

Actividad 11

2023-05-15

Ejercicio 1

Usando los datos del paquete nycflights13 responde las siguientes preguntas:

1.

Utilice flight2 como en clase que tiene un subconjunto de variables de flight como sigue

```
library(nycflights13)
library(magrittr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
data(flights)
flights2 <- flights %>%
  dplyr::select(year:day, hour, origin, dest, tailnum, carrier)
```

Usá left_join para relacionar flight2 con weather sin especificar el parámetro by. Qué variables usa para machear en flight2 y weather por defecto?

```
data(weather)
left_join(weather, flights2)
```

```
## Joining with `by = join_by(origin, year, month, day, hour)`
```

```
## # A tibble: 341,957 x 18
##   origin year month   day hour temp  dewp humid wind_dir wind_speed
##   <chr>   <int> <int> <int> <dbl> <dbl> <dbl> <dbl>    <dbl>    <dbl>
## 1 EWR    2013     1     1     1  39.0  26.1  59.4      270      10.4
## 2 EWR    2013     1     1     2  39.0  27.0  61.6      250       8.06
## 3 EWR    2013     1     1     3  39.0  28.0  64.4      240      11.5
## 4 EWR    2013     1     1     4  39.9  28.0  62.2      250      12.7
## 5 EWR    2013     1     1     5  39.0  28.0  64.4      260      12.7
## 6 EWR    2013     1     1     5  39.0  28.0  64.4      260      12.7
## 7 EWR    2013     1     1     6  37.9  28.0  67.2      240      11.5
## 8 EWR    2013     1     1     6  37.9  28.0  67.2      240      11.5
## 9 EWR    2013     1     1     6  37.9  28.0  67.2      240      11.5
## 10 EWR   2013     1     1     6  37.9  28.0  67.2      240      11.5
```

```
## # i 341,947 more rows
## # i 8 more variables: wind_gust <dbl>, precip <dbl>, pressure <dbl>,
## #   visib <dbl>, time_hour <dtm>, dest <chr>, tailnum <chr>, carrier <chr>
```

Utiliza todas las columnas que tienen en común (origin, year, month, day, hour)

2.

En flight2 una clave primaria puede ser (year, month, day, hour, origin)?

3.

Usando left_join relaciona flight2 con airports usando la variable dest y faa respectivamente.

4.

Si quisieramos dibujar (aproximadamente) la ruta de cada avión vuela desde el origen al destino. ¿Qué variables necesitaríamos? ¿Qué tablas necesitamos combinar? ¿Pensá cuál es el conjunto de datos necesarios para obtener el siguiente dibujo de 100 vuelos, obtené los datos pero no hacer el dibujo?