NYC Flights 2013 Analysis

- Data detail : https://cran.r-project.org/web/packages/nycflights13/nycflights13.pdf

```
# install.packages("nycflights13")
library(nycflights13)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0     v purrr   1.0.1
## v tibble  3.1.8     v dplyr   1.1.0
## v tidyr   1.3.0     v stringr 1.5.0
## v readr   2.1.3     v forcats 1.0.0
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
glimpse(weather)
```

```
## Rows: 26,115
## Columns: 15
## $ origin     <chr> "EWR", "EWR", "EWR", "EWR", "EWR", "EWR", "EWR", "EWR", "EW~
## $ year       <int> 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013, 2013,~
## $ month      <int> 1, 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, 1,~
## $ hour       <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17, 18, ~
## $ temp       <dbl> 39.02, 39.02, 39.02, 39.92, 39.02, 37.94, 39.02, 39.92, 39.~
## $ dewp       <dbl> 26.06, 26.96, 28.04, 28.04, 28.04, 28.04, 28.04, 28.04, 28.~
## $ humid      <dbl> 59.37, 61.63, 64.43, 62.21, 64.43, 67.21, 64.43, 62.21, 62.~
## $ wind_dir   <dbl> 270, 250, 240, 250, 260, 240, 240, 250, 260, 260, 260, 330,~
## $ wind_speed <dbl> 10.35702, 8.05546, 11.50780, 12.65858, 12.65858, 11.50780, ~
## $ wind_gust  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, 20.~
## $ precip     <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~
## $ pressure   <dbl> 1012.0, 1012.3, 1012.5, 1012.2, 1011.9, 1012.4, 1012.2, 101~
## $ visib      <dbl> 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10,~
## $ time_hour  <dttm> 2013-01-01 01:00:00, 2013-01-01 02:00:00, 2013-01-01 03:00~
```

```
glimpse(airports)
```

```
## Rows: 1,458
## Columns: 8
## $ faa   <chr> "04G", "06A", "06C", "06N", "09J", "0A9", "0G6", "0G7", "0P2", "~
## $ name  <chr> "Lansdowne Airport", "Moton Field Municipal Airport", "Schaumbur~
## $ lat   <dbl> 41.13047, 32.46057, 41.98934, 41.43191, 31.07447, 36.37122, 41.4~
## $ lon   <dbl> -80.61958, -85.68003, -88.10124, -74.39156, -81.42778, -82.17342~
## $ alt   <dbl> 1044, 264, 801, 523, 11, 1593, 730, 492, 1000, 108, 409, 875, 10~
## $ tz    <dbl> -5, -6, -6, -5, -5, -5, -5, -5, -5, -8, -5, -6, -5, -5, -5, -5, ~
## $ dst   <chr> "A", "A", "A", "A", "A", "A", "A", "A", "U", "A", "A", "U", "A",~
## $ tzone <chr> "America/New_York", "America/Chicago", "America/Chicago", "Ameri~
```

```
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~
```

```
glimpse(airlines)
```

```
## Rows: 16
## Columns: 2
## $ carrier <chr> "9E", "AA", "AS", "B6", "DL", "EV", "F9", "FL", "HA", "MQ", "O~
## $ name    <chr> "Endeavor Air Inc.", "American Airlines Inc.", "Alaska Airline~
```

```
glimpse(planes)
```

```
## Rows: 3,322
## Columns: 9
## $ tailnum      <chr> "N10156", "N102UW", "N103US", "N104UW", "N10575", "N105UW~
## $ year         <int> 2004, 1998, 1999, 1999, 2002, 1999, 1999, 1999, 1999, 199~
## $ type         <chr> "Fixed wing multi engine", "Fixed wing multi engine", "Fi~
## $ manufacturer <chr> "EMBRAER", "AIRBUS INDUSTRIE", "AIRBUS INDUSTRIE", "AIRBU~
## $ model        <chr> "EMB-145XR", "A320-214", "A320-214", "A320-214", "EMB-145~
## $ engines      <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, ~
## $ seats        <int> 55, 182, 182, 182, 55, 182, 182, 182, 182, 182, 55, 55, 5~
## $ speed        <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
## $ engine       <chr> "Turbo-fan", "Turbo-fan", "Turbo-fan", "Turbo-fan", "Turb~
```

The 10 Most-Carriers

```
flights %>%
  count(carrier) %>%
```

```
  arrange(desc(n)) %>%
  rename(total_carrier = n) %>%
  left_join(airlines, by="carrier") %>%
  head(10)
```

```
## # A tibble: 10 x 3
##    carrier total_carrier name
##    <chr>           <int> <chr>
##  1 UA              58665 United Air Lines Inc.
##  2 B6              54635 JetBlue Airways
##  3 EV              54173 ExpressJet Airlines Inc.
##  4 DL              48110 Delta Air Lines Inc.
##  5 AA              32729 American Airlines Inc.
##  6 MQ              26397 Envoy Air
##  7 US              20536 US Airways Inc.
##  8 9E              18460 Endeavor Air Inc.
##  9 WN              12275 Southwest Airlines Co.
## 10 VX               5162 Virgin America
```

