Energy’s Impact on Air Pollution

Andre Shearer, Erich Mitchell, Felix Pronove, Justynn Hammond,
Kim Christensen, Lisa Caruana

What we examined

Relationship between energy sources and air pollution.

QUESTION: Does an increase in renewable energy production reduce the air pollution levels?

Hypothesis: An increased use of renewable energies results in less air pollution.

Potential pitfall: Delayed effect of energy use on air pollution (i.e. energy used 2000 -> air pollution 2005)

Data Sources

- **EIA CSV file for energy production by source**
  - **Broken down by region & state**
- **EPA API -- Breaks down air pollution across US**

In [14]:
```
#Removing rows where all of the data is blank.
filtered_df2=filtered_df[((filtered_df['2001']!="--")&(filtered_df['2002']!="--")&
                         (filtered_df['2003']!="--")&(filtered_df['2004']!="--")&
                         (filtered_df['2005']!="--")&(filtered_df['2006']!="--")&
                         (filtered_df['2007']!="--")&(filtered_df['2008']!="--")&
                         (filtered_df['2009']!="--")&(filtered_df['2010']!="--")&
                         (filtered_df['2011']!="--")&(filtered_df['2012']!="--")&
                         (filtered_df['2013']!="--")&(filtered_df['2014']!="--")&
                         (filtered_df['2015']!="--")&(filtered_df['2016']!="--")&
                         (filtered_df['2017']!="--")&(filtered_df['2018']!="--")&
                         (filtered_df['2019']!="--"))]

filtered_df2
```

Out[14]:

3	Location	Energy_Source	units	source key	2001	2002	2003	2004	2005	2006	...	2011	2012	2013
7	United States	coal	thousand megawatthours	ELEC.GEN.COW-US-98.A	1882826	1910613	1952714	1957188	1992054	1969737	...	1717891	1500557	1567722
8	United States	petroleum liquids	thousand megawatthours	ELEC.GEN.PEL-US-98.A	110260	75079	98593	95770	95667	41665	...	15343	12649	13207
9	United States	petroleum coke	thousand megawatthours	ELEC.GEN.PC-US-98.A	8889	14654	15105	18908	20814	18043	...	12859	7423	11303
10	United States	natural gas	thousand megawatthours	ELEC.GEN.NG-US-98.A	554940	607683	567303	627172	683829	734417	...	926290	1132791	1028949
11	United States	other gases	thousand megawatthours	ELEC.GEN.OOG-US-98.A	586	1970	2647	3568	3777	4254	...	2939	2984	4322
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1540	Hawaii	wind	thousand megawatthours	ELEC.GEN.WND-HI-98.A	2	2	2	7	7	80	...	341	378	503
1544	Hawaii	geothermal	thousand megawatthours	ELEC.GEN.GEO-HI-98.A	207	73	178	213	222	212	...	224	261	275
1545	Hawaii	biomass	thousand megawatthours	ELEC.GEN.BIO-HI-98.A	164	169	339	141	134	129	...	39	22	29
1547	Hawaii	other biomass	thousand megawatthours	ELEC.GEN.WAS-HI-98.A	164	169	339	141	134	129	...	39	22	29
1549	Hawaii	other	thousand megawatthours	ELEC.GEN.OTH-HI-98.A	118	143	172	25	21	27	...	211	216	219

624 rows × 24 columns

```
mapper={' all fuels (utility-scale)': 'all',
        ' coal': 'non-renewable',
        ' petroleum liquids': 'non-renewable',
        ' petroleum coke':'non-renewable',
        ' natural gas':'non-renewable',
        ' other gases':'non-renewable',
        ' nuclear':'non-renewable',
        ' conventional hydroelectric':'renewable',
        ' other renewables':'renewable',
        ' wind':'renewable',
        ' all utility-scale solar':'all',
        ' utility-scale photovoltaic':'renewable',
        ' utility-scale thermal':'renewable',
        ' geothermal':'renewable',
        ' biomass':'renewable',
        ' wood and wood-derived fuels':'non-renewable',
        ' other biomass': 'renewable',
        ' hydro-electric pumped storage':'renewable',
        ' other':'non-renewable',
        ' all solar':'all',
        ' small-scale solar photovoltaic':'renewable'}

energy_df['Renewable']=energy_df['Energy_Source'].map(mapper)

energy_df
```

3	Location	Energy_Source	units	source key	2001	2002	2003	2004	2005	2006	...	2011	2012	2013
6	United States	all fuels (utility-scale)	thousand megawatthours	ELEC.GEN.ALL-US-98.A	3580053	3698458	3721159	3808360	3902192	3908077	...	3948186	3890358	3903715
7	United States	coal	thousand megawatthours	ELEC.GEN.COW-US-98.A	1882826	1910613	1952714	1957188	1992054	1969737	...	1717891	1500557	1567722
8	United States	petroleum liquids	thousand megawatthours	ELEC.GEN.PEL-US-98.A	110260	75079	98593	95770	95667	41665	...	15343	12649	13207
9	United States	petroleum coke	thousand megawatthours	ELEC.GEN.PC-US-98.A	8889	14654	15105	18908	20814	18043	...	12859	7423	11303
10	United States	natural gas	thousand megawatthours	ELEC.GEN.NG-US-98.A	554940	607683	567303	627172	683829	734417	...	926290	1132791	1028949

In [2]:
```
energy_df=df.drop([0, 1, 2, 4])
```

In [3]:
```
energy_df[['Net generation for electric power', 'Energy Source']] = energy_df['Net generation for electric pow
energy_df
```

