Analysis of US Renewable Energy Electricity Generation Trends

https://github.com/[redacted]/ENVIRON872L-EDA_Final-Project

[redacted]

Abstract

Electricity generation in the United States is dominated by fossil fuel-based technologies.Combustion of fossil fuelsis the main cause of human-driven emission of greenhouse gases (GHGs).Adoption of rebewable energy for electricity generation provides an opportunity to mitigate the negative environmental of fossil fuel fuel adoption. The US electricity generation has more than doubled since 2008. This paper analysed the trend in national and state level electricity generation from major renewable energy sources and found that post 2008 marks a season of significant change in the adoption trend of renewable energy resurces particularly wind and solar. It also found that policy changes could have contributed to this change in renewable energy adoption.

Contents

List of Tables

List of Figures

<Note: set up autoreferencing for figures and tables in your document>

1 Research Question and Rationale

Electricity generation in the United States is dominated by fossil fuel-based technologies. According to the U.S. Energy Information Administration, In 2018 63.5 % (EIA, 2018) of the electricity generated in the country was from fossil fuel based sources such as coal and natural gas. The combustion of fossil fuels, necessary for electricity generation, is the main cause of human-driven emission of greenhouse gases (GHGs) (Quadrelli & Peterson, 2007) particularly carbon dioxide. The alternative use of renewable energy provides an opportunity to decarbonize the electricity sector and mitigate the negative emission impacts. In the US there appears to be a significant shift in electricity generation from fossil fuels to renewable energy in the last decade. EIA estimates that the adoption of renewable energy to generate electricity has doubled from 2008 to 2019 to account for about 17% (EIA, 2018) of the total electricity generation mix.

This paper will aim to:

1. Research the trend in national level electricity generation from renewable energy sources disaggregated by the main renewable sources contributing to the growth in total proportion of renewable energy generation.
2. Investigate the specific trends in wind energy adoption in Kansas and solar energy in California to identify any policy change implications on renewable energy adoption.

Kansas state was selected for the second component of the analysis because wind energy contributes the largest proportion of the state's overall electricity generation (36% as of 2018) compared to other states. Kansas was also selected because in 2009 it adopted a mandatory renewable energy policy known as a Renewable Portfolio Standard (RPS) mandating that the state should generate 20% of its electricity from renewable sources by 2025. In 2015, Kansas repealed this policy and changing the renewable energy goal from mandatory to voluntary (Barbose, 2018). California was selected because solar energy contributes the largest proportion of the state's overall electricity generation (14% as of 2018) compared to other states. California also adopted an RPS policy in 2002. However, unlike Kansas, it is much more ambitious, it mandates that the state will generate 60% of its electricity from renewable sources by 2030 (Barbose, 2018). Also, unlike Kansas, California has made several revisions to its renewable energy policy to strengthen it including in 2006,2007,2011,2015 and 2018 (Barbose, 2018).

The dataset that shall be used for this analysis is the Energy Information Administration (EIA) data on Monthly total electric power generation data by State and energy source from 2001 - 2018 available here

Table 1: Summary table of raw dataset

	YEAR	MONTH	STATE	TYPE.OF.PRODUCER	ENERGY.SOURCE	GENERATION..Megawatthours.
	Min. :2001	Min. : 1.000	CA : 12945	Combined Heat and Power, Commercial Power : 43658	Total : 61824	Min. : -997855
	1st Qu.:2005	1st Qu.: 3.000	MI : 11882	Combined Heat and Power, Electric Power : 40674	Natural Gas : 56233	1st Qu.: 1814
	Median :2010	Median : 6.000	PA : 11102	Combined Heat and Power, Industrial Power : 65061	Petroleum : 53177	Median : 24701
	Mean :2010	Mean : 6.479	NY : 10654	Electric Generators, Electric Utilities : 77347	Coal : 40588	Mean : 1416908
	3rd Qu.:2014	3rd Qu.: 9.000	MN : 10393	Electric Generators, Independent Power Producers: 73257	Other Biomass : 38050	3rd Qu.: 285558
	Max. :2019	Max. :12.000	NC : 10314	Total Electric Power Industry :111800	Hydroelectric Conventional: 32771	Max. :421796659
	NA	NA	(Other):344507	NA	(Other) :129154	NA

2 Dataset Information

The dataset used for this analysis was Energy Information Administration (EIA) data on Monthly total electric power generation in Megawatt Hours (MWh) by State and energy source from 2001 - 2018 obtained from their site. The data contained excel sheets with monthly totals of electricity generation by type of power producer and type of energy source in 51 states in the US including national total from January 2001 to January 2019. The dataset included 14 tyoes of energy sources and 6 types of energy producers:

- Combined Heat and Power, Commercial Power
- Combined Heat and Power, Electric Power
- Combined Heat and Power, Industrial Power
- Electric Generators, Electric Utilities
- Electric Generators, Independent Power Producers
- Total Electric Power Industry

The U.S. Renewable Portfolio Standards 2018 Annual Status Report written by the Lawrence Berkeley National Laboratory was also used as the reerence source to obtain information on the years states, particularly Kansas and California, made major revisions to their RPS policies.

Table 2: First 10 rows of the raw dataset

YEAR	MONTH	STATE	TYPE.OF.PRODUCER	ENERGY.SOURCE	GENERATION..Megawatthours.
2001	1	AK	Total Electric Power Industry	Coal	46903
2001	1	AK	Total Electric Power Industry	Petroleum	71085
2001	1	AK	Total Electric Power Industry	Natural Gas	367521
2001	1	AK	Total Electric Power Industry	Hydroelectric Conventional	104549
2001	1	AK	Total Electric Power Industry	Wind	87
2001	1	AK	Total Electric Power Industry	Total	590145
2001	1	AK	Electric Generators, Electric Utilities	Coal	18410
2001	1	AK	Electric Generators, Electric Utilities	Petroleum	64883
2001	1	AK	Electric Generators, Electric Utilities	Natural Gas	305277
2001	1	AK	Electric Generators, Electric Utilities	Hydroelectric Conventional	104549

3 Exploratory Data Analysis and Wrangling

3.1 Initial data exploration

Explored the raw data set to determine the general structure of the dataset and variables within it as shown in the code below. This exploration also included displaying the first 10 rows as shown in Table 2 and last 10 rows of the dataset as shown in Table 3.

```
#Summary code
#Structure of the dataset
str(EIA.data.2001_2019)
```

```
## 'data.frame':    411797 obs. of  6 variables:
##  $ YEAR                         : int  2001 2001 2001 2001 2001 2001 2001 2001 2001 2001
##  $ MONTH                        : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ STATE                        : Factor w/ 53 levels "AK","AL","AR",..: 1 1 1 1 1 1 1 1
##  $ TYPE.OF.PRODUCER             : Factor w/ 6 levels "Combined Heat and Power, Commercia
##  $ ENERGY.SOURCE                : Factor w/ 14 levels "Coal","Geothermal",..: 1 9 4 3 13
##  $ GENERATION..Megawatthours.: int  46903 71085 367521 104549 87 590145 18410 64883 3
```

```
#Collumn names of the variables
colnames(EIA.data.2001_2019)
```

```
## [1] "YEAR"                        "MONTH"
## [3] "STATE"                       "TYPE.OF.PRODUCER"
## [5] "ENERGY.SOURCE"               "GENERATION..Megawatthours."
```

```
#First 10 rows of the dataset
kable(head(EIA.data.2001_2019, 10), caption = "First 10 rows of the raw dataset") %>%
  kable_styling(latex_options="scale_down")
```

```
#Last 10 rows of the dataset
kable(tail(EIA.data.2001_2019, 10), caption = "Last 10 rows of the raw dataset") %>%
  kable_styling(latex_options="scale_down")
```

