data analysis 2025 - dataviz

CYPRIEN 3/4/25 11:51AM

plots

♡ 0 ◯ 0
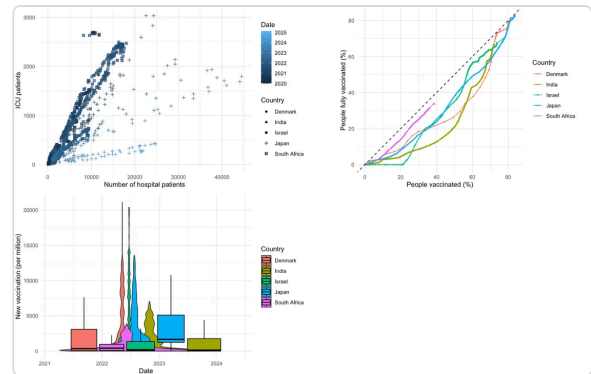
### VALIANT SHARK 3/4/25 12:04PM

Exercise 1

covid_data_subset%>% select(date,
country, icu_patients,hosp_patients)%>%
drop_na() %>% ggplot(.) +
geom_line(mapping = aes(
x = date,
y = icu_patients/hosp_patients,
colour = country
),size = 1.2)+ylab("Hospital/ICU")+
theme_classic()+
ggtitle("Hospital vs. ICU")+
scale_color_brewer(palette = "Set2") +
coord_cartesian(xlim = c(as.Date("2020-
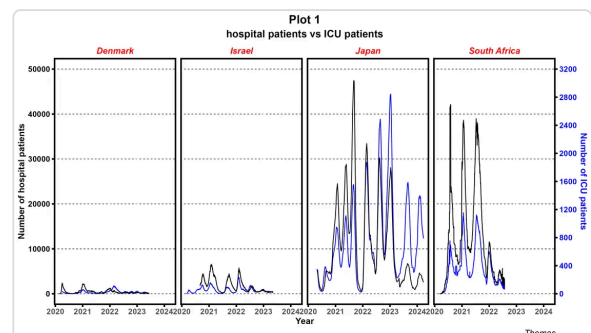05-01"), as.Date("2023-06-01")))+
facet_grid(country~.)

♡ 0 ◯ 0
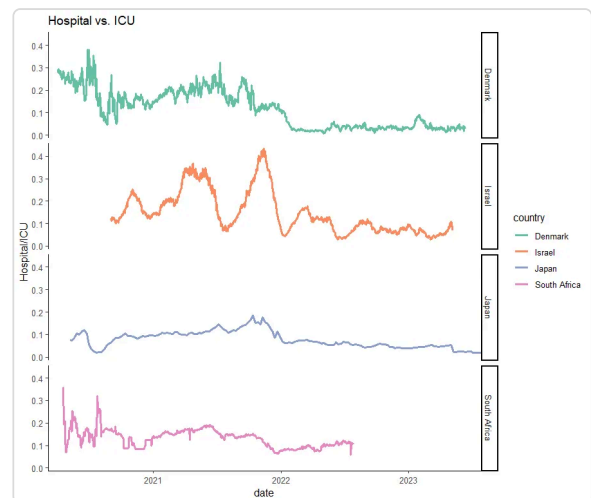
### THOMAS 3/4/25 12:10PM

Plot 1

♡ 0 ◯ 0

QUIRKY BARRACUDA 3/4/25 12:10PM

Not a scatterplot (whoops)...

```
covid_subset = covid_data %>%
  select(date, country,
icu_patients, hosp_patients) %>%
  pivot_longer(, cols = c("icu_patients",
"hosp_patients"), names_to = "type_patient",
values_to = "no._patients") %>%
  filter(!is.na(no._patients))

covid_subset %>%
  ggplot(., aes(x = date, y = no._patients,
fill = country, color = country)) +
  geom_area(stat = "identity",
position = "jitter", alpha = 0.3)+
  facet_wrap(type_patient ~ .,
        scales = "free") +
  theme_bw()+
  ylab("Number of Patients") +
  xlab("Time")+
  theme(axis.title.y = element_text(size
= 10, margin = margin(r=8)),
    axis.text.y = element_text(size=7),
    legend.text = element_text(size = 9),
    legend.title = element_text(size = 9))
```
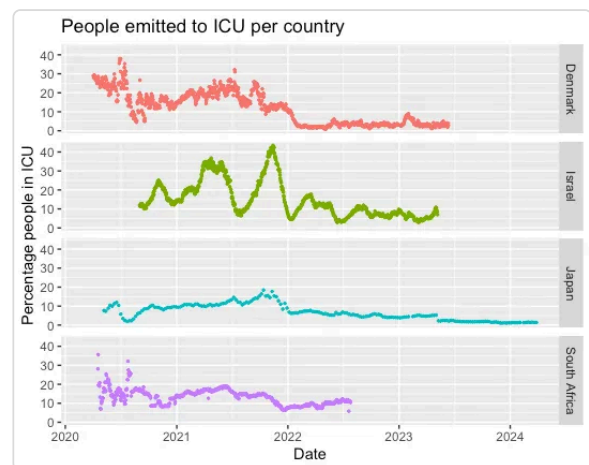

