***Tutorial 9– Data Manipulation (1)***

*flights* data set in the *nycflights13* package contains on-time data for all flights that departed NYC in 2013. The description of columns found in *flights* data set is given below. Use this dataset and *dplyr* package to find answers for the following questions:

|  |  |
| --- | --- |
| Column name(s) | Description |
| *year, month, day* | Date of departure. |
| *dep\_time, arr\_time* | Actual departure and arrival times (format HHMM or HMM), local tz. |
| *sched\_dep\_time, sched\_arr\_time* | Scheduled departure and arrival times (format HHMM or HMM), local tz. |
| *dep\_delay, arr\_delay* | Departure and arrival delays, in minutes. Negative times represent early departures/arrivals. |
| *carrier* | Two letter carrier abbreviation. |
| *flight* | Flight number. |
| *tailnum* | Plane tail number. |
| *origin, dest* | Origin and destination. |
| *air\_time* | Amount of time spent in the air, in minutes. |
| *distance* | Distance between airports, in miles. |
| *hour, minute* | Time of scheduled departure broken into hour and minutes. |
| *time\_hour* | Scheduled date and hour of the flight as a POSIXct date. |

* 1. Find all flights that had an arrival delay of three or more hours.
  2. Find all flights that arrived more than three hours late but didn’t leave late.
  3. Find all flights that were delayed by at least an hour but made up over 50 minutes in flight.
  4. Find all flights that departed between midnight and 5am.
  5. Find the most delayed flights.
  6. Find the flights travelled the farthest. **Hint**: use *air\_time* column.
  7. Find the flights travelled the shortest.
  8. Show the flights details for each flight: flight number, origin and destination.
  9. Show the flights details for each flight travelled in June 2013: flight number, Origin and destination.
  10. Find the carrier that has the maximum departure delay on 1st January 2013.