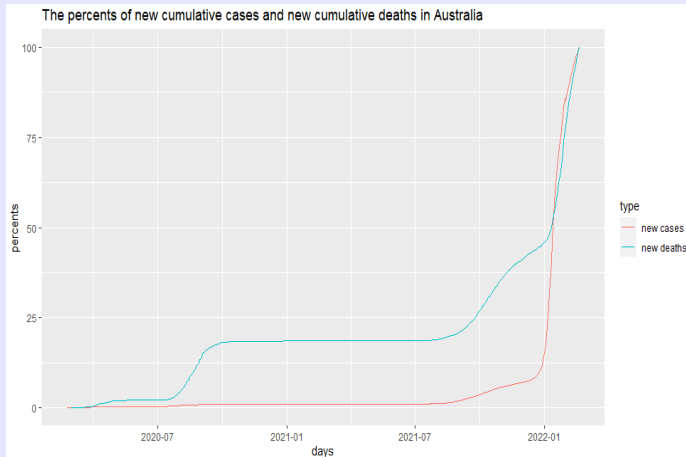
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Kết quả





Động cơ nghiên cứu

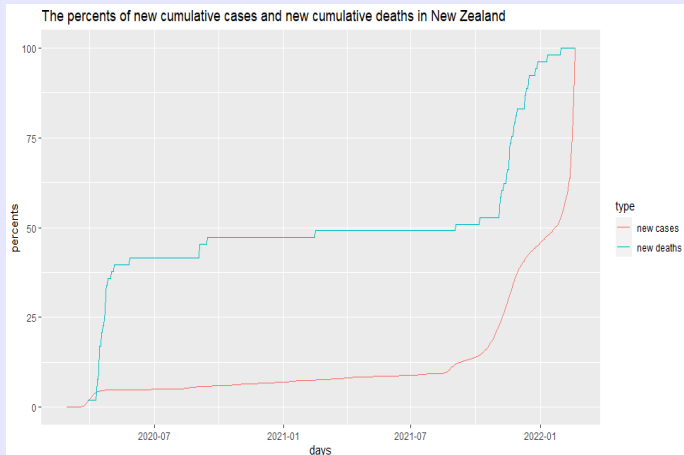
Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Kết quả





