

# R Notebook

[Code ▼](#)

Thanawat Chuntasadratsamee

## Data Visualization in R

[Hide](#)

```
library(tidyverse)
library(patchwork)
library(lubridate)
library(nycflights13)
library(dplyr)
library(glue)
library(ggthemes)
```

### Chart 1 : The average delay time for flights in each month of the year 2013

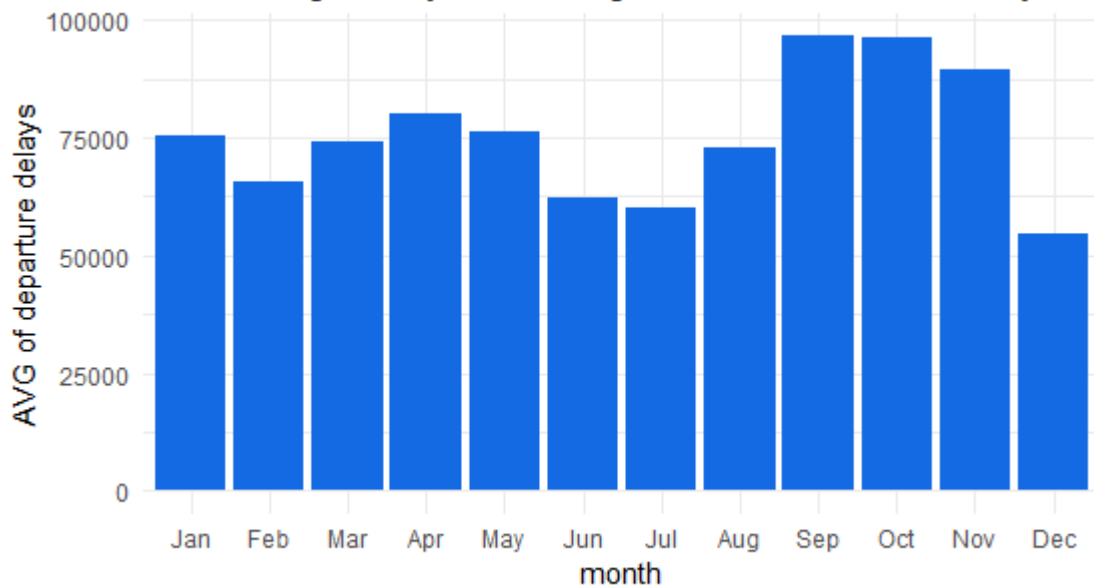
[Hide](#)

```
hw <- flights %>%
  tibble() %>%
  filter(year == 2013, dep_delay < 0 ) %>%
  mutate(date = month(time_hour, label=T)) %>%
  group_by(date) %>%
  summarise(avg_depDelay = abs(sum(dep_delay)),
            n = n()) %>%
  arrange((avg_depDelay)) %>%
  rename(Total_Flights = n)
```

[Hide](#)

```
ggplot(hw, aes(x = date, y = avg_depDelay)) +
  geom_bar(stat = "identity") +
  geom_col(fill="#146ae3") +
  theme_minimal() +
  labs(title = "The average delay time for flights in each month of the year 2013",
       caption = "Source : nycflights13",
       x = "month",
       y = "AVG of departure delays")
```

The average delay time for flights in each month of the year 2



Source : nycflights13

Chart 2 : Which carrier is the most delayed in 2013?

