**Packages Used:**

install.packages("here")

install.packages("skimr")

install.packages("tidyverse")

install.packages("ggplot2")

install.packages("janitor")

**Summary of Data:**

skim\_without\_charts(global\_power\_plant\_database)

── Data Summary ────────────────────────

Values

Name global\_power\_plant\_databa...

Number of rows 34936

Number of columns 36

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Column type frequency:

character 19

numeric 17

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Group variables None

── Variable type: character ────────────────────────────────────────────────────────────────────────

skim\_variable n\_missing complete\_rate min max empty n\_unique whitespace

1 country 0 1 3 3 0 1670

2 country\_long 0 1 4 32 0 1670

3 name 0 1 1 87 0 345280

4 gppd\_idnr 0 1 10 12 0 349360

5 primary\_fuel 0 1 3 14 0 150

6 other\_fuel1 0 1 0 12 32992 130

7 other\_fuel2 0 1 0 7 34660 120

8 other\_fuel3 0 1 0 7 34844 90

9 owner 0 1 0 140 14068 101450

10 source 0 1 0 73 15 8670

11 url 0 1 0 795 18 48710

12 geolocation\_source 0 1 0 83 419 290

13 wepp\_id 0 1 0 25 18702 152640

14 generation\_data\_source 0 1 0 73 23536 180

15 estimated\_generation\_note\_2013 0 1 7 15 0 50

16 estimated\_generation\_note\_2014 0 1 7 15 0 50

17 estimated\_generation\_note\_2015 0 1 7 15 0 50

18 estimated\_generation\_note\_2016 0 1 7 15 0 50

19 estimated\_generation\_note\_2017 0 1 7 18 0 60

── Variable type: numeric ────────────────────────────────────────────────────────────

skim\_variable n\_missing complete\_rate mean sd p0 p25 p50

1 capacity\_mw 0 1 163. 490. 1 4916.7

2 latitude 0 1 32.8 22.6 -77.8 29.3 39.7

3 longitude 0 1 -6.97 78.4 -180. -77.6 -2.13

4 commissioning\_year 17489 0.499 1997. 23.4 1896 1988 2007

5 year\_of\_capacity\_data 20049 0.426 2018. 1.61 2000 2017 2019

6 generation\_gwh\_2013 28519 0.184 593. 2175. -948. 1.95 23.4

7 generation\_gwh\_2014 27710 0.207 657. 2231. -990. 2.26 23.6

8 generation\_gwh\_2015 26733 0.235 762. 2533. -864. 2.66 26.1

9 generation\_gwh\_2016 25792 0.262 693. 2405. -769. 2.73 22.5

10 generation\_gwh\_2017 25436 0.272 662. 2369. -935. 2.47 17.9

11 generation\_gwh\_2018 25299 0.276 517. 1975. -983. 2.24 12.5

12 generation\_gwh\_2019 25277 0.276 424. 1699. -780. 2.75 11.5

13 estimated\_generation\_gwh\_2013 18816 0.461 239. 1191. 1.12 8.62 27.6

14 estimated\_generation\_gwh\_2014 18433 0.472 242. 1286. 0.87 8.68 28.2

15 estimated\_generation\_gwh\_2015 17886 0.488 236. 1267. 0.44 8.38 26.8

16 estimated\_generation\_gwh\_2016 17366 0.503 236. 1260. 0.3 8.32 27.6

17 estimated\_generation\_gwh\_2017 1798 0.949 716. 2484. 0 8.18 37.6

p75 p100

1 75.3 22500

2 46.3 71.3

3 49.5 179.

4 2014 2020

5 2019 2019

6 200. 50834

7 226. 32321.

8 286. 37434.

9 250. 32377.

10 215. 36449.

11 151. 35136

12 123. 31920.

13 107. 48675.

14 107. 58471.

15 103. 57113.

16 107. 60860.

17 230. 82811.

**Load File in R Studio:**

global\_power\_plant\_database <- read.csv("F:/Data Work/global\_power\_plant\_database\_v\_1\_3/global\_power\_plant\_database.csv", header=FALSE)

View(global\_power\_plant\_database)

**Structure of Data Set:**

'data.frame': 34936 obs. of 36 variables:

$ country : chr "AFG" "AFG" "AFG" "AFG" ...

$ country\_long : chr "Afghanistan" "Afghanistan" "Afghanistan" "Afghanistan" ...

$ name : chr "Kajaki Hydroelectric Power Plant Afghanistan" "Kandahar DOG" "Kandahar JOL" "Mahipar Hydroelectric Power Plant Afghanistan" ...

