Renewable Energy Trends Across Countries

Final Project

Roy Urbina

Introduction

The United States doesn't have an investment problem. Every administration has focused on at least one major fiscal policy that pumps billions of dollars into the economy, whether into specific markets or through direct cash inflows to the American people. Despite a turbulent political climate and a deeply embedded culture of fossil fuel reliance, the U.S. is not allergic to investing in renewable energy infrastructure and production. In fact, the federal government has implemented some form of tax policy targeting renewable energy since the Energy Tax Act of 1978. With the most recent extensions of the Investment Tax Credit (ITC) and Production Tax Credit (PTC), investment in this sector has remained remarkably consistent across decades and party lines.

Tax policy alone is not solely about generating revenue. In the context of renewable energy, it often serves a strategic purpose, offering short-term credits to stimulate long-term capital investments. The logic is clear, lower the cost of entry for private companies now, and enjoy broader economic and environmental benefits in the future. However, if the formula were as simple as x > y = investment, where x is the value of credits and incentives received by private companies and y is the tax revenue foregone by the government, then the United States should be on par with its international counterparts in terms of renewable energy growth. Yet, this is not the case. The inconsistency and temporary nature of these credits often create more uncertainty than stability, weakening their long-term impact.

While the U.S. has made significant strides in renewable energy development, these gains are often undermined by legislative uncertainty. The ITC and PTC, though effective, are frequently subject to expiration, renewal, or modification, creating unstable conditions for long-term project planning. In contrast, international counterparts such as Germany, Denmark, and Finland offer more predictable and durable policy environments through mechanisms like feed-in tariffs and carbon taxation. These models not only stimulate public and private investment but also reflect a broader cultural and political commitment to sustainability.

The above is the bases for my semester reasearch project in a different course. That analysis argued, in my opinion successfully, that while U.S. tax incentives have accelerated renewable energy growth they suffer from instability that undermines their long-term effectiveness. In contrast, international policies provide greater consistency and have successfully fostered sustained investment through mutual and popular support. This data set, while no all inclusive of the specific countries that I studied, will help see if these relationships are present.

So simply, **How well can a countries policy, infrastructure, and economic environment predict annual installed capacity?**

```r
library(tidyverse)
library(janitor)
library(knitr)
library(scales)
library(dplyr)
library(ranger)
library(ggplot2)
library(ggfortify)
library(dplyr)
library(mgcv)
```

```r
renewable <- read.csv("complete_renewable_energy_dataset.csv")

renewable <- renewable %>%
  clean_names()


renewable <- renewable %>%
  filter(year >= 2000 & year <= 2023) %>%
  mutate(
    YearGroup = case_when(
      year >= 2000 & year <= 2004 ~ "2000-2004",
      year >= 2005 & year <= 2009 ~ "2005-2009",
      year >= 2010 & year <= 2014 ~ "2010-2014",
      year >= 2015 & year <= 2019 ~ "2015-2019",
      year >= 2020 & year <= 2023 ~ "2020-2023"
    )
  )

numeric_cols <- names(renewable)[sapply(renewable, is.numeric)]
numeric_cols <- setdiff(numeric_cols, "year")

summary_list <- list()

for (col in numeric_cols) {
```

```
  result <- renewable %>%
    group_by(country, YearGroup, energy_type) %>%
    summarise(AverageValue = mean(.data[[col]], na.rm = TRUE),
              .groups = 'drop') %>%
    mutate(Metric = col)

  summary_list[[col]] <- result
}

renew2 <- bind_rows(summary_list)

renew2 <- renew2 %>%
  pivot_wider(
    names_from = Metric,
    values_from = AverageValue
  )

#renewable$gdp <- comma(renewable$gdp)
```

The data set used, has over 50 input variables, excluding our predicting variable - installed capacity. Because of this, I will approach my research question in two ways: (1) First I will use the PCA model on a broad and specified variable and then (2) follow my learning outcomes with a generalized additive model.

Grouping by YearGroup (such as 2000–2004, 2005–2009) helps smooth out year-to-year fluctuations and highlight clearer, more consistent trends in the data. This not only makes the dataset easier to understand and compare, especially across countries with missing years, but also aligns with how governments and organizations typically plan and report in 5-year intervals. In the context of PCA, grouping by YearGroup reduces noise and the total number of observations, helping PCA capture broader patterns rather than short-term variability. This leads to more meaningful insights when analyzing energy trends over time.

Explaining Variables

My research relies on the key fact that we can measure political stability within a government, plus the outcomes that come from political stability to include investment and research grants. For this reason, I will create a subset of all 50 variables to only include those that I can justify having some element that aid or helps political and economic stability.

```
rn_rq <- renew2 %>%
  select(country, YearGroup, energy_type, gdp, electricity_prices, energy_consumption, gover
```

```r
descriptions <- c(
  gdp = "Gross Domestic Product, an indicator of economic output",
  electricity_prices = "Average electricity price paid by consumers or industry",
  energy_consumption = "Total national energy usage",
  government_policies = "Coded indicator of renewable-related policy presence",
  renewable_energy_targets = "Whether the country has set specific renewable goals",
  energy_subsidies = "Financial support for energy (renewable or fossil)",
  energy_market_liberalization = "Status of deregulated energy markets",
  public_private_partnerships_in_energy = "Level of collaboration in energy projects",
  political_stability = "Risk of political instability",
  regulatory_quality = "Effectiveness of regulations for sustainability",
  corruption_perception_index = "Perceived corruption in the public sector",
  educational_level = "National average education level",
  renewable_energy_patents = "Number of patents related to renewable energy",
  technology_transfer_agreements = "Agreements to share clean energy technologies",
  number_of_research_institutions = "Institutions working on energy/environment R&D",
  grid_integration_capability = "Grid's ability to accept renewable energy",
  energy_storage_capacity = "Installed capacity for energy storage",
  solar_irradiance = "Average solar radiation potential",
  hydro_potential = "Total theoretical hydroelectric capacity"
)

rn_rq_table <- data.frame(
  "Variable Name" = names(descriptions),
  Description = unname(descriptions),
  stringsAsFactors = FALSE
)
```

```r
 kable(rn_rq_table,
       caption =
         "Variable Descriptions for Renewable Energy Analysis")
```

```
Warning in attr(x, "align"): 'xfun::attr()' is deprecated.
Use 'xfun::attr2()' instead.
See help("Deprecated")
```

```
Warning in attr(x, "format"): 'xfun::attr()' is deprecated.
Use 'xfun::attr2()' instead.
See help("Deprecated")
```