The 10 Most-Distance in December

```
flights %>%
  filter(month == 12) %>%
  group_by(carrier) %>%
  summarize(total_distance = sum(distance)) %>%
  arrange(desc(total_distance)) %>%
  left_join(airlines, by="carrier") %>%
  head(10)
```

```
## # A tibble: 10 x 3
##    carrier total_distance name
##    <chr>            <dbl> <chr>
##  1 UA             7609757 United Air Lines Inc.
##  2 B6             5185761 JetBlue Airways
##  3 DL             5109723 Delta Air Lines Inc.
##  4 AA             3671213 American Airlines Inc.
##  5 EV             2441056 ExpressJet Airlines Inc.
##  6 MQ             1237254 Envoy Air
##  7 VX             1190060 Virgin America
##  8 WN             1107649 Southwest Airlines Co.
##  9 US              960368 US Airways Inc.
## 10 9E              913992 Endeavor Air Inc.
```

The Most-Delay

- Departures Delay

```
flights %>%
  filter(dep_delay < 0) %>%
  group_by(carrier) %>%
```

```
  summarize(TotalTime_dep_delay = sum(dep_delay)) %>%
  arrange(TotalTime_dep_delay) %>%
  left_join(airlines, by="carrier") %>%
  head(10)
```

```
## # A tibble: 10 x 3
##    carrier TotalTime_dep_delay name
##    <chr>                 <dbl> <chr>
##  1 B6                  -147970 JetBlue Airways
##  2 EV                  -139752 ExpressJet Airlines Inc.
##  3 DL                  -127535 Delta Air Lines Inc.
##  4 UA                  -113920 United Air Lines Inc.
##  5 AA                  -102163 American Airlines Inc.
##  6 MQ                   -95194 Envoy Air
##  7 US                   -82649 US Airways Inc.
##  8 9E                   -54226 Endeavor Air Inc.
##  9 WN                   -14584 Southwest Airlines Co.
## 10 VX                   -10629 Virgin America
```

- Arrivals Delay

```
flights %>%
  filter(dep_delay < 0) %>%
  group_by(carrier) %>%
  summarize(TotalTime_arr_delay = sum(arr_delay)) %>%
  arrange(TotalTime_arr_delay) %>%
  left_join(airlines, by="carrier") %>%
  head(4)
```

```
## # A tibble: 4 x 3
##   carrier TotalTime_arr_delay name
##   <chr>                 <dbl> <chr>
## 1 HA                    -4141 Hawaiian Airlines Inc.
## 2 YV                    -3574 Mesa Airlines Inc.
## 3 F9                     -555 Frontier Airlines Inc.
## 4 OO                     -245 SkyWest Airlines Inc.
```

Top 10 Amount of the planes

```
planes %>%
  group_by(manufacturer) %>%
  summarize(total_amount = n()) %>%
  arrange(desc(total_amount)) %>%
  head(10)
```

```
## # A tibble: 10 x 2
##    manufacturer                total_amount
##    <chr>                              <int>
##  1 BOEING                              1630
##  2 AIRBUS INDUSTRIE                     400
```

```
##  3 BOMBARDIER INC                          368
##  4 AIRBUS                                  336
##  5 EMBRAER                                 299
##  6 MCDONNELL DOUGLAS                       120
##  7 MCDONNELL DOUGLAS AIRCRAFT CO           103
##  8 MCDONNELL DOUGLAS CORPORATION            14
##  9 CANADAIR                                  9
## 10 CESSNA                                    9
```

The 20 Flights Most Long Distance

```
flights %>%
  distinct(flight,tailnum,carrier,  distance) %>%
  arrange(desc(distance)) %>%
  left_join(airlines, by="carrier") %>%
  head(20)
```

```
## # A tibble: 20 x 5
##    flight tailnum carrier distance name
##     <int> <chr>   <chr>      <dbl> <chr>
##  1     51 N380HA  HA          4983 Hawaiian Airlines Inc.
##  2     51 N384HA  HA          4983 Hawaiian Airlines Inc.
##  3     51 N381HA  HA          4983 Hawaiian Airlines Inc.
##  4     51 N385HA  HA          4983 Hawaiian Airlines Inc.
##  5     51 N389HA  HA          4983 Hawaiian Airlines Inc.
##  6     51 N388HA  HA          4983 Hawaiian Airlines Inc.
##  7     51 N383HA  HA          4983 Hawaiian Airlines Inc.
##  8     51 N382HA  HA          4983 Hawaiian Airlines Inc.
##  9     51 N386HA  HA          4983 Hawaiian Airlines Inc.
## 10     51 N390HA  HA          4983 Hawaiian Airlines Inc.
## 11     51 N391HA  HA          4983 Hawaiian Airlines Inc.
## 12     51 N392HA  HA          4983 Hawaiian Airlines Inc.
## 13     51 N393HA  HA          4983 Hawaiian Airlines Inc.
## 14     51 N395HA  HA          4983 Hawaiian Airlines Inc.
## 15     15 N76065  UA          4963 United Air Lines Inc.
## 16     15 N77066  UA          4963 United Air Lines Inc.
## 17     15 N76064  UA          4963 United Air Lines Inc.
## 18     15 N69063  UA          4963 United Air Lines Inc.
## 19     15 N76055  UA          4963 United Air Lines Inc.
## 20     15 N68061  UA          4963 United Air Lines Inc.
```