**ĐẠI HỌC HUẾ**

# KHOA KỸ THUẬT VÀ CÔNG NGHỆ



🙠🙟🕮🙝🙢

****

**BÁO CÁO**

**ĐỒ ÁN**

**NĂM HỌC 2020-2021**

**Giảng viên hướng dẫn: Ts.Hồ Quốc Dũng**

**Lớp: KHDL&TTNT**

|  |
| --- |
| Số phách  *(Do hội đồng chấm thi ghi)* |

**Thừa Thiên Huế, ngày 12 tháng 7 năm 2021**

**ĐẠI HỌC HUẾ**

# KHOA KỸ THUẬT VÀ CÔNG NGHỆ



🙠🙟🕮🙝🙢

****

(MẪU BÌA PHỤ)

**BÁO CÁO**

**ĐỒ ÁN**

**NĂM HỌC 2020-2021**

**Giảng viên hướng dẫn: Ts. Hồ Quốc Dũng**

**Lớp: KHDL&TTNT**

**Sinh viên thực hiện: Trần Đăng Minh Vũ**

*(ký tên và ghi rõ họ tên)*

|  |
| --- |
| Số phách  *(Do hội đồng chấm thi ghi)* |

**Thừa Thiên Huế, ngày 12 tháng 7 năm 2021**

**MỤC LỤC**

Contents

[1](#_Toc76897039)[**.Cơ sở dữ liệu:** 5](#_Toc76897041)

[**2.** **Đồ thị:** 5](#_Toc76897042)

[**2.1.** **Đồ thị 1:** 5](#_Toc76897043)

[**2.2.** **Đồ thị 2:** 6](#_Toc76897044)

[**2.3.** **Đồ thị 3:** 7](#_Toc76897045)

[**2.4.** **Đồ thị 4:** 8](#_Toc76897046)

[**2.5.** **Đồ thị 5:** 9](#_Toc76897047)

[**2.6.** **Đồ thị 6:** 10](#_Toc76897048)

[**2.7.** **Đồ thị 7:** 11](#_Toc76897049)

[**2.8.** **Đồ thị 8:** 12](#_Toc76897050)

[**2.10.** **Đồ thị 10:** 14](#_Toc76897051)

[**2.11.** **Đồ thị 11:** 15](#_Toc76897052)

[**2.12.** **Đồ thị 12:** 16](#_Toc76897053)

[**2.13.** **Đồ thị 13:** 17](#_Toc76897054)

[**2.14.** **Đồ thị 14:** 18](#_Toc76897055)

[**2.15.** **Đồ thị 15:** 19](#_Toc76897056)

[**2.16.** **Đồ thị 16:** 20](#_Toc76897057)

[**2.17.** **Đồ thị 17:** 21](#_Toc76897058)

[**2.18.** **Đồ thị 18:** 22](#_Toc76897059)

[**2.19.** **Đồ thị 19:** 23](#_Toc76897060)

[**2.20.** **Đồ thị 20:** 24](#_Toc76897061)

1. **Cơ sở dữ liệu:**

* Sử dụng cơ sở dữ liệu: GitHub - CSSEGISandData/COVID-19: Novel Coronavirus (COVID-19) Cases, provided by JHU CSSE
* Đường link: <https://github.com/CSSEGISandData/COVID-19>

**Trước khi lập trình vẽ biểu đồ ta thêm đường dẫn và đọc file từ đường dẫn, sau đó sẽ tạo data frame để phân loại dữ liệu.**

**Mã nguồn R:**

*setwd("C:\\Users\\Administrator\\Desktop\\Doan\_R\\COVID-19-master\\csse\_covid\_19\_data\\csse\_covid\_19\_daily\_reports\_us ")*

*#Package*

*update.packages("tools")*

*install.packages("ggplot2", lib="C:/Users/Administrator/Documents/R/win-library/3.3")*

*update.packages("ggplot2")*

*update.packages("data.table")*

*library(data.table)*

*#Readfile*

*files <- list.files(pattern = ".csv")*

*temp <- lapply(files, fread, sep=",")*

*data <- rbindlist( temp, fill=TRUE)*

*names(data)*

*View(data)*

*#Dataframe*

*Alk <- data[Province\_State=="Alaska"]*

*Albm <- data[Province\_State=="Alabama"]*

*Albm$Case\_Fatality\_Ratio <- format(round(Albm$Case\_Fatality\_Ratio, 2), nsmall = 2)*

*Alk$Case\_Fatality\_Ratio <- format(round(Alk$Case\_Fatality\_Ratio, 2), nsmall = 2)*

*#readfile*

*#tai My ngay 01/01/2021*

*df1 <- read.table("01-01-2021.csv",*

*header = TRUE,*

*sep = ",")*

*names(df1)*

*#Chuyen sang so thap phan co 2 chu so sau dau phay*

*df1$Case\_Fatality\_Ratio <- format(round(df1$Case\_Fatality\_Ratio, 2), nsmall = 2)*

*df1$Case\_Fatality\_Ratio*

*#doc file du lieu ve covid cua cac bang/thanh pho*

*#tai My ngay 05/12/2020*

*df <- read.table("05-12-2020.csv",*

*header = TRUE,*

*sep = ",")*

*names(df)*

*View(df)*

*df$Mortality\_Rate <- format(round(df$Mortality\_Rate, 2), nsmall = 2)*

*#doc file du lieu ve covid cua cac bang/thanh pho*

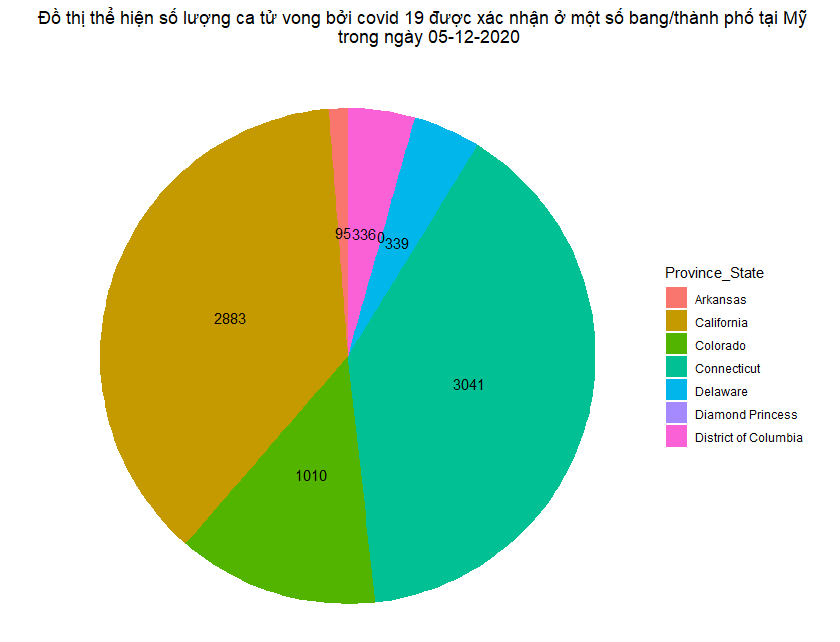
*#tai My ngay 05/12/2020 va ngay 01/01/2021*

*#dung thu vien ggplot2 de ve do thi*

*library("ggplot2", lib.loc="~/R/win-library/4.0")*

1. **Đồ thị:**

* 1. **Đồ thị 1:**

****

**Mã nguồn R:**

*ggplot(df[5:11,], aes(x='', y=Deaths, fill=Province\_State)) +*

*geom\_bar(stat="identity", width=1) +*

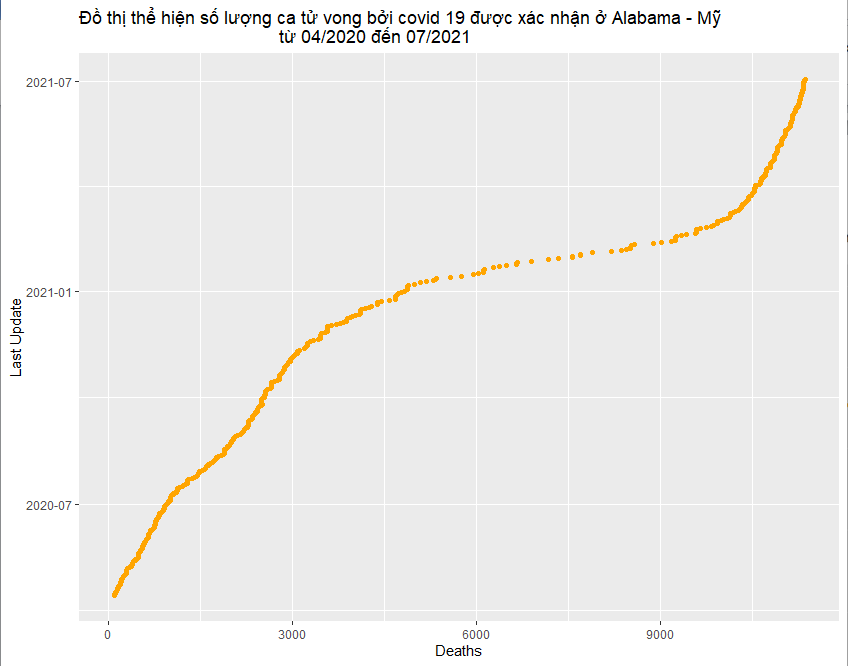
*coord\_polar("y", start=0)+*

*theme\_void() + labs(title="Đồ thị thể hiện số lượng ca tử vong bởi covid 19 được xác nhận ở một số bang/thành phố tại Mỹ*