Out[3]:

Net generation for electric power	Unnamed: 1	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	...	Unnamed: 13	Unname
3 description	units	source key	2001	2002	2003	2004	2005	2006	2007	...	2011	20
5 United States	NaN	ELEC.GEN..A	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	Na
6 United States	thousand megawatthours	ELEC.GEN.ALL-US-98.A	3580053	3698458	3721159	3808360	3902192	3908077	4005343	...	3948186	38903
7 United States	thousand megawatthours	ELEC.GEN.COW-US-98.A	1882826	1910613	1952714	1957188	1992054	1969737	1998390	...	1717891	15005
8 United States	thousand megawatthours	ELEC.GEN.PEL-US-98.A	110260	75079	98593	95770	95667	41665	46970	...	15343	126
... ...	...	...	...	...	...	...	...	...	...	...	...	
1550 Hawaii	thousand megawatthours	ELEC.GEN.TSN-HI-98.A	--	--	--	--	--	--	--	...	--	
1551 Hawaii	thousand megawatthours	ELEC.GEN.DPV-HI-98.A	--	--	--	--	--	--	--	...	--	
1552 Hawaii	thousand megawatthours	ELEC.GEN.SUN-HI-98.A	--	--	--	--	--	--	--	...	4	
1553 Hawaii	thousand megawatthours	ELEC.GEN.SPV-HI-98.A	--	--	--	--	--	--	--	...	4	
1554 Hawaii	thousand megawatthours	ELEC.GEN.STH-HI-98.A	--	--	--	--	--	--	--	...	--	

1551 rows × 23 columns

API CALL - EPA Pollution Data

```
# develop the query URL - Sample Data
sample_url=f'https://aqs.epa.gov/data/api/sampleData/byState?email={email}&key={apikey}&param={para
response=requests.get(sample_url)
sample_param=response.json()
sample_param['Data'][0].keys()
```

```
#annual summary data - pulling same params from previous kernel - Annual Data by state
annual_url=f'https://aqs.epa.gov/data/api/annualData/byState?email={email}&key={apikey}&param={para
response=requests.get(annual_url)
annual_param=response.json()
annual_param['Data'][0].keys()
```

```
# hard coded searches
param='42101'
bdate='20200131'
edate='20201231'

#query URL
query_url=f'https://aqs.epa.gov/data/api/annualData/byState?email={email}&key={apikey}&param={param

# response = requests.get(query_url + code).json()
```

```
# Set up empty lists for data in each column
states = []
cities = []
lat = []
long = []
mean = []
max_value = []
# Create a list of state codes so our request can loop through and pull each state
state_code = ["01", "02", "04", "05", "06", "08", "09", "10", "11", "12", "13", "15", "16", "17", "
              "23", "24", "25", "26", "27", "28", "29", "30", "31", "32", "33", "34", "35", "36", "
              "42", "44", "45", "46", "47", "48", "49", "50", "51", "53", "54", "55", "56", "66", "
# Loop through each state code and pull mean and max No2 values for all locations in each state
for code in state_code:
    response = requests.get(query_url + code).json()
```

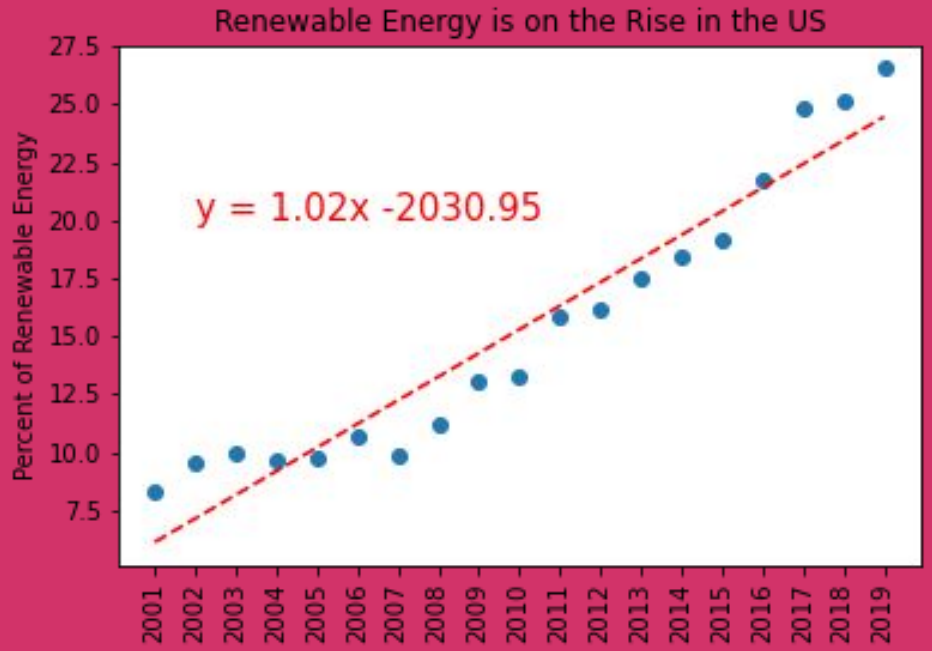
```
#grouby cities and remove duplicate city records
states20_grouped=states_20_df.groupby('City').first()
states19_grouped=states_19_df.groupby('City').first()
states18_grouped=states_18_df.groupby('City').first()
states17_grouped=states_17_df.groupby('City').first()
states16_grouped=states_16_df.groupby('City').first()
```

```
#merge the data frames
pol_merge_19_20=states20_grouped.merge(states19_grouped,how='outer',on=['City','State','Latitude','
pol_merge_19_20=pol_merge_19_20.rename(columns={'Mean Value_x':'2020 Mean','Max Value_x':'2020 Max'
pol_merge_18_20=pol_merge_19_20.merge(states18_grouped,how='outer',on=['City','State','Latitude','L
pol_merge_18_20=pol_merge_18_20.rename(columns={'Mean Value':'2018 Mean','Max Value':'2018 Max'})
pol_merge_17_20=pol_merge_18_20.merge(states17_grouped,how='outer',on=['City','State','Latitude','L
pol_merge_17_20=pol_merge_17_20.rename(columns={'Mean Value':'2017 Mean','Max Value':'2017 Max'})
pol_merge_16_20=pol_merge_17_20.merge(states16_grouped,how='outer',on=['City','State','Latitude','L
pol_merge_16_20=pol_merge_16_20.rename(columns={'Mean Value':'2016 Mean','Max Value':'2016 Max'})
#sort resulted merged data frame by state to have cleaner look
sorted_pol_merge=pol_merge_16_20.sort_values(by='State')
#drop null values to not skew data set
cleaned_sorted=sorted_pol_merge.dropna()
#drop lat and long, as not needed for further
dropped_clean=cleaned_sorted.drop(labels=['Latitude','Longitude'],axis=1)
#reindex to see line change in city
columns=['State','2020 Max','2019 Max','2018 Max','2017 Max','2016 Max','2020 Mean','2019 Mean','20

reindexed_df=dropped_clean.reindex(columns=columns)
reindexed_df.head(30)
```

