Course_Work_2



EUROPEAN FLIGHTS

Introduction

Covid_19 appeared as a turbulent stormy whirlwind across the globe in a pandemic rampage in the wake of the year 2020. The stormy whirlwind impact may have gone down, but the damage in the aviation industry in respect to the numbers of operated flights has not been fully recovered, particularly in Europe which is our case study.

This visualization project work will look at the various impact Covid_19 made in the 42 European Union (EU) continent countries air flights operations before, during and post Covid_19 era. This includes examining the "Departure flights", Arrival flights" and the combination of both the departure and arrival flights, known as Total flights. This data set comprises EU flights details from Jan 01, 2016 to May 31, 2022. This will be visualized with five different research questions. The dataset "flights" is available from; https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210914-1.

Research questions

- 1. Which EU countries operated the highest number of total flights before covid_19?
- 2. Which of the EU airports operated the highest numbers of flights before Covid_19, and the adopted/adjusted figures during and post covid 19 in respect of?
 - A. The airport with the highest number of total flights before Covid 19?
 - B. The airport with the highest number of total flights during Covid_19?
 - C. Which airport has the highest number of flights turnaround after covid 19?
- 3. Which country's airport quickly show the government's willingness to adjust to normal life by lifting or reducing Covid_19 restrictions?
- 4. Which country's airport recorded the highest flight per day recovery rate in post covid 19?
- 5. Has EU Aviation Industry fully recovered in the numbers of flights per day? setwd("~/DataViz/R Script")

library(tidyverse)

```
## Warning: package 'tidyverse' was built under R version 4.2.3
## Warning: package 'ggplot2' was built under R version 4.2.3
## Warning: package 'tibble' was built under R version 4.2.3
## Warning: package 'tidyr' was built under R version 4.2.3
## Warning: package 'readr' was built under R version 4.2.3
## Warning: package 'purrr' was built under R version 4.2.3
## Warning: package 'dplyr' was built under R version 4.2.3
## Warning: package 'stringr' was built under R version 4.2.3
## Warning: package 'forcats' was built under R version 4.2.3
## Warning: package 'lubridate' was built under R version 4.2.3
## — Attaching core tidyverse packages —
                                                               - tidyverse
2.0.0 -
## √ dplyr
               1.1.1
                         ✓ readr
                                     2.1.4
## √ forcats 1.0.0
                                     1.5.0

√ stringr

## √ ggplot2 3.4.1
                         √ tibble
                                     3.2.1
## ✓ lubridate 1.9.2
                         √ tidyr
                                     1.3.0
## √ purrr
               1.0.1
## — Conflicts -
tidyverse_conflicts() —
## X dplyr::filter() masks stats::filter()
## X dplyr::lag() masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
library(tidyr)
library(gghighlight)
## Warning: package 'gghighlight' was built under R version 4.2.3
library(gapminder)
## Warning: package 'gapminder' was built under R version 4.2.3
library(tinytex)
## Warning: package 'tinytex' was built under R version 4.2.3
```

Import flights data from the flights.Csv to work environment.

```
flights <- read.csv("flights.csv")
```

Retrieve the dimension of the flights data.

```
dim(flights)
## [1] 688099 15
```

Inspect the column names in the flight data to familiarise with them.

```
colnames(flights)
    [1] "X"
                                                           "MONTH_MON"
##
                         "YEAR"
                                          "MONTH NUM"
                         "APT ICAO"
                                          "APT NAME"
##
   [5] "FLT DATE"
                                                           "STATE NAME"
  [9] "FLT_DEP_1"
                         "FLT_ARR_1"
                                          "FLT_TOT_1"
                                                           "FLT_DEP_IFR_2"
##
## [13] "FLT_ARR_IFR_2" "FLT_TOT_IFR_2" "Pivot.Label"
```

The flights data frame for comprehensive view of the whole rows, columns and observations.

```
summary(flights)
                          YEAR
                                      MONTH NUM
                                                      MONTH MON
##
         Χ
##
   Min.
          :
                 1
                     Min.
                            :2016
                                           : 1.000
                                                     Length: 688099
                                    Min.
                                                     Class :character
   1st Qu.:172026
                     1st Qu.:2017
                                    1st Qu.: 3.000
##
## Median :344050
                     Median :2019
                                    Median : 6.000
                                                     Mode :character
##
   Mean
          :344050
                     Mean
                            :2019
                                    Mean
                                           : 6.301
   3rd Qu.:516075
                     3rd Qu.:2020
                                    3rd Qu.: 9.000
## Max.
          :688099
                     Max.
                            :2022
                                    Max.
                                           :12.000
##
##
      FLT DATE
                         APT ICAO
                                            APT NAME
                                                              STATE NAME
##
   Length: 688099
                       Length:688099
                                          Length:688099
                                                             Length: 688099
##
   Class :character
                       Class :character
                                          Class :character
                                                             Class :character
##
   Mode :character
                       Mode :character
                                          Mode :character
                                                             Mode :character
##
##
##
##
##
      FLT DEP 1
                       FLT ARR 1
                                        FLT TOT 1
                                                       FLT DEP IFR 2
                                                 0.0
                                                                  0.0
##
   Min.
          : 0.00
                     Min.
                               0.00
                                                       Min.
                                      Min.
##
   1st Qu.: 5.00
                     1st Qu.:
                               5.00
                                      1st Qu.:
                                                10.0
                                                       1st Qu.:
                                                                 38.0
   Median : 17.00
                     Median : 17.00
                                                35.0
                                                       Median: 91.0
##
                                      Median :
                                             : 126.5
##
           : 63.24
                            : 63.28
                                                              : 143.7
   Mean
                     Mean
                                      Mean
                                                       Mean
##
   3rd Qu.: 71.00
                     3rd Qu.: 71.00
                                      3rd Qu.: 141.0
                                                       3rd Qu.: 195.0
##
   Max.
           :847.00
                            :813.00
                                             :1628.0
                     Max.
                                      Max.
                                                       Max.
                                                              :1039.0
##
                                                       NA's
                                                              :479785
                     FLT TOT IFR 2
                                      Pivot.Label
## FLT ARR IFR 2
##
                     Min.
                                      Length: 688099
   Min.
           : 0.0
                                0.0
##
   1st Qu.: 38.0
                     1st Qu.:
                               76.0
                                      Class :character
                     Median : 182.0
   Median : 91.0
                                      Mode :character
##
## Mean
          :143.6
                     Mean
                           : 287.3
   3rd Qu.:195.0
                     3rd Qu.: 390.0
##
## Max.
         :817.0
                     Max. :1624.0
## NA's :479785
                     NA's :479785
```