Table 1: Variable Descriptions for Renewable Energy Analysis

Variable.Name	Description
gdp	Gross Domestic Product, an indicator of economic output
electricity_prices	Average electricity price paid by consumers or industry
energy_consumption	Total national energy usage
government_policies	Coded indicator of renewable-related policy presence
renewable_energy_targets	Whether the country has set specific renewable goals
energy_subsidies	Financial support for energy (renewable or fossil)
energy_market_liberalization	Status of deregulated energy markets
public_private_partnerships_in_energy	Level of collaboration in energy projects
political_stability	Risk of political instability
regulatory_quality	Effectiveness of regulations for sustainability
corruption_perception_index	Perceived corruption in the public sector
educational_level	National average education level
renewable_energy_patents	Number of patents related to renewable energy
technology_transfer_agreements	Agreements to share clean energy technologies
number_of_research_institutions	Institutions working on energy/environment R&D
grid_integration_capability	Grid's ability to accept renewable energy
energy_storage_capacity	Installed capacity for energy storage
solar_irradiance	Average solar radiation potential
hydro_potential	Total theoretical hydroelectric capacity

In Class Learning Model

PCA With All Variables

What is PCA and why use? PCA, Principle Component Analysis, is a form of analysis that helps simplify large data sets with numbers ranging on both extremes. Having too many variables makes simply analysis harder and basic statistical analysis on a cheap MacBook a tedious process. By using PCA, we compress all the information into one summary variable. The summary variable represents the biggest factors that influence the entire data set, so it captures most of what matters. However, 1 variable cannot capture everything, hence we can create multiple PCA columns, with each subsequent PCA variable capturing the left over information. This means that the first component is the most important and shows the strongest pattern with the data, the second component shows the second strongest pattern and so on.

```r
renewable_subset <- renew2 %>%
  select(-c(country, YearGroup, energy_type, installed_capacity_mw)) %>%
  mutate_at(1:52, as.numeric)

pca_renewables <- prcomp(renewable_subset, scale. = TRUE)
sort(pca_renewables$rotation[, "PC1"])
```

```
  renewable_energy_education_programs        number_of_research_institutions
                         -0.252808253                           -0.210072418
      international_aid_for_renewables                energy_storage_capacity
                         -0.178770169                           -0.161573076
          local_manufacturing_capacity    proportion_of_energy_from_renewables
                         -0.159756147                           -0.159238210
                        energy_exports                      political_stability
                         -0.157192137                           -0.150576386
                  geothermal_potential                           co2_emissions
                         -0.147410244                           -0.127116409
 regional_renewable_energy_cooperation                       natural_disasters
                         -0.104662782                           -0.104496037
                 renewable_energy_jobs                          energy_imports
                         -0.099801126                           -0.086080824
                            population              average_annual_temperature
                         -0.066893724                           -0.064686894
                    regulatory_quality        import_tariffs_on_energy_equipment
                         -0.063815975                           -0.040301006
            corruption_perception_index                         investments_usd
                         -0.035933503                           -0.019395513
            energy_efficiency_programs                     government_policies
                         -0.016162952                           -0.014987167
                  biomass_availability                        solar_irradiance
                         -0.014118872                           -0.001017639
                      innovation_index                   control_of_corruption
                          0.008308588                            0.011076753
                            wind_speed                         production_g_wh
                          0.033700105                            0.038714920
                     energy_subsidies                       electricity_prices
                          0.045780454                            0.066114992
           grid_integration_capability                 renewable_energy_patents
                          0.078234980                            0.081816699
             energy_sector_workforce                                       gdp
                          0.082776257                            0.087229878
               industrialization_rate  number_of_renewable_energy_conferences
```

```
                             0.088189103                              0.093658421
public_private_partnerships_in_energy                                 rule_of_law
                             0.093873947                              0.098965665