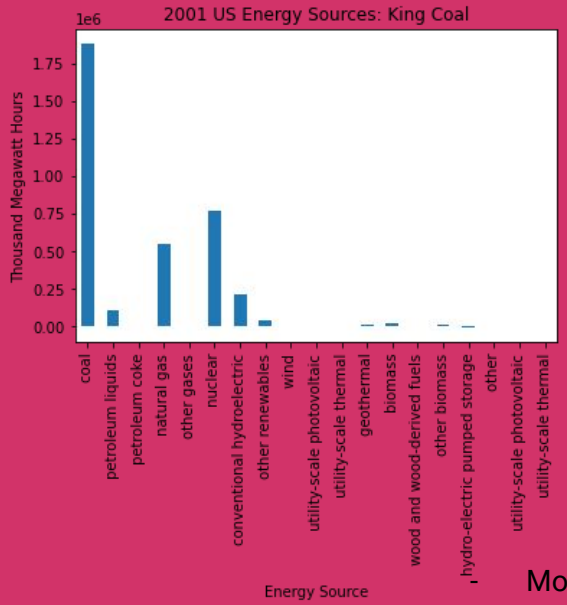
- Only able to pull one year at a time
- Had to pull each state by state code

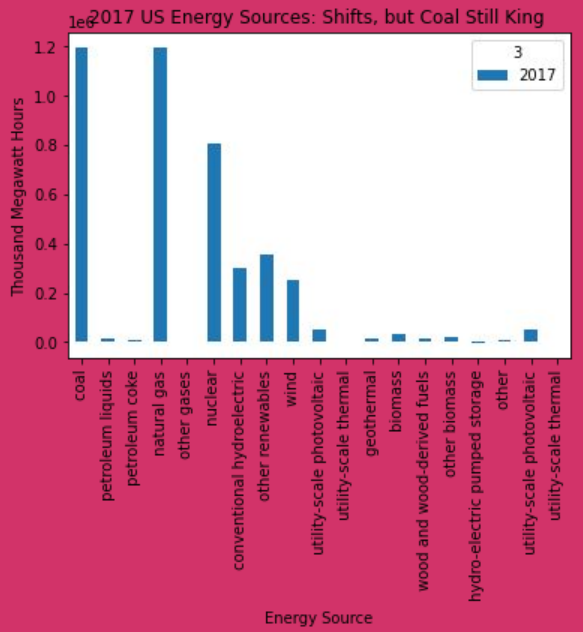
Renewable Energy Trends US

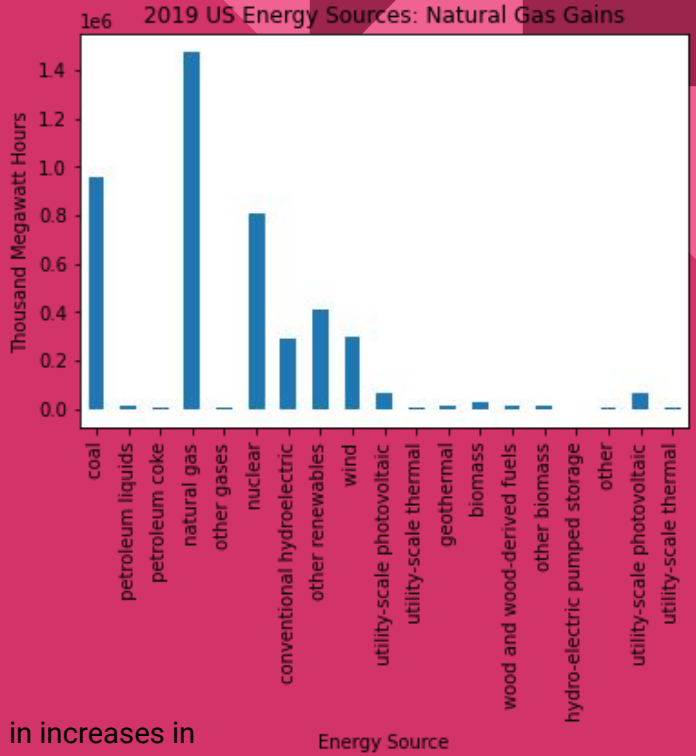


Renewable energy is increasing at 1.02% per year. At this rate, the US will reach 100% renewable energy in 2089. The forecasted renewable energy in 2020 is 25.47% and 26.49% in 2021.

