Untitled

BBosa Robert

# Load required packages

install.packages("ggrepel")  
install.packages("fBasics")  
install.packages("ggplot2")  
install.packages("dplyr")  
install.packages("vtable")

# Call the libraries

library(ggplot2)  
library(ggrepel)  
library(fBasics)

## Loading required package: timeDate

## Loading required package: timeSeries

library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:timeSeries':  
##   
## filter, lag

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(vtable)

## Loading required package: kableExtra

##   
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':  
##   
## group\_rows

# LOad the Data Set

naeedata1 <- read.csv("~/NAEE\_Data/naee\_data/data/data1.csv",   
 header = TRUE,  
 check.names = FALSE)  
  
naeedata1 <- naeedata1[1:336,1:24]  
  
View(naeedata1)  
names(naeedata1)

## [1] "Country"   
## [2] "Year"   
## [3] "Electricity installed capacity in Total renewable energy (MW)"   
## [4] "Electricity installed capacity in Hydropower (MW)"   
## [5] "Electricity installed capacity in Wind (MW)"   
## [6] "Electricity installed capacity in Solar (MW)"   
## [7] "Electricity installed capacity in Bioenergy (MW)"   
## [8] "Electricity installed capacity in Geothermal (MW)"   
## [9] "Electricity export (GWh)"   
## [10] "Electricity import (GWh)"   
## [11] "Electricity final consumption (GWh)"   
## [12] "Electricity generation, Total (GWh)"   
## [13] "Electricity generated from fossil fuels (GWh)"   
## [14] "Electricity generated from nuclear power (GWh)"   
## [15] "Electricity generated from renewable sources (GWh)"   
## [16] "Electricity generated from hydropower (GWh)"   
## [17] "Electricity generated from solar, wind, tide, wave and other sources (GWh)"  
## [18] "Electricity generated from biofuels and waste (GWh)"   
## [19] "Electricity generated from geothermal energy (GWh)"   
## [20] "Population access to electricity-National (% of population)"   
## [21] "Population access to electricity-Urban (% of population)"   
## [22] "Population access to electricity-Rural (% of population)"   
## [23] "Electricity generation per Capita (KWh)"   
## [24] "Electricity final consumption per capita (KWh)"

# Rename variables

colnames(naeedata1)[2:24] <- c("year",  
 "in\_tre",  
 "in\_hyd",  
 "in\_wind",  
 "in\_solar",  
 "in\_bio",  
 "in\_geo",  
 "exp",  
 "imp",  
 "cons",  
 "tgen",  
 "gen\_foss",  
 "gen\_nuc",  
 "gen\_ren",  
 "gen\_hyd",  
 "gen\_solar",  
 "gen\_bio",  
 "gen\_geo",  
 "access",  
 "acc\_urban",  
 "acc\_rural",  
 "perk\_gen",  
 "perk\_cons")  
   
names(naeedata1)

## [1] "Country" "year" "in\_tre" "in\_hyd" "in\_wind" "in\_solar"   
## [7] "in\_bio" "in\_geo" "exp" "imp" "cons" "tgen"   
## [13] "gen\_foss" "gen\_nuc" "gen\_ren" "gen\_hyd" "gen\_solar" "gen\_bio"   
## [19] "gen\_geo" "access" "acc\_urban" "acc\_rural" "perk\_gen" "perk\_cons"

round(fBasics::basicStats(naeedata1[,-c(1:2)]),2)