                         hydro_potential                          annual_rainfall
                             0.103555561                              0.110287846
                    ease_of_doing_business          energy_market_liberalization
                             0.126125413                              0.155801591
  export_incentives_for_energy_equipment                       energy_consumption
                             0.162724542                              0.169833832
               renewable_energy_targets number_of_renewable_energy_publications
                             0.180899013                              0.224352943
                        r_d_expenditure                        urbanization_rate
                             0.232730120                              0.234446142
         technology_transfer_agreements                    economic_freedom_index
                             0.238650454                              0.244532808
                       educational_level                        public_awareness
                             0.266378132                              0.267915505
```

```r
pca_renewables$x[,"PC1"]
```

```
  [1]  0.152405847 -0.476891353 -0.262356547 -1.827293273  0.617197976
  [6]  0.488623276 -0.728144002 -0.766274904  0.547308962 -1.290950547
 [11]  1.319706989 -0.115385990 -1.243125195 -0.738490186 -0.305214147
 [16] -0.052288111 -2.897092764 -0.232290942 -3.869580548  5.624552994
 [21]  2.183194312 -3.389044366  0.385632313  2.231504578 -0.163088278
 [26] -0.438872857 -1.033856398  0.974799277  2.285316838 -1.995401560
 [31] -0.772240840 -0.634890288 -1.684002088 -1.029055132 -2.873481911
 [36] -0.011782404  1.886130161  0.817640995 -0.740591274 -1.424470655
 [41]  0.769148459 -0.357772850  0.084856771  1.054958745  0.159422362
 [46]  0.067740867 -0.853229631  2.791897320 -0.695674640 -0.702635351
 [51]  1.153720547  0.349749797 -3.231356814  2.100816405 -1.321000861
 [56] -1.472288156 -0.359742242 -2.594676475 -0.926710654  1.274633192
 [61] -1.915846148 -2.362431402 -0.930671862 -1.042748358  1.223937277
 [66] -1.251020172 -1.190345536 -0.008291145  1.314130094  0.162486142
 [71] -1.040573603  0.337652523 -0.268981756 -0.273035321 -0.935524990
 [76] -0.564611260 -1.245586972  1.243428386 -1.176300996  0.003241533
 [81]  0.465039884  0.410546167 -2.006228778  0.897744390 -1.023030629
 [86]  2.611396731 -0.574446262 -1.464992664  0.892532699  0.684507456
 [91]  0.055206224 -0.137867940 -0.267645949  0.732835201  0.044111712
 [96] -0.128596890 -1.535776787  0.808140912 -0.044681311 -0.313678055
[101] -0.055649513 -1.533250347  2.066093922  0.483057962  0.472906791
[106]  0.205415180  0.459848959  0.868994269 -0.843059465 -1.742289536
```

```
[111]  0.050426842 -1.144957990  0.493314151 -0.947192424  1.103277448
[116] -0.404005418  0.740283810 -0.014713533  0.130043828 -0.001101851
[121] -0.092801553 -0.092443571  1.146500178 -1.039762363  1.385615094
[126] -0.135206347  0.074729664  0.544747168 -0.615397493 -0.707378692
[131] -0.226468014  0.524129222 -0.289585575 -0.252085136  0.455222994
[136]  2.261003456  1.183622915  0.607180463  0.234127972 -3.076197726
[141]  1.421832191 -0.287291380  0.632057498 -0.976129439 -0.084269963
[146]  2.149127041  1.225743151  0.166824082  2.063575449 -1.142975325
[151] -1.382172211 -0.510677295 -1.696565269  1.466049388 -0.460319340
[156] -0.112726658  1.648675553 -2.052008543  0.747257291 -0.046883369
[161]  0.121668958  0.180721218 -1.376178954 -2.132884044  0.750241988
[166]  0.529621121  4.214059201 -0.585964524  0.738361402 -0.765649976
[171] -1.040508706 -2.573429910  0.721974273  4.989442385 -0.432960767
[176]  1.823585786  0.206773792 -0.101642697  0.549787582  0.715486270
[181] -0.585276247  1.017543160 -0.729784453 -0.879992796  0.326577449
[186]  1.310084596  0.718576331 -1.153818482  2.654205256  0.055055816
[191] -0.962098239 -0.299280550 -1.962903441  0.231571313  1.299456793
[196]  2.354051945 -1.797454614  0.483937118 -1.737365862 -4.277517085
[201]  0.913725526  4.361570895  3.932024915  1.525319335  1.079670209
[206] -1.657051255 -0.646603622 -0.506230549 -1.885123929 -3.034752531
[211]  1.193743337 -1.815007548  0.982570513  0.814398580  1.822168146
[216]  1.667597116 -0.558458833 -0.834626663 -1.095003267  0.233205635
[221] -0.637539839  0.066533951 -0.168141056 -0.231526263  4.287882979
[226] -0.048847292 -0.342609724 -2.052948275 -0.126770726  0.963714640
[231]  0.917488996 -0.634075013 -0.404692210  2.700323499  2.733755921
[236]  4.315004462  0.024132517 -1.065397747 -0.360702863 -0.550253990
[241] -0.702358673 -0.714002847  3.527451337  1.032263103  2.325124268
[246] -0.302607275 -1.998530535 -2.298584059  0.957026944 -1.342178528
```

```r
renewable_w_PC <- renew2 %>%
  mutate(score_pc1 = pca_renewables$x[, "PC1"])

sort(pca_renewables$rotation[,"PC2"])
```

```
                     solar_irradiance         local_manufacturing_capacity
                         -0.400689323                         -0.284601866
number_of_renewable_energy_conferences                         co2_emissions
                         -0.253285777                         -0.243492470
                       production_g_wh           average_annual_temperature
                         -0.222729969                         -0.207426277
                  renewable_energy_jobs       technology_transfer_agreements
                         -0.162815524                         -0.158607931
```

