

# R Practice - dplyr and NYC Flights

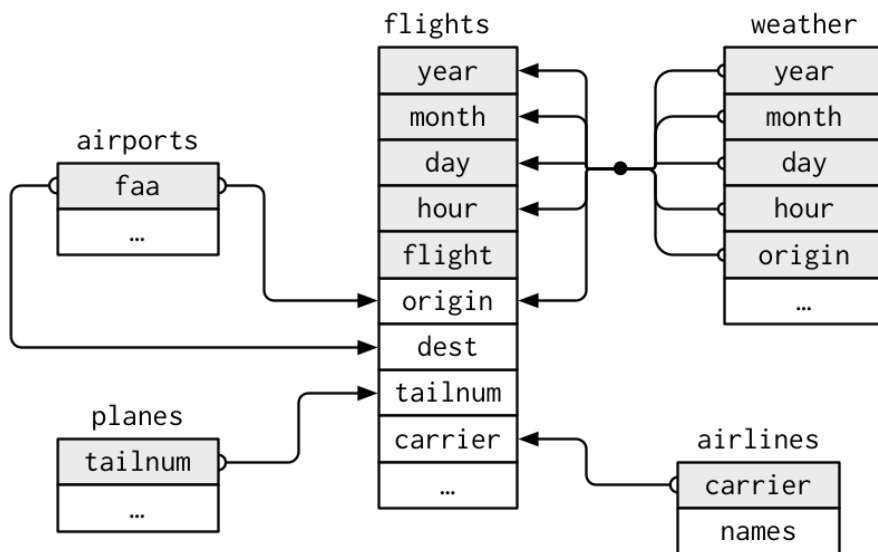
See if you can answer a few more questions about the NYC flights data set

## Accessing the data

Make sure you load the tidyverse package which includes dplyr as well as the nycflights13 package to get the flights data

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

Recall that the library includes the following tables



## Use dplyr to answer these questions

Note that these can all be answered in one dplyr chain.

- 1) What destination received most flights in June?
- 2) Which carrier had the greatest average distance per flight?
- 3) Which flight traveled the fastest (overall miles per hour)?

- 4) What day had the largest average arrival delay for all flights? (only count positive values)
- 5) What was the total distance for all flights in January? What was the average distance per flight?
- 6) What day of the week saw the most flights? (Hint, use the "time\_hour" column which is a date/time column and use google to find a function to give you the weekday for a date).
- 7) What was the average number of seats and engines on the planes that left from NYC on July 4? (The "seats" and "engines" data comes from the the planes table; Warning: Be careful with your joins!)
- 8) How many airlines do not have the word "air" somewhere in their name? (Hint, try googling for a function that can do string matching in R)
- 9) What was the most common plane model to fly out of NYC in October (there is a "model" column in the planes table)?
- 10) How many planes (tailnum) only flew one route (flight) but flew that route more than 10 times?
- 11) Which scheduled departure hour (use the "hour" column) had the largest proportion of flights delayed (dep\_delay) longer than 5 min?
- 12) Which flight(s) had the greatest scheduled length (time between scheduled departure and arrival)? (Warning: Take notice how the values of sched\_arr\_time and sched\_dep\_time are formatted in the table).

## Expected Values

- 1) Chicago O'Hare (ORD) with 1574
- 2) Hawaiian Airlines Inc. with an average of 4983 miles per flight
- 3) Flight 1499 to ATL flew 762 miles with average speed of 703.3846
- 4) July 10, with an average delay of 110 minutes
- 5) 27,188,805 total miles with an average of 1006.844 miles per flight
- 6) Monday with 50690
- 7) Seats: 140.6581 Engines: 1.991974
- 8) Only Virgin America
- 9) The A320-232 with 3717 flights
- 10) 12 planes only flew one route but did so more than 10 times
- 11) 9PM with 44.6%
- 12) Flight 51 with 6 hours 40 minutes