Variables might have different names in different tables

```
flights2 %>%
  left_join(airports, c("dest" = "faa"))
#> # A tibble: 336,776 x 15
#>    year month   day  hour origin dest  tailnum carrier name    lat   lon   alt
#>   <int> <int> <int> <dbl> <chr>  <chr> <chr>   <chr>   <chr> <dbl> <dbl> <dbl>
#> 1  2013     1     1     5 EWR    IAH   N14228  UA      Geor…  30.0 -95.3    97
#> 2  2013     1     1     5 LGA    IAH   N24211  UA      Geor…  30.0 -95.3    97
#> 3  2013     1     1     5 JFK    MIA   N619AA  AA      Miam…  25.8 -80.3     8
#> 4  2013     1     1     5 JFK    BQN   N804JB  B6      <NA>   NA    NA      NA
#> 5  2013     1     1     6 LGA    ATL   N668DN  DL      Hart…  33.6 -84.4  1026
#> 6  2013     1     1     5 EWR    ORD   N39463  UA      Chic…  42.0 -87.9   668
#> # … with 3.368e+05 more rows, and 3 more variables: tz <dbl>, dst <chr>,
#> #   tzone <chr>
```

You can choose the variables that are used with the `by` argument, which can be a list of pairs of strings that map variables to one another.

Exercise set 2

Compute the average delay by destination, then join on the `airports` data frame so you can show the spatial distribution of delays. Here's an easy way to draw a map of the United States:

```
airports %>%
  semi_join(flights, c("faa" = "dest")) %>%
  ggplot(aes(lon, lat)) +
    borders("state") +
    geom_point() +
    coord_quickmap()
```