More details from the flight dataset and the categories using glimpse.

```
glimpse(flights)
## Rows: 688,099
## Columns: 15
## $ X
                  <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15,
16, 1...
                  <int> 2016, 2016, 2016, 2016, 2016, 2016, 2016, 2016,
## $ YEAR
2016, 20...
## $ MONTH_NUM
                  1, 1,...
                  <chr> "JAN", "JAN", "JAN", "JAN", "JAN", "JAN", "JAN",
## $ MONTH_MON
"JAN", ...
                  <chr> "2016-01-01", "2016-01-01", "2016-01-01", "2016-01-
## $ FLT DATE
01", ...
                  <chr> "EBAW", "EBBR", "EBCI", "EBLG", "EBOS", "EDDB",
## $ APT_ICAO
"EDDC", ...
                  <chr> "Antwerp", "Brussels", "Charleroi", "Liège",
## $ APT NAME
"Ostend-Bru...
## $ STATE_NAME
                  <chr> "Belgium", "Belgium", "Belgium", "Belgium",
"Belgium", "...
## $ FLT_DEP_1
                  <int> 4, 174, 45, 6, 7, 98, 18, 1, 401, 3, 122, 92, 172,
276, ...
## $ FLT ARR 1
                  <int> 3, 171, 47, 7, 7, 99, 21, 1, 341, 4, 130, 90, 175,
236, ...
## $ FLT TOT 1
                  <int> 7, 345, 92, 13, 14, 197, 39, 2, 742, 7, 252, 182,
347, 5...
## $ FLT_DEP_IFR_2 <int> NA, 174, 45, NA, NA, NA, NA, NA, 401, NA, 125, NA,
NA, 2...
## $ FLT ARR IFR 2 <int> NA, 161, 45, NA, NA, NA, NA, NA, 306, NA, 129, NA,
NA, 2...
## $ FLT TOT IFR 2 <int> NA, 335, 90, NA, NA, NA, NA, NA, 707, NA, 254, NA,
NA, 5...
                 <chr> "Antwerp (EBAW)", "Brussels (EBBR)", "Charleroi
## $ Pivot.Label
(EBCI)",...
```

Data Dictionary

The characteristic of our variables are as stated below; what they stands for and what they mean

X is the serial number for the rows.

YEAR is an integer which reference the year flight was made.

MONTH_NUM is an numerical integer which stands for Month

MONTH_MON is a text character which represent month in a three letter code

FLT_DATE is a text character that represents the particular date of flights

APT_ICAO is a text character that represents the International Civil Aviation Organisations(ICAO) 4-letter airport designator name.

APT_NAME is a text character that the airport name.

STATE_NAME is a text character that represent the name of the country in which the airport is located.

FLT_DEP_1 is an integer that represents the number of instrument flight rules (IFR) departures as provided by network manager.

FLT_ARR_1 is an integer that represents the number of IFR arrival as provided by network manager.

FLT_TOT_1 is an integer that represents the total number of IFR movements as provided by network manager.

FLT_DEP_IFR_2 is an integer that represents the number of IFR departures as provided by the airport operator.

FLT_ARR_IFR_2 is an integer that represents the number of IFR arrivals as provided by the airport operator.

FLT_TOT_IFR_2 is an integer that represents the total number of IFR movement as provided by the airport operator.

Pivot.Label is a text character that shows the combination of name of country and airport.

Data Cleaning Let look at the number of NA (missing) values present in our data set

```
sum(is.na(flights))
## [1] 1439355
sum(is.na(flights$APT_NAME))
## [1] 0
sum(is.na(flights$STATE_NAME))
## [1] 0
sum(is.na(flights$FLT_TOT_1))
## [1] 0
sum(is.na(flights$YEAR))
## [1] 0
sum(is.na(flights$YEAR))
## [1] 0
```

There are 1,439,355 NAs in flights data set, however, there are no NA values in our variable of concern.

Convert YEAR from integer to a date factor using lubridate.

```
flights$FLT_DATE <- ymd(flights$FLT_DATE)

class(flights$FLT_DATE)

## [1] "Date"</pre>
```

Research questions

1. Which EU countries operated the highest number of total flights before covid_19?

Using the total numbers of flights (departures and arrivals), arranged in descending order and rounded up to one decimal place

Convert the STATE_NAME to a factor to allow us arrange our bars in descending order.

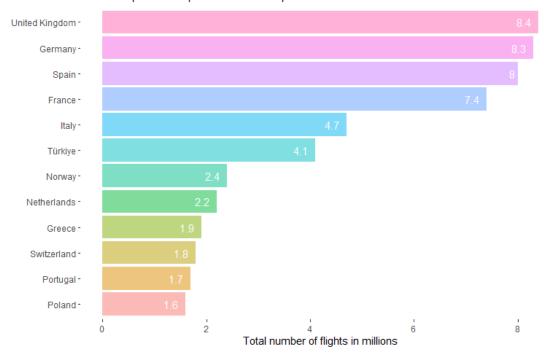
Select the list of the first 12 countries with highest total flights before Covid_19.

```
## 2 Germany
                              8.32
                                       8.3
                              8.01
## 3 Spain
                                       8
## 4 France
                              7.38
                                       7.4
                              4.73
## 5 Italy
                                       4.7
## 6 Türkiye
                              4.13
                                       4.1
## 7 Norway
                              2.39
                                       2.4
## 8 Netherlands
                              2.20
                                       2.2
## 9 Greece
                              1.91
                                       1.9
## 10 Switzerland
                              1.82
                                       1.8
## 11 Portugal
                              1.66
                                       1.7
## 12 Poland
                              1.61
                                       1.6
```

Graphical representation

```
ggplot(CountryTop12,
       aes(Round 1,
           STATE_NAME,
           fill = STATE_NAME)
       ) +
  geom_col(alpha = 0.5) +
  geom_text(aes(label =Round_1), stat = "identity", hjust = 1.5, colour =
"white")+
  labs(title = "The 12 countries with highest total number of flights before
covid_19",
       subtitle = "Source: https://ec.europa.eu/eurostat/web/products-
eurostat-news/-/ddn-20210914-1",
       x= "Total number of flights in millions",
       y = NULL) +
  theme(legend.position = NULL) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.background = element_blank()) +
  guides(fill = "none")
```

The 12 countries with highest total number of flights before covid_19 Source: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210914-1



United Kingdom is the country with the highest number of total flights (departure and arrival) before Covid_19, a little ahead of Spain and France respectively. Italy, Turkiye, Normay, netherlands, Greece, Switzerland, Portugal, Poland respectively followed the queue.

Research question

2A.

The airport with the highest number of total flights before Covid_19?