*trong ngày 05-12-2020") +*

*geom\_text(aes(label = paste0(Deaths)), position = position\_stack(vjust=0.5))*

* Mô tả đồ thị:
* Là biểu đồ bánh (Pie chart)
* Đồ thị so sánh 7 bang/thành phố ( Arkansas, California, Colorado, Connecticut, Delaware, Diamond Princess, District of Column) ở Mỹ
* 3 bang/thành phố California, Colorado, Connecticut chiếm tỉ lệ phần trăm về số lượng lớn nhất trong 7 bang/thành phố
* Diamond Princess không có ca tử vong nào trong ngày 05/12/2020
* Ý nghĩa của đồ thị: So sánh số lượng ca tử vong bởi covid của 7 bang/ thành phố tại Mỹ vào ngày 05/12/2020.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được về sự khác nhau và chênh lệch nhau về tỉ lệ phần trăm về số lượng ca tử vong vì covid 19 giữa 7 bang/thành phố ở Mỹ.
  1. **Đồ thị 2:**

****

**Mã nguồn R:**

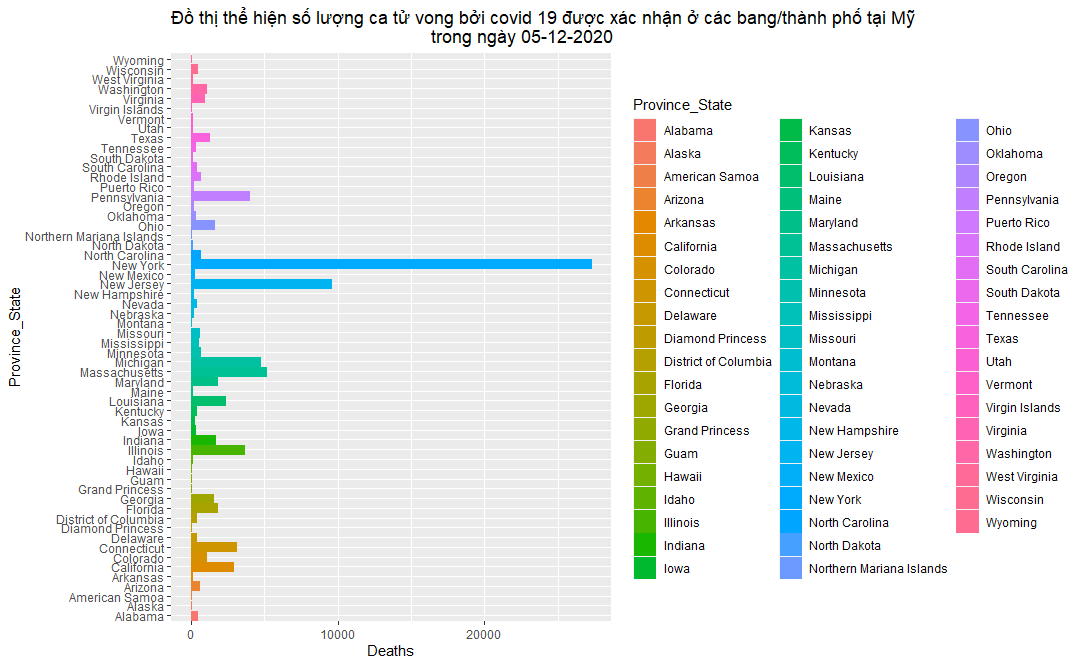
*ggplot(Albm, aes(x=Deaths, y=Last\_Update)) +*

*geom\_point(aes(colour = Deaths), colour = "orange") +*

*labs(title="Đồ thị thể hiện số lượng ca tử vong bởi covid 19 được xác nhận ở Alabama - Mỹ*

*từ 04/2020 đến 07/2021",x = "Deaths", y="Last Update")*

* Mô tả đồ thị:
* Là biểu đồ điểm
* Trục x ứng với “Deaths” (số lượng ca tử vong), trục y ứng với “Last\_Update” (thời gian dữ liệu được đưa lên)
* Điểm ứng với “Deaths” (Số lượng ca tử vong), điểm được vẽ với màu cam
* Mỗi điểm ứng với số lượng ca tử vong bởi covid 19 được xác nhận ở Alabama trong ngày tương ứng và số lượng tương ứng
* Ý nghĩa của đồ thị: Thể hiện và cho ta biết được về sự chuyển biến về số lượng ca tử vong bởi covid 19 mỗi ngày được xác định ở Alabama từ tháng 04/2020 đến tháng 07/2021.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được rằng số lượng ca tử vong bởi covid 19 được xác nhận ở Alabama từ tháng 04/2020 đến 07/2021 có xu hướng tăng nhanh mỗi ngày.
  1. **Đồ thị 3:**



**Mã nguồn R:**

*ggplot(df, aes(x=Deaths, color=Province\_State)) +*

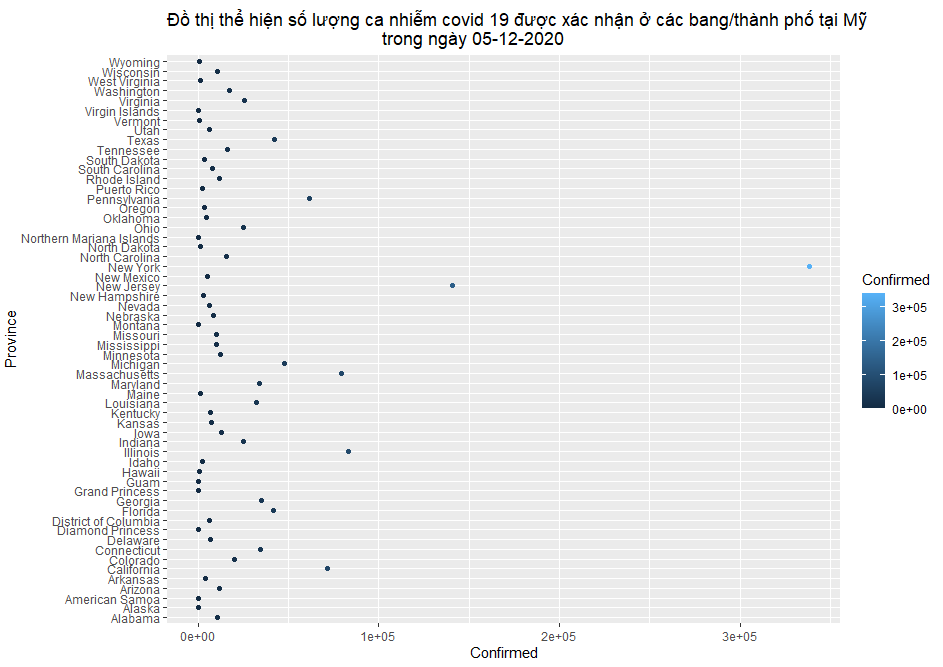
*geom\_col(aes(x=Deaths, y=Province\_State, fill =Province\_State)) +*

*theme\_grey() +*

*labs(title="Đồ thị thể hiện số lượng ca tử vong bởi covid 19 được xác nhận ở các bang/thành phố tại Mỹ*

*trong ngày 05-12-2020")*

* Mô tả đồ thị:
* Là biểu đồ cột
* Trục x ứng với “Deaths” ( Số lượng ca tử vong ứng với từng bang/thành phố), trục y ứng với “Province\_State” (Tên các bang/thành phố)
* Bang/ thành phố New York có số lượng ca tử vong cao nhất, chênh lệch khá lớn với các bang/thành phố khác
* American Samoa, Diamond Princess có số lượng ca tử vong thấp nhất, 0 ca tử vong
* Ý nghĩa của đồ thị: Đồ thị thể hiện và cho ta biết được số lượng ca tử vong bởi covid 19 được xác nhận của mỗi bang/thành phố ở Mỹ trong ngày 05/12/2020.
* Xu hướng dữ liệu của đồ thị: Đồ thị cho ta biết được về sự khác nhau và sự chênh lệch giữa số lượng ca tử vong bởi covid được xác nhận giữa các bang/thành phố ở Mỹ trong ngày 05/12/2020.
  1. **Đồ thị 4**