US Energy Production Trends

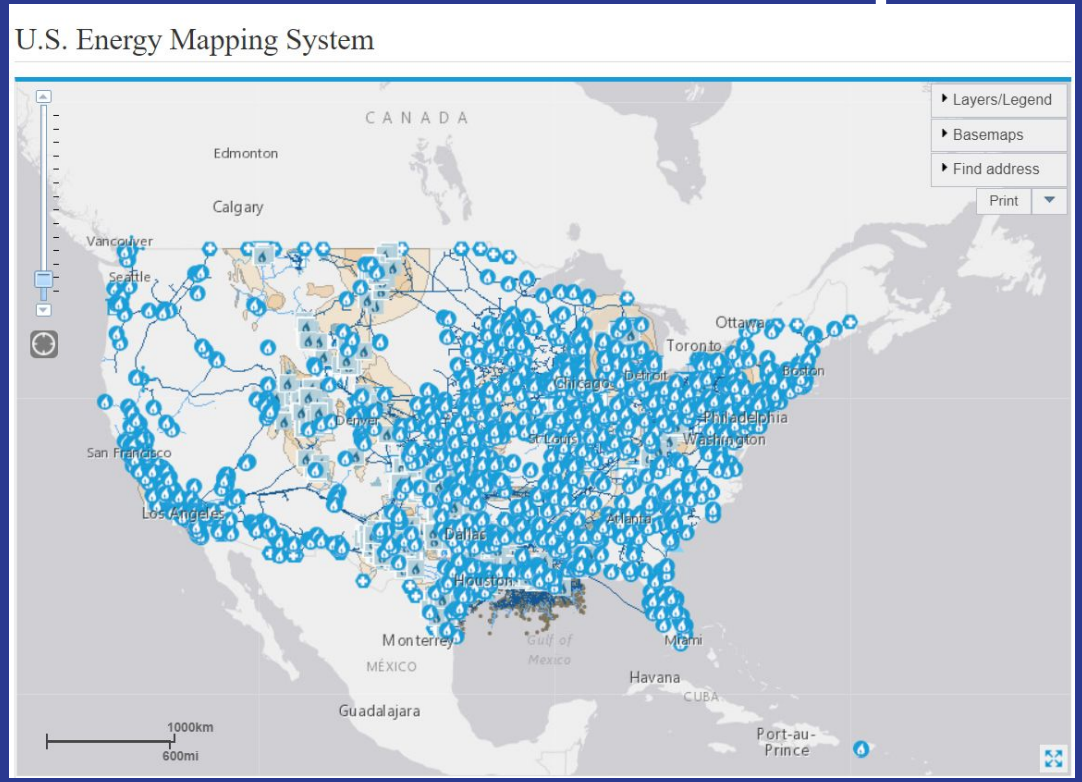






- More efficient extraction methods have resulted in increases in natural gas production from shale and other geologic sources that were previously inaccessible.
- While the US has made gains in renewable energy sources, production is still so heavily reliant on coal as to make it difficult to show all energy sources on the same chart.

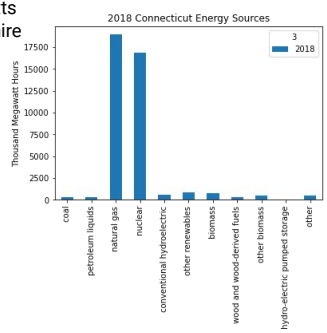
US Natural Gas Pipelines

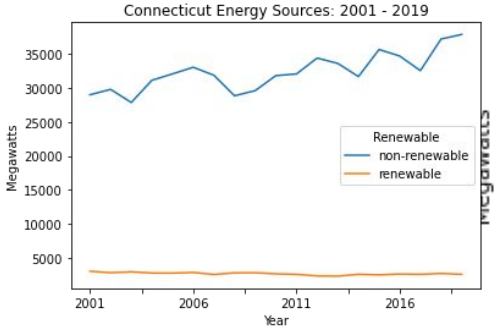


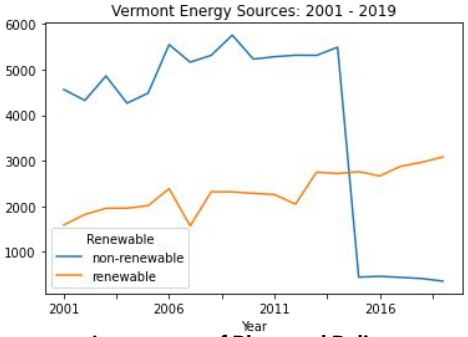
Source: https://www.eia.gov/state/maps.php

