Problem: Integrate data from multiple tables

Solution: Mutating joins

```
flights2 <- flights %>%
 select(year:day, hour, origin, dest, tailnum, carrier)
flights2
#> # A tibble: 336,776 x 8
#> year month day hour origin dest tailnum carrier
#> <int> <int> <int> <dbl> <chr> <chr>
                                                 <chr>
#> 1 2013
                                   IAH
                                         N14228
                                                UA
                          5 EWR
#> 2
     2013
                          5 LGA
                                   IAH
                                         N24211
                                                UA
#> 3
     2013
                          5 JFK
                                         N619AA
                                                 AA
                                   MIA
#> 4
     2013
                          5 JFK
                                   BON
                                         N804JB
                                                 B6
#> 5
     2013
                                         N668DN
                          6 LGA
                                   ATL
                                                 DL
#> 6
     2013
                          5 EWR
                                   ORD
                                         N39463
                                                 UA
#> # ... with 3.368e+05 more rows
```

```
flights2 %>%
 select(-origin, -dest) %>%
left join(airlines, by = "carrier")
#> # A tibble: 336,776 x 7
#> year month day hour tailnum carrier name
#> <int> <int> <int> <dbl> <chr> <chr>
                                        <chr>
#> 1 2013
                        5 N14228 UA United Air Lines Inc.
#> 2 2013
                        5 N24211 UA United Air Lines Inc.
#> 3 2013 1 1
                                        American Airlines Inc.
                        5 N619AA
                                AA
#> 4 2013
         1 1
                        5 N804JB
                                B6
                                        JetBlue Airways
#> 5 2013
                                        Delta Air Lines Inc.
                        6 N668DN
                                DL
#> 6 2013
                        5 N39463 UA
                                        United Air Lines Inc.
#> # ... with 3.368e+05 more rows
```

left join() matches two tables based on their keys and adds the new variables from the second one to the first one

Problem: Integrate data from multiple tables

Solution: Mutating joins

```
flights2 <- flights %>%
  select(year:day, hour, origin, dest, tailnum, carrier)
flights2
#> # A tibble: 336,776 x 8
      year month day hour origin dest tailnum carrier
     <int> <int> <int> <dbl> <chr> <chr>
                                                 <chr>
      2013
                           5 EWR
                                   IAH
                                         N14228
#> 2 2013
                           5 LGA
                                         N24211
                                   IAH
#> 3 2013 1
#> 4 2013 1
                           5 JFK
                                         N619AA
                                   MIA
                           5 JFK
                                         N804JB
                                   BQN
      2013
                           6 LGA
                                         N668DN
                                   ATL
#> 6 2013
                           5 EWR
                                         N39463
#> # ... with 3.368e+05 more rows
                                                              left join() matches two
flights2 %>%
  select(-origin, -dest) %>%
                                                             tables based on their keys and
  left_join(airlines, by = "carrier")
                                                            adds the new variables from the
#> # A tibble: 336,776 x 7
                                                              second one to the first one
     year month day hour tailnum carrier name
    <int> <int> <int> <dbl> <chr>
                                            <chr>
                          5 N14228 UA
                                            United Air Lines Inc.
     2013
                          5 N24211 UA
#> 2 2013
                                            United Air Lines Inc.
                          5 N619AA
#> 3 2013
                                            American Airlines Inc.
     2013
                          5 N804JB
                                            JetBlue Airways
#> 5 2013
                          6 N668DN
                                            Delta Air Lines Inc.
#> 6 2013
                          5 N39463
                                            United Air Lines Inc.
                                   UA
#> # ... with 3.368e+05 more rows
```

There are several types of join operations