Động cơ nghiên cứu

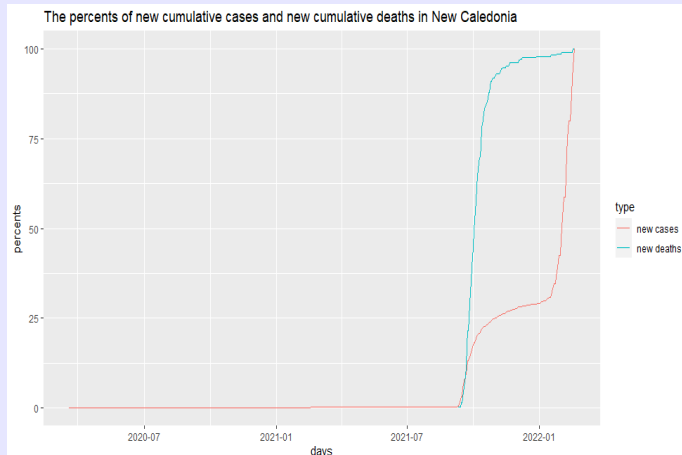
Mục tiêu

Mô tả dữ liệu

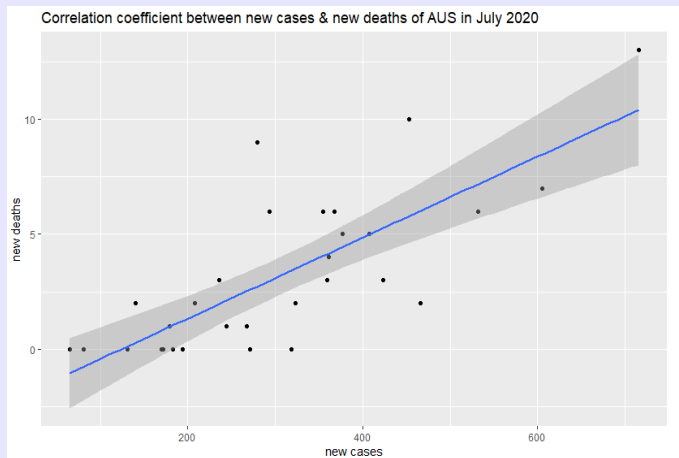
Nhiệm vụ

Tổng kết

Kết quả



Kết quả



Hệ số tương quan $r = 0.7736696$ (vì $|r| > 0.7$ nên số ca nhiễm mới và số ca tử vong có quan hệ tuyến tính mạnh (dương))



Động cơ nghiên cứu

Mục tiêu

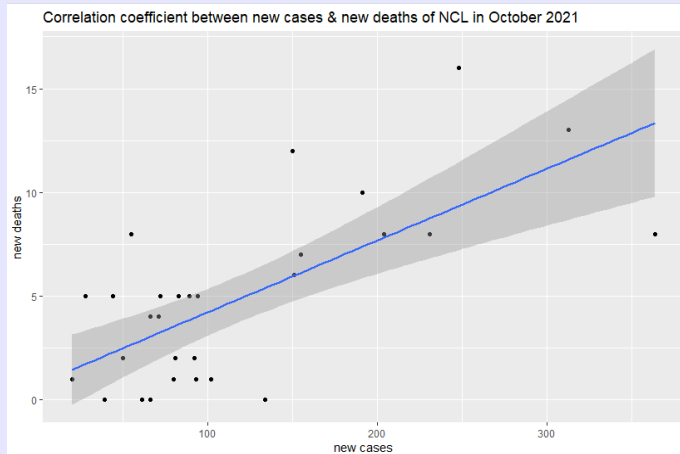
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 9

Kết quả

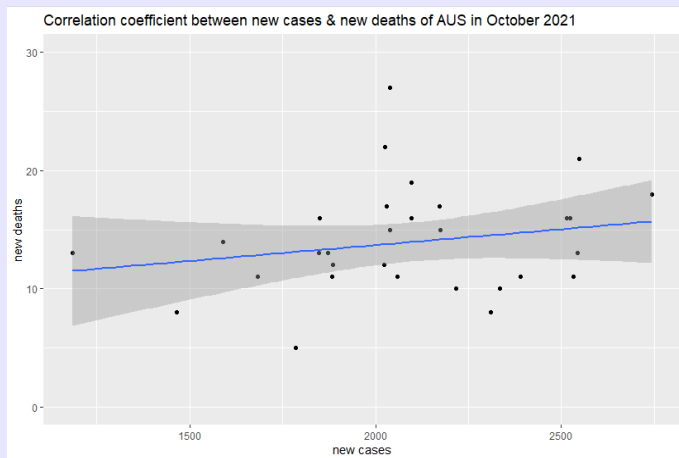


Hệ số tương quan $r = 0.6976481$ (vì $0.5 \leq |r| \leq 0.7$ nên số ca nhiễm mới và số ca tử vong có quan hệ tuyến tính trung bình (dương))



Task 9

Kết quả

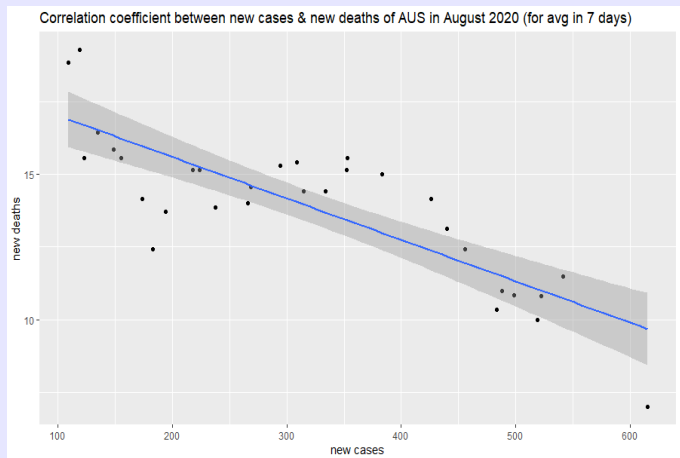


Hệ số tương quan $r = 0.2091307$ (vì $|r| \leq 0.3$ nên số ca nhiễm mới và số ca tử vong có quan hệ tuyến tính rất yếu (dương))