Using mean of the total numbers of flights (departures and arrivals), arranged in descending order and rounded up to one decimal place

```
summarise(Mean_TOT = mean(FLT_TOT_1)) %>%
arrange(desc(Mean_TOT))

A2016_2019 <-
    A2016_2019%>%
    mutate(Mean_TOT_2 = round(A2016_2019$Mean_TOT, digits = 0))
```

The new data frame (A2016_2019) now has both the mean and rounded up mean of the Total flights per airport.

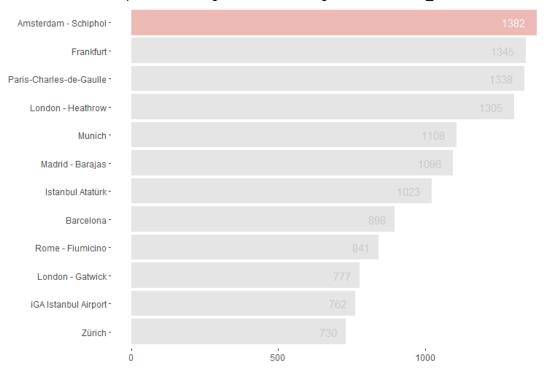
List the first 12 airports.

```
A2016 2019 Mean <-
  A2016_2019 [1:12,]
list(A2016 2019 Mean)
## [[1]]
## # A tibble: 12 × 3
      APT_NAME
##
                              Mean_TOT Mean_TOT_2
                                  <dbl>
                                             <dbl>
##
      <chr>>
## 1 Amsterdam - Schiphol
                                  1382.
                                              1382
## 2 Frankfurt
                                  1345.
                                              1345
## 3 Paris-Charles-de-Gaulle
                                  1338.
                                              1338
## 4 London - Heathrow
                                  1305.
                                              1305
## 5 Munich
                                  1108.
                                              1108
## 6 Madrid - Barajas
                                  1096.
                                              1096
## 7 Istanbul Atatürk
                                  1023.
                                              1023
## 8 Barcelona
                                   898.
                                               898
## 9 Rome - Fiumicino
                                   841.
                                               841
## 10 London - Gatwick
                                   777.
                                               777
## 11 iGA Istanbul Airport
                                   762.
                                               762
## 12 Zürich
                                               730
                                   730.
```

Convert the APT_NAME to a factor to allow us arrange our bars in descending order

Highlight the airport with the highest flights number before Covid_19.

Airport with the highest 'number of flights' before Covid_19



Amsterdam-Schiphol is the airport with the highest numbers of flights pre-covid_19. Closely followed by Frankfurt, and Paris-Charles-de-Gualle, London-Heathrow, respectively and the rest.

2B.

The airport with the highest number of total flights during Covid_19?

Using mean of the total numbers of flights (departures and arrivals), arrange in descending order and rounded up to whole number.

```
A2020_A <-
flights %>%
filter(YEAR == 2020) %>%
select(FLT_DATE,
```

List the first 12 airports for better comparison

```
A2020 Mean <-
  A2020 [1:12,]
list(A2020 Mean)
## [[1]]
## # A tibble: 12 × 3
                               Mean_TOT Mean_TOT_2
##
      APT NAME
##
                                  <dhl>
                                              <dbl>
      <chr>>
## 1 Amsterdam - Schiphol
                                   644.
                                                644
## 2 Paris-Charles-de-Gaulle
                                   603.
                                                603
## 3 Frankfurt
                                   580.
                                                580
## 4 London - Heathrow
                                   559.
                                                559
## 5 iGA Istanbul Airport
                                                504
                                   504.
## 6 Madrid - Barajas
                                                453
                                   453.
## 7 Munich
                                                394
                                   394.
## 8 Istanbul Sabiha Gökçen
                                   371.
                                                371
## 9 Barcelona
                                   335.
                                                335
## 10 Oslo - Gardermoen
                                   335.
                                                335
## 11 Athens
                                   298.
                                                298
## 12 Vienna
                                                295
                                   295.
```

Convert the APT_NAME to a factor to allow us arrange the bars in descending order

```
A2020_Mean$APT_NAME <- factor(A2020_Mean$APT_NAME)
A2020_Mean$APT_NAME <- fct_reorder(A2020_Mean$APT_NAME, A2020_Mean$Mean_TOT)

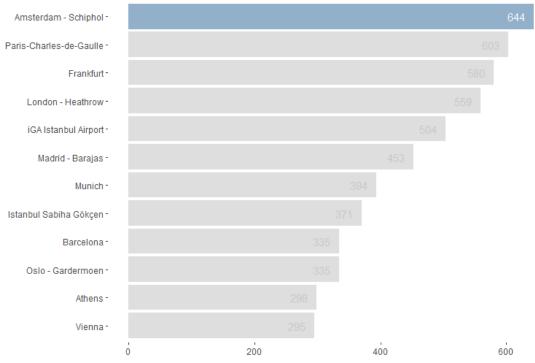
class(A2020_Mean$APT_NAME)

## [1] "factor"
```

Let plot the graph using ggplot.

```
ggplot(A2020 Mean,
       aes(Mean TOT,
           APT_NAME,
           fill = APT NAME)
       ) +
  geom_col(fill = "steelblue", alpha = 0.5) +
    geom_text(aes(label = Mean_TOT_2), stat = "identity", hjust = 1.5, colour
= "white") +
  labs(title = "Airport with highest 'number of flights' during Covid_19",
       x= NULL,
       y = NULL) +
  gghighlight(APT NAME == "Amsterdam - Schiphol") +
  theme(legend.position = NULL) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.background = element blank()) +
  guides(fill = "none")
```

Airport with highest 'number of flights' during Covid_19



Amsterdam-Schiphol continued its humanitarian services even in the face of the pandemic and standing out as the airport with the highest numbers of flights mean during covid_19. Frankfurt had dropped to third, which could mean that, Germany either placed high restrictions on flights or Frankfurt city was badly affected and people boycott the airport, while France became more welcoming.

2C.

Which airport has the highest number of flights turnaround after covid_19?