number_of_research_institutions	import_tariffs_on_energy_equipment
-0.157095900	-0.148235188
biomass_availability	export_incentives_for_energy_equipment
-0.137933990	-0.131429552
annual_rainfall	energy_efficiency_programs
-0.125790075	-0.123596799
educational_level	public_private_partnerships_in_energy
-0.121384060	-0.118258687
energy_consumption	grid_integration_capability
-0.097840905	-0.097470141
number_of_renewable_energy_publications	political_stability
-0.061796826	-0.050885107
energy_subsidies	energy_imports
-0.039075835	-0.034950100
regional_renewable_energy_cooperation	wind_speed
-0.034687632	-0.029940801
economic_freedom_index	natural_disasters
-0.025520014	-0.014915792
rule_of_law	control_of_corruption
-0.011850571	-0.006023198
population	proportion_of_energy_from_renewables
0.007377477	0.007575403
corruption_perception_index	energy_market_liberalization
0.009773755	0.011839967
urbanization_rate	energy_sector_workforce
0.014284012	0.054015551
electricity_prices	energy_exports
0.059329310	0.062464344
investments_usd	renewable_energy_patents
0.065123137	0.078013188
ease_of_doing_business	government_policies
0.082138066	0.094580513
regulatory_quality	hydro_potential
0.095022889	0.108269044
innovation_index	public_awareness
0.108650202	0.122552682
international_aid_for_renewables	industrialization_rate
0.124938044	0.129176794
geothermal_potential	r_d_expenditure
0.157085089	0.166227862
energy_storage_capacity	gdp
0.178269112	0.184167058
renewable_energy_education_programs	renewable_energy_targets

```
                    0.202129047                  0.202272571
```

```
renewable_w_PC <- renewable_w_PC %>%
  mutate(score_pc2 = pca_renewables$x[, "PC2"])
```

```
pc1_loadings <- round(pca_renewables$rotation[, "PC1"], 3)
pc2_loadings <- round(pca_renewables$rotation[, "PC2"], 3)

pc_table <- data.frame(
  PC1 = pc1_loadings,
  PC2 = pc2_loadings
) %>%
  arrange(desc(abs(PC1))) %>%
  head(10)
```

```
kable(head(pc_table, 10),
      caption = "Top 10 Variables Contributing to Principal Component 1 and 2",
      align = "lc")
```

Table 2: Top 10 Variables Contributing to Principal Component 1 and 2

	PC1	PC2
public_awareness	0.268	0.123
educational_level	0.266	-0.121
renewable_energy_education_programs	-0.253	0.202
economic_freedom_index	0.245	-0.026
technology_transfer_agreements	0.239	-0.159
urbanization_rate	0.234	0.014
r_d_expenditure	0.233	0.166
number_of_renewable_energy_publications	0.224	-0.062
number_of_research_institutions	-0.210	-0.157
renewable_energy_targets	0.181	0.202

Lets look at the table above:
Although the principal components from the PCA do not directly measure a country's renewable energy output, they uncover key themes that influence how effectively a country can expand its installed renewable energy capacity. The first principal component (PC1) is primarily driven by factors like public awareness,educational level,technology transfer agreements,urbanization rate, and R&D expenditure. These loadings suggest that PC1 captures a theme of national capacity-building and public readiness. In other words, countries that score

highly on PC1 tend to have populations that are more educated, more urbanized, and more aware of renewable energy issues (an assumption), paired with governments and institutions that are actively investing in research, innovation, and technology sharing. This mix reflects a broader knowledge-driven, innovation-consistent environment, which is often a prerequisite for developing renewable infrastructure effectively and sustainable.

The second principal component (PC2), meanwhile, shows a different kind of dynamic. It contrasts variables like renewable energy education programs and renewable energy targets (positive loadings) against others such as educational level and technology transfer agreements (negative loadings). This suggests that PC2 may represent a policy-driven versus institutional-capacity divide. Countries with high PC2 scores might be focusing more on government-led renewable energy programs and formal targets, possibly compensating for weaker existing institutional or educational infrastructure. Contrastly, countries with low PC2 scores may already have stronger baseline capacity, via education and technical projects, but rely less on strong government initiatives to drive change.