Table 3: Last 10 rows of the raw dataset

	YEAR	MONTH	STATE	TYPE.OF.PRODUCER	ENERGY.SOURCE	GENERATION..Megawatthours.
411788	2019	1	WY	Electric Generators, Independent Power Producers	Hydroelectric Conventional	0
411789	2019	1	WY	Electric Generators, Independent Power Producers	Natural Gas	285
411790	2019	1	WY	Electric Generators, Independent Power Producers	Solar Thermal and Photovoltaic	6850
411791	2019	1	WY	Electric Generators, Independent Power Producers	Wind	212257
411792	2019	1	WY	Electric Generators, Electric Utilities	Total	3750439
411793	2019	1	WY	Electric Generators, Electric Utilities	Coal	3492763
411794	2019	1	WY	Electric Generators, Electric Utilities	Hydroelectric Conventional	81780
411795	2019	1	WY	Electric Generators, Electric Utilities	Natural Gas	15131
411796	2019	1	WY	Electric Generators, Electric Utilities	Petroleum	2651
411797	2019	1	WY	Electric Generators, Electric Utilities	Wind	158114

Table 4: Factor types of Type of Producer Variable

	Freq	% Valid	% Valid Cum.	% Total	% Total Cum.
Combined Heat and Power, Commercial Power	43658	10.601826	10.60183	10.601826	10.60183
Combined Heat and Power, Electric Power	40674	9.877197	20.47902	9.877197	20.47902
Combined Heat and Power, Industrial Power	65061	15.799290	36.27831	15.799290	36.27831
Electric Generators, Electric Utilities	77347	18.782798	55.06111	18.782798	55.06111
Electric Generators, Independent Power Producers	73257	17.789590	72.85070	17.789590	72.85070
Total Electric Power Industry	111800	27.149299	100.00000	27.149299	100.00000
<NA>	0	NA	NA	0.000000	100.00000
Total	411797	100.000000	100.00000	100.000000	100.00000

Further exploration of the raw dataset was carried out to determine the different factor types in the variables 'TYPE.OF.PRODUCER' and 'ENERGY.SOURCE' and the number of unique factors in the STATE variable.

```r
#Showing the different factors in TYPE.OF.PRODUCER variable
kable(freq(EIA.data.2001_2019$TYPE.OF.PRODUCER),
      caption = "Factor types of Type of Producer Variable") %>%
  kable_styling(latex_options="scale_down")
```

```r
#Showing the different factors in ENERGY.SOURCE variable
kable(freq(EIA.data.2001_2019$ENERGY.SOURCE),
      caption = "Factor types of Energy Source Variable") %>%
  kable_styling(latex_options="scale_down")
```

```r
#Showing the number of unique state types in the STATE variable
unique(EIA.data.2001_2019$STATE)
```

```
##  [1] AK       AL       AR       AZ       CA       CO       CT
##  [8] DC       DE       FL       GA       HI       IA       ID
## [15] IL       IN       KS       KY       LA       MA       MD
## [22] ME       MI       MN       MO       MS       MT       NC
## [29] ND       NE       NH       NJ       NM       NV       NY
## [36] OH       OK       OR       PA       RI       SC       SD
## [43] TN       TX       UT       VA       VT       WA       WI
## [50] WV       WY       US-TOTAL US-Total
## 53 Levels: AK AL AR AZ CA CO CT DC DE FL GA HI IA ID IL IN KS KY LA ... US-Total
```

Table 5: Factor types of Energy Source Variable

	Freq	% Valid	% Valid Cum.	% Total	% Total Cum.
Coal	40588	9.8563127	9.856313	9.8563127	9.856313
Geothermal	3369	0.8181216	10.674434	0.8181216	10.674434
Hydroelectric Conventional	32771	7.9580473	18.632482	7.9580473	18.632482
Natural Gas	56233	13.6555147	32.287996	13.6555147	32.287996
Nuclear	14302	3.4730705	35.761067	3.4730705	35.761067
Other	29522	7.1690663	42.930133	7.1690663	42.930133
Other Biomass	38050	9.2399896	52.170123	9.2399896	52.170123
Other Gases	15852	3.8494695	56.019592	3.8494695	56.019592
Petroleum	53177	12.9134015	68.932994	12.9134015	68.932994
Pumped Storage	8499	2.0638810	70.996875	2.0638810	70.996875
Solar Thermal and Photovoltaic	12308	2.9888513	73.985726	2.9888513	73.985726
Total	61824	15.0132225	88.998948	15.0132225	88.998948
Wind	20041	4.8667183	93.865667	4.8667183	93.865667
Wood and Wood Derived Fuels	25261	6.1343332	100.000000	6.1343332	100.000000
<NA>	0	NA	NA	0.0000000	100.000000
Total	411797	100.0000000	100.000000	100.0000000	100.000000

```r
length(unique(EIA.data.2001_2019$STATE))
```

```
## [1] 53
```

The above data exploration revealed the following factors of interest, *Total Electric Power Industry* in the variable **Type of Producer** and *US-TOTAL* and *US-Total* in the **State** variable. Further data exploration was carried out to determine if the *Total Electric Power Industry* factor represents electricity generation values for the summation of all other types of producers and the difference between factor *US-TOTAL* and *US-Total* relevant to total electricity generation in all states (at the national level).

3.1.1 Initial data wrangling to determine variation on Producer Types and State factors

```r
#Creating a dataset of elec generation sum for of all types
#of producers and comparing it total electric power industry

EIA.data.producers <- EIA.data.2001_2019 %>%
  summarise(`First 5 producer types (GWh)` = sum(as.numeric(
    GENERATION..Megawatthours.[TYPE.OF.PRODUCER!=
                                    "Total Electric Power Industry"]))/1000,
            `Total Electric Power Industry (GWh)` = sum(as.numeric(
              GENERATION..Megawatthours.[TYPE.OF.PRODUCER==
                                    "Total Electric Power Industry"]))/1000)
```

Table 6: Total electricity generation by producer types.

First 5 producer types (GWh)	Total Electric Power Industry (GWh)
291739158	291739158

The total electricity generation of all producer types except 'Total Electric Power Industry' was equal to that of 'Total Electric Power Industry'

```
#Creating datasets with and without Us-Total and US-TOTAL factors
EIA.Data.only_states <- EIA.data.2001_2019 %>%
  filter(TYPE.OF.PRODUCER=="Total Electric Power Industry") %>%
  filter(ENERGY.SOURCE=="Total") %>%
  filter(STATE!= "US-TOTAL" & STATE!= "US-Total") %>%
  mutate(Date = make_date(YEAR,MONTH))

EIA.Data.only_states.sum <- EIA.Data.only_states %>%
  summarise(`Total Elec Gen for States (GWh)` = sum(
    as.numeric(
      GENERATION..Megawatthours.))/1000)

EIA.Data.US <- EIA.data.2001_2019 %>%
  filter(TYPE.OF.PRODUCER=="Total Electric Power Industry") %>%
  filter(ENERGY.SOURCE=="Total") %>%
  filter(STATE== "US-TOTAL" | STATE== "US-Total") %>%
  mutate(Date = make_date(YEAR,MONTH))

EIA.Data.US.sum <- EIA.Data.US %>%
  summarise(`Total Elec Gen US-TOTAL  (GWh)` = sum(
    as.numeric(
      GENERATION..Megawatthours.[STATE=="US-TOTAL"]))/1000,
            `Total Elec Gen US-Total  (GWh)` = sum(
              as.numeric(
                GENERATION..Megawatthours.[STATE=="US-Total"]))/1000,
            `Total Elec Gen US-Both  (GWh)` = sum(
              as.numeric(
                GENERATION..Megawatthours.))/1000)

#Table comparison of electricity generation summation of both datasets
EIA.Data.States.US.sum <- cbind(EIA.Data.only_states.sum,EIA.Data.US.sum)
kable(EIA.Data.States.US.sum,
      caption = "Sumation of Total Electricity generation for all STATES and both US fac
  kable_styling(latex_options="scale_down")
```

Plot of variation in electricity generation at the US-TOTAL and US-Total factor level

Table 7: Sumation of Total Electricity generation for all STATES and both US factors

Total Elec Gen for States (GWh)	Total Elec Gen US-TOTAL (GWh)	Total Elec Gen US-Total (GWh)	Total Elec Gen US-Both (GWh)
72934791	44021141	28913648	72934789

