



```
library(nycflights13)
```

```
flights
```

```
#> # A tibble: 336,776 x 19
```

```
#>   year month   day dep_time sched_dep_time dep_delay arr_time sched_arr_time  
#>   <int> <int> <int>   <int>         <int>         <dbl>      <int>         <int>
```

```
#> 1  2013     1     1     517           515           2        830           819
```

```
#> 2  2013     1     1     533           529           4        850           830
```

```
#> 3  2013     1     1     542           540           2        923           850
```

```
#> 4  2013     1     1     544           545          -1       1004          1022
```

```
#> 5  2013     1     1     554           600          -6        812           837
```

```
#> 6  2013     1     1     554           558          -4        740           728
```

```
#> # ... with 3.368e+05 more rows, and 11 more variables: arr_delay <dbl>,
```

```
#> #   carrier <chr>, flight <int>, tailnum <chr>, origin <chr>, dest <chr>,
```

```
#> #   air_time <dbl>, distance <dbl>, hour <dbl>, minute <dbl>, time_hour <dtm>
```

How does the weather affect delays?

Which carriers tend to buy certain types of aircraft?

```
library(nycflights13)
```

```
flights
```

```
#> # A tibble: 336,776 x 19
```

```
#>   year month   day dep_time sched_dep_time dep_delay arr_time sched_arr_time  
#>   <int> <int> <int>   <int>         <int>         <dbl>     <int>         <int>
```

```
#> 1  2013     1     1     517           515           2       830           819
```

```
#> 2  2013     1     1     533           529           4       850           830
```

```
#> 3  2013     1     1     542           540           2       923           850
```

```
#> 4  2013     1     1     544           545          -1      1004          1022
```

```
#> 5  2013     1     1     554           600          -6       812           837
```

```
#> 6  2013     1     1     554           558          -4       740           728
```

```
#> # ... with 3.368e+05 more rows, and 11 more variables: arr_delay <dbl>,
```

```
#> #   carrier <chr>, flight <int>, tailnum <chr>, origin <chr>, dest <chr>,
```

```
#> #   air_time <dbl>, distance <dbl>, hour <dbl>, minute <dbl>, time_hour <dtm>
```

How does the weather affect delays?

Which carriers tend to buy certain types of aircraft?

## Problem: Integrate data from multiple tables

```
airlines
#> # A tibble: 16 x 2
#>   carrier name
#>   <chr>     <chr>
#> 1 9E      Endeavor Air Inc.
#> 2 AA      American Airlines Inc.
#> 3 AS      Alaska Airlines Inc.
#> 4 B6      JetBlue Airways
#> 5 DL      Delta Air Lines Inc.
#> 6 EV      ExpressJet Airlines Inc.
#> # ... with 10 more rows
```