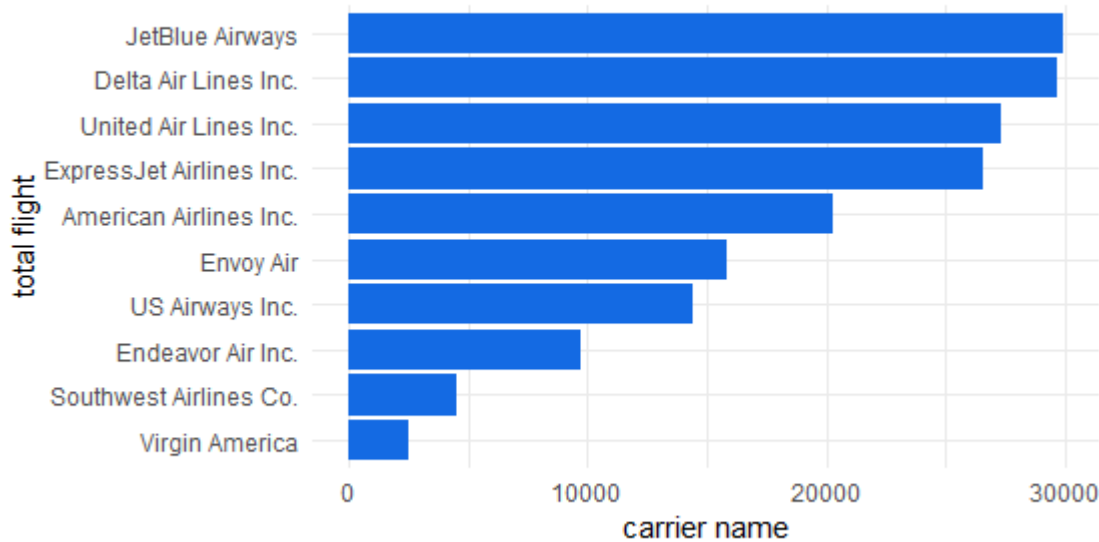
Hide

```
hw2 <- flights %>%
  filter(year == 2013, dep_delay < 0 ) %>%
  count(carrier) %>%
  arrange(desc(n)) %>%
  left_join(airlines, by = c("carrier")) %>%
  rename(total_flight = n, carrier_name = name) %>%
  head(10)
```

Hide

```
ggplot(hw2, aes(x = reorder(carrier_name, total_flight), y = total_flight)) +
  geom_bar(stat = "identity", fill = "#146ae3") +
  coord_flip() +
  theme_minimal() +
  labs(title = "The top 10 most flights for delay in 2013")
",
  caption = "Source : nycflights13",
  x = "total flight",
  y = "carrier name")
```

The top 10 most flights for delay in 2013



Source : nycflights13

Chart 3 : Which carrier is the most depart on time or the earliest in 2013?

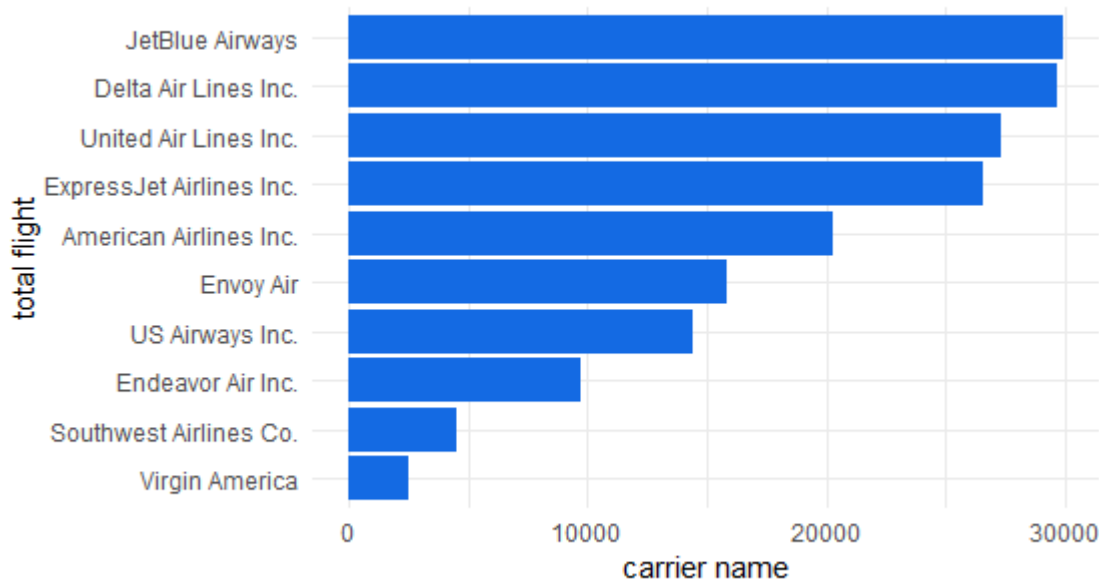
Hide

```
hw3 <- flights %>%
  filter(year == 2013, dep_delay > 0 ) %>%
  count(carrier) %>%
  arrange(desc(n)) %>%
  left_join(airlines, by = c("carrier")) %>%
  rename(total_flight = n, carrier_name = name) %>%
  head(10)
```

Hide

```
ggplot(hw2, aes(x = reorder(carrier_name, +total_flight), y = total_flight)) +
  geom_bar(stat = "identity", fill = "#146ae3") +
  coord_flip() +
  theme_minimal() +
  labs(title = "The top 10 most flights for depart on time or the earliest in 2013",
       caption = "Source : nycflights13",
       x = "total flight",
       y = "carrier name")
```