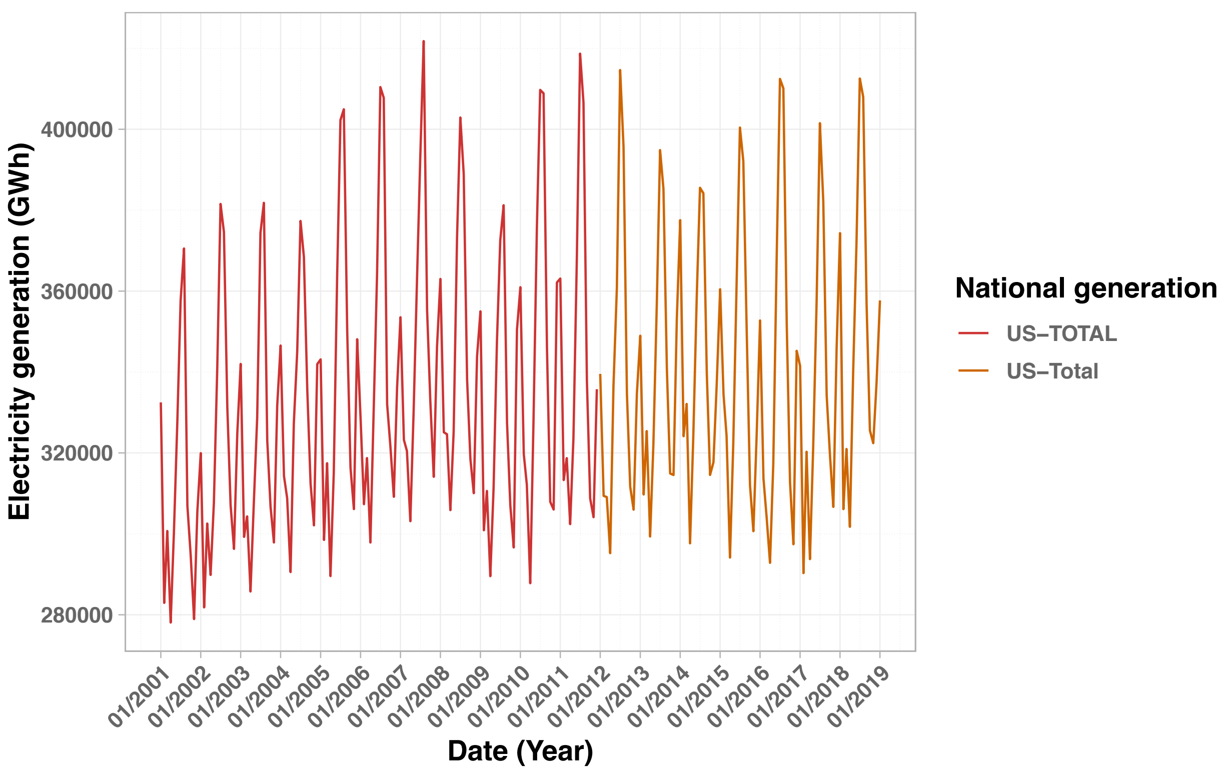

Figure 1: National electricity generation

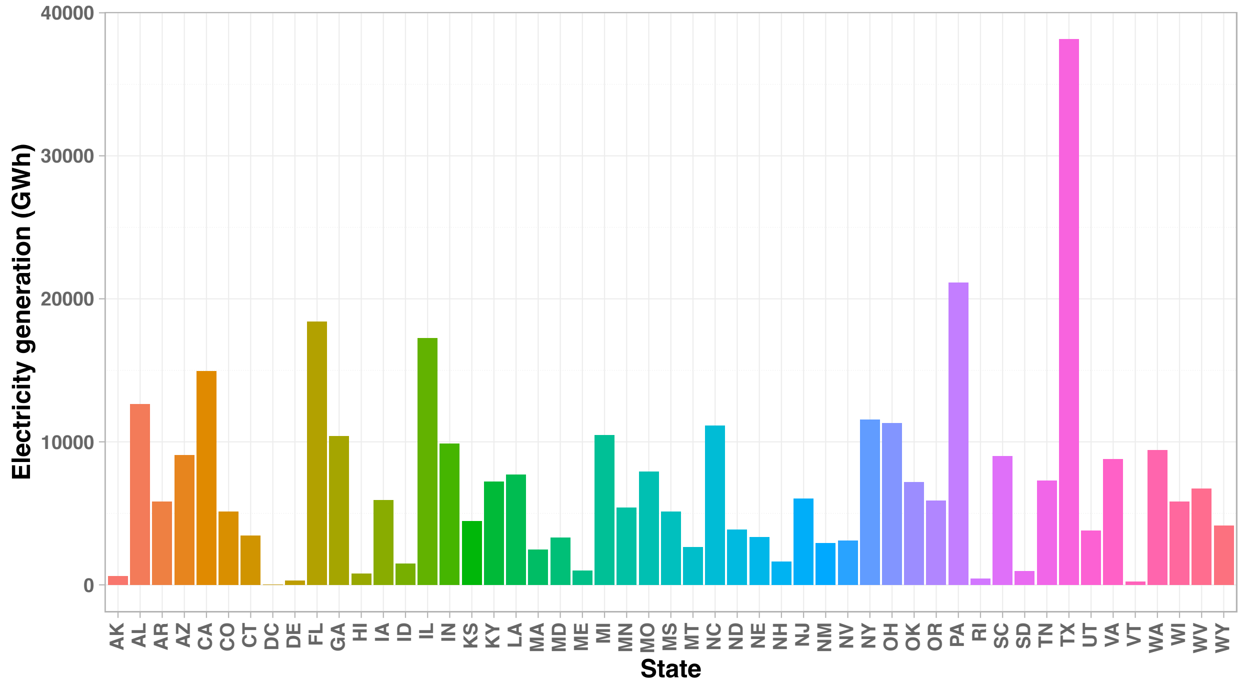

Figure 2: Total electricity generation in Jan 2019 per state

There was a small variation between the total states electricity generation and the US summation as shown in the Table 7. A dataset using only the states data was therefore used for the state level analysis and a dataset including summation of generation across states was used for the national level analysis. Date at the US-TOTAL and US-Total factor level were therefore not used because of the slight discrepacy with aggregated state values.

Plot of variation in electricity generation at the State level

3.2 Data wrangling to create the main datasets for national and state level

3.2.1 Creation of first dataset - EIA.data.by_state

To obtain the first dataset for data analysis, EIA.data.by_state, the EIA.data.2001_2019 raw dataset was wrangled as shown in the code below to include only the Total Electric Power Industry which is a summation of all producer codes as shown in the Table 6 and to remove US national totals. The data wrangling was also done to spread data according to MWh generation of each energy source and to combine the year and month collumn into one date variable. The final dataset, EIA.data.by_state, includes the following variables of interest: Date, State, Geothermal, Hydro,Other.Biomass, Solar, Wind, Wood.products, Renewable.total, Coal, Nat.Gas, Petroleum, Total

```
EIA.data.by_state <- EIA.data.2001_2019 %>%
  filter(TYPE.OF.PRODUCER=="Total Electric Power Industry") %>%
  filter(STATE!= "US-TOTAL" & STATE!= "US-Total") %>%
  spread(ENERGY.SOURCE,GENERATION..Megawatthours.) %>%
  mutate(Geothermal =ifelse(is.na(Geothermal),0,Geothermal)) %>%
  mutate(Hydro = ifelse(is.na(`Hydroelectric Conventional`),
                      0,`Hydroelectric Conventional`)) %>%
  mutate(Other.Biomass = ifelse(is.na(`Other Biomass`),
                              0,`Other Biomass`)) %>%
  mutate(Solar = ifelse(is.na(`Solar Thermal and Photovoltaic`),
                      0,`Solar Thermal and Photovoltaic`)) %>%
  mutate(Wind = ifelse(is.na(Wind),0,Wind)) %>%
  mutate(Wood.products = ifelse(is.na(`Wood and Wood Derived Fuels`),
                              0,`Wood and Wood Derived Fuels`)) %>%
  mutate(Renewable.total = Geothermal+Hydro+Other.Biomass+
           Solar+Wind+Wood.products) %>%
  mutate(Coal = ifelse(is.na(Coal),0,Coal)) %>%
  mutate(Nat.Gas = ifelse(is.na(`Natural Gas`),
                        0,`Natural Gas`)) %>%
  mutate(Petroleum = ifelse(is.na(Petroleum),
                          0,Petroleum)) %>%
  mutate(Nuclear = ifelse(is.na(Nuclear),
                        0,Nuclear)) %>%
  mutate(Date = make_date(YEAR,MONTH)) %>%
  select(Date, STATE, Geothermal,Hydro,
         Other.Biomass,Solar,Wind, Wood.products,
         Coal,Nat.Gas,Petroleum,
         Nuclear,Renewable.total,Total)
```