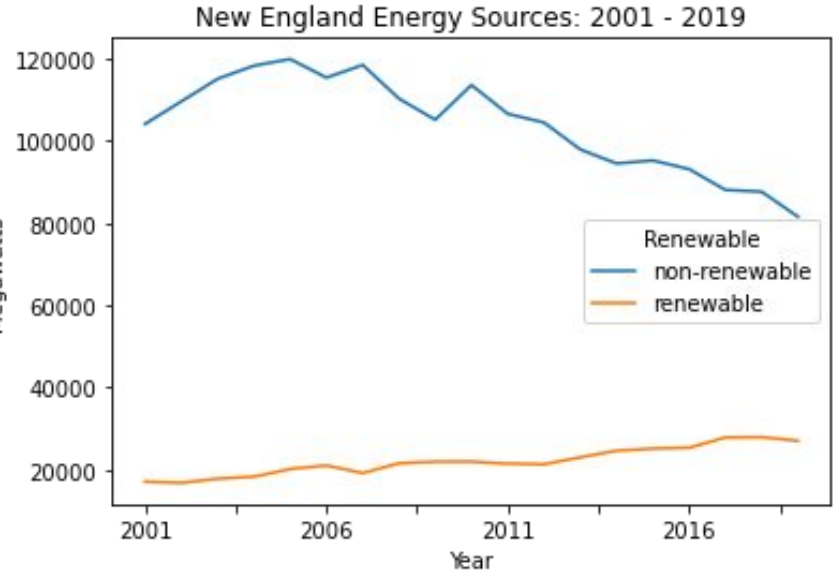
New England Region Energy Trends

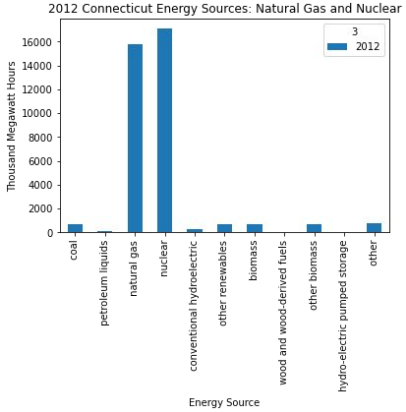
- Connecticut
- Maine
- Massachusetts
- New Hampshire
- Rhode Island
- Vermont





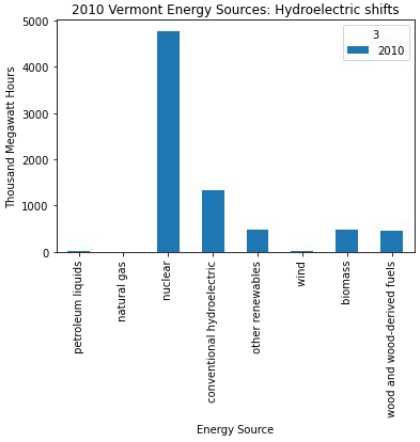


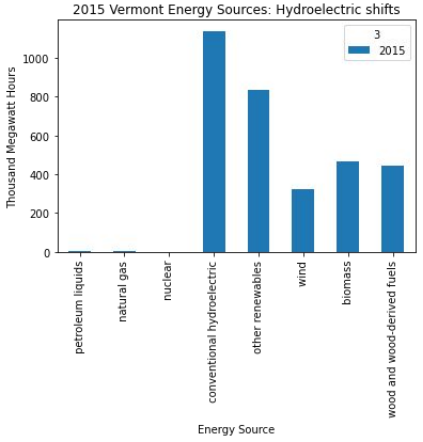




Importance of Place and Policy

- Vermont and Connecticut do not have any fossil fuel energy reserves.
- 2016 Vermont Comprehensive Energy Plan and the shutdown of the Yankee Nuclear Power Station in 2014 resulted in a drastic shift in energy sources.
- Largest share of electricity in Vermont comes from hydropower in Canada, which has its own environmental impacts.





Mountain Region Energy Trends

The Mountain region is classified in by EIA as containing the following states:

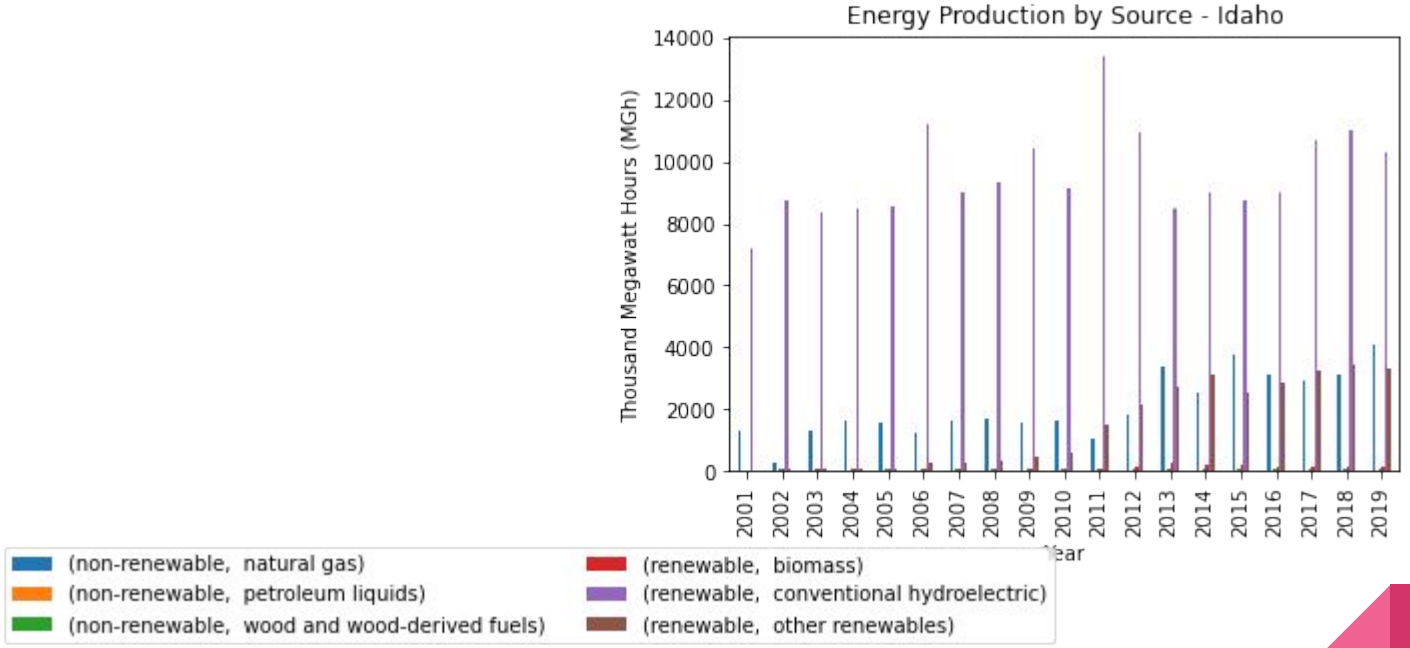
- Arizona
- Colorado
- Idaho
- Montana
- Nevada
- New Mexico
- Utah
- Wyoming