****

**Mã nguồn R:**

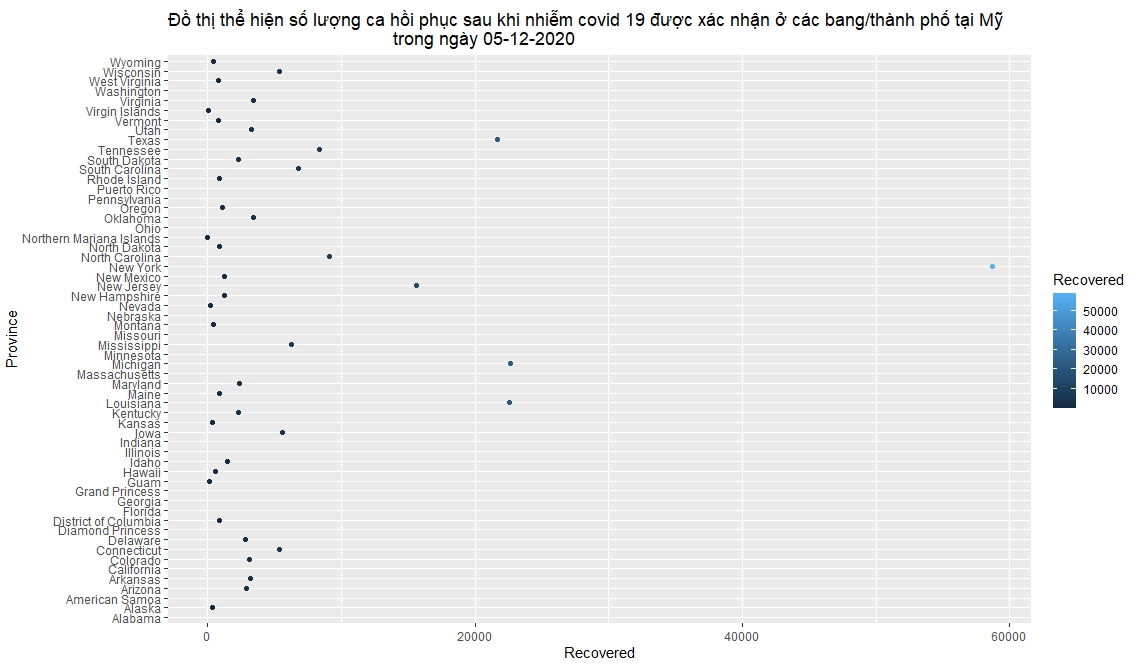
*ggplot(df, aes(x=Confirmed, y=Province\_State, fill= Confirmed)) +*

*geom\_point(aes(color=Confirmed)) +*

*labs(title="Đồ thị thể hiện số lượng ca nhiễm covid 19 được xác nhận ở các bang/thành phố tại Mỹ*

*trong ngày 05-12-2020",x = "Confirmed", y="Province")*

* Mô tả đồ thị
* Là biểu đồ điểm
* Trục x ứng với “Confirmed” (Số ca nhiễm được xác nhận), trục y ứng với “Province” (Tên bang/ thành phố)
* Mỗi điểm tương ứng với số lượng ca được xác nhận nhiễm covid 19 trong ngày 05/12/2020 ở các bang/thành phố tại Mỹ
* Số ca được xác nhận từ số lượng ít đến số lượng nhiều theo thứ tự từ màu xanh đậm chuyển dần về màu xanh nhạt.
* Bang/ thành phố New York có số lượng ca nhiễm covid19 cao nhất, chênh lệch khá lớn với các bang/thành phố khác
* American Samoa có số lượng ca tử vong thấp nhất, 0 ca nhiễm
* Ý nghĩa của đồ thị: Đồ thị thể hiện và cho ta biết được về số lượng ca nhiễm covid 19 của từng bang/thành phố tại Mỹ trong ngày 05/12/2020.
* Xu hướng dữ liệu của đồ thị: Thể hiện sự khác nhau và chênh lệch về số lượng ca nhiễm covid 19 được xác nhận ở các bang/ thành phố tại Mỹ ngày 05/12/2020.
  1. **Đồ thị 5**

****

**Mã nguồn R:**

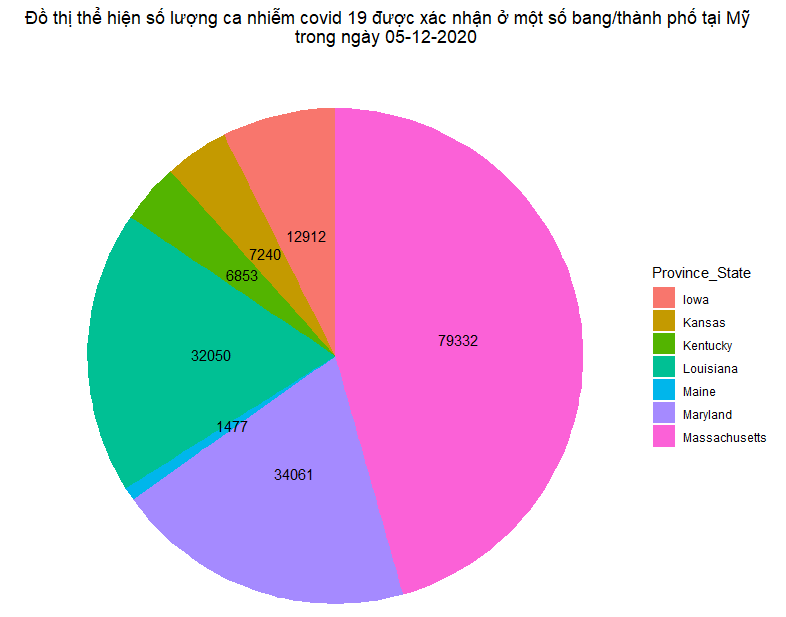
*ggplot(df, aes(x=Recovered, y=Province\_State, fill= Recovered)) +*

*geom\_point(aes(color=Recovered)) +*

*labs(title="Đồ thị thể hiện số lượng ca hồi phục sau khi nhiễm covid 19 được xác nhận ở các bang/thành phố tại Mỹ*

*trong ngày 05-12-2020",x = "Recovered", y="Province")*

* Mô tả đồ thị
* Là biểu đồ điểm
* Trục x ứng với “Recovered” (Số ca hồi phục được xác nhận), trục y ứng với “Province” (Tên bang/ thành phố)
* Mỗi điểm tương ứng với số lượng ca được xác nhận đã hồi phục sau khi nhiễm covid 19 trong ngày 05/12/2020 ở các bang/thành phố tại Mỹ
* Số ca đã hồi phục được xác nhận từ số lượng ít đến số lượng nhiều theo thứ tự từ màu xanh đậm chuyển dần về màu xanh nhạt.
* New York có số lượng ca hồi phục cao nhất
* Ý nghĩa của đồ thị: Đồ thị thể hiện và cho ta biết được về số lượng ca đã hồi phục sau khi nhiễm covid 19 được xác nhận của từng bang/thành phố tại Mỹ trong ngày 05/12/2020.
* Xu hướng dữ liệu của đồ thị: Thể hiện sự khác nhau và chênh lệch về số lượng ca hồi phục sau khi nhiễm covid 19 được xác nhận ở các bang/ thành phố tại Mỹ ngày 05/12/2020.
  1. **Đồ thị 6:**



**Mã nguồn R:**

*ggplot(df[20: 26,], aes(x='', y=Confirmed, fill=Province\_State)) +*

*geom\_bar(stat="identity", width=1) +*

*theme\_void() +*

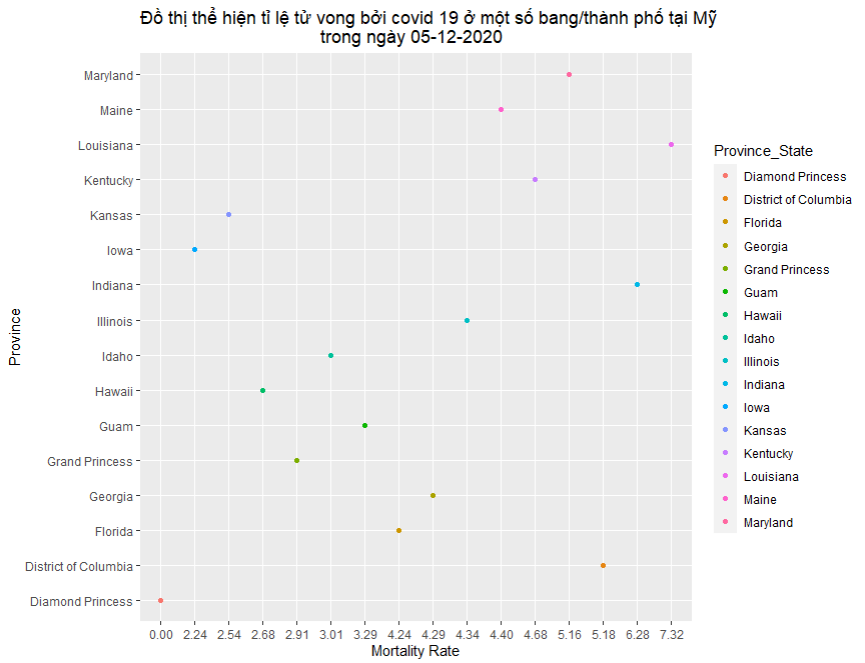
*coord\_polar("y", start=0) +*

*labs(title="Đồ thị thể hiện số lượng ca nhiễm covid 19 được xác nhận ở một số bang/thành phố tại Mỹ*

*trong ngày 05-12-2020") +*

*geom\_text(aes(label = paste0(Confirmed)), position = position\_stack(vjust=0.5))*

