Data Wrangling

STAT 231 Blog Project

Gloria Wu, Justyce Williams, Ben Snyderman

2025-05-06

Variables

Wrangling

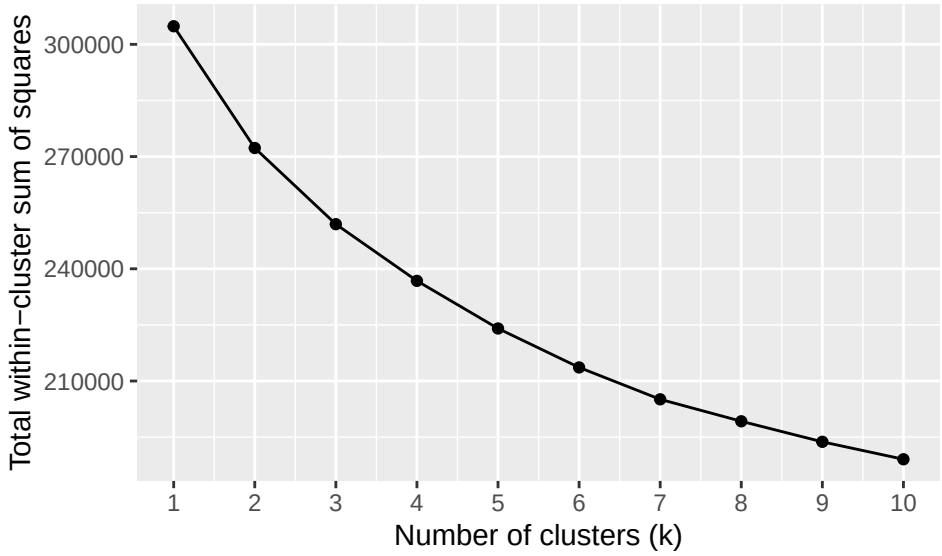
```
[1] TRUE
```

```
[1] TRUE
```

```
[1] TRUE
```

```
[1] TRUE
```

```
Warning: There were 11 warnings in `mutate()`.
The first warning was:
i In argument: `kmeans_results = purrr::map(k,
  ~kmeans(spotify_kmeans_standardized, .x, nstart = 20))`.
Caused by warning:
! Quick-TRANSfer stage steps exceeded maximum (= 1172500)
i Run `dplyr::last_dplyr_warnings()` to see the 10 remaining warnings.
```




```
Warning in inner_join(songs_top100, ht, by = join_by(track_artist)): Detected an unexpected r
i Row 5 of `x` matches multiple rows in `y`.
i Row 5 of `y` matches multiple rows in `x`.
i If a many-to-many relationship is expected, set `relationship =
  "many-to-many"` to silence this warning.
```

```
[1] 91
```

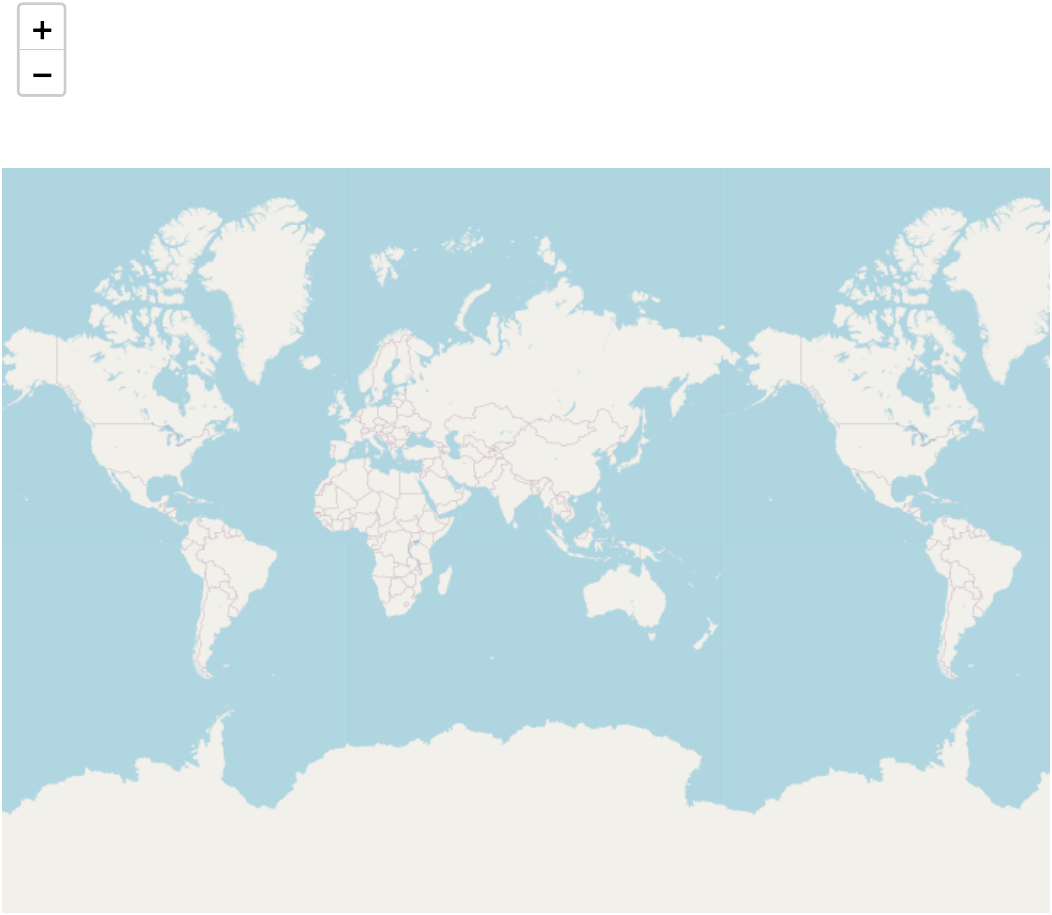
```
Warning in left_join(network, genre, by = join_by(track_artist)): Detected an unexpected many
i Row 2 of `x` matches multiple rows in `y`.
i Row 10 of `y` matches multiple rows in `x`.
i If a many-to-many relationship is expected, set `relationship =
  "many-to-many"` to silence this warning.
```

```
# A tibble: 2,237 x 3
   track_artist  collabArtist genre
   <chr>         <chr>        <chr>
 1 ava max       mariah carey pop
 2 sam smith     mariah carey pop
 3 sam smith     mariah carey pop
 4 justin bieber mariah carey pop
```

```
 5 justin bieber mariah carey pop
 6 justin bieber mariah carey pop
 7 sam smith     mariah carey pop
 8 sam smith     mariah carey pop
 9 post malone   mariah carey pop
10 post malone   mariah carey pop
# i 2,227 more rows
```

```
Warning in right_join(hometownAdd, songs_top100, by = ("track_artist")): Detected an unexpec
i Row 5 of `x` matches multiple rows in `y`.
i Row 5 of `y` matches multiple rows in `x`.
i If a many-to-many relationship is expected, set `relationship =
  "many-to-many"` to silence this warning.
```

```
Warning in CPL_crs_from_input(x): GDAL Message 1: CRS EPSG:2163 is deprecated.
Its non-deprecated replacement EPSG:9311 will be used instead. To use the
original CRS, set the OSR_USE_NON_DEPRECATED configuration option to NO.
```


+
−
Leaflet | © OpenStreetMap, ODbL

Sources

- https://www.rdocumentation.org/packages/geodata/versions/0.6-2 For Canada Data