Similar to other regions, the Mountain region is dependent primarily on coal (51% of production), and natural gas (21% of production)

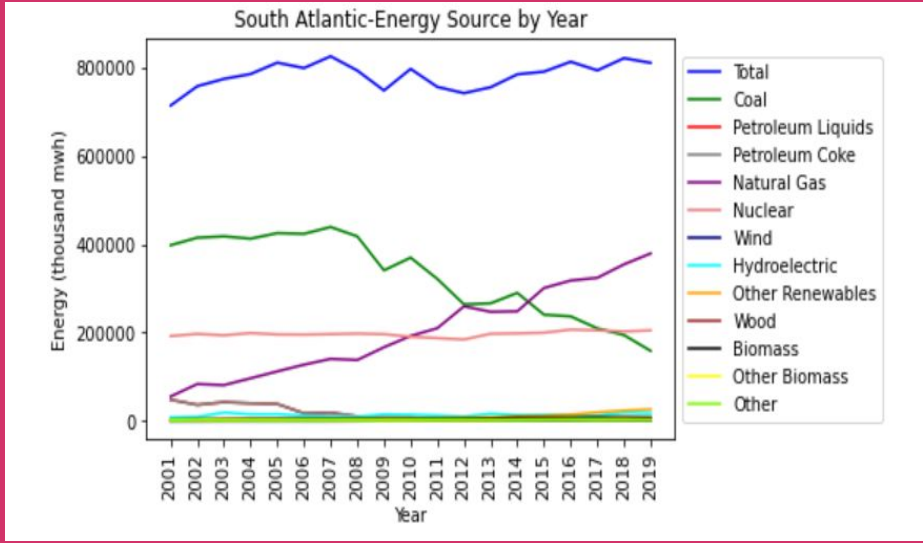
Over the 2001-19 period, the Mountain region had higher renewable energy as a percent of total production (19%), than the nation as a whole (16%).

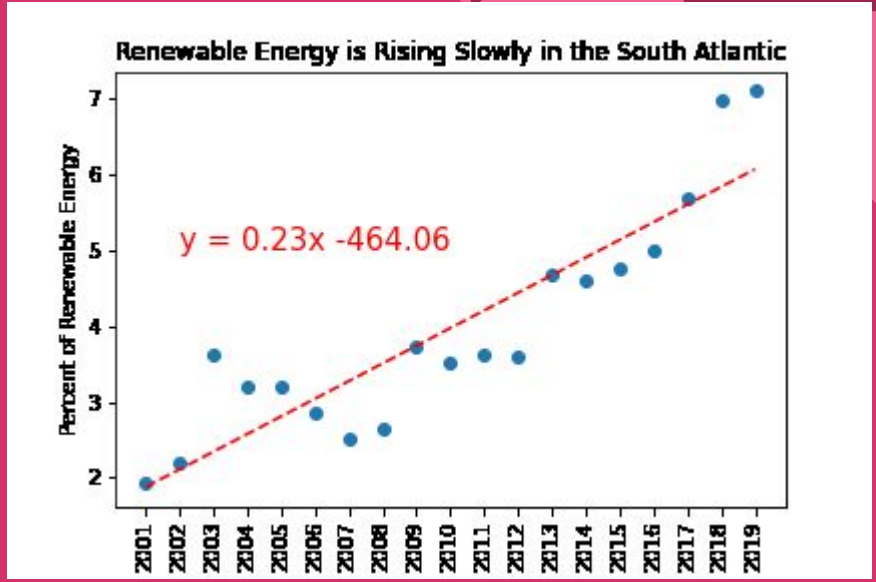
The Mountain state with the highest percentage renewable energy production is **Idaho**, with about **85%** of its energy being renewable. This is due to hydroelectric power infrastructure.

