Workflow: code style

Gabriel Montes

2025-07-17

1. Restyle the following pipelines following the guidelines above.

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.2     v tibble    3.3.0
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.4
## -- Conflicts ------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(dplyr)
library(nycflights13)
flights|>filter(dest=="IAH")|>group_by(year,month,day)|>summarize(n=n(),
delay=mean(arr_delay,na.rm=TRUE))|>filter(n>10)
```

```
## `summarise()` has grouped output by 'year', 'month'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 365 x 5
## # Groups:   year, month [12]
##     year month   day     n delay
##    <int> <int> <int> <int> <dbl>
##  1  2013     1     1    20 17.8
##  2  2013     1     2    20  7
##  3  2013     1     3    19 18.3
##  4  2013     1     4    20 -3.2
##  5  2013     1     5    13 20.2
##  6  2013     1     6    18  9.28
##  7  2013     1     7    19 -7.74
##  8  2013     1     8    19  7.79
##  9  2013     1     9    19 18.1
## 10  2013     1    10    19  6.68
## # i 355 more rows
```

```r
flights|>filter(carrier=="UA",dest%in%c("IAH","HOU"),sched_dep_time>
0900,sched_arr_time<2000)|>group_by(flight)|>summarize(delay=mean(
arr_delay,na.rm=TRUE),cancelled=sum(is.na(arr_delay)),n=n())|>filter(n>10)
```

```
## # A tibble: 74 x 4
##    flight delay cancelled     n
```

```
##      <int> <dbl>      <int> <int>
## 1       53 12.5          2    18
## 2      112 14.1          0    14
## 3      205 -1.71         0    14
## 4      235 -5.36         0    14
## 5      255 -9.47         0    15
## 6      268 38.6          1    15
## 7      292  6.57         0    21
## 8      318 10.7          1    20
## 9      337 20.1          2    21
## 10     370 17.5          0    11
## # i 64 more rows
```

restyled

```r
houston_daily_summary <- flights |>
  filter(dest == "IAH") |>
  group_by(year, month, day) |>
  summarize(
    n     = n(),
    delay = mean(arr_delay, na.rm = TRUE)
  ) |>
  filter(n > 10)
```

```
## `summarise()` has grouped output by 'year', 'month'. You can override using the
## `.groups` argument.
```

```r
ua_daytime_summary <- flights |>
  filter(
    carrier == "UA",
    dest %in% c("IAH", "HOU"),
    sched_dep_time >  0900,
    sched_arr_time < 2000
  ) |>
  group_by(flight) |>
  summarize(
    delay     = mean(arr_delay, na.rm = TRUE),
    cancelled = sum(is.na(arr_delay)),
    n         = n()
  ) |>
  filter(n > 10)
```