* Mô tả đồ thị:
* Là biểu đồ bánh (Pie chart)
* Đồ thị so sánh tỉ lệ giữa 7 bang/thành phố: Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts ở Mỹ
* Mỗi màu ứng với từng bang/thành phố được thể hiện ở trong đồ thị
* Massachusetts chiếm tỉ lệ cao nhất về số lượng giữa 7 bang/thành phố
* Ý nghĩa của đồ thị: So sánh Số lượng ca nhiễm covid 19 được xác nhận của 7 bang/ thành phố tại Mỹ vào ngày 05/12/2020.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được về sự khác nhau và chênh lệch giữa số lượng ca nhiễm covid 19 giữa 7 bang/thành phố ở Mỹ.
  1. **Đồ thị 7:**

****

**Mã nguồn R:**

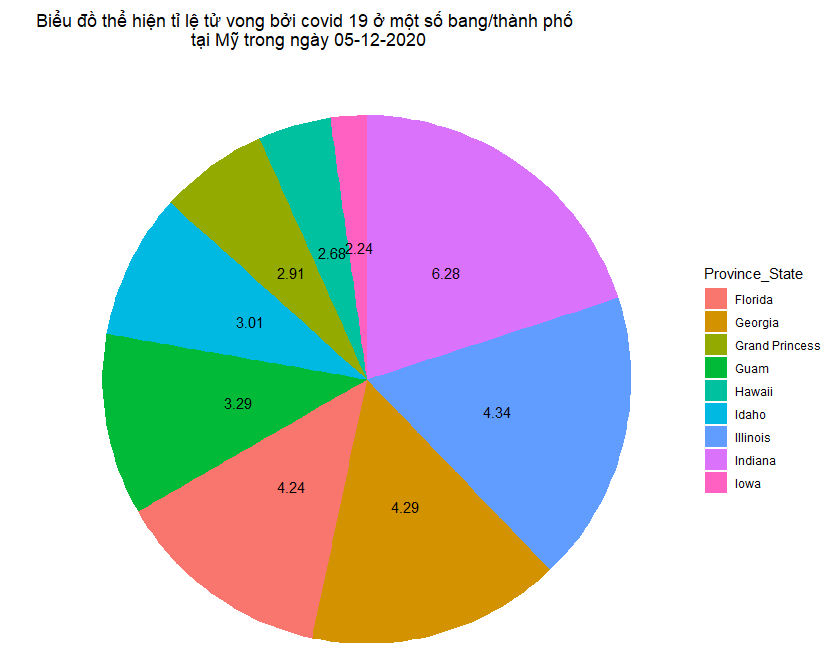
*ggplot(df[10:25,], aes(x=Mortality\_Rate, y=Province\_State, color=Province\_State)) +*

*geom\_point() +*

*labs(title="Đồ thị thể hiện tỉ lệ tử vong bởi covid 19 ở một số bang/thành phố tại Mỹ*

*trong ngày 05-12-2020",x="Mortality Rate", y="Province")*

* Mô tả đồ thị
* Là biểu đồ điểm
* Trục x ứng với “Mortality” (Tỉ lệ tử vong vì nhiễm covid 19), trục y ứng với “Province” (Tên bang/ thành phố)
* 16 điểm tương ứng với tỉ lệ tử vong vì nhiễm covid 19 trong ngày 05/12/2020 ở 16 bang/thành phố tại Mỹ
* Mỗi bang/ thành phố được thể hiện bằng mỗi màu riêng biệt
* Louisiana có tỉ lệ tử vong cao nhất trong 16 bang/thành phố
* Diamond Princess có tỉ lệ tử vong thấp nhất, tỉ lệ tử vong bằng 0
* Ý nghĩa của đồ thị: Đồ thị thể hiện và cho ta biết được về tỉ lệ tử vong vì nhiễm covid 19 được xác nhận của 16 bang/thành phố tại Mỹ trong ngày 05/12/2020.
* Xu hướng dữ liệu của đồ thị: Thể hiện sự khác nhau và chênh lệch về tỉ lệ tử vong vì nhiễm covid 19 được xác nhận ở 16 bang/ thành phố tại Mỹ ngày 05/12/2020.
  1. **Đồ thị 8:**



**Mã nguồn R:**

*ggplot(df[12:20,], aes(x='', y=Mortality\_Rate, fill=Province\_State)) +*

*geom\_bar(stat="identity", width=1) +*

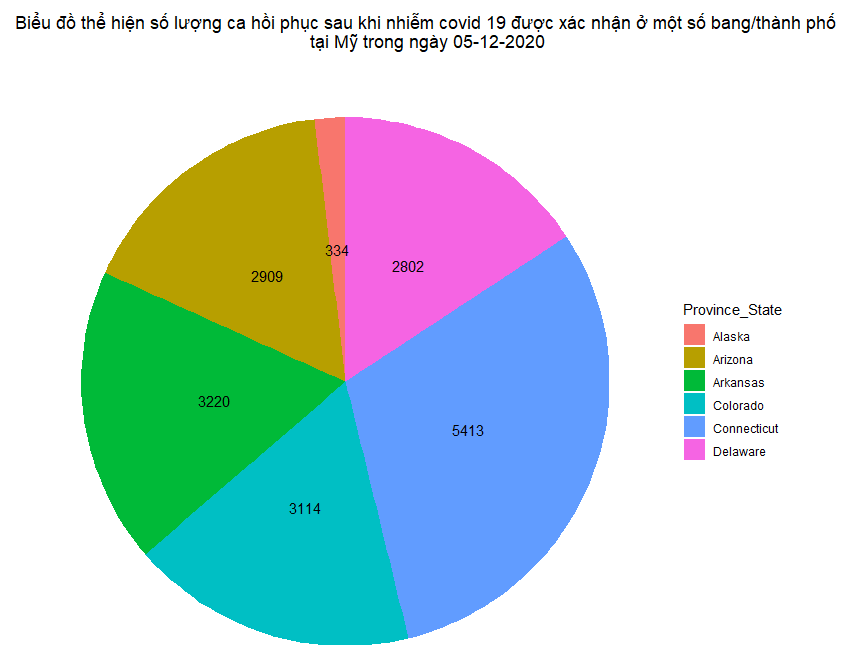
*theme\_void() + coord\_polar("y", start=0) +*

*labs(title="Biểu đồ thể hiện tỉ lệ tử vong bởi covid 19 ở một số bang/thành phố*

*tại Mỹ trong ngày 05-12-2020") +*

*geom\_text(aes(label = paste0(Mortality\_Rate)), position = position\_stack(vjust=0.5))*

* Mô tả đồ thị:
* Là biểu đồ bánh (Pie chart)
* Đồ thị so sánh tỉ lệ tử vong giữa 9 bang/thành phố: Florida, Georgia, Grand Princess, Guam, Hawaii, Idaho, Illinois, Indiana, Iowa
* Mỗi màu ứng với từng bang/thanh phố được thể hiện ở trong đồ thị
* Ý nghĩa của đồ thị: So sánh tỉ lệ tử vong do nhiễm covid 19 được xác nhận của 9 bang/ thành phố tại Mỹ vào ngày 05/12/2020.
* Xu hướng số liệu của đồ thị: Đồ thị so sánh tỉ lệ tử vong giữa các bang/thành phố, từ đó cho ta biết được về sự khác nhau, sự chênh lệch về tỉ lệ tử vong do nhiễm covid 19 giữa 9 bang/thành phố ở Mỹ ngày 05/12/2020.
  1. **Đồ thị 9:**



**Mã nguồn R:**

*ggplot(df[1:10,], aes(x='', y=Recovered, fill=Province\_State)) +*

*geom\_bar(stat="identity", width=1) +*

*theme\_void() +*

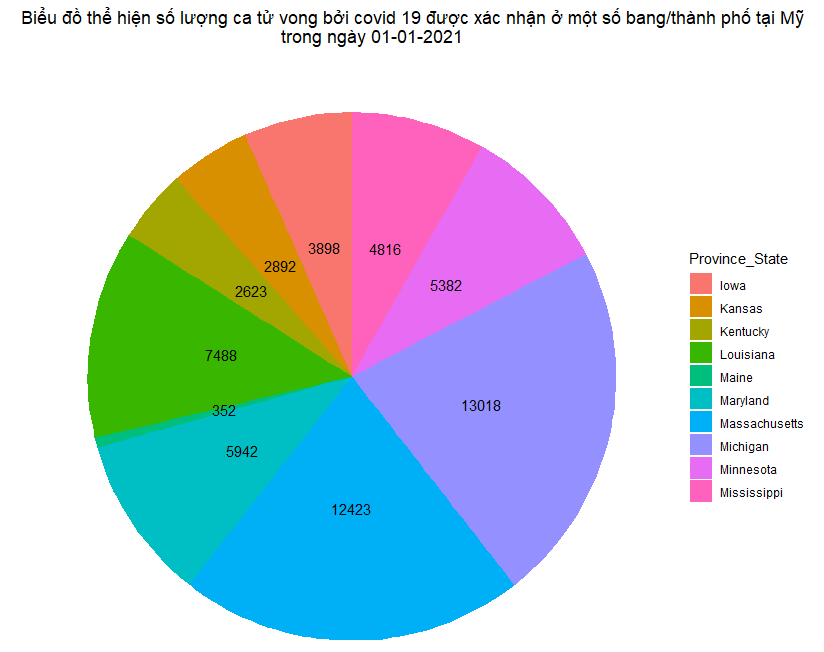
*coord\_polar("y", start=0) +*

*labs(title="Biểu đồ thể hiện số lượng ca hồi phục sau khi nhiễm covid 19 được xác nhận ở một số bang/thành phố*