Idaho Energy Use



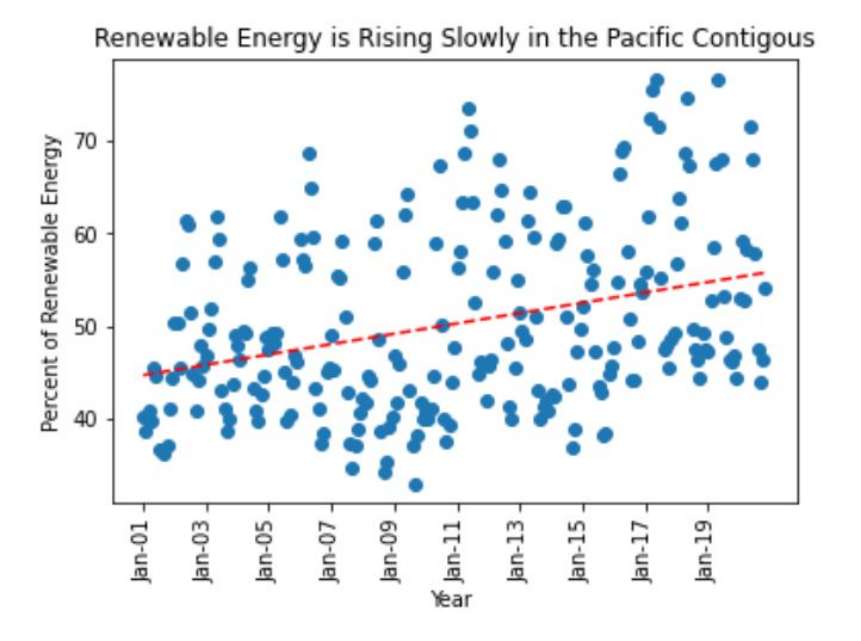
South Atlantic Energy Trends

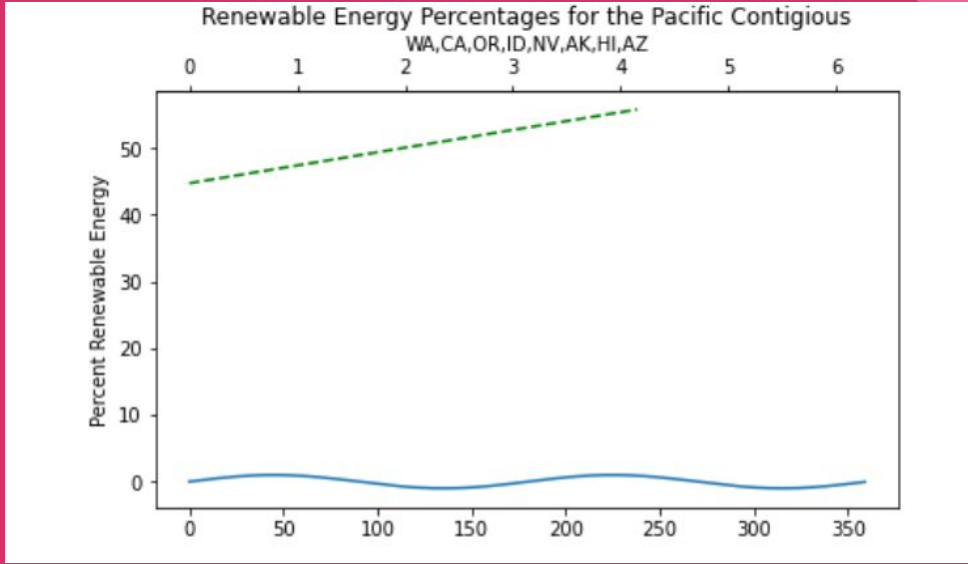




- Moving from coal-powered to natural gas in this region, but renewables are rising slowly.
- At this rate, it will take until 2452 for the South Atlantic region to reach 100% Renewable Energy.