Using mean of the total numbers of flights (departures and arrivals), arrange in descending order and rounded up to whole number.

```
A2021_2022 <-
  flights %>%
  filter(YEAR >= 2021 & YEAR <= 2022) %>%
  select(FLT DATE,
         MONTH_MON,
         YEAR,
         FLT_DEP_1,
         FLT_ARR_1,
         FLT_TOT_1,
         STATE NAME,
         APT NAME
         ) %>%
  group by (APT NAME) %>%
  summarise(Mean_TOT = mean(FLT_TOT_1)) %>%
  arrange(desc(Mean_TOT))
A2021_2022 <-
  A2021 2022 %>%
  mutate(Mean_TOT_2 = round(A2021_2022$Mean_TOT,
                             digits = 0))
```

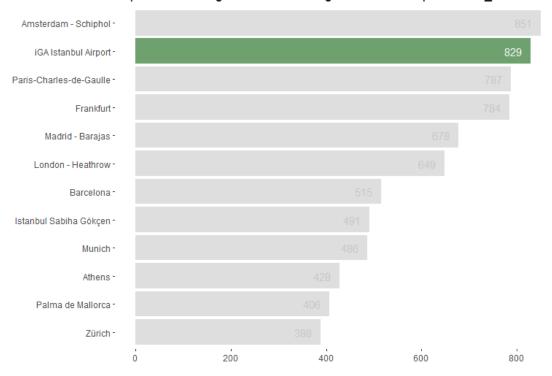
Let choose the first 12 airports for better representations & evaluation

```
A2021 2022 Mean <-
  A2021_2022 [1:12,]
list(A2021_2022_Mean)
## [[1]]
## # A tibble: 12 × 3
                               Mean TOT Mean TOT 2
##
      APT NAME
##
      <chr>>
                                  <dbl>
                                             <dbl>
## 1 Amsterdam - Schiphol
                                   851.
                                               851
## 2 iGA Istanbul Airport
                                   829.
                                               829
## 3 Paris-Charles-de-Gaulle
                                   787.
                                               787
## 4 Frankfurt
                                               784
                                   784.
## 5 Madrid - Barajas
                                   678.
                                               678
## 6 London - Heathrow
                                               649
                                   649.
## 7 Barcelona
                                               515
                                   515.
## 8 Istanbul Sabiha Gökçen
                                   491.
                                               491
## 9 Munich
                                               486
                                   486.
## 10 Athens
                                   428.
                                               428
## 11 Palma de Mallorca
                                   406.
                                               406
## 12 Zürich
                                               388
                                   388.
```

Let convert the APT_NAME to a factor to allow us arrange our bars in descending order

```
A2021 2022 Mean$APT NAME <- factor(A2021 2022 Mean$APT NAME)
A2021_2022_Mean$APT_NAME <- fct_reorder(A2021_2022_Mean$APT_NAME,
                                        A2021_2022_Mean$Mean_TOT)
class(A2021_2022_Mean$APT_NAME)
## [1] "factor"
ggplot(A2021_2022_Mean,
       aes(Mean_TOT_2,
           APT_NAME,
           fill = APT NAME)
       ) +
  geom_col(fill = "darkgreen", alpha = 0.5) +
  geom_text(aes(label = Mean_TOT_2), stat = "identity", hjust = 1.5, colour =
"white") +
  gghighlight(APT_NAME == "iGA Istanbul Airport") +
  labs(title = "Airport with the highest 'number of flights' turnaround post
Covid_19",
       x= NULL,
       V = NULL) +
  theme(legend.position = NULL) +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element blank(),
        panel.background = element_blank()) +
  guides(fill = "none")
```

Airport with the highest 'number of flights' turnaround post Covid_19



Amsterdam-Schiphol maintained the leading in all the era under considerations. Surprisingly, IGA Istanbul (Turkey airport) made 2nd position which was 7th on the queue before Covid_19. Turkey Civil Aviation Industry (TCAI) made research on reason while passengers lost confidence in the flights operations. "Another important aspect for demand recovery is that airlines have to investigate the reasons behind the decrease in passenger confidence and how to regain it". (Muhammet Deveci, & et al., 2022). This obviously paid off. Also, the government provided a 24/7 5,000 square meter PCR testing center with 20,000 capacity daily inside the Istanbul Airport terminal serving the passengers. All test can do quickly with results out with delay, according of Istanbul Airport report (online)

Research question

3. Which country's airport shows the government quick willingness to adjust to normal life by lifting or reducing Covid_19 restrictions?

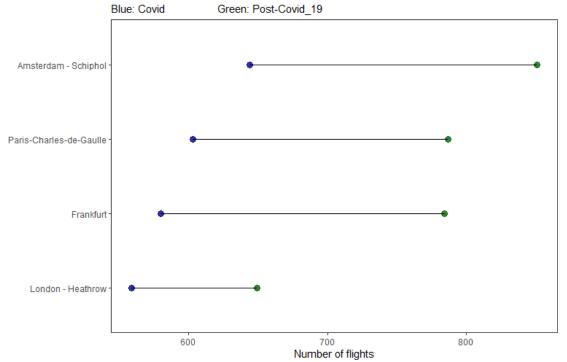
Choose the top 4 airports in pre_covid_19 and consider their performance during, and after covid_19 using lollypop graph.

```
A2016 2019 T4 <-
  A2016_2019_Mean [1:4,]
A2020 T4 <-
  A2020_Mean [1:4,]
A2021 2022 T4 <-
  A2021_2022_Mean [1:6,]
list(A2016 2019 T4)
## [[1]]
## # A tibble: 4 × 3
## APT NAME
                             Mean_TOT Mean_TOT_2
##
     <fct>
                                 <dbl>
                                            <dbl>
## 1 Amsterdam - Schiphol
                                 1382.
                                             1382
## 2 Frankfurt
                                 1345.
                                             1345
## 3 Paris-Charles-de-Gaulle
                                 1338.
                                             1338
## 4 London - Heathrow
                                 1305.
                                             1305
New flights <-
  merge(A2016 2019 T4,
        A2020 T4,
        by = "APT_NAME")
New flights2 <-
  merge(New_flights,
        A2021 2022 T4,
        by = "APT NAME")
New flights3 <-
 New flights2 %>%
```

Let convert the Airport to a factor to allow arranging the lollipop in proper order

```
New_flights3$Airport <- factor(New_flights3$Airport)</pre>
New_flights3$Airport <- fct_reorder(New_flights3$Airport,</pre>
                                     New flights3$Covid)
class(New_flights3$Airport)
## [1] "factor"
ggplot(New flights3,
       aes(x = Airport)
       ) +
  geom_point(aes
             (y = Covid),
             colour = "darkblue",
             size = 3,
             alpha = 0.8) +
  geom_point(aes
             (y = Post_Covid),
             colour = "darkgreen",
             size = 3,
             alpha = 0.8) +
  geom_segment(aes
               (x = Airport,
                 y = Covid
                 xend = Airport,
                 yend = Post Covid)) +
  labs(title = "Flights resuscitation relationship after covid_19",
       subtitle = "Blue: Covid
                                                 Green: Post-Covid 19",
        y = "Number of flights",
       x = NULL) +
  theme bw() +
  theme(panel.grid.major = element blank(),
        panel.grid.minor = element blank(),
        panel.background = element_blank()) +
  coord flip()
```

Flights resuscitation relationship after covid_19



Netherlands (Amstedam - Schiphol airport) shows quicker willingness to adjust back to normal life ahead of others.