*tại Mỹ trong ngày 05-12-2020") +*

*geom\_text(aes(label = paste0(Recovered)), position = position\_stack(vjust=0.5))*

* Mô tả đồ thị:
* Là biểu đồ bánh (Pie chart)
* Đồ thị miêu tả về sự chênh lệch về số lượng ca hồi phục sau khi nhiễm covid 19 giữa các bang/thành phố
* Đồ thị so sánh số lượng ca hồi phục giữa 6 bang/thành phố: Alaska, Arizona, Arkansas, Colorado, Connecticut, Delaware
* Mỗi màu ứng với từng bang/thanh phố được thể hiện ở trong đồ thị
* Connecticut chiếm tỉ lệ phần trăm lớn nhất về số lượng ca hồi phục sau khi nhiễm covid 19 trong 6 bang/thành phố
* Alaska chiếm tỉ lệ thấp nhất
* Ý nghĩa của đồ thị: So sánh Số lượng ca hồi phục sau khi nhiễm covid 19 được xác nhận của 6 bang/ thành phố tại Mỹ vào ngày 05/12/2020.
* Xu hướng số liệu của đồ thị: Đồ thị so sánh tỉ lệ phần trăm số lượng ca hồi phục, từ đó cho ta biết được về sự khác nhau và sự chênh lệch giữa số lượng ca hồi phục sau khi nhiễm covid 19 giữa 6 bang/thành phố ở Mỹ ngày 05/12/2020.
  1. **Đồ thị 10:**



**Mã nguồn R:**

*ggplot(df1[20:29,], aes(x='', y=Deaths, fill=Province\_State)) +*

*geom\_bar(stat="identity", width=1) +*

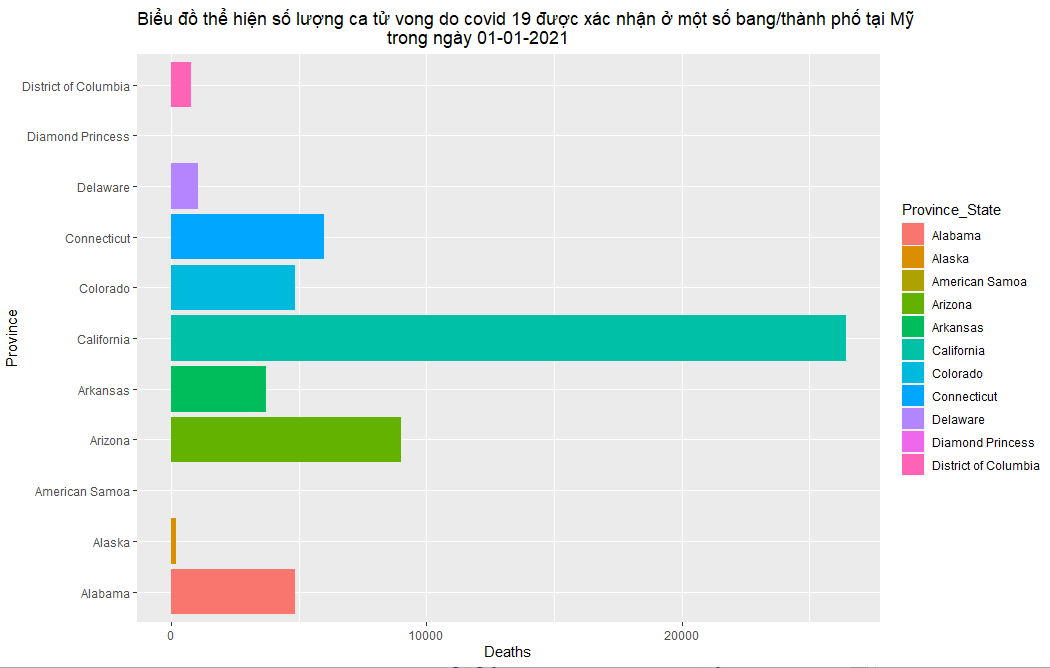
*coord\_polar("y", start=0)+*

*theme\_void() + labs(title="Biểu đồ thể hiện số lượng ca tử vong bởi covid 19 được xác nhận ở một số bang/thành phố tại Mỹ*

*trong ngày 01-01-2021") +*

*geom\_text(aes(label = paste0(Deaths)), position = position\_stack(vjust=0.5))*

* Mô tả đồ thị:
* Là biểu đồ bánh (Pie chart)
* Đồ thị miêu tả về sự chênh lệch về số lượng ca tử vong sau khi nhiễm covid 19 giữa các bang/thành phố
* Đồ thị so sánh số lượng ca tử vong giữa 10 bang/thành phố: Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi
* Mỗi màu ứng với từng bang/thanh phố được thể hiện ở trong đồ thị
* Ý nghĩa của đồ thị: Thể hiện và cho ta biết được về Số lượng ca tử vong do nhiễm covid 19 được xác nhận của 10 bang/ thành phố tại Mỹ vào ngày 01/01/2021.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được về sự khác nhau, sự chênh lệch giữa số lượng ca tử vong do nhiễm covid 19 giữa 10 bang/thành phố ở Mỹ ngày 01/01/2021.
  1. **Đồ thị 11:**

****

**Mã nguồn R:**

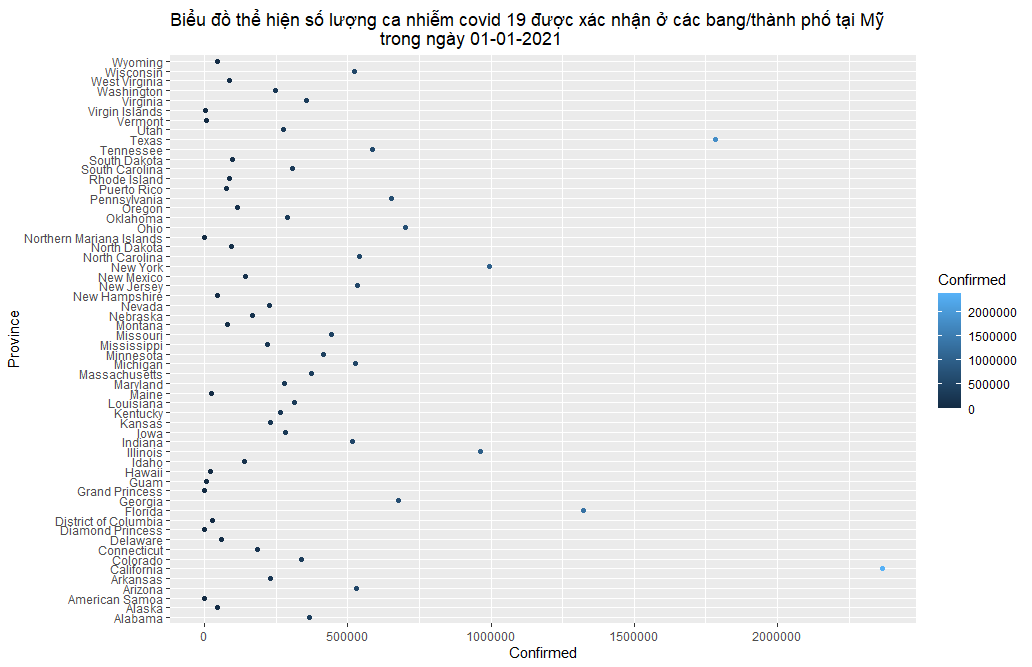
*ggplot(df1[1:11,], aes(x=Deaths, y=Province\_State)) +*

*geom\_col(aes(x=Deaths, y=Province\_State, fill = Province\_State)) +*

*labs(title="Biểu đồ thể hiện số lượng ca tử vong do covid 19 được xác nhận ở một số bang/thành phố tại Mỹ*

*trong ngày 01-01-2021",x = "Deaths", y="Province")*

* Mô tả đồ thị:
* Là biểu đồ cột
* Trục x ứng với “Deaths” ( Số lượng ca tử vong ứng với từng bang/thành phố), trục y ứng với “Province\_State” (Tên các bang/thành phố)
* Mỗi một cột màu thể hiện số lượng ca tử vong bởi covid 19 được xác nhận của mỗi bang/thành phố ứng với màu của nó
* California có số lượng ca tử vong do covid 19 gây ra lớn nhất
* Ý nghĩa của đồ thị: Đồ thị thể hiện và cho ta biết được số lượng ca tử vong bởi covid 19 được xác nhận của 11 bang/thành phố ở Mỹ trong ngày 01/01/2021.
* Xu hướng dữ liệu của đồ thị: Đồ thị cho ta biết được về sự khác nhau giữa số lượng ca tử vong bởi covid được xác nhận giữa 11 bang/thành phố ở Mỹ trong ngày 01/01/2021.
  1. **Đồ thị 12:**