3.2.2 Creation of second dataset - EIA.data.national

A second analysis dataset was created to represent national electricity generation across various energy sources by summing state level values as shown in the wrangling code below.

```
EIA.data.national <- EIA.data.by_state %>%
  group_by(Date) %>%
  summarise(National.Geothermal = sum(Geothermal),
            National.Hydro = sum(Hydro),
            National.Other.Biomass = sum(Other.Biomass),
            National.Solar = sum(Solar),
            National.Wind = sum(Wind),
```

Table 8: Summary statistics for electricity generation (MWh) by energy source for all states

	N.Valid	Pct.Valid	Mean	Std.Dev	Min	Median	Max	IQR
Geothermal	11067	100	25052.51	147441.44	-140	0	1147970	0.0
Hydro	11067	100	440516.99	1076092.36	0	113029	11210102	250229.5
Other.Biomass	11067	100	30067.67	47596.23	-535	8299	265105	35747.0
Solar	11067	100	20349.35	134828.70	-3	0	3144661	1057.0
Wind	11067	100	174190.36	494949.85	0	7651	7548279	135000.0
Wood.products	11067	100	63761.44	89673.07	-390	19501	685925	102163.5
Coal	11067	100	2788475.91	2906988.91	-5559	2157406	15815851	3475819.5
Nat.Gas	11067	100	1627877.87	3063518.75	-903	568015	28568297	1555542.0
Petroleum	11067	100	94456.31	310868.65	-3417	11224	5317378	55609.5
Nuclear	11067	100	1294177.77	1719842.22	-25629	798993	8871262	2213163.5
Renewable.total	11067	100	753938.32	1349140.89	0	328044	11691617	547045.0
Total	11067	100	6590294.65	6253557.87	-1094	4738322	49047812	6287579.5

```
            National.wood.products = sum(Wood.products),
            National.Coal = sum(Coal),
            National.Nat.Gas = sum(Nat.Gas),
            National.Petroleum = sum(Petroleum),
            National.Nuclear =sum(Nuclear),
            National.Renewable.total = sum(Renewable.total),
            National.Total = sum(Total))
```

3.2.3 Data exploration of created datasets

Table 8 shows the summary statistics for the electricity generation from all the selected energy sources including total renewable energy (Renewable.total variable) and all the energy sources (Total variable).

