Analysis NYCFight13

Date: 29 Dec 2022

Author: Pongphisan Pisuttiwat (Seng)

Course: R Data Transformation

```
install.packages(c("tidyverse",
                   "nycflights13"))
```

```
     Installing packages into ‘/usr/local/lib/R/site-library’
     (as ‘lib’ is unspecified)
```

```
library(tidyverse)
library(nycflights13)
```

```
data("flights")
glimpse(flights)
```

```
     Rows: 336,776
     Columns: 19
     $ year           <int> 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2…
     $ month          <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1…
     $ day            <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1…
     $ dep_time       <int> 517, 533, 542, 544, 554, 554, 555, 557, 557, 558, 558, …
     $ sched_dep_time <int> 515, 529, 540, 545, 600, 558, 600, 600, 600, 600, 600, …
     $ dep_delay      <dbl> 2, 4, 2, -1, -6, -4, -5, -3, -3, -2, -2, -2, -2, -2, -1…
     $ arr_time       <int> 830, 850, 923, 1004, 812, 740, 913, 709, 838, 753, 849,…
     $ sched_arr_time <int> 819, 830, 850, 1022, 837, 728, 854, 723, 846, 745, 851,…
     $ arr_delay      <dbl> 11, 20, 33, -18, -25, 12, 19, -14, -8, 8, -2, -3, 7, -1…
     $ carrier        <chr> "UA", "UA", "AA", "B6", "DL", "UA", "B6", "EV", "B6", "…
     $ flight         <int> 1545, 1714, 1141, 725, 461, 1696, 507, 5708, 79, 301, 4…
     $ tailnum        <chr> "N14228", "N24211", "N619AA", "N804JB", "N668DN", "N394…
     $ origin         <chr> "EWR", "LGA", "JFK", "JFK", "LGA", "EWR", "EWR", "LGA",…
     $ dest           <chr> "IAH", "IAH", "MIA", "BQN", "ATL", "ORD", "FLL", "IAD",…
     $ air_time       <dbl> 227, 227, 160, 183, 116, 150, 158, 53, 140, 138, 149, 1…
     $ distance       <dbl> 1400, 1416, 1089, 1576, 762, 719, 1065, 229, 944, 733, …
     $ hour           <dbl> 5, 5, 5, 5, 6, 5, 6, 6, 6, 6, 6, 6, 6, 6, 6, 5, 6, 6, 6…
     $ minute         <dbl> 15, 29, 40, 45, 0, 58, 0, 0, 0, 0, 0, 0, 0, 0, 0, 59, 0…
     $ time_hour      <dttm> 2013-01-01 05:00:00, 2013-01-01 05:00:00, 2013-01-01 0…
```

Q1: Top 5 Routes in May 2013

```
flights %>%
  filter(month == 5) %>%
  group_by(origin, dest) %>%
  summarise(n = n()) %>%
  arrange(desc(n)) %>%
  head(5)
```

```
`summarise()` has grouped output by 'origin'. You can override using the
`.groups` argument.
```

A grouped_df: 5 × 3

origin	dest	n
<chr>	<chr>	<int>
JFK	LAX	960
LGA	ATL	925
LGA	ORD	816
JFK	SFO	696
EWR	ORD	549

Q2: Top 5 Airlines in Oct 2013

```
flights %>%
  filter(month == 10) %>%
  count(carrier) %>%
  arrange(desc(n)) %>%
  left_join(airlines, by = "carrier") %>%
  head(5)
```

A tibble: 5 × 3

carrier	n	name
<chr>	<int>	<chr>
UA	5060	United Air Lines Inc.
EV	4908	ExpressJet Airlines Inc.
B6	4361	JetBlue Airways
DL	4093	Delta Air Lines Inc.
AA	2715	American Airlines Inc.

Q3: How many long flights & Short flight in Feb 2013 (long flight distance more than 1000 km)

```
flights %>%
  filter(month == 2) %>%
  mutate(flight_type = if_else(distance > 1000, "long_flight", "short_flight")) %>%
  count(flight_type)
```

A tibble: 2 × 2

flight_type	n
<chr>	<int>
long_flight	10767
short_flight	14184

Q4: How many arrival delay and ontime flight for each Origin in Nov 2013 (NA = Ontime)

```
flights %>%
  filter(month == 11) %>%
  mutate(arr_delay_clean = replace_na(arr_delay, 0),
         ARR = if_else (arr_delay_clean > 0, "Delay", "On Time")) %>%
  group_by(origin, ARR) %>%
  summarise(n = n())
```

```
`summarise()` has grouped output by 'origin'. You can override using the
`.groups` argument.
```

A grouped_df: 6 × 3

origin	ARR	n
<chr>	<chr>	<int>
EWR	Delay	3352
EWR	On Time	6355
JFK	Delay	2943
JFK	On Time	5767
LGA	Delay	3344
LGA	On Time	5507

Q5:Top 5 Airlines Highnest departure late in 2013 (drop NA rows)

```
flights_clean <- drop_na(flights)
```

```
flights_clean %>%
    filter(dep_delay > 0) %>%
    count(carrier) %>%
    arrange(desc(n)) %>%
```

```
left_join(airlines) %>%
head(5)
```

```
Joining, by = "carrier"
```

A tibble: 5 × 3

carrier	n	name
<chr>	<int>	<chr>
UA	27125	United Air Lines Inc.
EV	22976	ExpressJet Airlines Inc.
B6	21372	JetBlue Airways
DL	15186	Delta Air Lines Inc.
AA	10105	American Airlines Inc.

ผลิตภัณฑ์ Colab แบบมีค่าใช้จ่าย - ยกเลิกสัญญาที่นี่

0 วินาที เสร็จสมบูรณ์เมื่อ 15:23