**Mã nguồn R:**

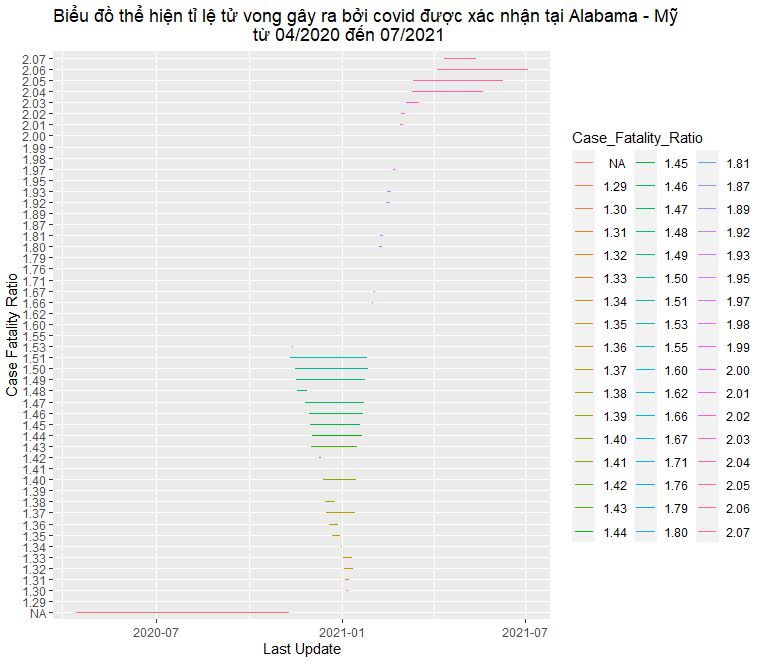
*ggplot(df1, aes(x=Confirmed, y=Province\_State, fill= Confirmed)) +*

*geom\_point(aes(color=Confirmed)) +*

*labs(title="Biểu đồ thể hiện số lượng ca nhiễm covid 19 được xác nhận ở các bang/thành phố tại Mỹ*

*trong ngày 01-01-2021",x = "Confirmed", y="Province")*

* Mô tả đồ thị
* Là biểu đồ điểm
* Trục x ứng với “Confirmed” (Số ca nhiễm được xác nhận), trục y ứng với “Province” (Tên bang/ thành phố)
* Mỗi điểm tương ứng với số lượng ca nhiễm covid 19 được xác nhận trong ngày 01/01/2021 ở các bang/thành phố tại Mỹ
* Số ca nhiễm covid 19 được xác nhận từ số lượng ít đến số lượng nhiều theo thứ tự từ màu xanh đậm chuyển dần về màu xanh nhạt.
* California có số lượng ca nhiễm covid 19 cao nhất
* Ý nghĩa của đồ thị: Đồ thị thể hiện và cho ta biết được về số lượng ca nhiễm covid 19 được xác nhận của từng bang/thành phố tại Mỹ trong ngày 01/01/2021.
* Xu hướng dữ liệu của đồ thị: Thể hiện sự khác nhau và chênh lệch về số lượng ca nhiễm covid 19 được xác nhận ở các bang/ thành phố tại Mỹ ngày 01/01/2021.
  1. **Đồ thị 13:**



**Mã nguồn R:**

*ggplot(Albm) +*

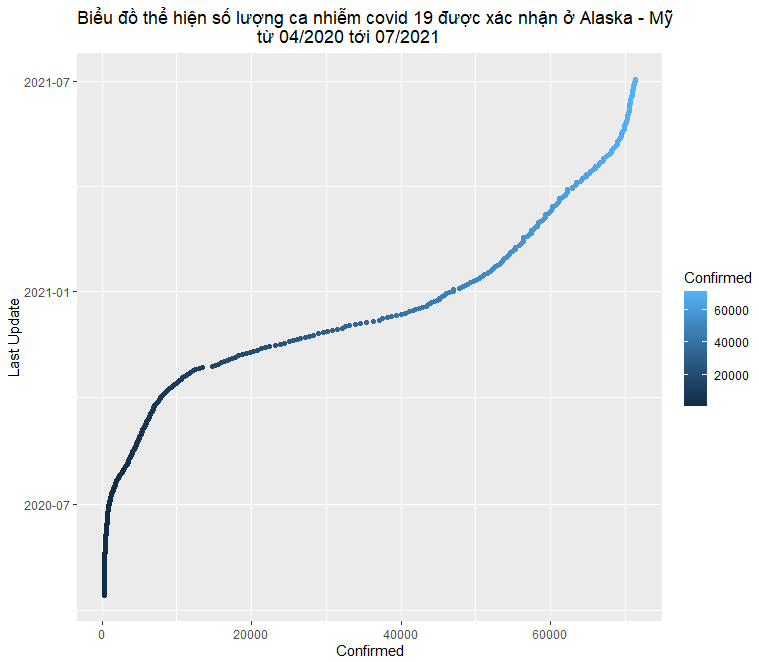
*geom\_line(aes(y=Case\_Fatality\_Ratio, x=Last\_Update, color=Case\_Fatality\_Ratio)) +*

*theme\_gray() +*

*labs(title="Biểu đồ thể hiện tỉ lệ tử vong gây ra bởi covid được xác nhận tại Alabama - Mỹ*

*từ 04/2020 đến 07/2021", x="Last Update", y="Case Fatality Ratio")*

* Mô tả đồ thị:
* Là biểu đồ đường thẳng
* Trục x ứng với “Last\_Update” (Thời điểm dữ liệu được đưa lên), trục y ứng với “Case Fatality Ratio” (Tỉ lệ tử vong)
* Mỗi đường ứng với “Case Fatality Ratio” (Tỉ lệ tử vong)
* Mỗi đường ứng với tỉ lệ tử vong bởi covid 19 được xác nhận ở Alabama trong ngày tương ứng và số lượng và màu tương ứng
* Ý nghĩa của đồ thị: Thể hiện và cho ta biết được về sự thay đổi về tỉ lệ tử vong bởi covid 19 được xác định ở Alabama từ tháng 04/2020 đến tháng 07/2021.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được rằng tỉ lệ tử vong gây ra bởi covid 19 được xác nhận ở Alabama từ tháng 04/2020 đến 07/2021 có xu hướng tăng mỗi ngày.
  1. **Đồ thị 14:**



**Mã nguồn R:**

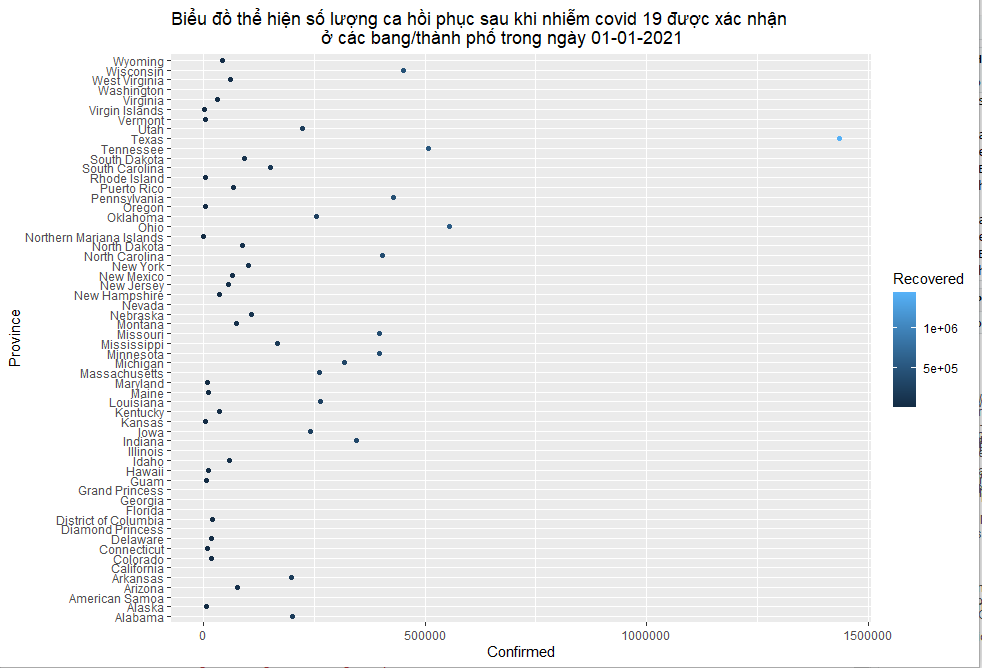
*ggplot(Alk, aes(x=Confirmed, y=Last\_Update, fill = Confirmed)) +*

*geom\_point(aes(colour = Confirmed)) +*

*labs(title="Biểu đồ thể hiện số lượng ca nhiễm covid 19 được xác nhận ở Alaska - Mỹ*

*từ 04/2020 tới 07/2021", x = "Confirmed", y="Last Update")*

* Mô tả đồ thị:
* Là biểu đồ điểm
* Trục x ứng với “Confirmed” (Số lượng ca nhiễm), trục y ứng với “Last\_Update” (Thời gian dữ liệu cập nhập)
* Mỗi điểm tương ứng với số lượng ca nhiễm covid 19 được xác nhận từng ngày từ tháng 04/2020 đến tháng 07/2021 ở Alaska – Mỹ
* Số ca nhiễm covid 19 được xác nhận từ số lượng ít đến số lượng nhiều theo thứ tự từ màu xanh đậm chuyển dần về màu xanh nhạt.
* Ý nghĩa của đồ thị: Thể hiện và cho ta biết được về sự chuyển biến về số lượng ca nhiễm covid 19 mỗi ngày được xác định ở Alaska từ tháng 04/2020 đến tháng 07/2021.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được rằng số lượng ca nhiễm covid 19 được xác nhận ở Alaska từ tháng 04/2020 đến 07/2021 có xu hướng tăng mỗi ngày.
  1. **Đồ thị 15:**

