



# Dataset “NYC flights 13”

Date	@March 20, 2024
Status	Done

## NYC flights 13

6 questions about this database

```
library(nycflights13)
library(tidyverse)    # dplyr
```

```
data("flights")
data("airlines")
data("airports")
View(flights)
view(airlines)
view(airports)
```

### Q1. Most flight carrier in Sep 2013

```

flights %>%
  filter(month == 9) %>%
  count(carrier) %>%
  arrange(-n) %>%
  left_join(airlines, by = "carrier")

```

```

# A tibble: 16 × 3
  carrier     n name
  <chr>    <int> <chr>
1 EV        4725 ExpressJet Airlines Inc.
2 UA        4694 United Air Lines Inc.
3 B6        4291 JetBlue Airways
4 DL        3883 Delta Air Lines Inc.
5 AA        2614 American Airlines Inc.
6 MQ        2206 Envoy Air
7 US        1698 US Airways Inc.
8 9E        1540 Endeavor Air Inc.
9 WN        1010 Southwest Airlines Co.
10 VX       453 Virgin America
11 FL       255 AirTran Airways Corporation
12 AS       60 Alaska Airlines Inc.
13 F9       58 Frontier Airlines Inc.
14 YV       42 Mesa Airlines Inc.
15 HA       25 Hawaiian Airlines Inc.
16 00       20 SkyWest Airlines Inc.

```

ในเดือนกันยายน ปี 2013 สายการบินที่มีรอดับบันมากที่สุดได้แก่ ExpressJet Airlines, United Air Lines , JetBlue Airways, Delta Air Lines และ American Airlines ตามลำดับ

## Q2. Find average departure delay and average arrival delay for each airline

```

flights %>%
  drop_na(dep_delay, arr_delay) %>%

```

```

group_by(carrier) %>%
summarise(avg_dep_delay = mean(dep_delay),
avg_arr_delay = mean(arr_delay)) %>%
arrange(avg_dep_delay, avg_arr_delay) %>%
left_join(airlines, by = "carrier") %>%
select(name, avg_dep_delay, avg_arr_delay)

```

	name	avg_dep_delay	avg_arr_delay
1	US Airways Inc.	3.74	2.13
2	Hawaiian Airlines Inc.	4.90	-6.92
3	Alaska Airlines Inc.	5.83	-9.93
4	American Airlines Inc.	8.57	0.364
5	Delta Air Lines Inc.	9.22	1.64
6	Envoy Air	10.4	10.8
7	United Air Lines Inc.	12.0	3.56
8	SkyWest Airlines Inc.	12.6	11.9
9	Virgin America	12.8	1.76
10	JetBlue Airways	13.0	9.46
11	Endeavor Air Inc.	16.4	7.38
12	Southwest Airlines Co.	17.7	9.65
13	AirTran Airways Corporation	18.6	20.1
14	Mesa Airlines Inc.	18.9	15.6
15	ExpressJet Airlines Inc.	19.8	15.8
16	Frontier Airlines Inc.	20.2	21.9

### Q3. Which NY city destinations were the most popular?

```

flights %>%
  group_by(dest) %>%
  count(dest) %>%
  arrange(-n) %>%

```

```
left_join(airports, by = c("dest" = "faa")) %>%
  select(dest, full_name = name, n)
```

```
# A tibble: 105 × 3
# Groups:   dest [105]
  dest   full_name                      n
  <chr> <chr>                         <int>
1 ORD    Chicago Ohare Intl            17283
2 ATL    Hartsfield Jackson Atlanta Intl 17215
3 LAX    Los Angeles Intl              16174
4 BOS    General Edward Lawrence Logan Intl 15508
5 MCO    Orlando Intl                 14082
6 CLT    Charlotte Douglas Intl       14064
7 SFO    San Francisco Intl           13331
8 FLL    Fort Lauderdale Hollywood Intl 12055
9 MIA    Miami Intl                  11728
10 DCA   Ronald Reagan Washington Natl 9705
```

จุดหมายปลายทางในเมือง New York ที่นิยมมากที่สุดได้แก่ Chicago Ohare โดยมีจำนวนรอบบินก็งหนด 17,283 ครั้ง รองลงมาคือ Hartsfield Jackson Atlanta จำนวน 17,215 ครั้ง และ Los Angeles 16,174 ครั้ง ตามลำดับ

## Q4. Which airlines fly the longest distances?

```
flights %>%
  group_by(carrier) %>%
  summarise(sum_dis = sum(distance)) %>%
  arrange(desc(sum_dis)) %>%
  left_join(airlines, by = "carrier") %>%
  select(name, sum_dis) %>%
  head(5)
```

```
# A tibble: 5 × 2
  name          sum_dis
  <chr>        <dbl>
1 AA             111000
2 UA             108000
3 DL             107000
4 YV             106000
5 B6             105000
```

```
<chr><dbl>
1 United Air Lines Inc.      89705524
2 Delta Air Lines Inc.       59507317
3 JetBlue Airways           58384137
4 American Airlines Inc.    43864584
5 ExpressJet Airlines Inc.   30498951
```

สายการบินที่มีระบบรวมมากที่สุด ได้แก่ United Air Lines

## Q5. Most frequent route flying

```
flights %>%
  group_by(origin, dest) %>%
  mutate(route = paste(origin, dest, sep = " - ")) %>%
  count(route) %>%
  arrange(-n) %>%
  select(route, times = n)
```

```
# A tibble: 224 × 4
# Groups:   origin, dest [224]
  origin dest route     times
  <chr><chr><chr><int>
1 JFK   LAX   JFK - LAX 11262
2 LGA   ATL   LGA - ATL 10263
3 LGA   ORD   LGA - ORD 8857
4 JFK   SFO   JFK - SFO 8204
5 LGA   CLT   LGA - CLT 6168
6 EWR   ORD   EWR - ORD 6100
7 JFK   BOS   JFK - BOS 5898
8 LGA   MIA   LGA - MIA 5781
9 JFK   MCO   JFK - MCO 5464
10 EWR  BOS   EWR - BOS 5327
```

ในปี 2013 เส้นทางที่มีการบินมากที่สุด 3 อันดับแรก ได้แก่ JFK - LAX จำนวน 11,262 ครั้ง, LGA - ATL จำนวน 10,263 ครั้ง และ LGA - ORD จำนวน 8,857 ครั้ง

## Q6. Which are the top 5 airlines that fly to Chicago Ohare Intl (ORD) between January and June 2013?

```
flights %>%
  filter(dest == "ORD", month >=1 & month <= 6) %>%
  count(carrier) %>%
  arrange(-n) %>%
  left_join(airlines, by = "carrier") %>%
  head(5)
```

```
# A tibble: 5 × 3
  carrier     n name
<chr><int><chr>
1 UA        3180 United Air Lines Inc.
2 AA        2908 American Airlines Inc.
3 MQ        1300 Envoy Air
4 9E         542 Endeavor Air Inc.
5 B6         421 JetBlue Airways
```

5 อันดับ สายการบินที่บินไป Chicago Ohare Intl ในช่วง มกราคม - มิถุนายน 2013 มากที่สุด ได้แก่ United Air Lines, American Airlines, Envoy Air, Endeavor Air และ JetBlue Airways