PCA With Selected Variables

```
rn_rq_subset <- rn_rq %>%
  select(-c(country, YearGroup, energy_type)) %>%
  mutate_at(1:19, as.numeric)

pca_rn_rq <- prcomp(rn_rq_subset, scale. = TRUE)
sort(pca_rn_rq$rotation[, "PC1"])
```

```
                solar_irradiance                political_stability
                     -0.39381280                        -0.35235404
      grid_integration_capability public_private_partnerships_in_energy
                     -0.27478860                        -0.26564303
                 energy_subsidies      number_of_research_institutions
                     -0.24127805                        -0.11437059
               regulatory_quality                 electricity_prices
                     -0.10493030                        -0.06383595
      corruption_perception_index       technology_transfer_agreements
                     -0.06142294                        -0.03748321
                  hydro_potential                                gdp
                      0.04858494                         0.05854665
              government_policies            energy_storage_capacity
                      0.11156789                         0.15389767
     energy_market_liberalization                 energy_consumption
                      0.16559690                         0.27065245
```

```
              educational_level          renewable_energy_targets
                     0.27569615                        0.29694620
      renewable_energy_patents
                     0.41993289
```

```r
pca_rn_rq$x[,"PC1"]
```

```
  [1]  1.055316554 -0.135415192  0.244085822 -2.640283623  2.788119533
  [6]  0.659518016 -0.606832330  0.151878018 -1.493310597  1.078086238
 [11] -0.668186820  2.026520579 -0.079716782 -2.512822430 -0.063027806
 [16] -1.433774862 -1.820270996 -0.211281247  1.284548272 -0.562911609
 [21] -1.644656630 -1.588237781  2.743725553  0.602897651  1.016446222
 [26] -0.845015096 -2.162322246 -1.911871119  1.810794871 -2.453773260
 [31] -2.430431387 -1.144146879 -1.736862905 -1.990192304 -1.433459075
 [36]  0.615089471  1.101121961 -0.917957025  2.652002303 -0.388205806
 [41] -1.610645375 -0.961925488  0.356652877  0.959193434 -1.889461831
 [46] -0.165435717 -1.129373435  2.703102407  0.162918144 -1.244467937
 [51]  1.190780281  0.233373589 -2.855010832  0.681850498 -1.902176040
 [56] -1.048996694 -1.008612648 -2.054668077 -0.250646248 -0.909809574
 [61] -0.200009416 -1.062848148  0.346372763  0.500066424 -0.276304876
 [66] -0.265673601  0.197571506 -0.229921394  1.156524130  1.559094202
 [71] -1.871374328  0.206313841 -1.426458434  1.447011615  0.442886173
 [76] -0.643355625  0.383701285  0.604321862  1.205979232 -0.684209852
 [81]  0.188951985  0.922135278 -1.489772873  0.890763542 -1.038818758
 [86]  0.470918957 -1.365000727 -0.530441925  1.116322133 -0.541398649
 [91]  0.784448010 -0.504458826 -0.296805462  1.030641530 -1.233043443
 [96]  0.878847403 -0.238570561  0.062892489 -4.214202477 -0.445487407
[101]  0.352100107 -0.439173804 -0.743938644 -0.135520212  0.440753755
[106]  1.649612864  0.036170460 -0.175052859 -0.895667417  1.199441146
[111]  0.720403715 -0.084035884 -0.231165200 -0.691233479 -1.042011783
[116]  0.702341661  0.019988690  0.133833946 -0.138859054  1.299983707
[121]  0.925014366  0.166392364  1.339700132  0.193564903 -1.370194195
[126] -0.384263759 -1.773511219 -0.302927764 -0.422435730  0.098135631
[131] -0.464548267 -0.294548218  1.281521103 -0.598476236  0.096307793
[136] -0.623286848  1.159272301  0.915750170  0.483649971 -1.903333324
[141] -0.619495423 -0.600452948 -0.491722525  0.324268394 -0.172509120
[146] -1.839538266  1.211321429  0.300596443  1.234259717  0.140198031
[151] -3.491170907  0.188860016  3.115813669  1.209480877 -0.137369573
[156]  0.522871009 -0.789129651 -1.345948898 -0.251908465  0.508161269
[161] -0.520073858  1.168595208  0.963840738 -0.348233298 -0.356908263
[166]  0.708252505  2.391124806 -0.423214254 -0.154987863 -1.036522343
[171]  0.787149782 -0.152558093 -0.817047528  1.589525780  0.161703813
```

```
[176]  2.250171613 -0.402493740  1.630755655 -0.616268235 -0.170279362
[181]  0.438107685  0.996542421 -0.170748565 -0.465074139  1.458303566
[186] -0.338669012 -0.496392045  2.288232168  1.576912049  0.131551528
[191] -1.863402997  2.009924481 -2.726506381  0.585399050  0.524667931
[196]  1.043672227  1.658994938  1.880532014  1.597054347 -0.073370219
[201]  0.782155245 -1.981791484 -0.168079224  0.037205420  1.712555299
[206] -0.191511106 -0.820863313  0.403062973 -0.691812083 -1.567211944
[211]  2.029072100 -0.662972247  0.001886527 -0.437308491  0.580075516
[216] -0.085416055 -0.404287973  0.046553200  1.493711640  1.942372703
[221] -0.130517149 -0.597201300  0.863549420 -0.971094614 -1.010622051
[226] -0.613658840 -0.773893550 -0.324070911 -0.107045696  0.792763661
[231] -1.475284605 -0.076182069  0.964013244 -0.469126741  1.667891923
[236] -1.707100517  1.303161457 -0.026519443 -0.098794704  1.058045299
[241] -0.377562009  1.085928834  3.018574065  0.582686327  7.170968793
[246]  0.415766775  0.610082444 -0.519862220 -0.210539426 -0.335791354
```

```r
rn_rq_w_PC <- renew2 %>%
  mutate(score_pc1 = pca_rn_rq$x[, "PC1"])

sort(pca_rn_rq$rotation[,"PC2"])
```

```
                     educational_level                      solar_irradiance
                           -0.44776349                           -0.34768602
public_private_partnerships_in_energy                    energy_consumption
                           -0.26264534                           -0.23445932
       technology_transfer_agreements                    electricity_prices
                           -0.22590684                           -0.17753111
            renewable_energy_targets                      energy_subsidies
                           -0.16319926                           -0.14596595
         grid_integration_capability          energy_market_liberalization
                           -0.13487100                           -0.07922359
            renewable_energy_patents           corruption_perception_index
                           -0.05827924                           -0.03934628
                                   gdp                       hydro_potential
                            0.09846109                            0.11692957
     number_of_research_institutions                   government_policies
                            0.14235167                            0.18052577
                   regulatory_quality                   political_stability
                            0.22624529                            0.25311709
              energy_storage_capacity
                            0.45087723
```

```r
rn_rq_w_PC <- rn_rq_w_PC %>%
  mutate(score_pc2 = pca_rn_rq$x[, "PC2"])


pc1_loadings_rn_rq <- round(pca_rn_rq$rotation[, "PC1"], 3)
pc2_loadings_rn_rq <- round(pca_rn_rq$rotation[, "PC2"], 3)

pc_rn_rq_table <- data.frame(
  PC1 = pc1_loadings_rn_rq,
  PC2 = pc2_loadings_rn_rq
) %>%
  arrange(desc(abs(PC1))) %>%
  head(10)
```

```r
kable(head(pc_rn_rq_table, 10),
      caption = "Subset Top 10 Variables Contributing to
      Principal Component 1 and 2",
      align = "lc")
```