♡ 0 ⌵ 0
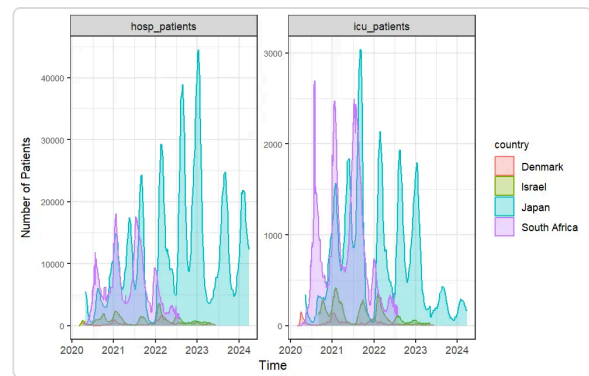
CAROLINE 3/4/25 12:15PM

Exercise 1

```
covid_data_subset_noindia
<- covid_data_subset %>%
  filter(!is.na(hosp_patients))
ggplot(covid_data_subset_noindia) +
  geom_point(mapping = aes(
    x = date,
    y = icu_patients/hosp_patients*100,
    color = country), size = 0.5) +
  facet_grid(country ~ .) +
  xlab("Date") +
  ylab("Percentage people in ICU") +
  ggtitle("People emitted to ICU per country") +
  theme(legend.position = "none")
```


Plot 1

♡ 0 ⌵ 0

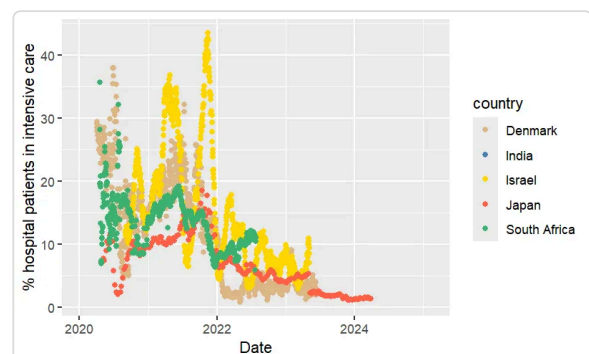
RASMUS 3/4/25 12:15PM

Ex01

```
colors <- c("burlywood", "steelblue",
"gold", "tomato", "mediumseagreen")
names(colors) <- countries
ggplot(covid_data_subset) +
  geom_point(mapping = aes(
    x = date,
    y =
icu_patients_per_million/hosp_patients_per
_million*100,
    color = country
```
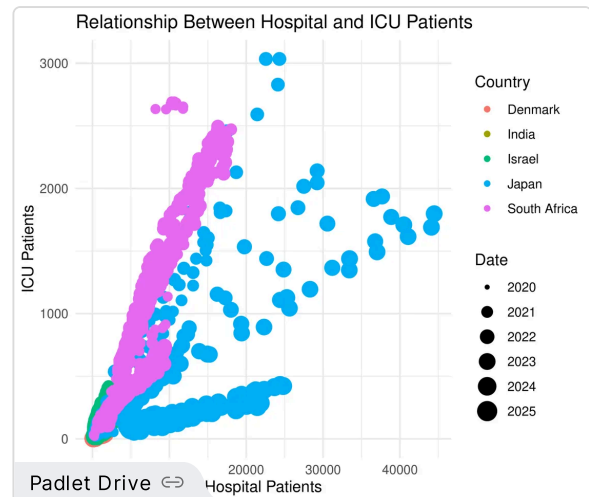
```
))+ labs(x="Date",y="% hospital
patients in intensive care") +
  scale_color_manual(values = colors)
```