****

**Mã nguồn R:**

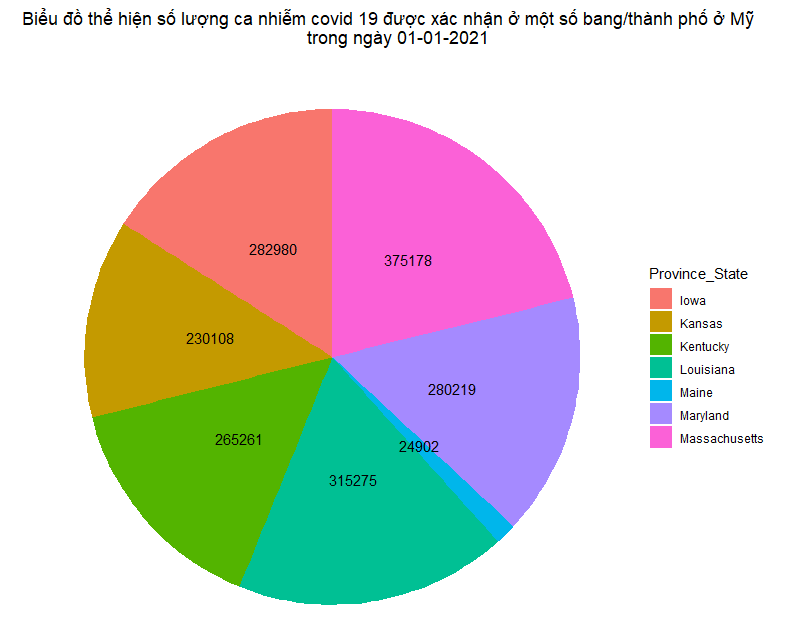
*ggplot(df1, aes(x=Recovered, y=Province\_State, fill= Recovered)) +*

*geom\_point(aes(color=Recovered)) +*

*labs(title="Biểu đồ thể hiện số lượng ca hồi phục sau khi nhiễm covid 19 được xác nhận*

*ở các bang/thành phố trong ngày 01-01-2021",x = "Confirmed", y="Province")*

* Mô tả đồ thị
* Là biểu đồ điểm
* Trục x ứng với “Recovered” (Số ca hồi phục được xác nhận), trục y ứng với “Province” (Tên bang/ thành phố)
* Mỗi điểm tương ứng với số lượng ca được xác nhận đã hồi phục sau khi nhiễm covid 19 trong ngày 01/01/2021 ở các bang/thành phố tại Mỹ
* Số ca đã hồi phục được xác nhận từ số lượng ít đến số lượng nhiều theo thứ tự từ màu xanh đậm chuyển dần về màu xanh nhạt.
* Texas có số lượng ca hồi phục cao nhất
* Ý nghĩa của đồ thị: Đồ thị thể hiện và cho ta biết được về số lượng ca đã hồi phục sau khi nhiễm covid 19 được xác nhận của từng bang/thành phố tại Mỹ trong ngày 01/01/2021.
* Xu hướng dữ liệu của đồ thị: Thể hiện sự khác nhau và chênh lệch về số lượng ca hồi phục sau khi nhiễm covid 19 được xác nhận ở các bang/ thành phố tại Mỹ ngày 01/01/2021.
  1. **Đồ thị 16:**

****

**Mã nguồn R:**

*ggplot(df1[20: 26,], aes(x='', y=Confirmed, fill=Province\_State)) +*

*geom\_bar(stat="identity", width=1) +*

*theme\_void() +*

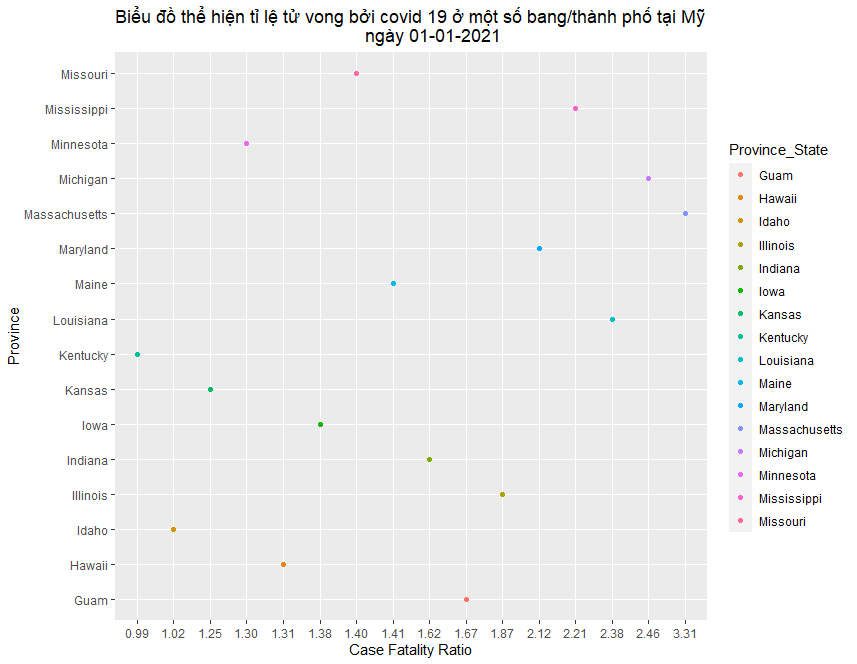
*coord\_polar("y", start=0) +*

*labs(title="Biểu đồ thể hiện số lượng ca nhiễm covid 19 được xác nhận ở một số bang/thành phố ở Mỹ*

*trong ngày 01-01-2021") +*

*geom\_text(aes(label = paste0(Confirmed)), position = position\_stack(vjust=0.5))*

* Mô tả đồ thị:
* Là biểu đồ bánh (Pie chart)
* Đồ thị so sánh số lượng ca tử vong giữa 7 bang/thành phố: Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts
* Mỗi màu ứng với từng bang/thành phố được thể hiện ở trong đồ thị
* Iowa, Louisiana, Massachusett là 3 bang/thành phố có số lượng ca nhiễm covid cao và gần bằng nhau trong 7 bang/thành phố
* Maine có số lượng ca tử vong thấp nhất
* Ý nghĩa của đồ thị: So sánh về Số lượng ca nhiễm covid 19 được xác nhận của 7 bang/ thành phố tại Mỹ vào ngày 01/01/2021.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được về sự khác nhau, sự chênh lệch giữa số lượng ca nhiễm covid 19 giữa 7 bang/thành phố ở Mỹ ngày 01/01/2021.
  1. **Đồ thị 17:**



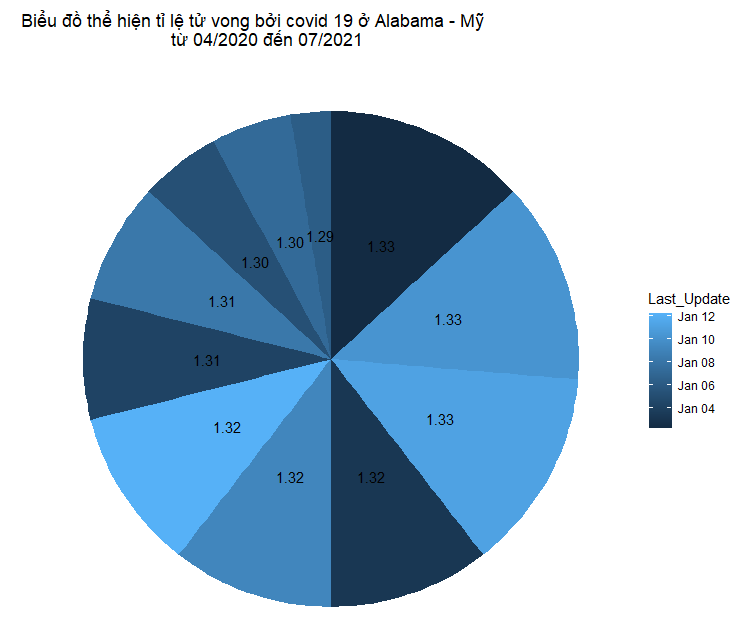
**Mã nguồn R:**

*ggplot(df1[15:30,], aes(x=Case\_Fatality\_Ratio, y=Province\_State, color=Province\_State)) +*

*geom\_point() + labs(title="Biểu đồ thể hiện tỉ lệ tử vong bởi covid 19 ở một số bang/thành phố tại Mỹ*