$ gppd\_idnr : chr "GEODB0040538" "WKS0070144" "WKS0071196" "GEODB0040541" ...

$ capacity\_mw : num 33 10 10 66 100 ...

$ latitude : num 32.3 31.7 31.6 34.6 34.6 ...

$ longitude : num 65.1 65.8 65.8 69.5 69.7 ...

$ primary\_fuel : chr "Hydro" "Solar" "Solar" "Hydro" ...

$ other\_fuel1 : chr "" "" "" "" ...

$ other\_fuel2 : chr "" "" "" "" ...

$ other\_fuel3 : chr "" "" "" "" ...

$ commissioning\_year : num NA NA NA NA NA ...

$ owner : chr "" "" "" "" ...

$ source : chr "GEODB" "Wiki-Solar" "Wiki-Solar" "GEODB" ...

$ url : chr "http://globalenergyobservatory.org" "https://www.wiki-solar.org" "https://www.wiki-solar.org" "http://globalenergyobservatory.org" ...

$ geolocation\_source : chr "GEODB" "Wiki-Solar" "Wiki-Solar" "GEODB" ...

$ wepp\_id : chr "1009793" "" "" "1009795" ...

$ year\_of\_capacity\_data : int 2017 NA NA 2017 2017 2017 2017 2017 2017 NA ...

$ generation\_gwh\_2013 : num NA NA NA NA NA NA NA NA NA NA ...

$ generation\_gwh\_2014 : num NA NA NA NA NA NA NA NA NA NA ...

$ generation\_gwh\_2015 : num NA NA NA NA NA NA NA NA NA NA ...

$ generation\_gwh\_2016 : num NA NA NA NA NA NA NA NA NA NA ...

$ generation\_gwh\_2017 : num NA NA NA NA NA NA NA NA NA NA ...

$ generation\_gwh\_2018 : num NA NA NA NA NA NA NA NA NA NA ...

$ generation\_gwh\_2019 : num NA NA NA NA NA NA NA NA NA NA ...

$ generation\_data\_source : chr "" "" "" "" ...

$ estimated\_generation\_gwh\_2013 : num 123.8 18.4 18.6 225.1 406.2 ...

$ estimated\_generation\_gwh\_2014 : num 162.9 17.5 17.6 203.6 357.2 ...

$ estimated\_generation\_gwh\_2015 : num 97.4 18.2 19.1 146.9 271 ...

$ estimated\_generation\_gwh\_2016 : num 137.8 17.7 17.6 230.2 395.4 ...

$ estimated\_generation\_gwh\_2017 : num 119.5 18.3 18.7 174.9 350.8 ...

$ estimated\_generation\_note\_2013: chr "HYDRO-V1" "SOLAR-V1-NO-AGE" "SOLAR-V1-NO-AGE" "HYDRO-V1" ...

$ estimated\_generation\_note\_2014: chr "HYDRO-V1" "SOLAR-V1-NO-AGE" "SOLAR-V1-NO-AGE" "HYDRO-V1" ...

$ estimated\_generation\_note\_2015: chr "HYDRO-V1" "SOLAR-V1-NO-AGE" "SOLAR-V1-NO-AGE" "HYDRO-V1" ...

$ estimated\_generation\_note\_2016: chr "HYDRO-V1" "SOLAR-V1-NO-AGE" "SOLAR-V1-NO-AGE" "HYDRO-V1" ...

$ estimated\_generation\_note\_2017: chr "HYDRO-V1" "SOLAR-V1-NO-AGE" "SOLAR-V1-NO-AGE" "HYDRO-V1" ...

**Column Name:**

[1] "country" "country\_long"

[3] "name" "gppd\_idnr"

[5] "capacity\_mw" "latitude"

[7] "longitude" "primary\_fuel"

[9] "other\_fuel1" "other\_fuel2"

[11] "other\_fuel3" "commissioning\_year"

[13] "owner" "source"

[15] "url" "geolocation\_source"

[17] "wepp\_id" "year\_of\_capacity\_data"

[19] "generation\_gwh\_2013" "generation\_gwh\_2014"

[21] "generation\_gwh\_2015" "generation\_gwh\_2016"

[23] "generation\_gwh\_2017" "generation\_gwh\_2018"

[25] "generation\_gwh\_2019" "generation\_data\_source"

[27] "estimated\_generation\_gwh\_2013" "estimated\_generation\_gwh\_2014"

[29] "estimated\_generation\_gwh\_2015" "estimated\_generation\_gwh\_2016"

[31] "estimated\_generation\_gwh\_2017" "estimated\_generation\_note\_2013"

[33] "estimated\_generation\_note\_2014" "estimated\_generation\_note\_2015"