Research Question

4. Which country's airport recorded the highest flight per day after (post) Covid_19.

What is the behavioural pattern of the European aviation operations before the appearance of Covid_19. in respect of

```
flights_4 <-
  flights %>%
  filter(APT_NAME == c("Amsterdam - Schiphol",
                          "Frankfurt",
                          "London - Heathrow",
                          "Paris-Charles-de-Gaulle")
           ) %>%
  select(FLT_DATE,
         YEAR,
         APT_NAME,
         FLT_TOT_1
         ) %>%
  drop_na(
  ) %>%
  rename_with(.cols = 3,
              ~"Airports")
```

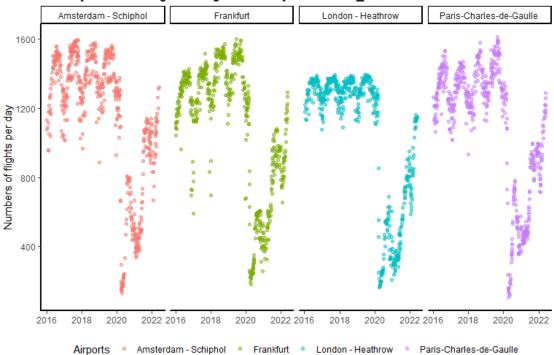
```
## Warning: There was 1 warning in `filter()`.
## i In argument: `==...`.
## Caused by warning in `APT_NAME == c("Amsterdam - Schiphol", "Frankfurt",
"London - Heathrow",
       "Paris-Charles-de-Gaulle")`:
## ! longer object length is not a multiple of shorter object length
ggplot(flights_4,
       aes(
         x = FLT_DATE
         y = FLT_TOT_1
         colour = Airports)
       ) +
  geom_point(alpha = 0.5) +
  theme_test() +
  theme(legend.position = "bottom") +
  labs(title = "Country with the hightest flights per day after Covid_19",
x = NULL
y = "Numbers of flights per day")
```

Country with the hightest flights per day after Covid_19



Using facet_wrap to segment each airport separately for better visualization.

Country With The Hightest Flights Per Day After Covid 19



Netherlands recorded the highest numbers of flights by day post Covid_19. (Seveno Victoria, 2022). According to Seveno V. (2022) thanks to the EU digital covid_19 certificate and the relaxation of various international travel restrictions. London - Heathrow has a 100,000 daily passengers limits up until October 2022. (Cunningham Ed, 2022), obviously the reason why they were left behind.

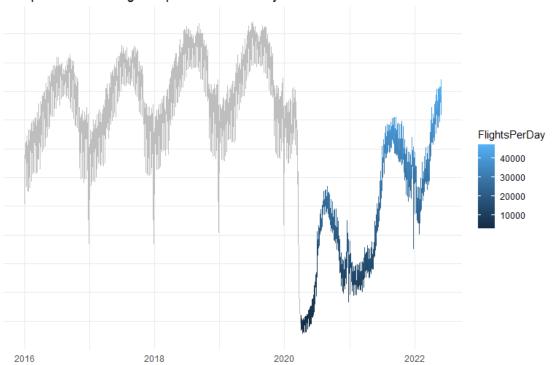
Research question

5. Has EU Aviation Industry fully recovered in the numbers of flights per day?

```
flights_Day <-
  flights %>%
  group_by(FLT_DATE) %>%
  summarise(FlightsPerDay = sum(FLT_TOT_1))

flights_Day %>%
  ggplot(aes(
    x = FLT_DATE,
    y = FlightsPerDay
```

European Aviation Flights Operations Recovery Pattern



EU Aviation Industry has not fully recovered in it total number of flights per day compared to pre-Covid_19. There is still about 20,000 daily flights difference between pre and post-Covid_19

Conclusion

EU aviation industry is not yet where they were before Covid_19 emergence, but there have been great improvement from 2021. and there seems to be possibilities of getting there in the next 2 years or 3 years if the growth rate continues. Although, one of the imprint of covid_19 is the new wave of virtual system of meeting, learning, participation etc. if this continues, we might not experience the volumes like it use to be in EU and the globe at large.

Currently, London - Heathrow has reclaimed the position of the busiest airport as at the time of this report (April 2023), (Cunningham Ed, 2022), and has taken over from

Amsterdam - Schiphol. There will need for further investigation to get to the root cause of this. However,

Reference:

Instanbul Airport, Instanbul Airport Test Centre is at your services 24/7! online article, [viewed 24 April 2023] Available from: https://www.istairport.com/en/announcements/istanbul-airport-test-center-is-at-your-service-24-7?locale=en

Muhammet Deveci, & et al., (2022) Impact of COVID-19 pandemic on the Turkish civil aviation industry, Sustainable Operations and Computers, Volume 3, Pages 93-102, ISSN 2666-4127, [viewed 24 April 2023] Available from: https://doi.org/10.1016/j.susoc.2021.11.002. , https://www.sciencedirect.com/science/article/pii/S2666412721000490)), Impact of COVID-19 pandemic on the Turkish civil aviation industry - PMC (nih.gov)

Cunningham Ed, 2022, These are officially the Europe busiest airport right now. Online article [viewed: 24 April 2023] Available from: Europe's Busiest Airports include London and Amsterdam (timeout.com)

Seveno Victoria, 2022. Schiphol once again one of the Eupore's five busiest airport. Online article [viewed 24 April 2023] Available from: Schiphol once again one of Europe's five busiest airports (iamexpat.nl)

```
title: "Course_Work_2"
author: "Olugbenga Solesi_S2242452"
date: "2023-04-03"
output:
   word_document:
     fig_width: 8
     fig_height: 5.5
   html_document: default
   pdf_document: default
editor_options:
   markdown:
     wrap: 72
---
   ```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)

EUROPEAN FLIGHTS
```

### ## Introduction

Covid\_19 appeared as a turbulent stormy whirlwind across the globe in a pandemic rampage in the wake of the year 2020. The stormy whirlwind impact may have gone down, but the damage in the aviation industry in respect to the numbers of operated flights has not been fully recovered, particularly in Europe which is our case study.

This visualization project work will look at the various impact Covid\_19 made in the 42 European Union (EU) continent countries air flights operations before, during and post Covid\_19 era. This includes examining the "Departure flights", Arrival flights" and the combination of both the departure and arrival flights, known as Total flights. This data set comprises EU flights details from Jan 01, 2016 to May 31, 2022. This will be visualized with five different research questions. The dataset "flights" is available from;

<a href="https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210914-1">https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210914-1</a>.

### ## Research questions

- 1. Which EU countries operated the highest number of total flights before covid 19?
- 2. Which of the EU airports operated the highest numbers of flights before Covid\_19, and the adopted/adjusted figures during and post\_covid\_19 in respect of?
  - A. The airport with the highest number of total flights before Covid 19?
  - B. The airport with the highest number of total flights during Covid 19?
  - C. Which airport has the highest number of flights turnaround after covid 19?
- 3. Which country's airport quickly show the government's willingness to adjust to normal life by lifting or reducing Covid\_19 restrictions?
- 4. Which country's airport recorded the highest flight per day recovery rate in post covid\_19?
- 5. Has EU Aviation Industry fully recovered in the numbers of flights per day?