Table 3: Subset Top 10 Variables Contributing to Principal Component 1 and 2

	PC1	PC2
renewable_energy_patents	0.420	-0.058
solar_irradiance	-0.394	-0.348
political_stability	-0.352	0.253
renewable_energy_targets	0.297	-0.163
educational_level	0.276	-0.448
grid_integration_capability	-0.275	-0.135
energy_consumption	0.271	-0.234
public_private_partnerships_in_energy	-0.266	-0.263
energy_subsidies	-0.241	-0.146
energy_market_liberalization	0.166	-0.079

Lets take a look at this second PCA Analysis

This PCA analysis reveals two main patterns in how countries support renewable energy. The first pattern (PC1) highlights countries that are strong in innovation and policy commitment, those investing in renewable energy patents, setting ambitious targets, and relying less on natural solar resources. The second pattern (PC2) contrasts countries that benefit from political stability and natural advantages (like high solar irradiance) with those that invest more in education, public-private partnerships, and efficient energy systems. Together, these findings suggest that successful renewable energy development doesn't rely solely on geography or political stability but also on innovation, clear policy goals, and strong institutions.

Generalized Additive Model

A Generalized Additive Model (GAM) is a flexible statistical tool that works similar to linear regression. While traditional regression fits straight lines between variables, GAM allows for smooth, curved relationships that better reflect the complexity of real-world data. This means it can present patterns that wouldn't show up if we assumed everything followed a simple straight line. For this renewable energy analysis, where the goal is to predict a country's installed renewable energy capacity based on factors like policy, infrastructure, and economic conditions, GAM is a clear choice. These factors don't always affect outcomes in a straight-line.

```r
gam_model <- gam(installed_capacity_mw ~
                   s(gdp) +
                   s(electricity_prices) +
                   s(energy_consumption) +
                   s(government_policies) +
                   s(renewable_energy_targets) +
                   s(energy_subsidies) +
                   s(energy_market_liberalization) +
                   s(public_private_partnerships_in_energy) +
                   s(political_stability) +
                   s(regulatory_quality) +
                   s(corruption_perception_index) +
                   s(educational_level) +
                   s(renewable_energy_patents) +
                   s(technology_transfer_agreements) +
                   s(number_of_research_institutions) +
                   s(grid_integration_capability) +
                   s(energy_storage_capacity) +
                   s(solar_irradiance) +
                   s(hydro_potential),
                 data = renew2)

summary(gam_model)
```

```
Family: gaussian
```

```
Link function: identity

Formula:
installed_capacity_mw ~ s(gdp) + s(electricity_prices) + s(energy_consumption) +
    s(government_policies) + s(renewable_energy_targets) + s(energy_subsidies) +
    s(energy_market_liberalization) + s(public_private_partnerships_in_energy) +
    s(political_stability) + s(regulatory_quality) + s(corruption_perception_index) +
    s(educational_level) + s(renewable_energy_patents) + s(technology_transfer_agreements) +
    s(number_of_research_institutions) + s(grid_integration_capability) +
    s(energy_storage_capacity) + s(solar_irradiance) + s(hydro_potential)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  25067.0      260.3   96.29   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
                                          edf Ref.df     F p-value
s(gdp)                                  5.731  6.884 2.814 0.00997 **
s(electricity_prices)                   5.029  6.136 3.404 0.00379 **
s(energy_consumption)                   2.256  2.892 1.482 0.24246
s(government_policies)                  1.000  1.000 0.881 0.34918
s(renewable_energy_targets)             1.000  1.000 3.095 0.08010 .
s(energy_subsidies)                     1.000  1.000 0.795 0.37361
s(energy_market_liberalization)         4.085  5.032 1.056 0.38990
s(public_private_partnerships_in_energy) 1.000  1.000 1.296 0.25638
s(political_stability)                  1.000  1.000 4.049 0.04556 *
s(regulatory_quality)                   1.000  1.000 1.971 0.16189
s(corruption_perception_index)          1.000  1.000 1.151 0.28464
s(educational_level)                    5.475  6.554 1.645 0.15596
s(renewable_energy_patents)             1.000  1.000 0.136 0.71305
s(technology_transfer_agreements)       8.227  8.815 2.341 0.01359 *
s(number_of_research_institutions)      3.500  4.357 0.794 0.53829
s(grid_integration_capability)          4.879  5.974 1.778 0.11533
s(energy_storage_capacity)              1.000  1.000 4.044 0.04570 *
s(solar_irradiance)                     1.026  1.050 1.153 0.29254
s(hydro_potential)                      2.502  3.166 1.224 0.27791
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.228   Deviance explained = 38.8%
GCV = 2.1468e+07  Scale est. = 1.6942e+07  n = 250
```

Explaining GAM Terms

For each individual factor added to the GAM, we get four output variables: edf, Ref.df, F, p-value.

edf - Estimated Degree of Freedom which simply means how curvy the line is. (shapes the line, 1 = linear relationship)

Ref.df - The reference of the degree of freedom helps make sure we are fairly testing each variable to form the line. Ref.df gives us a rounded, more standard number to compare against, which helps make the p-values accurate.