♡ 0   ⬭ 0

🔗 **EMMA** 3/4/25 12:17PM

1
```
ggplot(covid_data_subset) +
  geom_point(mapping = aes(
    x = hosp_patients,
    y = icu_patients,
    color = country,
    size = date
  )) +
  labs(
    title = "Relationship Between
Hospital and ICU Patients",
    x = "Hospital Patients",
    y = "ICU Patients",
    color = "Country",
    size = "Date"
  ) +
  theme_minimal()
```
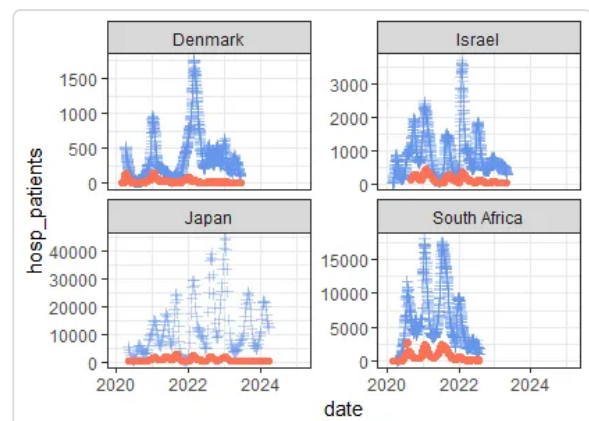

plot_1

♡ 0   ⬭ 0

🔗 **VICTOR** 3/4/25 12:19PM

```
covid_data_subset%>%
  filter(country != "India")%>%
  ggplot(aes(x = date, y = hosp_patients))+
  geom_point(alpha = 0.4, shape
= 3, color = "cornflowerblue")+
  geom_point(data = covid_data_subset%>%
filter(country != "India"), aes(x = date, y =
icu_patients), color = "coral1")+
  facet_wrap(~country, scales = "free_y")+
  theme_bw()
```
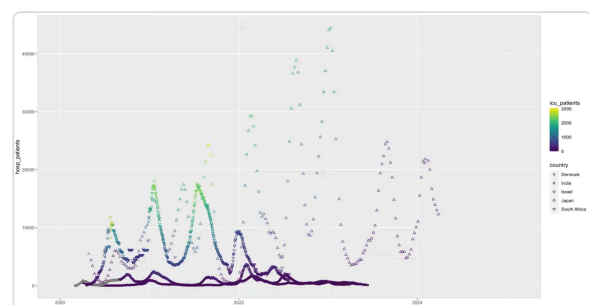

♡ 0   ⬭ 0

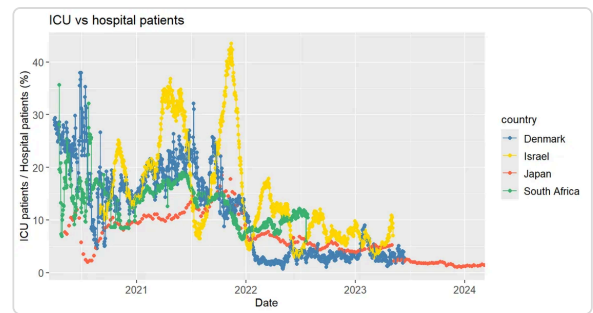🔗 **CHEERFUL HUMMINGBIRD** 3/4/25 12:22PM

Exercise 1
```
ggplot(covid_data_subset) +
  geom_point(mapping = aes(
    x = date,
    y = hosp_patients,
    color = icu_patients,
    shape = country
  )) +
  scale_color_viridis() +
  scale_shape_manual(values = shapes)
```


♡ 0   ⬭ 0

WENXI 3/4/25 12:24 PM

plot1

```
ggplot(covid_data_subset) +
    geom_point(mapping = aes(
        x = hosp_patients,
        y = icu_patients,
        color = country,  # differentiate countries
by color
        shape = year       # use different shapes
for different years
    ), size = 4) +  # increase the size of each
point
    scale_color_brewer(palette = "Set1") +  # use a
colorblind-friendly palette
    labs(
        x = "Number of Hospital Patients",
        y = "Number of ICU Patients",
        title = "COVID-19: Hospital vs ICU
Patients",
        color = "Country",
        shape = "Year"    # label for shape legend
    ) +
    theme_bw() +  # clean theme
    theme(
        panel.grid.minor = element_blank(),  #
reduce visual clutter
        legend.position = "right"
    )
```