## in\_tre in\_hyd in\_wind in\_solar in\_bio in\_geo exp  
## nobs 336.00 336.00 336.00 336.00 336.00 336.00 336.00  
## NAs 0.00 1.00 0.00 0.00 0.00 0.00 0.00  
## Minimum 0.00 0.00 0.00 0.00 0.00 0.00 0.00  
## Maximum 9638.70 3700.90 2636.00 5989.60 265.70 823.80 16894.80  
## 1. Quartile 8.93 0.00 0.00 1.15 0.00 0.00 0.00  
## 3. Quartile 742.10 701.70 1.30 22.70 33.80 0.00 42.83  
## Mean 696.73 449.80 92.08 113.70 28.29 14.20 821.35  
## Median 169.00 66.60 0.00 6.15 0.00 0.00 0.00  
## Sum 234099.70 150682.20 30938.30 38201.70 9504.30 4772.70 275972.30  
## SE Mean 70.25 39.05 18.24 30.92 3.02 5.25 167.57  
## LCL Mean 558.53 372.98 56.19 52.87 22.35 3.89 491.73  
## UCL Mean 834.92 526.61 127.97 174.52 34.22 24.52 1150.96  
## Variance 1658371.02 510868.31 111834.53 321305.01 3060.85 9247.09 9434269.78  
## Stdev 1287.78 714.75 334.42 566.84 55.32 96.16 3071.53  
## Skewness 3.36 2.09 4.59 7.73 2.52 6.99 4.23  
## Kurtosis 14.44 4.05 23.15 65.02 6.44 48.77 16.52  
## imp cons tgen gen\_foss gen\_nuc  
## nobs 336.00 3.360000e+02 3.360000e+02 3.360000e+02 336.00  
## NAs 0.00 0.000000e+00 0.000000e+00 0.000000e+00 0.00  
## Minimum 0.00 1.250000e+01 2.850000e+01 7.000000e-01 0.00  
## Maximum 16369.00 2.089137e+05 2.781408e+05 2.473010e+05 15814.40  
## 1. Quartile 0.00 3.636000e+02 4.062200e+02 2.204300e+02 0.00  
## 3. Quartile 508.32 8.249750e+03 9.723800e+03 4.053100e+03 0.00  
## Mean 985.90 1.382159e+04 1.756565e+04 1.469055e+04 348.85  
## Median 18.00 1.729450e+03 2.294350e+03 5.850500e+02 0.00  
## Sum 331263.60 4.644055e+06 5.902059e+06 4.936025e+06 117214.00  
## SE Mean 140.54 2.108210e+03 2.674300e+03 2.450790e+03 122.41  
## LCL Mean 709.44 9.674600e+03 1.230512e+04 9.869680e+03 108.06  
## UCL Mean 1262.37 1.796858e+04 2.282619e+04 1.951142e+04 589.64  
## Variance 6636959.72 1.493364e+09 2.403029e+09 2.018135e+09 5034831.28  
## Stdev 2576.23 3.864407e+04 4.902070e+04 4.492366e+04 2243.84  
## Skewness 3.59 3.960000e+00 3.980000e+00 3.930000e+00 6.27  
## Kurtosis 13.23 1.491000e+01 1.538000e+01 1.486000e+01 37.70  
## gen\_ren gen\_hyd gen\_solar gen\_bio gen\_geo access  
## nobs 336.00 336.00 336.00 336.00 336.00 336.00  
## NAs 0.00 0.00 0.00 0.00 0.00 42.00  
## Minimum 0.00 0.00 0.00 0.00 0.00 3.60  
## Maximum 16574.20 16398.20 8792.80 625.50 5343.00 100.00  
## 1. Quartile 11.18 0.90 0.20 0.00 0.00 32.08  
## 3. Quartile 3189.57 2597.00 29.10 12.25 0.00 76.33  
## Mean 2526.25 2055.78 323.27 43.79 103.41 54.23  
## Median 562.25 226.45 5.10 0.00 0.00 52.45  
## Sum 848820.60 690741.00 108618.60 14715.10 34744.60 15944.20  
## SE Mean 225.06 194.37 67.89 6.05 37.13 1.67  
## LCL Mean 2083.53 1673.44 189.72 31.89 30.37 50.95  
## UCL Mean 2968.97 2438.11 456.82 55.70 176.44 57.51  
## Variance 17019794.92 12693861.87 1548860.26 12309.36 463182.96 817.23  
## Stdev 4125.51 3562.84 1244.53 110.95 680.58 28.59  
## Skewness 1.97 2.29 5.09 3.08 6.60 0.11  
## Kurtosis 3.03 4.75 27.21 9.51 42.69 -1.09  
## acc\_urban acc\_rural perk\_gen perk\_cons  
## nobs 336.00 336.00 336.00 336.00  
## NAs 42.00 42.00 0.00 0.00  
## Minimum 1.90 0.90 10.90 11.40  
## Maximum 100.00 100.00 6120.40 6467.50  
## 1. Quartile 61.90 8.25 87.70 64.17  
## 3. Quartile 91.25 61.10 710.18 949.98  
## Mean 74.93 35.73 830.08 701.54  
## Median 79.05 23.50 337.90 251.10  
## Sum 22029.50 10503.20 278908.00 235716.60  
## SE Mean 1.26 1.96 73.91 57.87  
## LCL Mean 72.46 31.87 684.70 587.70  
## UCL Mean 77.40 39.58 975.47 815.37  
## Variance 463.65 1126.55 1835357.03 1125268.18  
## Stdev 21.53 33.56 1354.75 1060.79  
## Skewness -1.12 0.87 2.49 2.65  
## Kurtosis 1.25 -0.72 5.63 8.36

