MergeScriptQB

Quinn Bankson

2023-01-30

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   1.0.1
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(janitor)
```

```
##
## Attaching package: 'janitor'
##
## The following objects are masked from 'package:stats':
##
##     chisq.test, fisher.test
```

```
rawdata <- read_csv(file = "/Users/mac/Documents/R Assignments PLCY 715/DarkOrchid/Bankson/InitialData/
```

```
## Rows: 13491 Columns: 37
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (27): Tracker ID, TrackerLOC, ParentID, Wiki page, Country, Subnational ...
## dbl  (8): Capacity (MW), RETIRED, Planned Retire, Latitude, Longitude, Annua...
## num  (2): Heat rate (Btu per kWh), Emission factor (kg of CO2 per TJ)
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
USrawdata <- rawdata %>% filter(Country == "United States")
USdata <- clean_names(USrawdata)
names(USdata)
```

```
##  [1] "tracker_id"                      "tracker_loc"
##  [3] "parent_id"                       "wiki_page"
##  [5] "country"                         "subnational_unit_province_state"
```

```
##  [7] "unit"                              "plant"
##  [9] "chinese_name"                      "other_names"
## [11] "owner"                             "parent"
## [13] "capacity_mw"                       "status"
## [15] "year"                              "retired"
## [17] "planned_retire"                    "combustion_technology"
## [19] "coal_type"                         "coal_source"
## [21] "location"                          "local_area_taluk_county"
## [23] "major_area_prefecture_district"    "region"
## [25] "latitude"                          "longitude"
## [27] "accuracy"                          "permits"
## [29] "captive"                           "captive_industry_use"
## [31] "captive_residential_use"           "heat_rate_btu_per_k_wh"
## [33] "emission_factor_kg_of_co2_per_tj" "capacity_factor"
## [35] "annual_co2_million_tonnes_annum"   "lifetime_co2"
## [37] "remaining_plant_lifetime_years"
```

```
USdata <- USdata %>%
    rename("state" = "subnational_unit_province_state",
           "county" = "local_area_taluk_county")
us_coal <- USdata %>% select("state", "year","county", "status", "combustion_technology", "emission_fac

write_csv(us_coal, file = "/Users/mac/Documents/R Assignments PLCY 715/DarkOrchid/Bankson/coal_dta_clea
```