[35] "estimated\_generation\_note\_2016" "estimated\_generation\_note\_2017"

**Filtering Data:**

Solar <- filter(global\_power\_plant\_database, primary\_fuel == "Solar")

View(Solar)

Hydro <- filter(global\_power\_plant\_database, primary\_fuel == " Hydro ")

View(Hydro)

Gas <- filter(global\_power\_plant\_database, primary\_fuel == "Gas")

View(Gas)

Wind <- filter(global\_power\_plant\_database, primary\_fuel == "Wind")

View(Wind)

Biomass <- filter(global\_power\_plant\_database, primary\_fuel == "Biomass")

View(Biomass)

Oil <- filter(global\_power\_plant\_database, primary\_fuel == "Oil")

View(Oil)

Coal <- filter(global\_power\_plant\_database, primary\_fuel == "Coal")

View(Coal)

Nuclear <- filter(global\_power\_plant\_database, primary\_fuel == "Nuclear")

View(Nuclear)

Geothermal <- filter(global\_power\_plant\_database, primary\_fuel == "Geothermal")

View(Geothermal)

Waste <- filter(global\_power\_plant\_database, primary\_fuel == "Waste")

View(Waste)

Solar <- filter(global\_power\_plant\_database, primary\_fuel == "Solar")

View(Solar)

Asia <- filter(global\_power\_plant\_database, country\_long == 'Pakistan' | country\_long == 'China' |

country\_long == 'Indonesia' | country\_long == 'Iran' | country\_long == 'Israel'

| country\_long == 'Japan' | country\_long == 'Japan' | country\_long == 'Singapore'

| country\_long == 'Philippines' | country\_long == 'Bangladesh' | country\_long == 'Uzbekistan'

| country\_long == 'Kazakhstan' | country\_long == 'Bahrain' | country\_long == 'Myanmar'

| country\_long == 'Jordan' | country\_long == 'Afganistan' | country\_long == 'Oman'

| country\_long == 'Qatar' | country\_long == 'Syria' | country\_long == 'Tajikistan'

| country\_long == 'Thailand' | country\_long == 'Yemen' | country\_long == 'Cambodia'

| country\_long == 'Vietnam' | country\_long == 'Laos' | country\_long == 'Turmenistan'

| country\_long == 'Saudia Arabia' | country\_long == 'Kuwait' | country\_long == 'Lebanon'

| country\_long == 'Iraq' | country\_long == 'United Arab Emirates' | country\_long == 'Nepal'

| country\_long == 'Sri Lanka' | country\_long == 'Malaysia' | country\_long == 'South Korea'

| country\_long == 'Maldives' | country\_long == 'Mongolia' | country\_long == 'Bhutan'

| country\_long == 'Palestine' | country\_long == 'North Korea' | country\_long == 'Armenia'

| country\_long == 'Azerbaijan' | country\_long == 'Turkiye' | country\_long == 'Georgia'

| country\_long == 'Taiwan' | country\_long == 'Russia' | country\_long == 'Macao' | country\_long == 'Hong Kong')

Africa <- filter(global\_power\_plant\_database, country\_long == 'Nigeria' | country\_long == 'Ethiopia' |

country\_long == 'Egypt' | country\_long == 'Tanzania' | country\_long == 'South Africa'

| country\_long == 'Kenya' | country\_long == 'Uganda' | country\_long == 'Algeria'

| country\_long == 'Angola' | country\_long == 'Morocco' | country\_long == 'Sudan'

| country\_long == 'Mozambique' | country\_long == 'Ghana' | country\_long == 'Madagascar'

| country\_long == 'Burkina Faso' | country\_long == 'Niger' | country\_long == 'Cameroon'

| country\_long == 'Mali' | country\_long == 'Malawi' | country\_long == 'Zambia'

| country\_long == 'Somalia' | country\_long == 'Chad' | country\_long == 'Senegal'

| country\_long == 'Zimbabwe' | country\_long == 'Guinea' | country\_long == 'Rwanda'

| country\_long == 'Tunisia' | country\_long == 'Burundi' | country\_long == 'Benin'

| country\_long == 'South Sudan' | country\_long == 'Togo' | country\_long == ' Sierra Leone'

| country\_long == 'Liberia' | country\_long == 'Congo' | country\_long == 'Libya'

| country\_long == 'Central African Republic' | country\_long == 'Mauritania' | country\_long == 'Eritrea'

| country\_long == 'Botswana' | country\_long == 'Gambia' | country\_long == 'Namibia'

| country\_long == 'Gabon' | country\_long == 'Lesotho' | country\_long == 'Equatorial Guinea'

| country\_long == 'Seychelles' | country\_long == 'Comoros' | country\_long == 'Cabo Verde' | country\_long == 'Mauritius')