?tibble

## starting httpd help server ... done

st(tibble(naeedata1))

Summary Statistics

Variable

N

Mean

Std. Dev.

Min

Pctl. 25

Pctl. 75

Max

year

336

2016.5

2.295

2013

2014.75

2018.25

2020

in\_tre

336

696.725

1287.778

0

8.925

742.1

9638.7

in\_hyd

335

449.798

714.751

0

0

701.7

3700.9

in\_wind

336

92.078

334.417

0

0

1.3

2636

in\_solar

336

113.696

566.838

0

1.15

22.7

5989.6

in\_bio

336

28.287

55.325

0

0

33.8

265.7

in\_geo

336

14.204

96.162

0

0

0

823.8

exp

336

821.346

3071.526

0

0

42.825

16894.8

imp

336

985.904

2576.23

0

0

508.325

16369

cons

336

13821.592

38644.065

12.5

363.6

8249.75

208913.7

tgen

336

17565.653

49020.7

28.5

406.225

9723.8

278140.8

gen\_foss

336

14690.549

44923.659

0.7

220.425

4053.1

247301

gen\_nuc

336

348.851

2243.843

0

0

0

15814.4

gen\_ren

336

2526.252

4125.505

0

11.175

3189.575

16574.2

gen\_hyd

336

2055.777

3562.845

0

0.9

2597

16398.2

gen\_solar

336

323.27

1244.532

0

0.2

29.1

8792.8

gen\_bio

336

43.795

110.948

0

0

12.25

625.5

gen\_geo

336

103.407

680.575

0

0

0

5343

access

294

54.232

28.587

3.6

32.075

76.325

100

acc\_urban

294

74.93

21.532

1.9

61.9

91.25

100

acc\_rural

294

35.725

33.564

0.9

8.25