3.2.4 Data visualization of created datasets

Plot of national level electricity generation by main energy sources across time

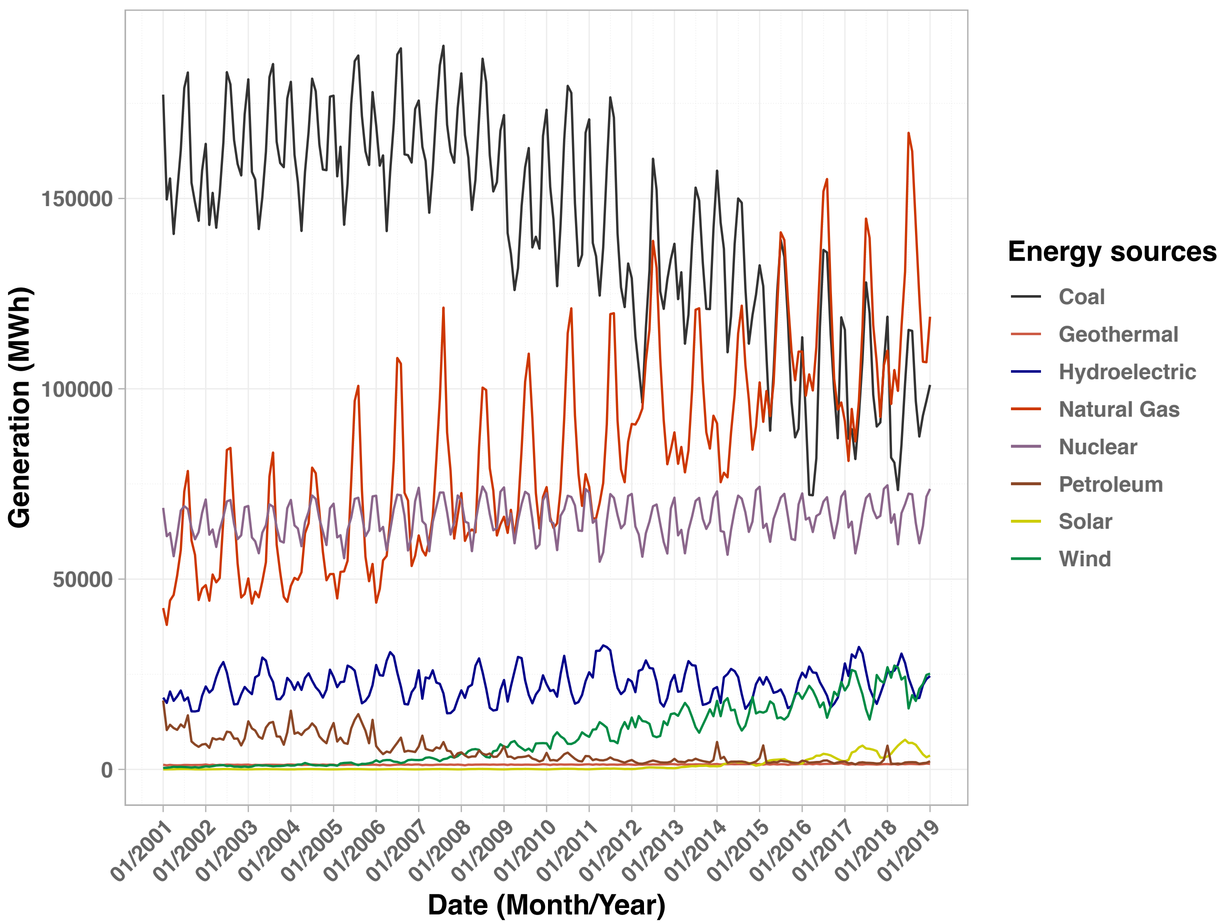

Figure 3: National electricity generation by energy source

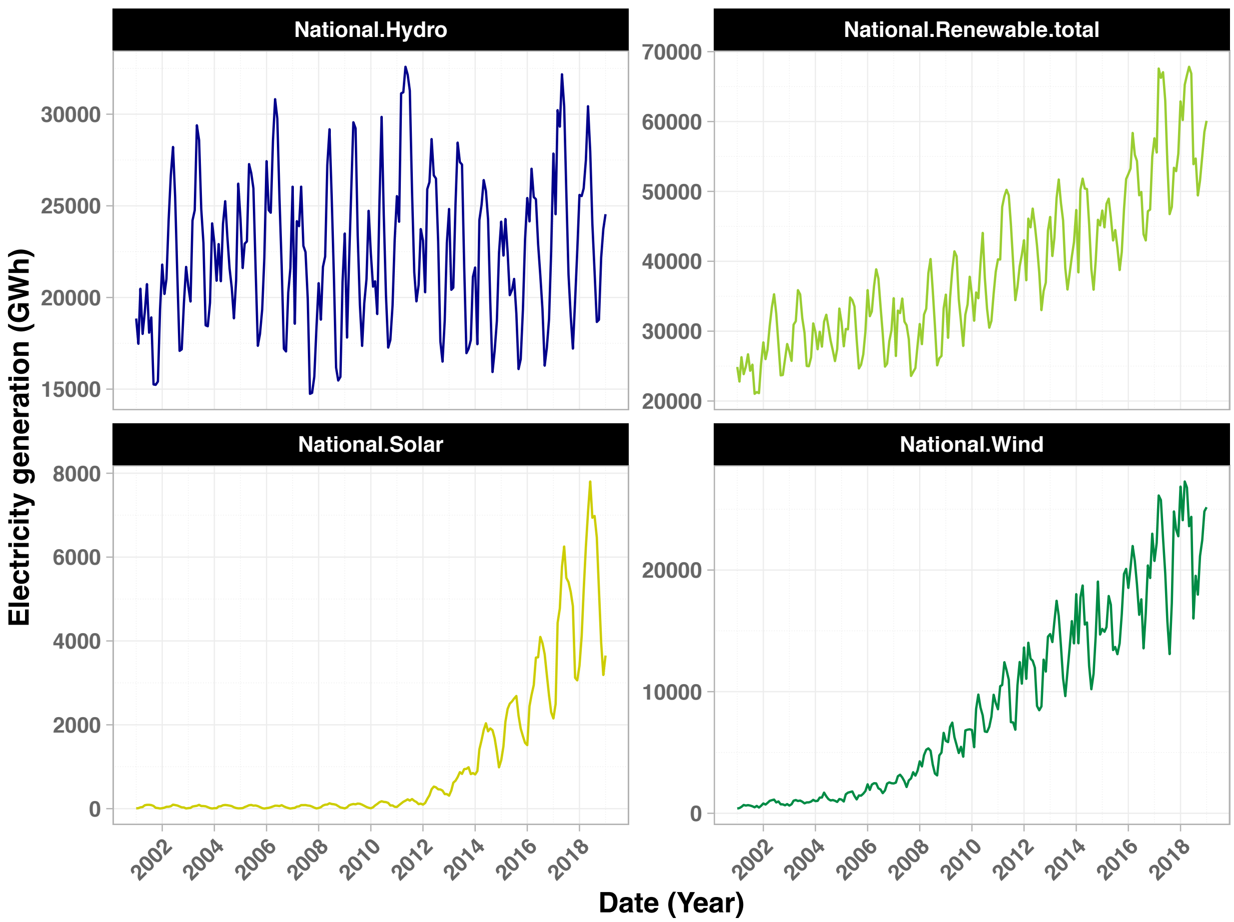

Figure 4: Electricity generation by major renewable sources

4 Analysis

4.1 National level trend analysis

4.1.1 Visualization of trends of major renewable and non-renewable energy sources

Based on the data exploration and specifically figure 3, The major renewable energy sources appear to be Hydroelectric power, wind energy and solar energy. The major non renewable energy sources are coal energy, natural gas energy and nuclear energy.

Based on the series of the major renewable energy sources, hydroelectric energy does not appear to have any inclеasing trend therefore it is unlikely that it has contributed to the visibly increasing growth in total renewable energy generation at the national level. Solar and wind energy however appear to have a similar

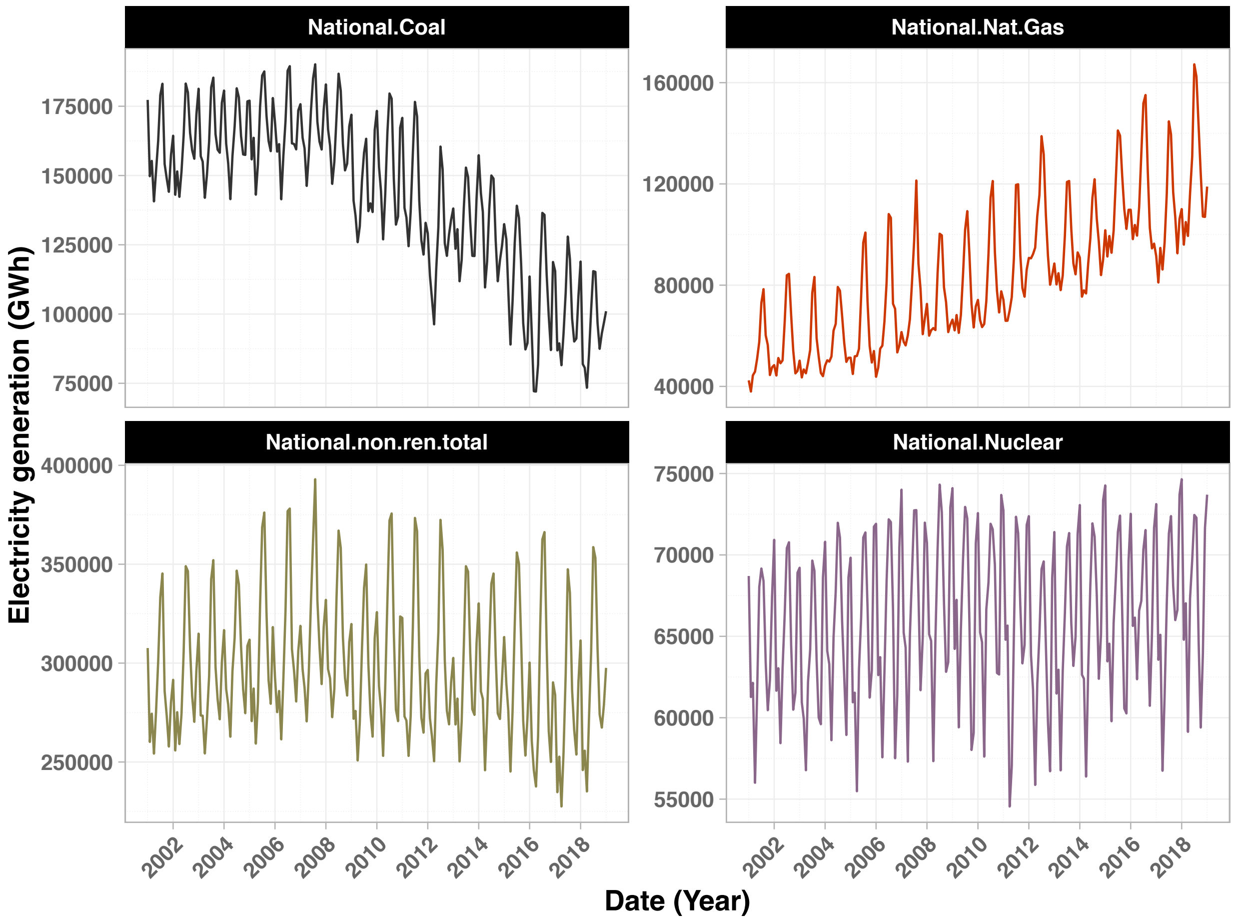

Figure 5: Electricity generation by major renewable sources

increasing trend. The trend analysis at the national level was therefore carried out on the solar and wind series as well as the total renewable energy series.

Based on the series of the major non-renewable energy sources, nuclear energy does not appear to have as significant a trend as coal and Natural gas. It is therefore unlikely that it has contributed significantly to the slight decrease in total non renewable energy generation at the national level. The trend analysis at the national level was therefore carried out on the coal and natural gas series.