F - how strongly the variable explains variation

p-value - Significance of variable (if less than 0.05 then it is significant)

R-sq. (adj.) - shows the percent of variation explained by the model.

Explaining GAM Results

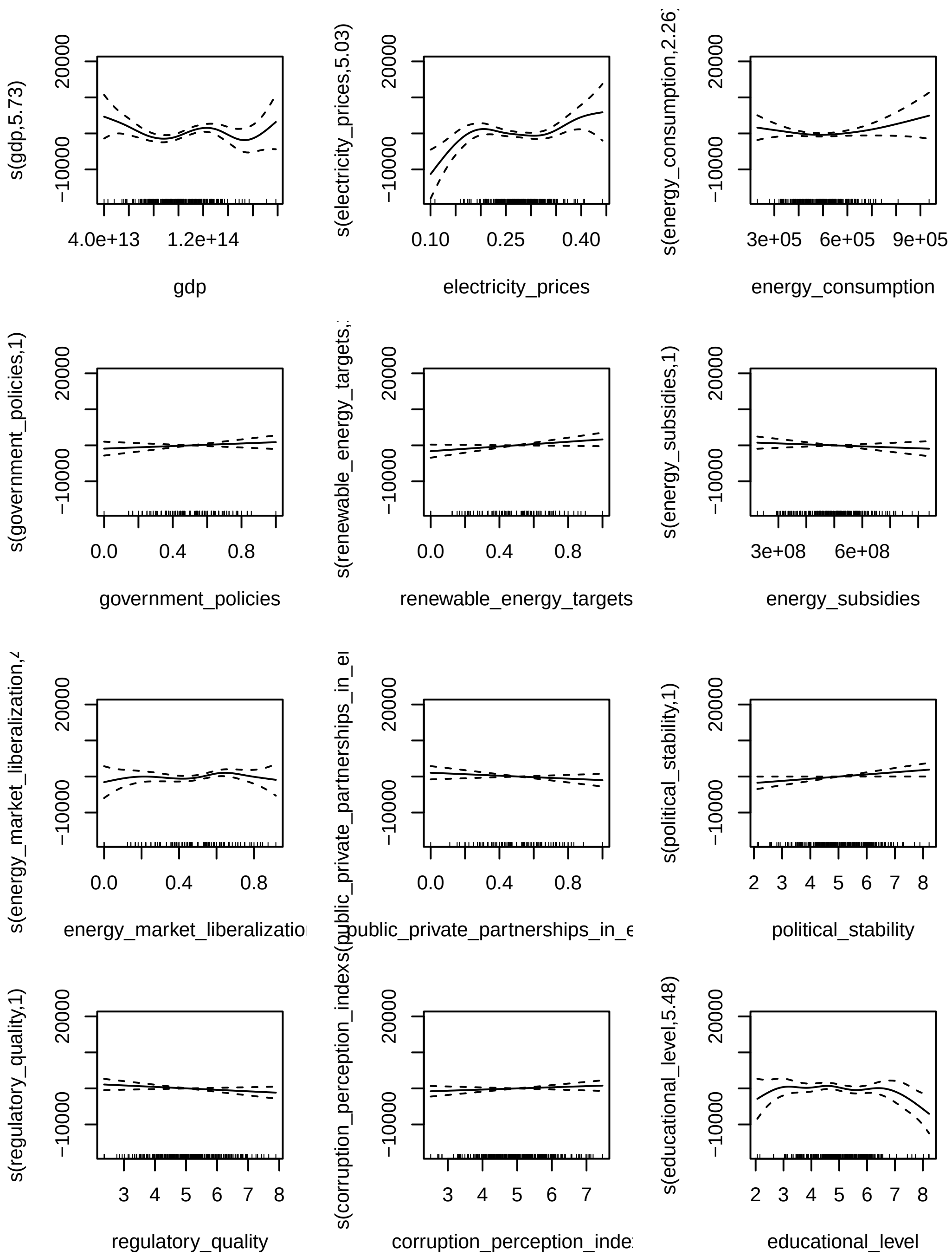
According to the above, a countries GDP, electricity price, and political stability (seen through the p value), have the most significant effect on installed capacity. Countries with higher GDPs and certain electricity price patterns tend to install more renewable energy, but political stability play a role which can be infered to be that tmore stable governments tend to do better in renewable energy expansion.