61.1

100

perk\_gen

336

830.083

1354.753

10.9

87.7

710.175

6120.4

perk\_cons

336

701.538

1060.787

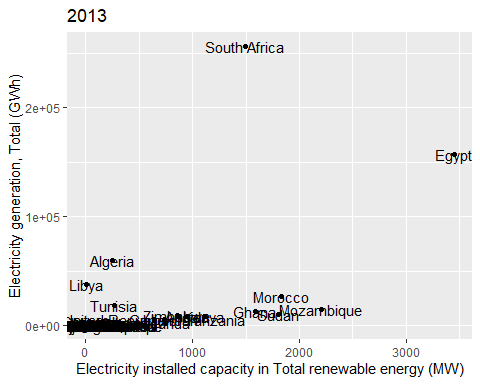
11.4

64.175

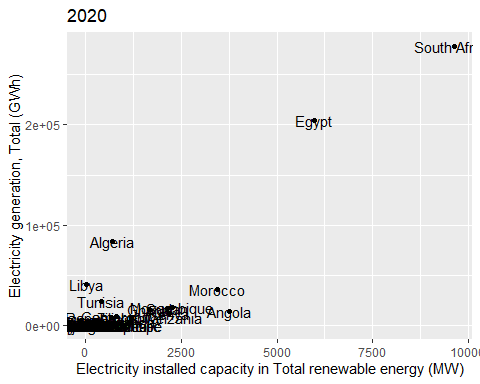
949.975

6467.5

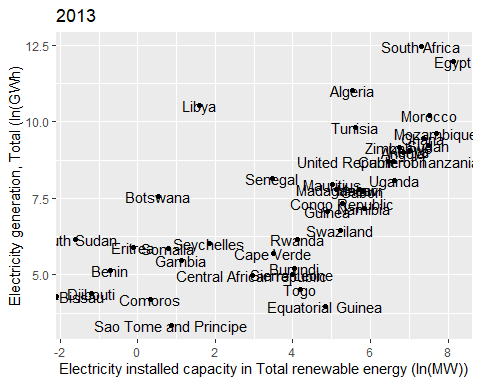
naeedata1 %>%   
 filter(year == "2013") %>%  
 ggplot(aes(x = in\_tre, y = tgen)) +  
 geom\_point()+  
 geom\_text(aes(label = Country), size = 4)+  
 labs(title = "2013",  
 x="Electricity installed capacity in Total renewable energy (MW)",  
 y="Electricity generation, Total (GWh)")



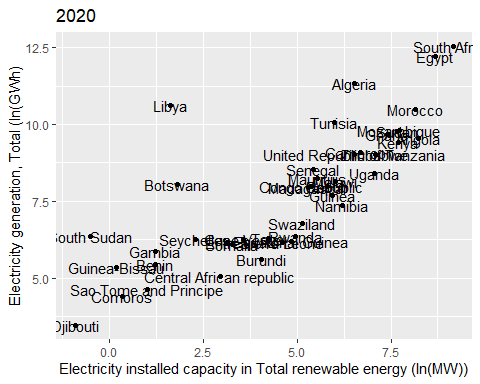
naeedata1 %>%   
 filter(year == "2020") %>%  
 ggplot(aes(x = in\_tre, y = tgen)) +  
 geom\_point()+  
 geom\_text(aes(label = Country), size = 4)+  
 labs(title = "2020",  
 x="Electricity installed capacity in Total renewable energy (MW)",  
 y="Electricity generation, Total (GWh)")



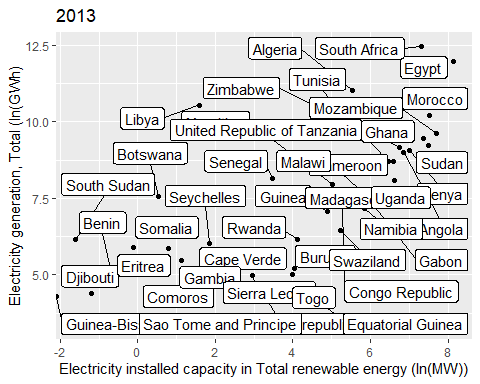
naeedata1 %>%   
 filter(year == "2013") %>%  
 mutate(ln\_intre=log(in\_tre)) %>%   
 mutate(ln\_tgen=log(tgen)) %>%   
 ggplot(aes(x = ln\_intre, y = ln\_tgen)) +  
 geom\_point()+  
 geom\_text(aes(label = Country), size = 4)+  
 labs(title = "2013",  
 x="Electricity installed capacity in Total renewable energy (ln(MW))",  
 y="Electricity generation, Total (ln(GWh)")