4.1.2 Trend analysis of major renewable and non-renewable energy series.

4.1.2.1 Changing to timeseries data and removing seasonality

It was necessary to change the energy series into time series and to remove the obvious seasonality so as to be able to carry out trend analysis tests.

```r
#Changing series of interest into timeseries data
Nat.Solar.ts.data <- ts(EIA.data.national[,5],
                        start = c(2001,01), end=c(2019,1),frequency = 12)
Nat.Wind.ts.data <- ts(EIA.data.national[,6],
                       start = c(2001,01), end=c(2019,1),frequency = 12)
Nat.Renewable.ts.data <- ts(EIA.data.national[,12],
                            start = c(2001,01), end=c(2019,1),frequency = 12)
Nat.Coal.ts.data <- ts(EIA.data.national[,8],
                       start = c(2001,01), end=c(2019,1),frequency = 12)
Nat.Gas.ts.data <- ts(EIA.data.national[,9],
                      start = c(2001,01), end=c(2019,1),frequency = 12)

##Removing seasonality
##Removing seasonality in solar series
Nat.Solar.decomp <- decompose(Nat.Solar.ts.data)
deseasonal_Nat.Solar <- seasadj(Nat.Solar.decomp)

##Removing seasonality in wind series
Nat.Wind.decomp <- decompose(Nat.Wind.ts.data)
deseasonal_Nat.Wind <- seasadj(Nat.Wind.decomp)

##Removing seasonality in renewable series
Nat.Renewable.decomp <- decompose(Nat.Renewable.ts.data)
deseasonal_Nat.Renewable <- seasadj(Nat.Renewable.decomp)

##Removing seasonality in coal series
Nat.Coal.decomp <- decompose(Nat.Coal.ts.data)
deseasonal_Nat.Coal <- seasadj(Nat.Coal.decomp)

##Removing seasonality in natural gas series
Nat.Gas.decomp <- decompose(Nat.Gas.ts.data)
deseasonal_Nat.Gas <- seasadj(Nat.Gas.decomp)
```

4.1.2.2 Carrying out ADF and Mankandal trend test

4.1.2.2.1 ADF trend tests on series

To test for a stochastic trend in the series, and ADF trend test was carried out

```r
#Checking for a stochastic trend with ADF test on solar series
adf.test(deseasonal_Nat.Solar, alternative = "stationary")
```

```
##
##  Augmented Dickey-Fuller Test
```

```
##
## data:  deseasonal_Nat.Solar
## Dickey-Fuller = -0.90921, Lag order = 5, p-value = 0.9506
## alternative hypothesis: stationary
```

```r
#Checking for a stochastic trend with ADF test on wind series
adf.test(deseasonal_Nat.Wind, alternative = "stationary")
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  deseasonal_Nat.Wind
## Dickey-Fuller = -3.2801, Lag order = 5, p-value = 0.07573
## alternative hypothesis: stationary
```

```r
#Checking for a stochastic trend with ADF test on total renewable energy series
adf.test(deseasonal_Nat.Renewable, alternative = "stationary")
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  deseasonal_Nat.Renewable
## Dickey-Fuller = -3.2402, Lag order = 5, p-value = 0.08238
## alternative hypothesis: stationary
```

```r
#Checking for a stochastic trend with ADF test on coal series
adf.test(deseasonal_Nat.Coal, alternative = "stationary")
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  deseasonal_Nat.Coal
## Dickey-Fuller = -2.6226, Lag order = 5, p-value = 0.3148
## alternative hypothesis: stationary
```

```r
#Checking for a stochastic trend with ADF test on natural gas series
adf.test(deseasonal_Nat.Gas, alternative = "stationary")
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  deseasonal_Nat.Gas
## Dickey-Fuller = -3.9301, Lag order = 5, p-value = 0.01354
## alternative hypothesis: stationary
```

The null hypothesis of the ADF test is that a stochastic trend is present in the series. The test results indicate that the following series have a stochastic trend.

- The national solar energy series (p-value = 0.9506)

- The national wind energy series (p-value = 0.07573)
- The national total renewable energy series (p-value = 0.08238)
- The national coal energy series (p-value = 0.3148)

  The national natural gas series has a low p value (p-value = 0.01354) indicating it does not have a stochastic trend. A Man Kendall test to test for a deterministic trend in the series was therefore carried out.

4.1.2.2.2 Mann-Kendall trend tests on natural gas series

```
#Checking for a monotonic trend using the Mann Kendall Test
MannKendall(deseasonal_Nat.Gas)
```

```
## tau = 0.817, 2-sided pvalue =< 2.22e-16
```

The null hypothesis of the Mann-Kendall test is that there is no monotonic trend. The low p value of the test (pvalue =< 2.22e-16) indicates that we should reject the null hypothesis for the alternative that a monotonic trend is present in the natural gas series. The trend is increasing (tau = 0.817)

4.1.3 Changepoint analysis of major renewable energy sourcess.

4.1.3.1 Changepoint analysis for solar energy series

```
#analysis of first change point for solar series
pettitt.test(deseasonal_Nat.Solar)
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_Nat.Solar
## U* = 10750, p-value < 2.2e-16
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                             130
```

Low p-value (< 2.2e-16) indicates that there is a significant changepoint at observation 130 which coresponds to the date 2011-10-01.

```
#analysis of second change point for solar series
pettitt.test(deseasonal_Nat.Solar[131:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_Nat.Solar[131:217]
```

```
## U* = 1880, p-value = 2.978e-14
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                               40
```

Low p-value (2.978e-14) indicates that there is a significant changepoint at observation 40 (170) which coresponds to the date 2015-02-01.

```
#analysis of third change point for solar series
pettitt.test(deseasonal_Nat.Solar[171:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_Nat.Solar[171:217]
## U* = 532, p-value = 2.216e-07
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                               24
```

Low p-value (2.216e-07) indicates that there is a significant changepoint at observation 24 (194) which coresponds to the date 2017-02-01.

```
#analysis of fourth change point for solar series
pettitt.test(deseasonal_Nat.Solar[195:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_Nat.Solar[195:217]
## U* = 64, p-value = 0.2886
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K                            <NA>
##                               12                              13
```

the high p-value (0.2886) of the test indicates that there no other significant changepoint.

Plot of National solar changepoints

4.1.3.2 Changepoint analysis for wind energy series

```
#analysis of first change point for wind series
pettitt.test(deseasonal_Nat.Wind)
```

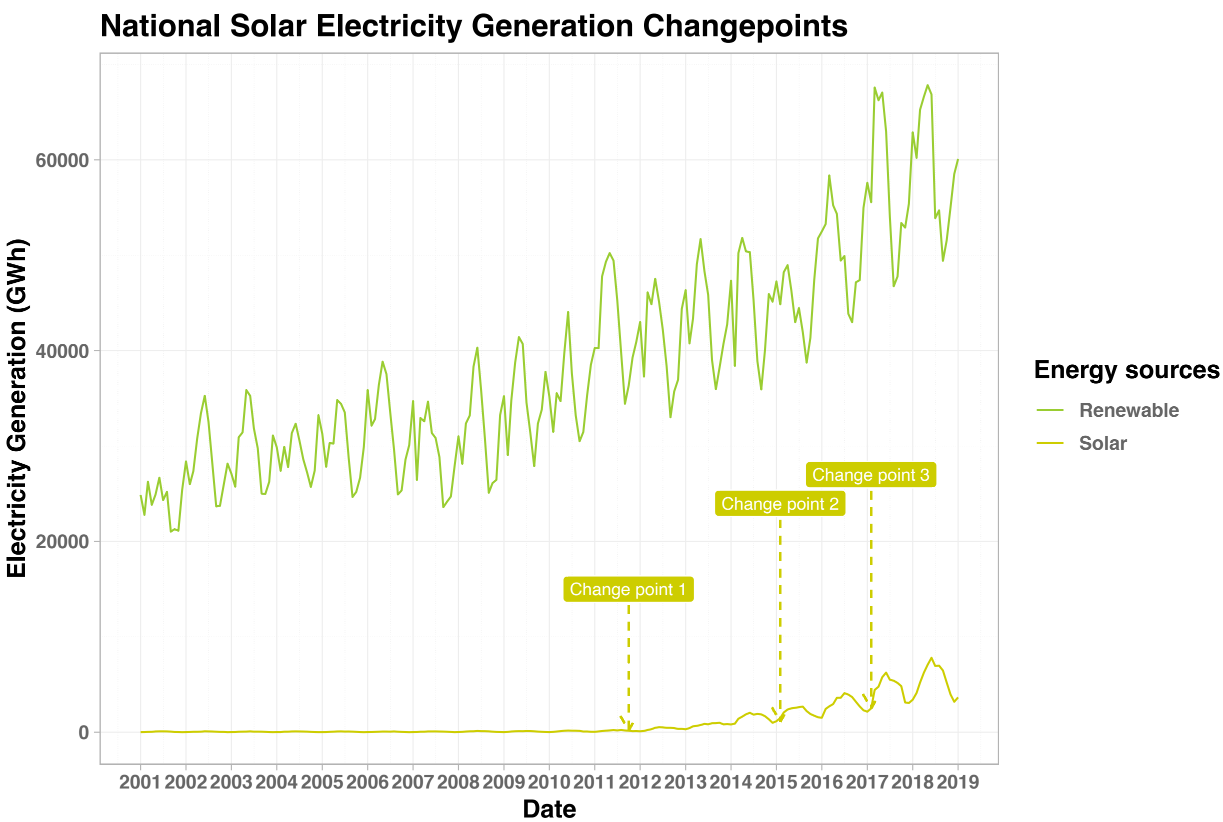

Figure 6: National Solar Electricity Generation Changepoints

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_Nat.Wind
## U* = 11768, p-value < 2.2e-16
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                             110
```

Low p-value (< 2.2e-16) indicates that there is a significant changepoint at observation 110 which coresponds to the date 2010-02-01.