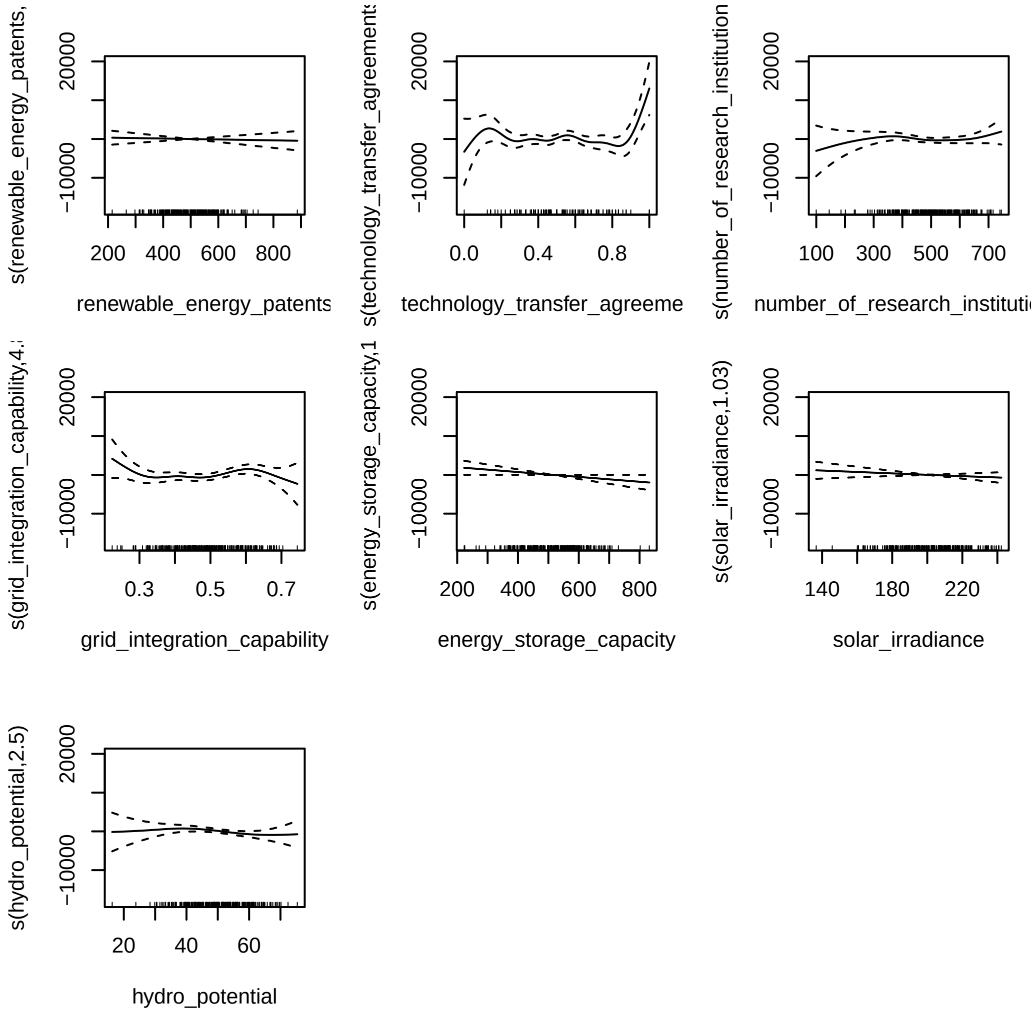
When relating it to our PCA analysis from above, we can see how the principal components align with the influential variables identified in the GAM analysis. While it can be a stretch to say that the first principal component captured much of the variance associated with economic indicators, particularly GDP and electricity prices, and those indicators fall within the category of innovation and policy commitment. A clearer pattern is the second component which incorporated governance and political variables, which supports our interpretation that political stability, although perhaps subtler in its statistical weight, plays a meaningful role in shaping renewable energy outcomes. Together, the PCA reinforces the idea that wealthier and more politically stable countries with moderate to high electricity prices are generally more successful in scaling up renewable energy infrastructure.

Without the background knowledge of renewable investment, it would seem that investing in renewable when also having high energy prices in counter intuitive. Why wouldn't a country simply subsidies energy prices with the investment funds. However, over the last several years, solar energy became cheaper to produce and export to households compared to electricity generated from fossil fuels, so countries and companies with the capital to invest in renewable energy infrastructure are doing so.

The below is a depiction of the GAM analysis and the shaping of the graph (smooth versus linear)

```r
plot(gam_model, pages = 4, shade = TRUE)
```




s(renewable_energy_patents,
20000
−10000
200 400 600 800
renewable_energy_patents
s(technology_transfer_agreements
20000
−10000
0.0 0.4 0.8
technology_transfer_agreeme
s(number_of_research_institution
20000
−10000
100 300 500 700
number_of_research_institutic
s(grid_integration_capability,4.
20000
−10000
0.3 0.5 0.7
grid_integration_capability
s(energy_storage_capacity,1
20000
−10000
200 400 600 800
energy_storage_capacity
s(solar_irradiance,1.03)
20000
−10000
140 180 220
solar_irradiance
s(hydro_potential,2.5)
20000
−10000
20 40 60
hydro_potential

Conclusion

While my introduction relied on having reliable data for the US and European questions, the reality of coming up with a research questions before the data set is that we don't always get what we want.

The data provided can create problems when generalizing my analysis. My analysis focused on separating between developed wealthy countries and less developed countries. However, generally speaking we can categorized all the countries provided within a similar bubble. My counter is that not all countries in the data set were the same across years and 2 decades gives a lot of variation between investments and renewable development.

The second critique is that I did not collapse the data to interpret similar energy types (i.e. solar versus biomass). If I looked at individual levels, I could have found different relationships within a country depending on type, for instance political stability might be important for solar but not as important for biomass. While that level of analysis is appropriate, my counter is that my question focused on general investment in infrastructure rather than a specific subset.

For those reason my research question went from **How well can a countries policy, infrastructure, and economic environment predict annual installed capacity,** to **What relationships are present in terms of policy, infrastructure, and economic environment in a countries total annual installed capacity of renewable energy infrastructure.**

It is difficult to predict total amounts as a quantitative variable, but if we can infer a qualitative variable, as above or below average, then we can see how similar countries are doing in comparison to the international system. The project is the starting steps to see if we can come up with a qualitative formula to make some sort of predictions. For instance if we have two counties with a similar government then we can assume that they are similar in installed capacity, however if one has a higher GDP with average or higher electricity prices combined with some aspect of of higher innovation score; then that country should have a higher rate of installed capacity.