Europe <- filter(global\_power\_plant\_database, country\_long == 'Germany' | country\_long == 'United Kingdom' |

country\_long == 'France' | country\_long == 'Itlay' | country\_long == 'Spain'

| country\_long == 'Ukraine' | country\_long == 'Poland' | country\_long == 'Romania'

| country\_long == 'Netherlands' | country\_long == 'Belgium' | country\_long == 'Czech Republic'

| country\_long == 'Sweden' | country\_long == 'Portugal' | country\_long == 'Greece'

| country\_long == 'Hungry' | country\_long == 'Belarus' | country\_long == 'Austria'

| country\_long == 'Serbia' | country\_long == 'Switzerland' | country\_long == 'Bulgaria'

| country\_long == 'Denmark' | country\_long == 'Finland' | country\_long == 'Slovakia'

| country\_long == 'Norway' | country\_long == 'Ireland' | country\_long == 'Croatia'

| country\_long == 'Albania' | country\_long == 'Bosnia and Herzegovina' | country\_long == 'Moldova'

| country\_long == 'Lithuania' | country\_long == 'Russia' | country\_long == 'North Macedonia'

| country\_long == 'Estonia' | country\_long == 'Latvia' | country\_long == ' Slovenia'

| country\_long == 'Montenegro' | country\_long == 'Luxembourg' | country\_long == 'Malta'

| country\_long == 'Monaco' | country\_long == 'Andorra' | country\_long == 'Iceland'

| country\_long == 'Liechtenstein' | country\_long == 'San Marino' | country\_long == 'Holy See'

| country\_long == 'Taiwan' | country\_long == 'Russia' | country\_long == 'Macao' | country\_long == 'Hong Kong')

Europe <- filter(global\_power\_plant\_database, country\_long == 'United States of America' | country\_long == 'Mexico' |

country\_long == 'Haiti' | country\_long == 'Guatemala' | country\_long == 'Canada'

| country\_long == 'Cuba' | country\_long == 'Dominican Republic' | country\_long == 'Honduras'

| country\_long == 'Costa Rica' | country\_long == 'El Salvador' | country\_long == 'Nicaragua'

| country\_long == 'Panama' | country\_long == 'Jamaica' | country\_long == 'Puerto Rico'

| country\_long == 'Belize' | country\_long == ' Guadeloupe' | country\_long == 'Trinidad and Tobago'

| country\_long == 'Bahamas' | country\_long == 'Martinique' | country\_long == 'Barbados'

| country\_long == 'Saint Vincent and the Grenadines' | country\_long == 'Grenada' | country\_long == 'Saint Lucia'

| country\_long == 'Aruba' | country\_long == 'United States Virgin Islands' | country\_long == 'Antigua and Barbuda'

| country\_long == 'Bermuda' | country\_long == 'Cayman Islands' | country\_long == 'Dominica'

| country\_long == 'Greenland' | country\_long == 'Saint Kitts and Nevis' | country\_long == 'Sint Maarten'

| country\_long == 'British Virgin Islands' | country\_long == 'Saint Martin' | country\_long == 'Turks and Caicos Islands'

| country\_long == 'Caribbean Netherlands' | country\_long == 'Anguilla' | country\_long == 'Saint Pierre and Miquelon'

| country\_long == 'Montserrat')

South\_America <- filter(global\_power\_plant\_database, country\_long == 'Brazil' | country\_long == 'Colombia' |

country\_long == 'Venezuela' | country\_long == 'Peru' | country\_long == 'Argentina'

| country\_long == 'Chile' | country\_long == 'Ecuador' | country\_long == 'Bolivia'

| country\_long == 'Guyana' | country\_long == 'Uruguay' | country\_long == 'Paraguay'

| country\_long == 'Suriname' | country\_long == 'Portugal' | country\_long == 'French Guiana'

| country\_long == 'Falkland Islands')

**Sorting Each Filtered Data Set:**

arrage\_Hydro <- global\_power\_plant\_database %>%

arrange(capacity\_mw)

View(arrage\_Hydro)

groupByCountry <- global\_power\_plant\_database %>% group\_by(country\_long, primary\_fuel) %>%

summarize(max\_capacity\_mw = max(capacity\_mw), min\_capacity\_mw = min(capacity\_mw))

**Visualizations:**

**Europe:**

ggplot(data = Europe, aes(x = latitude, y = longitude))+

geom\_point(aes(color = country\_long))+facet\_wrap(~primary\_fuel)