課題解説

HRY

2019年5月22日

では、5月24日（金）までの指定課題(p.67)の解説をします。 まず、準備段階として[Issuesページ](https://github.com/Ylab-Shiny/lesson_190524/issues/1)述べた install.packages(c("tidyverse", "nycflights13"))をRconsoleで実行し、外部パッケージをインストール してください。

# データ(flights)の確認

とりあえず、データセットの確認を行ってみましょう。

# パッケージの読み込み  
library(tidyverse)

## -- Attaching packages ------------------------------------------------------------------- tidyverse 1.2.1 --

## √ ggplot2 2.2.1 √ purrr 0.2.4  
## √ tibble 1.4.1 √ dplyr 0.7.4  
## √ tidyr 0.7.2 √ stringr 1.2.0  
## √ readr 1.1.1 √ forcats 0.2.0

## -- Conflicts ---------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

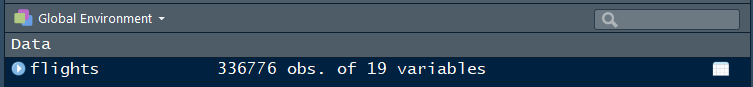
library(nycflights13)

## Warning: package 'nycflights13' was built under R version 3.4.4

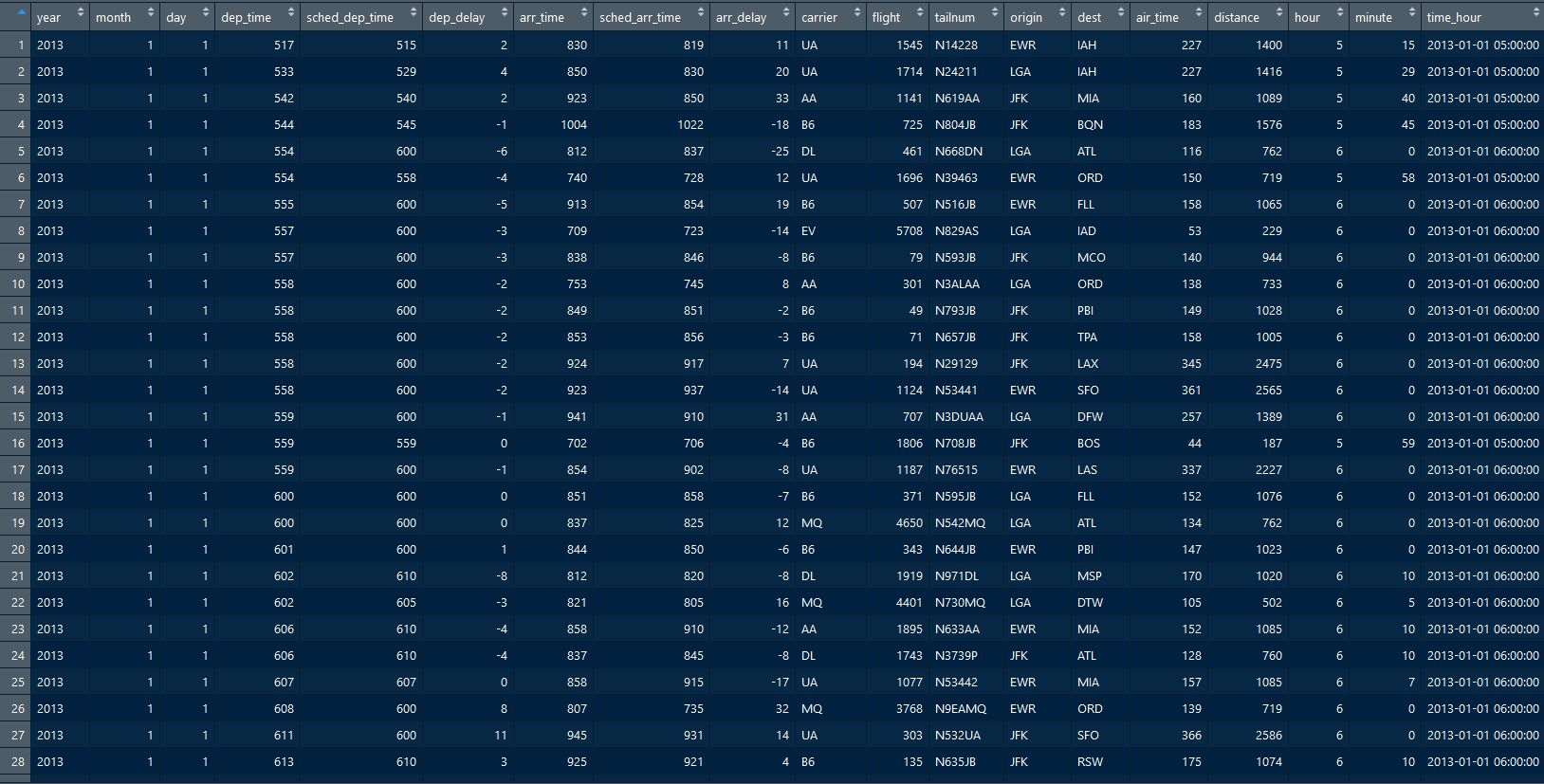
# RstudioのGlobal Environmentにflightsオブジェクトを作成  
flights <- flights  
# 上から5行までのデータの表示  
head(flights)

## # A tibble: 6 x 19  
## year month day dep\_t~ sche~ dep\_~ arr\_~ sche~ arr\_~ carr~ flig~ tail~  
## <int> <int> <int> <int> <int> <dbl> <int> <int> <dbl> <chr> <int> <chr>  
## 1 2013 1 1 517 515 2.00 830 819 11.0 UA 1545 N142~  
## 2 2013 1 1 533 529 4.00 850 830 20.0 UA 1714 N242~  
## 3 2013 1 1 542 540 2.00 923 850 33.0 AA 1141 N619~  
## 4 2013 1 1 544 545 -1.00 1004 1022 -18.0 B6 725 N804~  
## 5 2013 1 1 554 600 -6.00 812 837 -25.0 DL 461 N668~  
## 6 2013 1 1 554 558 -4.00 740 728 12.0 UA 1696 N394~  
## # ... with 7 more variables: origin <chr>, dest <chr>, air\_time <dbl>,  
## # distance <dbl>, hour <dbl>, minute <dbl>, time\_hour <dttm>

flights <- flightsというコードをRstudioのConsoleで実行することによって、RstudioのGlobal Environmentにデータの行数(obs.)と列数(variables)が表示されます。



また、View(flights)を同様に実行することでRstudio上で文字通りデータを下図のように眺めることができます。



データの系列（変数）の確認については、?flightsを実行することで確認できます。



* year,month,day：文字通りの意味で出発日に関する系列(Date of departure)
* dep\_time,arr\_time：実際の出発時刻と到着時刻の系列(Actual departure and arrival times (format HHMM or HMM), local tz.)
* sched\_dep\_time,sched\_arr\_time：予定出発時刻と予定到着時刻の系列(Scheduled departure and arrival times (format HHMM or HMM), local tz.)
* dep\_delay,arr\_delay：出発と到着の分単位の遅れの系列。負値は早い出発及び到着をあらわす(Departure and arrival delays, in minutes. Negative times represent early departures/arrivals.)
* hour,minute：出発予定時刻を時間と分に分割した系列。(Time of scheduled departure broken into hour and minutes.)
* carrier：2文字の業者の略語(Two letter carrier abbreviation. See airlines() to get name)