# dplyr Lab

*2019-01-31*

- The following exercises are all from Section 5.7.1 of RDS.

- Use the `flights` data frame from the nycflights13 package.

1. Which plane (`tailnum`) has the worst departure delay record?

2. What time of day should you fly if you want to avoid delays as much as possible?

3. For each destination, compute the total minutes of arrival delay. For each flight, compute the proportion of the arrival delay for its destination.

4. Delays are typically temporally correlated: even once the problem that caused the initial delay has been resolved, later flights are delayed to allow earlier flights to leave. Using `lag()`, explore how the departure delay of a flight is related to the delay of the immediately preceding flight.

5. Look at each destination. Can you find flights that are suspiciously fast? (i.e. flights that represent a potential data entry error). Compute the air time of a flight relative to the shortest flight to that destination. Which flights were most delayed in the air?

6. Find all destinations that are flown by at least two carriers. (hint: use `n_distinct()`)