VALIANT SHARK 3/4/25 12:29 PM

E2

covid_data_subset
ggplot(aes(x =
people_vaccinated_per_hundred, y
=people_fully_vaccinated_per_hundred,
label=date),data=covid_data_subset) +
  geom_point(mapping =
aes(colour = country), size =1.5)+
  geom_line(aes(group=country,colour
= country))+
  geom_abline(intercept = 0,
slope=1, linetype="dashed", size=2)+
 theme_classic()+
  ggtitle("Fully vaccinated vs. vaccinated")
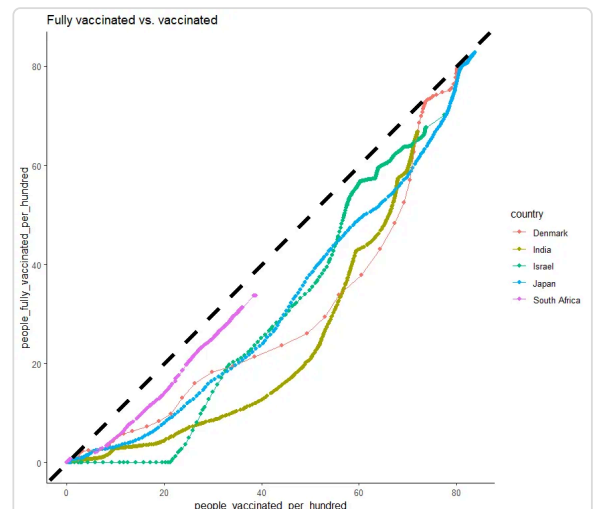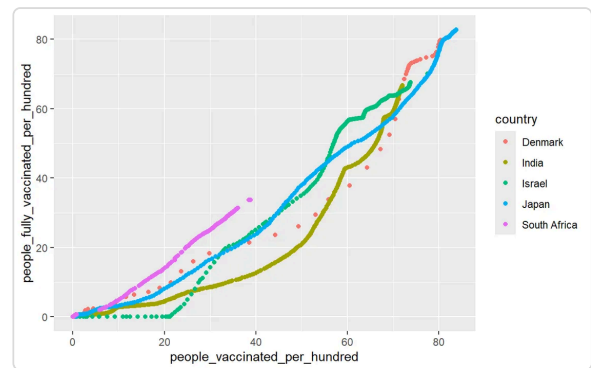
CAROLINE 3/4/25 12:35PM

Exercise 2

ggplot(covid_data_subset) +
 geom_point(mapping = aes(
  x = people_vaccinated_per_hundred,
  y = people_fully_vaccinated_per_hundred,
  color = country
 )) +
 geom_abline(aes(intercept = 0, slope = 1),
      linetype = 2) +
 xlab("People vaccinated per hundred") +
 ylab("People fully vaccinated per hundred") +
 ggtitle("Ratio of people fully
vaccinated per country") +
 scale_color_manual(
  name = "Country",
  values = colors)
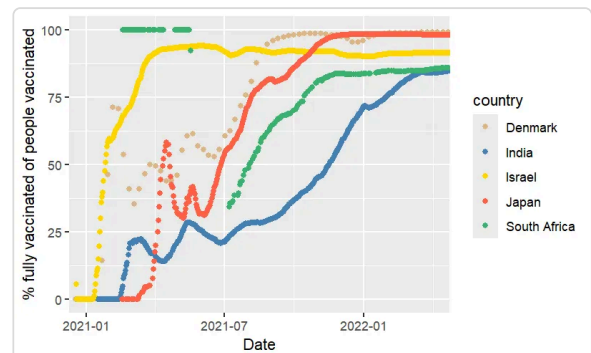
♡ 0 ◯ 0
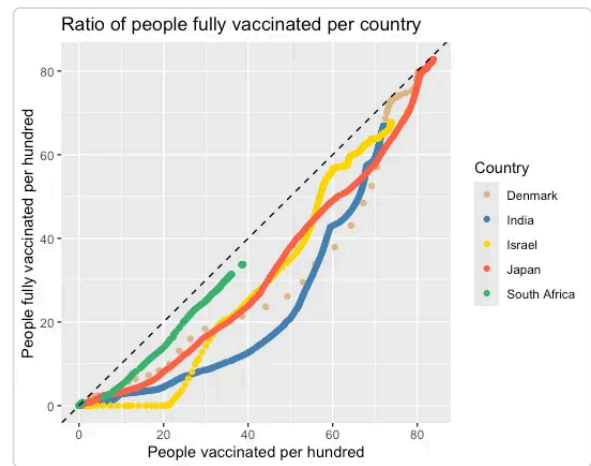
Plot 2