```
#analysis of second change point for wind series
pettitt.test(deseasonal_Nat.Wind[111:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_Nat.Wind[111:217]
## U* = 2672, p-value = 1.8e-15
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                              56
```

Low p-value (1.8e-15) indicates that there is a significant changepoint at observation 56 (166) which coresponds to the date 2014-10-01.

```
#analysis of third change point for wind series
pettitt.test(deseasonal_Nat.Wind[167:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_Nat.Wind[167:217]
## U* = 542, p-value = 4.379e-06
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                              23
```

Low p-value (4.379e-06) indicates that there is a significant changepoint at observation 23 (189) which coresponds to the date 2016-09-01.

```
#analysis of fourth change point for solar series
pettitt.test(deseasonal_Nat.Wind[190:217])
```

```
##
```

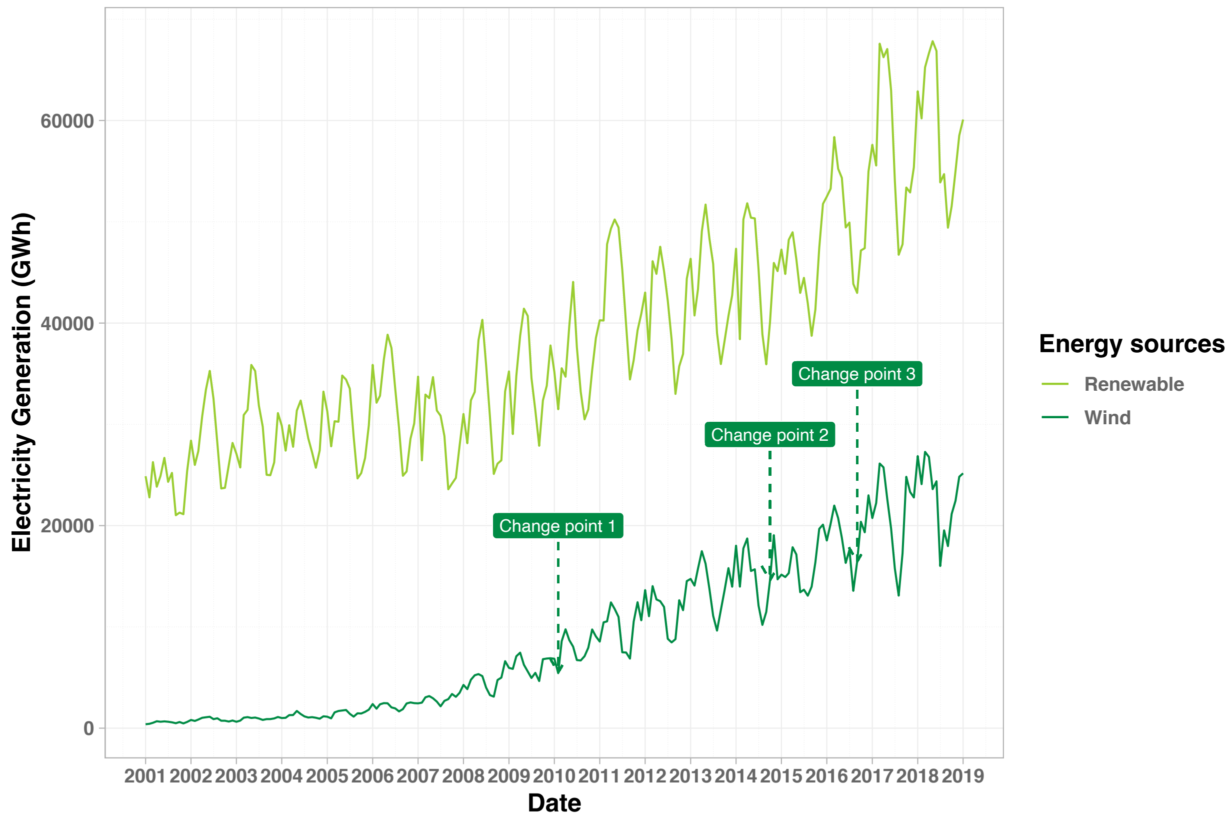

Figure 7: National Wind Electricity Generation Changepoints

```
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_Nat.Wind[190:217]
## U* = 108, p-value = 0.09209
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                               12
```

the high p-value (0.09209) of the test indicates that there no other significant changepoint.

Plot of National wind changepoints

4.2 State level trend analysis

4.2.1 Trend analysis of wind energy in Kansas.

A changepoint trend analysis was caried out for Kansas state to determine whether renewable energy from wind follows a similar trend to the national wind series

4.2.1.0.1 Kansas dataset

```
#Creating Kansas Dataset
data.KS.state <- EIA.data.by_state %>%
  filter(STATE=="KS") %>%
  select(Date,Wind,Renewable.total,Total)

#Converting into time series
KS.Wind <- ts(data.KS.state[,2], start = c(2001,01),
              end=c(2019,1),frequency = 12)
KS.renewable <- ts(data.KS.state[,3], start = c(2001,01),
                   end=c(2019,1),frequency = 12)
KS.total <- ts(data.KS.state[,4], start = c(2001,01),
               end=c(2019,1),frequency = 12)
```

4.2.1.1 Removing seasonal component out in Kansas series

```
#Removing seasonal component from wind series
KS.Wind.decomp <- decompose(KS.Wind)
deseasonal_KS.Wind <- seasadj(KS.Wind.decomp)

#Removing seasonal component from the renewable energy series
KS.renewable.decomp <- decompose(KS.renewable)
deseasonal_KS.renewable <- seasadj(KS.renewable.decomp)

#Checking for a stochastic trend with ADF test
adf.test(deseasonal_KS.Wind, alternative = "stationary")
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  deseasonal_KS.Wind
## Dickey-Fuller = -1.5031, Lag order = 5, p-value = 0.7845
## alternative hypothesis: stationary
```

```
adf.test(deseasonal_KS.renewable, alternative = "stationary")
```

```
##
```

```
##  Augmented Dickey-Fuller Test
##
## data:  deseasonal_KS.renewable
## Dickey-Fuller = -1.5036, Lag order = 5, p-value = 0.7843
## alternative hypothesis: stationary
```

The result of the ADF test of the Wind series (p=0.7845) indicates that the series has a stochastic trend, this result is similar for the total renewable energy series (p=0.7843).

4.2.1.2 Changepoint analysis of wind energy series

```
#analysis of first change point
pettitt.test(deseasonal_KS.Wind)
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_KS.Wind
## U* = 11694, p-value < 2.2e-16
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                              102
```

small pvalues (< 2.2e-16) indicates a significant change point was detected in both series in 2009-06-01 (102) for the wind series.

```
#analysis of second change point
pettitt.test(deseasonal_KS.Wind[103:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_KS.Wind[103:217]
## U* = 3202, p-value < 2.2e-16
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                               50
```

small p value (< 2.2e-16) indicates a significant changepoint at obs 50 (152). This changepoint occured at 2013-08-01.