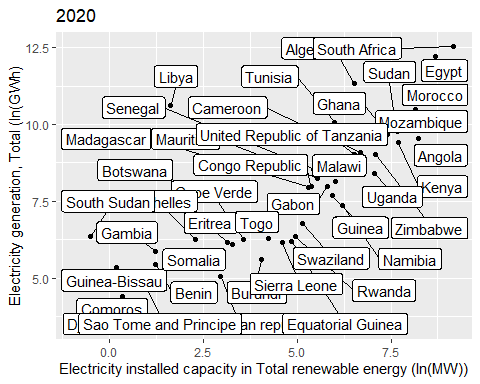
naeedata1 %>%   
 filter(year == "2020") %>%  
 mutate(ln\_intre=log(in\_tre)) %>%   
 mutate(ln\_tgen=log(tgen)) %>%   
 ggplot(aes(x = ln\_intre, y = ln\_tgen)) +  
 geom\_point()+  
 geom\_text(aes(label = Country), size = 4)+  
 labs(title = "2020",  
 x="Electricity installed capacity in Total renewable energy (ln(MW))",  
 y="Electricity generation, Total (ln(GWh)")



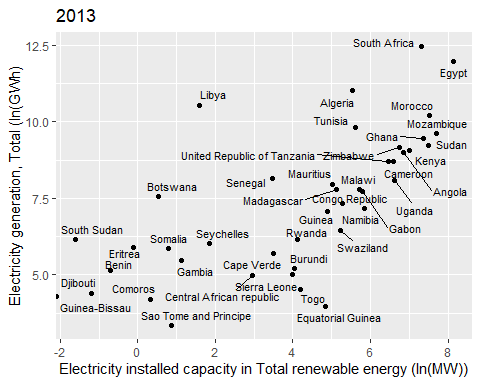
naeedata1 %>%   
 filter(year == "2013") %>%   
 mutate(ln\_intre=log(in\_tre)) %>%   
 mutate(ln\_tgen=log(tgen)) %>%   
 ggplot(aes(x = ln\_intre, y = ln\_tgen)) +  
 geom\_point()+  
 geom\_label\_repel(aes(label = Country), size = 4, max.overlaps = Inf)+  
 labs(title = "2013",  
 x="Electricity installed capacity in Total renewable energy (ln(MW))",  
 y="Electricity generation, Total (ln(GWh)")



naeedata1 %>%   
 filter(year == "2020") %>%   
 mutate(ln\_intre=log(in\_tre)) %>%   
 mutate(ln\_tgen=log(tgen)) %>%   
 ggplot(aes(x = ln\_intre, y = ln\_tgen)) +  
 geom\_point()+  
 geom\_label\_repel(aes(label = Country), size = 4, max.overlaps = Inf)+  
 labs(title = "2020",  
 x="Electricity installed capacity in Total renewable energy (ln(MW))",  
 y="Electricity generation, Total (ln(GWh)")



naeedata1 %>%   
 filter(year == "2013") %>%   
 mutate(ln\_intre=log(in\_tre)) %>%   
 mutate(ln\_tgen=log(tgen)) %>%   
 ggplot(aes(x = ln\_intre, y = ln\_tgen)) +  
 geom\_point() +  
 geom\_text\_repel(aes(label = Country), size = 3, max.overlaps = Inf)+  
 labs(title = "2013",  
 x="Electricity installed capacity in Total renewable energy (ln(MW))",  
 y="Electricity generation, Total (ln(GWh)")



naeedata1 %>%   
 filter(year == "2020") %>%   
 mutate(ln\_intre=log(in\_tre)) %>%   
 mutate(ln\_tgen=log(tgen)) %>%   
 ggplot(aes(x = ln\_intre, y = ln\_tgen, col = Country)) +  
 geom\_point(aes(color= Country)) +  
 geom\_label\_repel(aes(label = Country), size = 3, max.overlaps = Inf)+  
 labs(title = "2020",  
 x="Electricity installed capacity in Total renewable energy (ln(MW))",  
 y="Electricity generation, Total (ln(GWh)") +  
 scale\_color\_discrete() +  
 theme(legend.position = "none")