Task 9

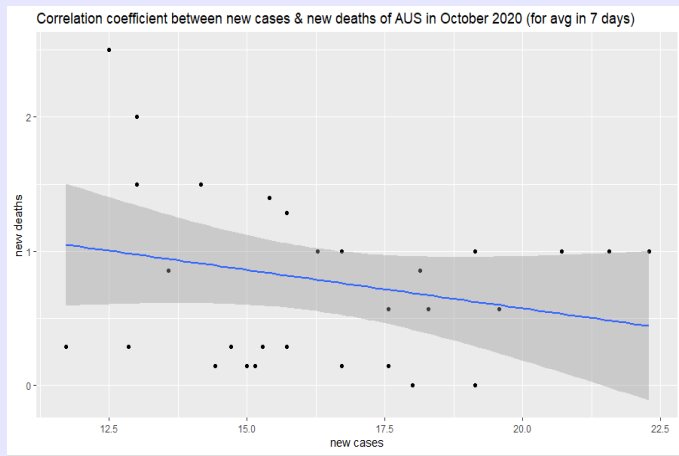
Kết quả



Hệ số tương quan $r = -0.822898$ (vì $|r| > 0.7$ nên số ca nhiễm mới và số ca tử vong có quan hệ tuyến tính mạnh (âm))



Kết quả



Hệ số tương quan $r = -0.249297$ (vì $|r| \leq 0.3$ nên số ca nhiễm mới và số ca tử vong có quan hệ tuyến tính rất yếu (âm))



Động cơ nghiên cứu

Mục tiêu

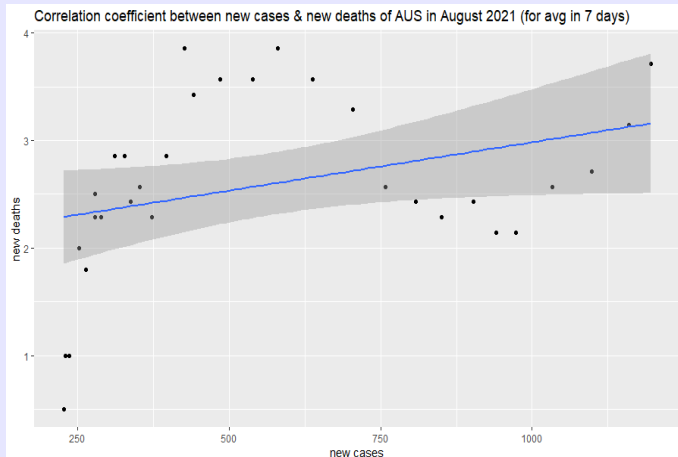
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Task 9

Kết quả

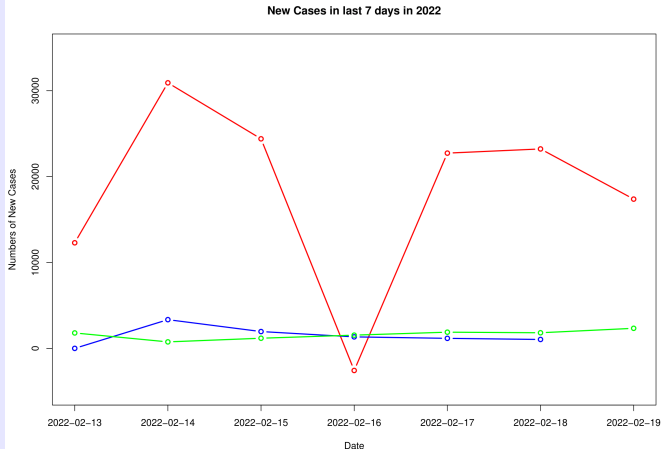


Hệ số tương quan $r = 0.3436588$ (vì $0.3 \leq |r| \leq 0.5$ nên số ca nhiễm mới và số ca tử vong có quan hệ tuyến tính trung bình (dương))