```
> setwd("\~/DataViz/R Script")
```{r}
library(tidyverse)
library(tidyr)
library(gghighlight)
library(gapminder)
```

```
library(tinytex)
Import flights data from the flights.Csv to work environment.
```{r}
flights <- read.csv("flights.csv")</pre>
Retrieve the dimension of the flights data.
```{r}
dim(flights)
Inspect the column names in the flight data to familiarise with them.
```{r}
colnames (flights)
The flights data frame for comprehensive view of the whole rows, columns
and observations.
```{r}
summary(flights)
More details from the flight dataset and the categories using glimpse.
```{r}
glimpse(flights)
Data Dictionary
The characteristic of our variables are as stated below; what they
stands for and what they mean
X is the serial number for the rows.
YEAR is an integer which reference the year flight was made.
MONTH NUM is an numerical integer which stands for Month
```

MONTH MON is a text character which represent month in a three letter

FLT DATE is a text character that represents the particular date of flights

APT ICAO is a text character that represents the International Civil Aviation Organisations (ICAO) 4-letter airport designator name.

APT NAME is a text character that the airport name.

STATE NAME is a text character that represent the name of the country in which the airport is located.

FLT DEP 1 is an integer that represents the number of instrument flight rules (IFR) departures as provided by network manager.

FLT ARR 1 is an integer that represents the number of IFR arrival as provided by network manager.

FLT TOT 1 is an integer that represents the total number of IFR movements as provided by network manager.

FLT DEP IFR 2 is an integer that represents the number of IFR departures

as provided by the airport operator. FLT\_ARR\_IFR\_2 is an integer that represents the number of IFR arrivals as provided by the airport operator.  ${\tt FLT\_TOT\_IFR\_2}$  is an integer that represents the total number of  ${\tt IFR}$ movement as provided by the airport operator. Pivot.Label is a text character that shows the combination of name of country and airport. Data Cleaning Let look at the number of NA (missing) values present in our data set ```{r} sum(is.na(flights)) sum(is.na(flights\$APT NAME)) sum(is.na(flights\$STATE NAME)) sum(is.na(flights\$FLT TOT 1)) sum(is.na(flights\$YEAR)) sum(is.na(flights\$FLT DATE)) There are 1,439,355 NAs in flights data set, however, there are no NA values in our variable of concern. Convert YEAR from integer to a date factor using lubridate. ```{r} flights\$FLT\_DATE <- ymd(flights\$FLT DATE)</pre> class(flights\$FLT DATE) ### Research questions 1. Which EU countries operated the highest number of total flights before covid 19? Using the total numbers of flights (departures and arrivals), arranged in descending order and rounded up to one decimal place ```{r} CountryTop <flights %>% filter(YEAR >= 2016 & YEAR <= 2019) %>% group by (STATE NAME) %>% summarise(flights Total = sum(FLT TOT 1)) %>% arrange(desc(flights Total)) %>% mutate(flights Total = flights\_Total/1000000) CountryTop <-CountryTop %>% mutate(Round 1 = round(CountryTop\$flights Total, digits = 1)Convert the STATE NAME to a factor to allow us arrange our bars in descending order. ```{r} CountryTop\$STATE NAME <- factor(CountryTop\$STATE NAME)</pre> CountryTop\$STATE NAME <- fct reorder(CountryTop\$STATE NAME, CountryTop\$flights Total) class(CountryTop\$STATE NAME)

Select the list of the first 12 countries with highest total flights before  $Covid\_19$ .

```
```{r}
CountryTop12 <-
  CountryTop [1:12,]
list(CountryTop12)
Graphical representation
```{r}
ggplot(CountryTop12,
 aes (Round 1,
 STATE NAME,
 fill = STATE NAME)
) +
 geom\ col(alpha = 0.5) +
 geom_text(aes(label =Round 1), stat = "identity", hjust = 1.5, colour = "white")+
 labs(title = "The 12 countries with highest total number of flights before covid 19",
 subtitle = "Source: https://ec.europa.eu/eurostat/web/products-eurostat-
news/-/ddn-20210914-1",
 x= "Total number of flights in millions",
 y = NULL) +
 theme(legend.position = NULL) +
 theme(panel.grid.major = element blank(),
 panel.grid.minor = element blank(),
 panel.background = element blank()) +
 guides(fill = "none")
United Kingdom is the country with the highest number of total flights
(departure and arrival) before Covid 19, a little ahead of Spain and
France respectively. Italy, Turkiye, Normay, netherlands, Greece,
Switzerland, Portugal, Poland respectively followed the queue.
Research question
2A.
The airport with the highest number of total flights before Covid 19?
Using mean of the total numbers of flights (departures and arrivals), arranged in
descending order and rounded up to one decimal place
```{r}
A2016 2019 <-
  flights %>%
  filter(YEAR >= 2016 & YEAR <= 2019) %>%
  select (FLT DATE,
         MONTH MON,
         YEAR,
         FLT DEP 1,
         FLT ARR 1,
         FLT TOT 1,
         STATE NAME,
         APT NAME
         ) %>%
  group by (APT NAME) %>%
  summarise(Mean TOT = mean(FLT_TOT_1)) %>%
 arrange(desc(Mean TOT))
```{r}
A2016 2019 <-
 A2016 2019%>%
 mutate(Mean TOT 2 = round(A2016 2019$Mean TOT, digits = 0))
The new data frame (A2016 2019) now has both the mean and rounded up mean of the Total
```

flights per airport.