```
#analysis of third change point
pettitt.test(deseasonal_KS.Wind[153:217])
```

```
##
```

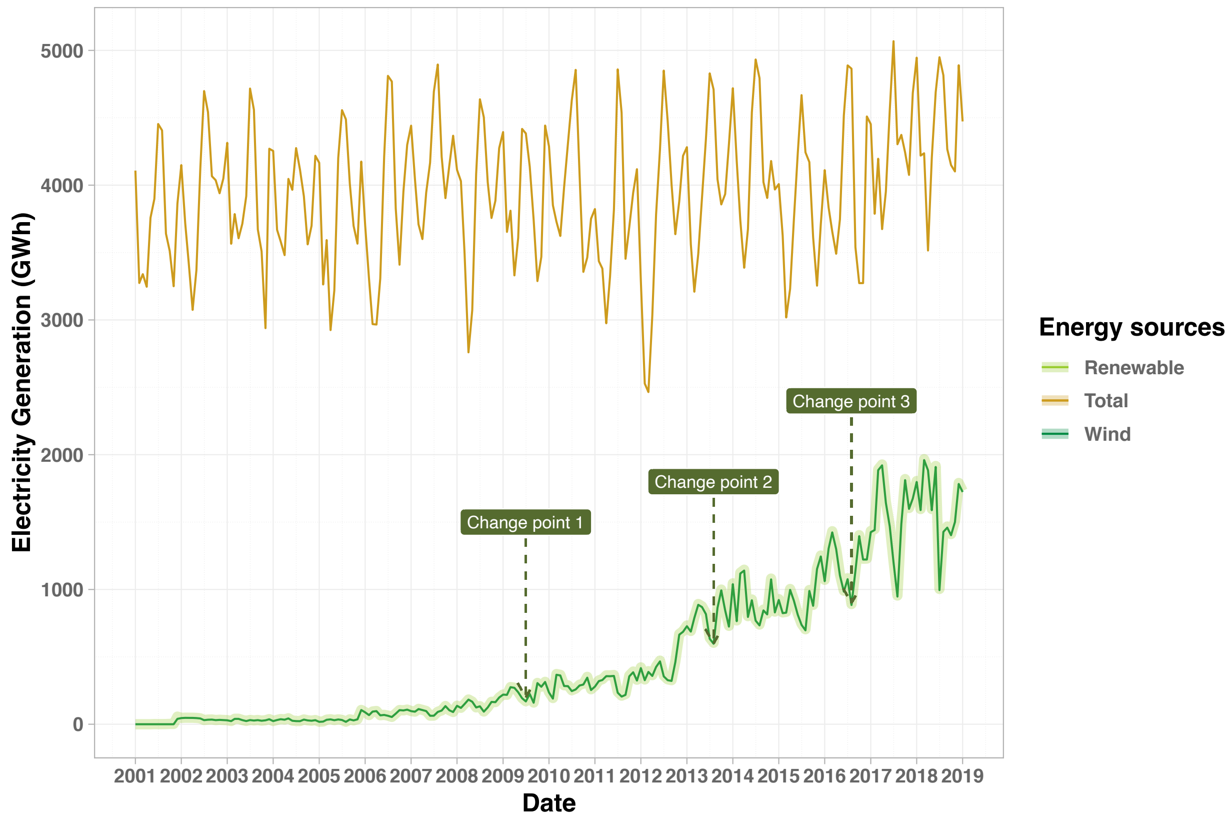

Figure 8: Kansas wind electricity generation

```
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_KS.Wind[153:217]
## U* = 994, p-value = 1.17e-09
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                               36
```

small p value (1.17e-09) indicates a significant changepoint at obs 36 (188). This changepoint occured at 2016-08-01.

Plot of Kansas wind changepoints

4.2.2 Trend analysis of solar energy in California.

4.2.2.0.1 California dataset

```
#Creating California Dataset
data.CA.state <- EIA.data.by_state %>%
  filter(STATE=="CA") %>%
  select(Date,Solar,Renewable.total,Total)

#Converting into time series
CA.Solar <- ts(data.CA.state[,2], start = c(2001,01),
               end=c(2019,1),frequency = 12)
CA.renewable <- ts(data.CA.state[,3], start = c(2001,01),
                   end=c(2019,1),frequency = 12)
CA.total <- ts(data.CA.state[,4], start = c(2001,01),
               end=c(2019,1),frequency = 12)
```

4.2.2.1 Removing seasonal component out of California series

```
#Removing seasonal component from solar series
CA.Solar.decomp <- decompose(CA.Solar)
deseasonal_CA.Solar <- seasadj(CA.Solar.decomp)

#Removing seasonal component from the renewable energy series
CA.renewable.decomp <- decompose(CA.renewable)
deseasonal_CA.renewable <- seasadj(CA.renewable.decomp)

#Checking for a stochastic trend with ADF test
adf.test(deseasonal_CA.Solar, alternative = "stationary")
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  deseasonal_CA.Solar
## Dickey-Fuller = -1.6276, Lag order = 5, p-value = 0.7322
## alternative hypothesis: stationary
```

```
adf.test(deseasonal_CA.renewable, alternative = "stationary")
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  deseasonal_CA.renewable
## Dickey-Fuller = -2.7292, Lag order = 5, p-value = 0.2701
## alternative hypothesis: stationary
```

The result of the ADF test of the solar series (p=0.7322) indicates that the series has a stochastic trend. This result is similar for the total renewable energy series (p=0.2701).

4.2.2.2 Changepoint analysis of solar energy series

```r
#analysis of first change point
pettitt.test(deseasonal_CA.Solar)
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_CA.Solar
## U* = 10398, p-value < 2.2e-16
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                              142
```

small p value (< 2.2e-16) indicates there a change point was detected in the solar series in observation 142 corresponding to 2012-10-01.

```r
#analysis of second change point
pettitt.test(deseasonal_CA.Solar[143:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_CA.Solar[143:217]
## U* = 1342, p-value = 2.106e-11
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                               39
```

small p value (2.106e-11) indicates there a changepoint was detected at observation 39 (181) corresponding to 2016-01-01.

```r
#analysis of third change point
pettitt.test(deseasonal_CA.Solar[182:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_CA.Solar[182:217]
## U* = 209, p-value = 0.00846
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                               13
```

small p value (0.00846)indicates there a changepoint was detected at observation 13 (194) corresponding to 2017-02-01.

Figure 9: California solar electricity generation

```
#analysis of fourth change point
pettitt.test(deseasonal_CA.Solar[195:217])
```

```
##
##  Pettitt's test for single change-point detection
##
## data:  deseasonal_CA.Solar[195:217]
## U* = 50, p-value = 0.6137
## alternative hypothesis: two.sided
## sample estimates:
## probable change point at time K
##                              13
```

large p value (0.6137)indicates there is no other significant changepoint.

Plot of California solar changepoints

5 Summary and Conclusions

The table below summarizes the results of the changepoint analysis tests for the electricity generation from solar energy at the national level and in California and electricity generation from wind energy at the national level and in Kansas.

Dataset	Changepoint 1	Changepoint 2	Changepoint 3
National solar energy	10/2011	02/2015	02/2017
National wind energy	02/2010	10/2014	09/2016
Kansas wind energy	06/2009	08/2013	08/2016
California solar energy	10/2012	01/2016	02/2017

all analysed energy generation series had three changepoints. All changepoints occur post 2008 indicating the rate of adoption of renewable energy for electricity generation has changed in the last decade, consistent with the literature. All changepoints indicate also an increase in electricty generation from the renewable resource.

Changepoint 1 and two at the national and state level do not appear to be related for both wind energy and solar. Change point 3 for solar and wind energy at the state and national level however coincide. This could indicate that significant increase in renewable energy levels is being driven by specific states and cannot neccessaritly be generalized to all.

The changepoint 1 in Kansas wind energy series coresponds to the year the state adopted the RPS policy. Changepoint 2 does not correspond to any policy change however it is the year Kansas reached its renewable energy target of 20% of its electricity generation being from renewable resources. It was expected that a changepoint would be detected in 2015 or early 2016 due to the Kansa's decision to repeal its RPS policy and change it from a mandatory goal to a voluntary one. A lack of clear changepoint related to this indicates negative policy changes may not have an immediate impact on adoption rates of renewable energy resource as compared to positive ones.

For California, changepoint 1 and 3 does not correspond to any major renewable energy policy changes. A major policy revison however took place in 2015 and could have resulted in a shift in the solar energy trend as shown in changepoint 2.

It is therefore probable that polciy changes have impacted the adoption trends for wind and solar energy in Kansas and California, however more analysis needs to be done in this area to be able to definitively determine the causal relationship between policy changes and renewable energy adoption.