*ngày 01-01-2021", x="Case Fatality Ratio", y="Province")*

* Mô tả đồ thị
* Là biểu đồ điểm
* Trục x ứng với “Case Fatality Ratio” (Tỉ lệ tử vong vì nhiễm covid 19), trục y ứng với “Province” (Tên bang/ thành phố)
* 16 điểm tương ứng với 16 giá trị tỉ lệ tử vong vì nhiễm covid 19 trong ngày 01/01/2021 ở 16 bang/thành phố tại Mỹ
* Mỗi bang/ thành phố được thể hiện bằng mỗi màu riêng biệt
* Massachusett có tỉ lệ tử vong cao nhất trong 16 bang/thành phố
* Kentucky có tỉ lệ thấp nhất
* Ý nghĩa của đồ thị: Đồ thị thể hiện và cho ta biết được về sự chênh lệch tỉ lệ tử vong vì nhiễm covid 19 được xác nhận của 16 bang/thành phố tại Mỹ trong ngày 01/01/2021.
* Xu hướng dữ liệu của đồ thị: Thể hiện sự khác nhau và chênh lệch về tỉ lệ tử vong vì nhiễm covid 19 được xác nhận ở 16 bang/ thành phố tại Mỹ ngày 01/01/2021.
  1. **Đồ thị 18:**



**Mã nguồn R:**

*ggplot(Albm[1:11,], aes(x='', y=Case\_Fatality\_Ratio, fill=Last\_Update)) +*

*geom\_bar(stat="identity", width=1) +*

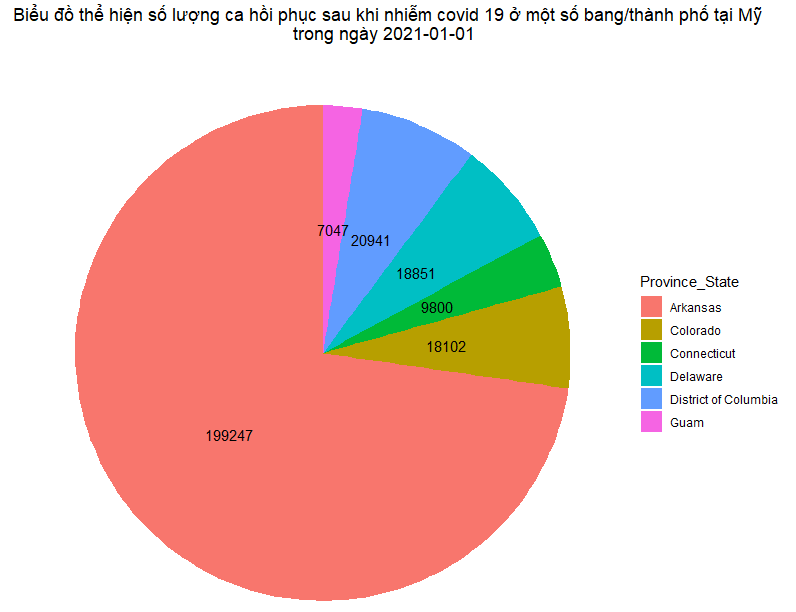
*theme\_void() + coord\_polar("y", start=0) +*

*labs(title="Biểu đồ thể hiện tỉ lệ tử vong bởi covid 19 ở Alabama - Mỹ*

*từ 04/2020 đến 07/2021") +*

*geom\_text(aes(label = paste0(Case\_Fatality\_Ratio)), position = position\_stack(vjust=0.5))*

* Mô tả đồ thị:
* Là biểu đồ bánh (Pie chart)
* Đồ thị so sánh tỉ lệ tử vong giữa 11 ngày tính từ ngày 02 đến ngày 12/ 01/2021
* Mỗi màu xanh đậm, nhạt ứng với mỗi ngày từ 02 đến 12/01/2021
* Ý nghĩa của đồ thị: So sánh về tỉ lệ tử vong do nhiễm covid 19 được xác nhận trong 11 ngày ở Alabama – Mỹ.
* Xu hướng số liệu của đồ thị: Đồ thị so sánh về tỉ lệ tử vong do covid 19 từ 02-12/01/2021, từ đó cho ta biết được về sự chênh lệch, sự thay đổi giữa tỉ lệ tử vong do nhiễm covid 19 giữa các ngày từ 02-12/01/2021. Giảm rồi tăng trở lại.
  1. **Đồ thị 19:**



**Mã nguồn R:**

*ggplot(df1[5:15,], aes(x='', y=Recovered, fill=Province\_State)) +*

*geom\_bar(stat="identity", width=1) +*

*theme\_void() +*

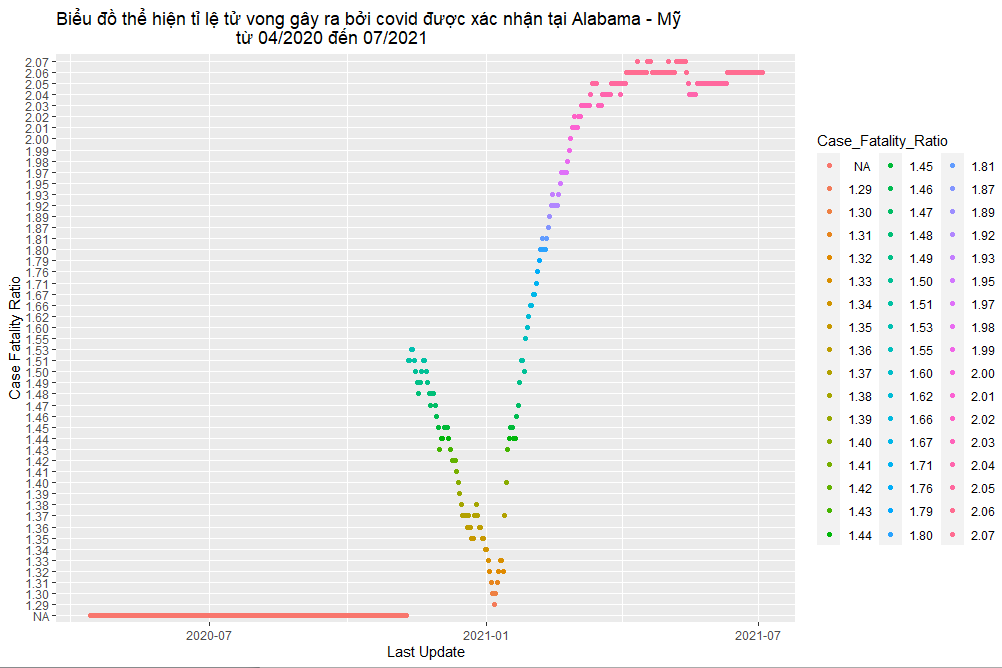
*coord\_polar("y", start=0) +*

*labs(title="Biểu đồ thể hiện số lượng ca hồi phục sau khi nhiễm covid 19 ở một số bang/thành phố tại Mỹ*

*trong ngày 2021-01-01") +*

*geom\_text(aes(label = paste0(Recovered)), position = position\_stack(vjust=0.5))*

* Mô tả đồ thị:
* Là biểu đồ bánh (Pie chart)
* Đồ thị so sánh số lượng ca hồi phục giữa 6 bang/thành phố: Arkansas, Colorado, Connecticut, Delaware, District of Columbia, Guam
* Mỗi màu ứng với từng bang/thanh phố được thể hiện ở trong đồ thị
* Arkansas có số lương ca hồi phục sau khi nhiễm covid 19 nhiều nhất
* Guam có số lượng ca hồi phục ít nhất
* Ý nghĩa của đồ thị: So sánh về Số lượng ca hồi phục sau khi nhiễm covid 19 được xác nhận của 6 bang/ thành phố tại Mỹ vào ngày 01/01/2021.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được về sự khác nhau, sự chênh lệch giữa số lượng ca hồi phục sau khi nhiễm covid 19 giữa 6 bang/thành phố ở Mỹ ngày 01/01/2021.
  1. **Đồ thị 20:**

****

**Mã nguồn R:**

*ggplot(Albm, aes(y=Case\_Fatality\_Ratio, x=Last\_Update, fill= Case\_Fatality\_Ratio)) +*

*geom\_point(aes(color=Case\_Fatality\_Ratio)) +*

*labs(title="Biểu đồ thể hiện tỉ lệ tử vong gây ra bởi covid được xác nhận tại Alabama - Mỹ*

*từ 04/2020 đến 07/2021", x="Last Update", y="Case Fatality Ratio")*

* Mô tả đồ thị:
* Là biểu đồ điểm
* Trục x ứng với “Last\_Update” (Thời điểm dữ liệu được đưa lên), trục y ứng với “Case Fatality Ratio” (Tỉ lệ tử vong)
* Mỗi điểm ứng với “Case Fatality Ratio” (Tỉ lệ tử vong)
* Mỗi điểm ứng với tỉ lệ tử vong bởi covid 19 được xác nhận ở Alabama trong ngày tương ứng và số lượng và màu tương ứng
* Ý nghĩa của đồ thị: Thể hiện và cho ta biết được về sự thay đổi về tỉ lệ tử vong bởi covid 19 được xác định ở Alabama từ tháng 04/2020 đến tháng 07/2021.
* Xu hướng số liệu của đồ thị: Đồ thị cho ta biết được rằng tỉ lệ tử vong gây ra bởi covid 19 được xác nhận ở Alabama từ tháng 04/2020 đến 07/2021 có xu hướng giảm những tháng cuối năm 2020 nhưng có xu hướng tăng kể từ đầu tháng 01/2021.