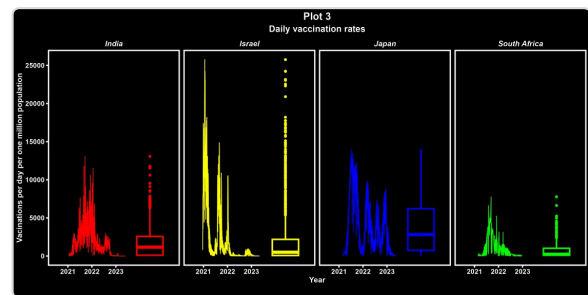
♡ 0 ◯ 0

RASMUS 3/4/25 12:38PM

Ex02

ggplot(covid_data_subset) +
 geom_point(mapping = aes(
  x = date,
  y =
people_fully_vaccinated/people_vaccinated*1
00,
  color = country
 ))+ labs(x="Date",y="% fully
vaccinated of people vaccinated") +
 scale_color_manual(values = colors) +
 coord_cartesian(xlim = c(ymd("2021-
01-01"), ymd("2022-04-01")))

♡ 0 ◯ 0

Plot 3 - had fun with colors

```
plot_df <- covid_data_subset %>%
  select (country, date,
new_vaccinations, population) %>%
  mutate (vaccinations_pr_pop =
new_vaccinations/(population/1000000)) %>%
  filter (!is.na(vaccinations_pr_pop))
ggplot () +
  geom_line(data = plot_df, aes (x = date, y
= vaccinations_pr_pop, color = country)) +
  geom_boxplot(data = plot_df,
          aes(x = as.Date("2024-06-
01"),  # Change to a relevant date
              y = vaccinations_pr_pop,
              color = country),
            width = 400,
linewidth = 1.2, fill = "black") +
  scale_x_date(breaks = seq(as.Date("2021-
01-01"), as.Date("2023-06-01"), by = "1 year"),
            limits = c(as.Date("2020-
06-01"), as.Date("2025-06-01")),
            date_labels = "%Y") +
  scale_y_continuous(breaks
= seq(0, 25000, 5000)) +
  scale_color_manual(values =
c("red", "yellow", "blue", "green")) +
  labs (title = "Plot 3",
      subtitle = "Daily vaccination rates",
      caption = "Thomas") +
  ylab ("Vacinations per day
per one million population") +
  xlab ("Year") +
  theme(axis.text.x = element_text(size=11,
face = "bold", color = "white",
                      family = font_fam),
      axis.title.y.left = element_text(size=12,
color = "white", face = "bold"),
      axis.text.y.left = element_text(size=11,
color = "white", face = "bold"),
      axis.title.y.right = element_text(size=12,
color = "blue", face = "bold"),
      axis.text.y.right = element_text(size=11,
color = "blue", face = "bold"),
      axis.title.x = element_text(size
= 12, face = "bold", color = "white",
                      family = font_fam),
      axis.ticks = element_line
(color = "white", linewidth = 1),
      rect = element_blank(),
      plot.title = element_text(size=16,
color="white", face="bold",
                      hjust
= 0.5, family = font_fam),
      plot.subtitle = element_text(size=14,
color="white",face="bold",
                      hjust
= 0.5, family = font_fam),
      plot.caption =
element_text(size = 12, face = "italic",
                      family = font_fam),
```

```
    panel.border = element_rect(color
= "white", fill = "NA", linewidth = 2),
    panel.background =
element_rect(fill = "black"),
    plot.background =
element_rect(fill = "black"),
    panel.grid = element_blank(),
    strip.text.x = element_text(size
= 12, color = "white", face = "bold.italic"),
    legend.position = "") +
  facet_grid(~ country)
```

⊝ **WENXI** 3/4/25 12:57PM

```
ggplot(covid_data_subset) +
  geom_point(mapping = aes(
    x = people_vaccinated_per_hundred,
    y =
people_fully_vaccinated_per_hundred,
    color = country
  )) +
  geom_line(mapping = aes(
    x = people_vaccinated_per_hundred,
    y =
people_fully_vaccinated_per_hundred,
    group = country,
    color = country
  )) +
  geom_abline(slope = 1, intercept =
0, linetype = "dashed", color = "black") +
  coord_fixed(ratio = 1) +
  labs(
    x = "Rate of People
Vaccinated per Hundred",
    y = "Rate of People Fully
Vaccinated per Hundred",
    title = "Vaccination Rates:
Vaccinated vs Fully Vaccinated",
    color = "Country"
  ) +
  theme_bw()
```

🕊️ 🕊️ 🕊️