Pacific Coast Energy Trends



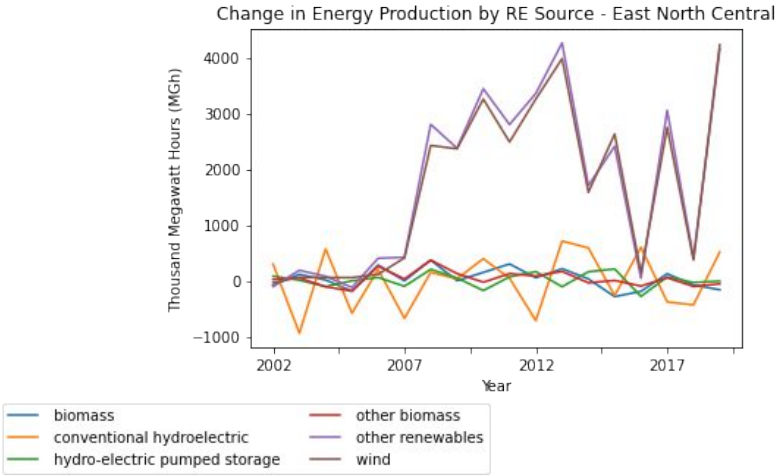


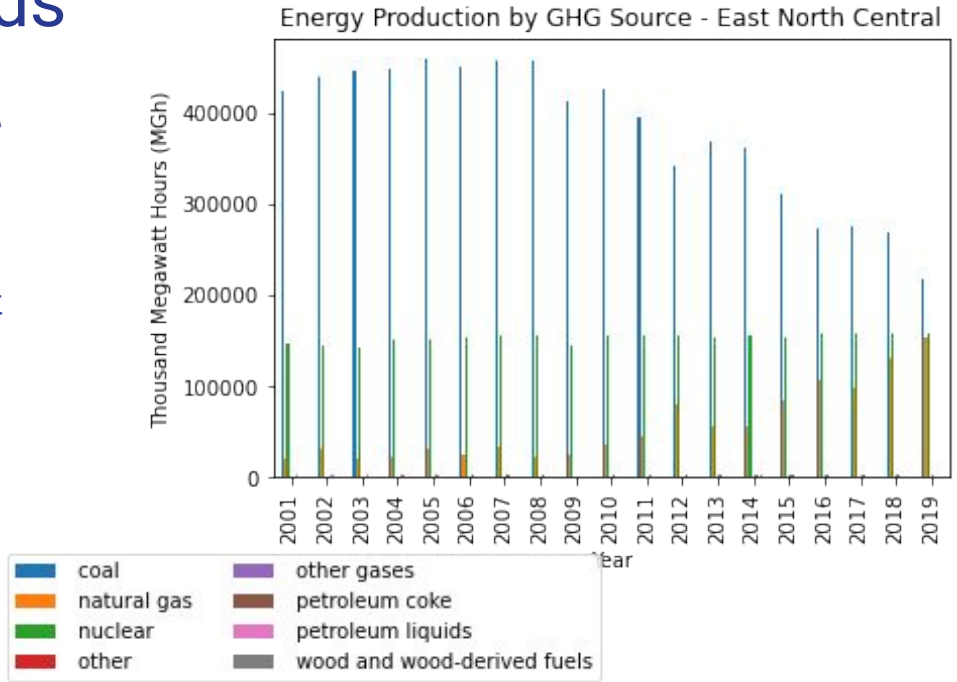
East Midwest Energy Trends

Coal overwhelmingly primary energy source

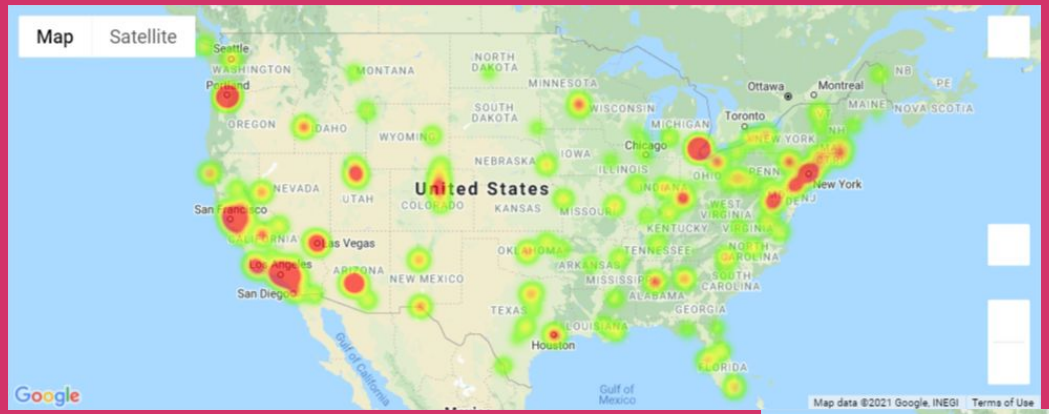
Natural gas 2nd

Wind & other renewables (solar, etc.) largest renewable





Air Pollution Trends



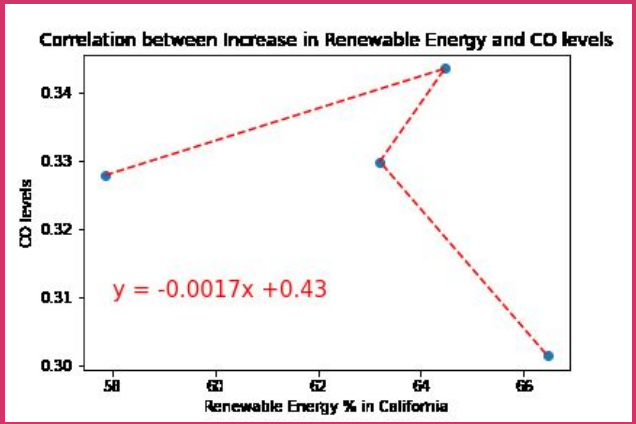
Largest source of CO is vehicles.

2020 Heatmap of US Carbon Monoxide readings in PPM. Heatmap readout can appear skewed due to monitoring locations.

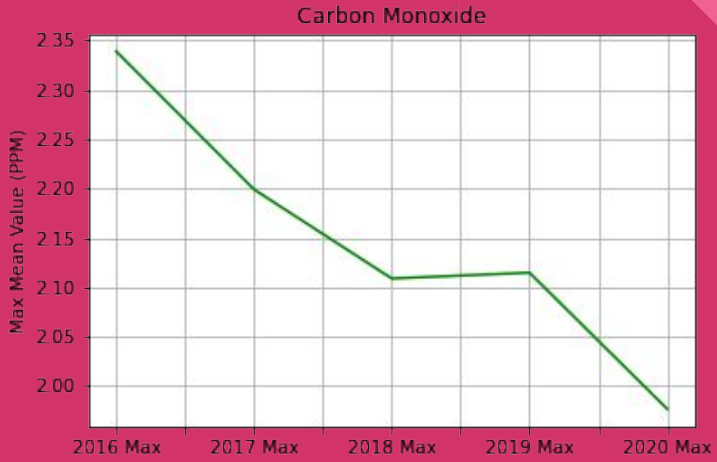
Does energy source impact air pollution?

CONCLUSION

Although, US has made strides in moving towards renewables, still heavily dependent on Coal and Natural Gas



Correlation between increase in Renewable Energy and CO levels in CA is -0.351, indicating a weak negative relationship.



2016 V 2020

Analysis & Challenges

- Challenges:
  - API calls for 1 year at a time
  - CSV file required a massive amount of cleaning/data formatting
- Additional Analysis:
  - COVID changes to environmental pollution
  - Look at different pollutants
  - Move to renewable energy -- challenges with energy storage