Task 10 - Subtask 1

So sánh tình trạng nhiễm bệnh của các quốc gia trong 7 ngày cuối của năm cuối cùng



CẤU TRÚC RỜI RẠC
CHO KHMT

Huỳnh Tường Nguyễn,
Nguyễn Ngọc Lê



Động cơ nghiên cứu

Mục tiêu

Mô tả dữ liệu

Nhiệm vụ

Tổng kết



Nhận xét

- Trong 7 ngày cuối trong số liệu của năm cuối cùng (từ 13/02-19/02/2022):
 - New Caledonia và New Zealand khá may mắn vì số lượng ca nhiễm bệnh mới khá thấp, cho thấy nước này đảm bảo được sự an toàn trước đại dịch.
 - Australia tuy nhiên lại có số lượng ca nhiễm mới rất lớn (đạt số đều trên 10000 ca/ngày và ngày có số lượng ca mắc mới nhiều nhất là ngày 14/02/2022 với hơn 30000 ca).
- > Những điều này chứng tỏ những ảnh hưởng nặng nề mà đại dịch này để lại cho con người, những con virus mang lại hậu quả sức khỏe hết sức nguy hiểm không thể chủ quan.

Task 10 - Subtask 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?

Ngày có số lượng tử vong lớn nhất ở Australia là
"2022-01-28"

Ngày có số lượng tử vong lớn nhất ở New Caledonia là
"2021-09-22" "2021-10-13"

Ngày có số lượng tử vong lớn nhất ở New Zealand là
"2020-04-14"

Từ đó ta suy ra được khoảng thời gian bùng phát tử vong lớn nhất giữa các quốc gia không chồng lên nhau.

Task 10 - Subtask 10



Động cơ nghiên cứu

Mục tiêu

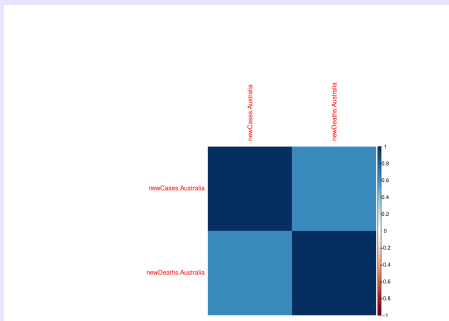
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Hãy mô tả mối quan hệ tuyến tính giữa nhiễm bệnh và tử vong bằng cách đo độ kết hợp của mỗi quan hệ dùng correlation r (correlation coefficient) và hướng kết hợp

- Australia



Task 10 - Subtask 10



Động cơ nghiên cứu

Mục tiêu

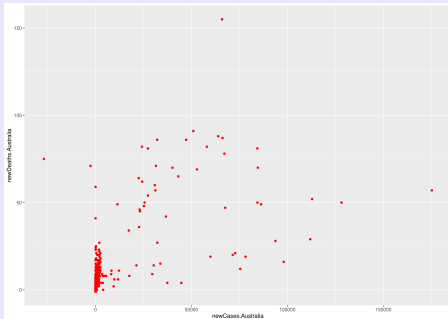
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Hãy mô tả mối quan hệ tuyến tính giữa nhiễm bệnh và tử vong bằng cách đo độ kết hợp của mối quan hệ dùng correlation r (correlation coefficient) và hướng kết hợp

- Australia



Task 10 - Subtask 10



Động cơ nghiên cứu

Mục tiêu

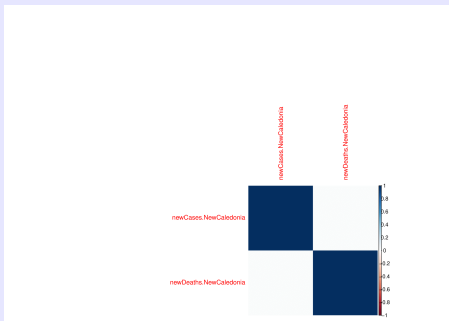
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Hãy mô tả mối quan hệ tuyến tính giữa nhiễm bệnh và tử vong bằng cách đo độ kết hợp của mỗi quan hệ dùng correlation r (correlation coefficient) và hướng kết hợp

