

select () picks variables by their names

```
# Select all columns except those from year to day (inclusive)
select(flights, -(year:day))
#> # A tibble: 336,776 x 16
#>   dep_time sched_dep_time dep_delay arr_time sched_arr_time arr_delay carrier
#>   <int>         <int>         <dbl>   <int>         <int>         <dbl> <chr>
#> 1      517           515           2      830           819          11 UA
#> 2      533           529           4      850           830          20 UA
#> 3      542           540           2      923           850          33 AA
#> 4      544           545          -1     1004          1022         -18 B6
#> 5      554           600          -6      812           837         -25 DL
#> 6      554           558          -4      740           728          12 UA
#> # ... with 3.368e+05 more rows, and 9 more variables: flight <int>,
#> #   tailnum <chr>, origin <chr>, dest <chr>, air_time <dbl>, distance <dbl>,
#> #   hour <dbl>, minute <dbl>, time_hour <dtm>
```

mutate () creates new variables from existing ones

```
flights_sml <- select(flights,
  year:day,
  ends_with("delay"),
  distance,
  air_time
)
mutate(flights_sml,
  gain = dep_delay - arr_delay,
  speed = distance / air_time * 60
```

```
#> # A tibble: 336,776 x 9
#>   year month   day dep_delay arr_delay distance air_time   gain speed
#>   <int> <int> <int>   <dbl>   <dbl>   <dbl>   <dbl> <dbl> <dbl>
#> 1  2013     1     1         2        11    1400     227     -9   370.
#> 2  2013     1     1         4        20    1416     227    -16   374.
#> 3  2013     1     1         2        33    1089     160    -31   408.
#> 4  2013     1     1        -1       -18    1576     183     17   517.
#> 5  2013     1     1        -6       -25     762     116     19   394.
#> 6  2013     1     1        -4        12     719     150    -16   288.
#> # ... with 3.368e+05 more rows
```