```
List the first 12 airports.
```{r}
A2016_2019_Mean <-
 A2016 2019 [1:12,]
list(A2016 2019 Mean)
Convert the APT NAME to a factor to allow us arrange our bars in
descending order
```{r}
A2016 2019 Mean$APT NAME <- factor(A2016 2019 Mean$APT NAME)
A2016 2019 Mean$APT NAME <- fct reorder(A2016 2019 Mean$APT NAME,
 A2016 2019 Mean$Mean TOT)
class(A2016 2019 Mean$APT NAME)
Highlight the airport with the highest flights number.
```{r}
ggplot(A2016 2019 Mean,
       aes (Mean TOT,
           APT NAME,
           fill = APT NAME)
       ) +
  geom\ col(alpha = 0.4) +
  geom text(aes(label = Mean TOT 2), stat = "identity", hjust = 1.5, colour = "white")
  gghighlight(APT NAME %in% c("Amsterdam - Schiphol")) +
  labs(title = "Airport with the highest 'number of flights' before Covid 19",
       x = NULL,
       y = NULL) +
  theme(legend.position = "remove") +
  theme(panel.grid.major = element blank(),
       panel.grid.minor = element blank(),
        panel.background = element blank())
Amsterdam-Schiphol is the airport with the highest numbers of flights pre-covid 19.
Closely followed by Frankfurt, and Paris-Charles-de-Gualle, London-Heathrow,
respectively and the rest.
### 2B.
The airport with the highest number of total flights during Covid 19?
Using mean of the total numbers of flights (departures and arrivals), arrange in
descending order and rounded up to whole number.
```{r}
A2020 A <-
 flights %>%
 filter(YEAR == 2020) %>%
 select(FLT DATE,
 MONTH MON,
 YEAR,
 FLT DEP 1,
 FLT ARR 1,
 FLT TOT 1,
 STATE NAME,
 APT NAME
) %>%
 group by (APT NAME) %>%
 summarise(Mean TOT = mean(FLT TOT 1)) %>%
 arrange(desc(Mean TOT))
```

```
A2020 <-
 A2020 A %>%
 mutate(Mean_TOT_2 = round(A2020_A$Mean_TOT,
 digits = 0))
List the first 12 airports for better comparison
```{r}
A2020 Mean <-
 A2020 [1:12,]
list(A2020 Mean)
Convert the APT NAME to a factor to allow us arrange the bars in descending order
```{r}
A2020 Mean$APT NAME <- factor(A2020 Mean$APT NAME)
A2020 Mean$APT NAME <- fct reorder(A2020 Mean$APT NAME, A2020 Mean$Mean TOT)
class(A2020 Mean$APT NAME)
Let plot the graph using ggplot.
```{r}
ggplot (A2020 Mean,
       aes (Mean TOT,
           APT NAME,
           fill = APT NAME)
  geom col(fill = "steelblue", alpha = 0.5) +
    geom text(aes(label = Mean TOT 2), stat = "identity", hjust = 1.5, colour =
  labs(title = "Airport with highest 'number of flights' during Covid 19",
       x= NULL,
       y = NULL) +
  gghighlight(APT NAME == "Amsterdam - Schiphol") +
  theme(legend.position = NULL) +
  theme(panel.grid.major = element blank(),
        panel.grid.minor = element blank(),
        panel.background = element blank()) +
  guides(fill = "none")
Amsterdam-Schiphol continued its humanitarian services even in the face
of the pandemic and standing out as the airport with the highest numbers
of flights mean during covid 19. Frankfurt had dropped to third, which
could mean that, Germany either placed high restrictions on flights or
Frankfurt city was badly affected and people boycott the airport, while
France became more welcoming.
Which airport has the highest number of flights turnaround after covid 19?
Using mean of the total numbers of flights (departures and arrivals), arrange in
descending order and rounded up to whole number.
```{r}
A2021 2022 <-
 flights %>%
 filter(YEAR >= 2021 & YEAR <= 2022) %>%
 select(FLT DATE,
 MONTH MON,
 YEAR,
 FLT DEP 1,
 FLT ARR 1,
 FLT TOT 1,
 STATE NAME,
```

```
APT NAME
) %>%
 group_by(APT_NAME) %>%
 summarise(Mean_TOT = mean(FLT_TOT_1)) %>%
 arrange(desc(Mean_TOT))
```{r}
A2021 2022 <-
  A2021 2022 %>%
  mutate (Mean TOT 2 = \text{round} (A2021 2022\$Mean TOT,
                             digits = 0))
Let choose the first 12 airports for better representations & evaluation
```{r}
A2021 2022 Mean <-
 A2021 2022 [1:12,]
list(A2021 2022 Mean)
Let convert the APT NAME to a factor to allow us arrange our bars in
descending order
```{r}
A2021 2022 Mean$APT NAME <- factor(A2021 2022 Mean$APT NAME)
A2021 2022 Mean$APT NAME <- fct reorder(A2021 2022 Mean$APT NAME,
                                         A2021 2022 Mean$Mean TOT)
class(A2021 2022 Mean$APT NAME)
```{r}
ggplot (A2021 2022 Mean,
 aes (Mean TOT 2,
 APT NAME,
 fill = APT NAME)
 geom\ col(fill = "darkgreen", alpha = 0.5) +
 geom text(aes(label = Mean TOT 2), stat = "identity", hjust = 1.5, colour = "white")
 gghighlight(APT NAME == "iGA Istanbul Airport") +
 labs(title = "Airport with the highest 'number of flights' turnaround post Covid 19",
 x = NULL
 y = NULL) +
 theme(legend.position = NULL) +
 theme(panel.grid.major = element blank(),
 panel.grid.minor = element blank(),
 panel.background = element blank()) +
 guides(fill = "none")
Amsterdam-Schiphol maintained the leading in all the era under considerations.
```

Amsterdam-Schiphol maintained the leading in all the era under considerations. Surprisingly, IGA Istanbul (Turkey airport) made 2nd position which was 7th on the queue before Covid\_19. Turkey Civil Aviation Industry (TCAI) made research on reason while passengers lost confidence in the flights operations. "Another important aspect for demand recovery is that airlines have to investigate the reasons behind the decrease in passenger confidence and how to regain it". (Muhammet Deveci, & et al., 2022). This obviously paid off. Also, the government provided a 24/7 5,000 square meter PCR testing center with 20,000 capacity daily inside the Istanbul Airport terminal serving the passengers. All test can do quickly with results out with delay, according of Istanbul Airport report (online)

#### ## Research question

3. Which country's airport shows the government quick willingness to adjust to normal life by lifting or reducing Covid\_19 restrictions?