- New Caledonia



Task 10 - Subtask 10



Động cơ nghiên cứu

Mục tiêu

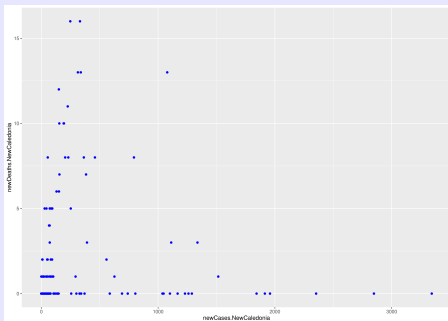
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Hãy mô tả mối quan hệ tuyến tính giữa nhiễm bệnh và tử vong bằng cách đo độ kết hợp của mỗi quan hệ dùng correlation r (correlation coefficient) và hướng kết hợp

- New Caledonia



Task 10 - Subtask 10



Động cơ nghiên cứu

Mục tiêu

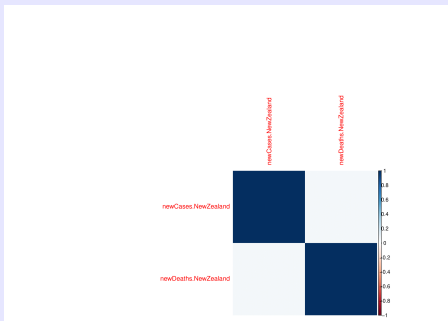
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Hãy mô tả mối quan hệ tuyến tính giữa nhiễm bệnh và tử vong bằng cách đo độ kết hợp của mỗi quan hệ dùng correlation r (correlation coefficient) và hướng kết hợp

- New Zealand



Task 10 - Subtask 10



Động cơ nghiên cứu

Mục tiêu

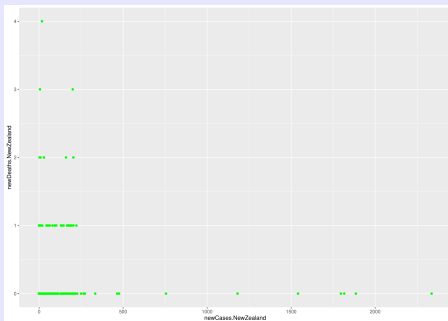
Mô tả dữ liệu

Nhiệm vụ

Tổng kết

Hãy mô tả mối quan hệ tuyến tính giữa nhiễm bệnh và tử vong bằng cách đo độ kết hợp của mỗi quan hệ dùng correlation r (correlation coefficient) và hướng kết hợp

- New Zealand





[Động cơ nghiên cứu](#)

[Mục tiêu](#)

[Mô tả dữ liệu](#)

[Nhiệm vụ](#)

[Tổng kết](#)

Tóm lại

Phân tích dữ liệu về dịch bệnh Covid có ý nghĩa quan trọng trong việc đánh giá mức độ nghiêm trọng của đại dịch trên toàn thế giới. Ngoài ra, những dữ liệu riêng của từng châu lục, hay từng quốc gia giúp ta dễ dàng xác định được những khu vực có dịch bệnh bùng nổ, hay những nơi có dấu hiệu xuất hiện của dịch bệnh để tìm ra những biện pháp xử lý, phòng tránh phù hợp. Những kết quả ta phân tích được trong quá khứ có ý nghĩa thực tế trong việc dự đoán những con số sẽ xuất hiện trong tương lai. Qua bài tập lớn này, ngoài việc tiếp xúc và rèn luyện với ngôn ngữ mới là R và Latex, ta còn được nâng cao kỹ năng phân tích số liệu cũng như biểu diễn các số liệu thống kê một cách dễ hiểu cho người xem.