The top 10 most flights for depart on time or the ea



Source : nycflights13

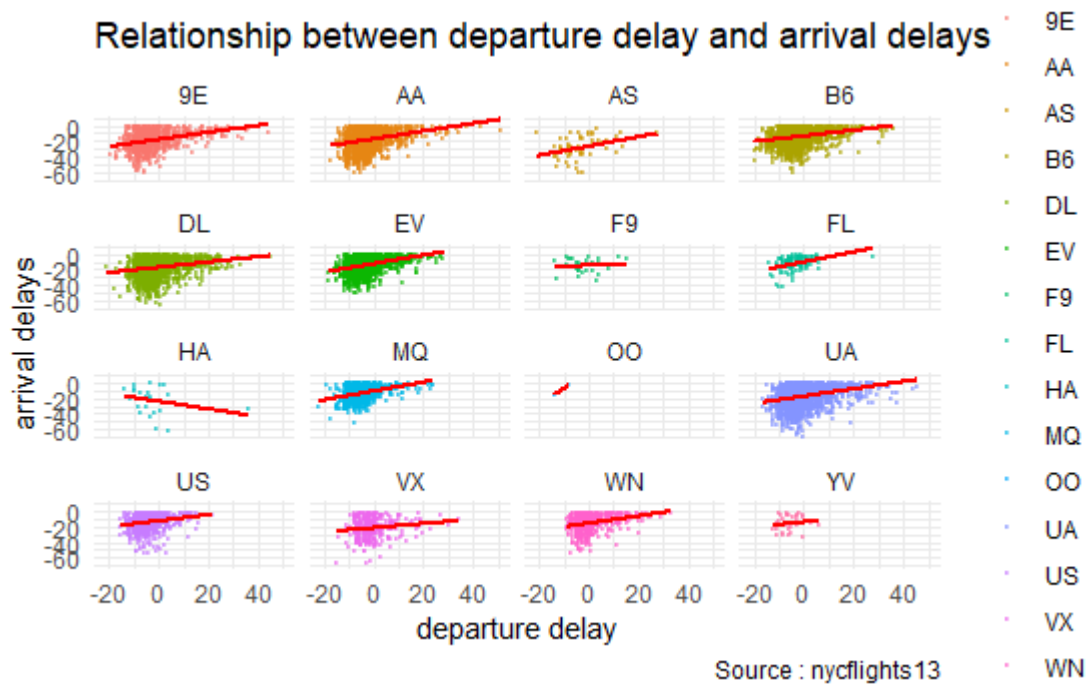
Chart 4 : Relationship between departure delay and arrival delays at carrier

Hide

```
hw4 <- flights %>%
  select(arr_delay, dep_delay, carrier) %>%
  filter(arr_delay < 0) %>%
  left_join(airlines, by = c("carrier")) %>%
  rename(carrier_name = name)
```

Hide

```
ggplot(sample_n(hw4, 20000), aes(dep_delay, arr_delay, color=carrier)) +
  geom_point(size = 0.2,
    alpha = 0.5) +
  geom_smooth(method="lm", se=F, color="RED") +
  theme_minimal() +
  facet_wrap(~ carrier, ncol=4) +
  labs(title = "Relationship between departure delay and arrival delays",
    caption = "Source : nycflights13",
    x = "departure delay",
    y = "arrival delays")
```



carrier <chr>	carrier_name <chr>
9E	Endeavor Air Inc.
AA	American Airlines Inc.
AS	Alaska Airlines Inc.
B6	JetBlue Airways
DL	Delta Air Lines Inc.
EV	ExpressJet Airlines Inc.
F9	Frontier Airlines Inc.
FL	AirTran Airways Corporation
HA	Hawaiian Airlines Inc.
MQ	Envoy Air
1-10 of 16 rows	
<a href="#">Previous</a> <a href="#">1</a> <a href="#">2</a> <a href="#">Next</a>	

Chart 5 : Total Flight in 2013 by month

Hide

```

flights %>%
  group_by(month) %>%
  summarise(flights = n()) %>%
  ggplot( aes(x = month, y = flights)) +
  geom_line(stat = "identity", color = 'sky blue', size = 1.5) +
  scale_x_continuous(breaks = seq(1,12,by=1), minor_breaks = seq(1,12, by = 1)) +
  geom_text(aes(label=flights), color='black') +
  labs(
    title = "Total Flight in 2013 by month",
    caption = "Source : nycflights13") +
  theme_minimal()

```