```
Choose the top 4 airports in pre_covid_19 and consider their performance
during and after covid_19 using lollypop graph.
```{r}
A2016_2019 T4 <-
  A2016 2019 Mean [1:4,]
A2020 T4 < -
  A2020 Mean [1:4,]
A2021 2022 T4 <-
  A20\overline{2}1 \ 20\overline{2}2 \ Mean \ [1:6,]
list(A2016 2019 T4)
```{r}
New_flights <-</pre>
 merge(A2016_2019_T4,
 A2020 T4,
 by = \overline{"}APT NAME")
New flights2 <-
 merge (New flights,
 A2021 2022 T4,
 by = "APT \overline{N}AME")
New flights3 <-
 New flights2 %>%
 select (APT NAME,
 Mean TOT 2.x,
 Mean TOT 2.y,
 Mean TOT 2)
 colnames(New flights3) <-</pre>
 c("Airport",
 "Pre_Covid",
 "Covid",
 "Post Covid")
Let convert the Airport to a factor to allow arranging the lollipop in
proper order
```{r}
New flights3$Airport <- factor(New flights3$Airport)</pre>
New flights3$Airport <- fct reorder(New flights3$Airport,
                                        New flights3$Covid)
class(New flights3$Airport)
```{R}
ggplot(New flights3,
 aes(x = Airport)
) +
 geom_point(aes
 (y = Covid),
 colour = "darkblue",
 size = 3,
 alpha = 0.8) +
 geom_point(aes
 (y = Post_Covid),
 colour = "darkgreen",
 size = 3,
 alpha = 0.8) +
 geom segment (aes
 (x = Airport,
```

y = Covid,

```
xend = Airport,
 yend = Post_Covid)) +
 labs(title = "Flights resuscitation relationship after covid_19",
 subtitle = "Blue: Covid
 Green: Post-Covid 19",
 y = "Number of flights",
 x = NULL) +
 theme bw() +
 theme(panel.grid.major = element_blank(),
 panel.grid.minor = element blank(),
 panel.background = element blank()) +
 coord flip()
Netherlands (Amstedam - Schiphol airport) shows quicker willingness to
adjust back to normal life ahead of others.
Research Question
4. Which country's airport recorded the highest flight per day after
 (post) Covid 19.
What is the behavioural pattern of the European aviation operations
before the appearance of Covid 19. in respect of
```{r}
flights 4 <-
  flights %>%
  filter(APT NAME == c("Amsterdam - Schiphol",
                         "Frankfurt",
                         "London - Heathrow",
                         "Paris-Charles-de-Gaulle")
           select (FLT DATE,
         YEAR,
         APT NAME,
         FLT TOT 1
         ) 응>응
  drop na (
  ) %>%
  rename with (.cols = 3,
              ~"Airports")
```{r}
ggplot(flights 4,
 aes(
 x = FLT DATE,
 y = FLT TOT 1,
 colour = Airports)
) +
 geom\ point(alpha = 0.5) +
 theme test() +
 theme(legend.position = "bottom") +
 labs(title = "Country with the hightest flights per day after Covid 19",
y = "Numbers of flights per day")
Using facet wrap to segment each airport separately for better visualization.
```{r}
ggplot(flights_4,
       aes(x = FLT DATE,
           y = FLT TOT 1,
           colour = Airports, )
       ) +
  geom\ point(alpha = 0.5) +
  facet wrap(~ Airports,
             ncol = 4) +
```

```
theme_classic() +
  theme(legend.position = "bottom", names(NULL)) +
  labs(title = "Country With The Hightest Flights Per Day After Covid_19",
x = NULL,
y = "Numbers of flights per day")
```

Netherlands recorded the highest numbers of flights by day post Covid_19. (Seveno Victoria, 2022). According to Seveno V. (2022) thanks to the EU digital covid_19 certificate and the relaxation of various international travel restrictions. London - Heathrow has a 100,000 daily passengers limits up until October 2022. (Cunningham Ed, 2022), obviously the reason why they were left behind.

Research question

5. Has EU Aviation Industry fully recovered in the numbers of flights per day?

```
```{r}
flights Day <-
 flights %>%
 group by (FLT DATE) %>%
 summarise(FlightsPerDay = sum(FLT TOT 1))
```{r}
flights Day %>%
  ggplot(aes(
   x = FLT DATE,
    y = FlightsPerDay
    , colour = FlightsPerDay)
    ) +
  geom line(alpha = 1) +
 gghighlight(FLT DATE \geq "2020-04-01") +
 theme minimal() +
  theme (axis.text.y = element blank(),
       axis.ticks.y = element blank()) +
 labs(title = "European Aviation Flights Operations Recovery Pattern",
x = NULL
y = NULL)
```

EU Aviation Industry has not fully recovered in it total number of flights per day compared to pre-Covid_19. There is still about 20,000 daily flights difference between pre and post-Covid_19

Conclusion

EU aviation industry is not yet where they were before Covid_19 emergence, but there have been great improvement from 2021. and there seems to be possibilities of getting there in the next 2 years or 3 years if the growth rate continues. Although, one of the imprint of covid_19 is the new wave of virtual system of meeting, learning, participation etc. if this continues, we might not experience the volumes like it use to be in EU and the globe at large.

Currently, London - Heathrow has reclaimed the position of the busiest airport as at the time of this report (April 2023), (Cunningham Ed, 2022), and has taken over from Amsterdam - Schiphol. There will need for further investigation to get to the root cause of this. However,

Reference:

Instanbul Airport, Instanbul Airport Test Centre is at your services
24/7! online article, [viewed 24 April 2023] Available from:
<https://www.istairport.com/en/announcements/istanbul-airport-test-center-is-at-your-service-24-7?locale=en>

Muhammet Deveci, & et al., (2022) Impact of COVID-19 pandemic on the Turkish civil aviation industry, Sustainable Operations and Computers, Volume 3, Pages 93-102, ISSN 2666-4127, [viewed 24 April 2023] Available from: https://doi.org/10.1016/j.susoc.2021.11.002, [https://www.sciencedirect.com/science/article/pii/S2666412721000490)] (https://www.sciencedirect.com/science/article/pii/S2666412721000490)), [Impact of COVID-19 pandemic on the Turkish civil aviation industry - PMC (nih.gov)](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8627591/)

Cunningham Ed, 2022, These are officially the Europe busiest airport right now. Online article [viewed: 24 April 2023] Available from: [Europe\'s Busiest Airports include London and Amsterdam (timeout.com)](https://www.timeout.com/news/these-are-officially-europes-busiest-airports-right-now-101422)

Seveno Victoria, 2022. Schiphol once again one of the Eupore's five busiest airport. Online article [viewed 24 April 2023] Available from: [Schiphol once again one of Europe\'s five busiest airports (iamexpat.nl)](https://www.iamexpat.nl/expat-info/dutch-expat-news/schiphol-once-again-one-europes-five-busiest-airports)