National Solar Power Usage

Project 1 for UCSD Extension Data Science and Analytics Bootcamp

"JFG" Team Members:
Dominic Gonzales
Gage Stout
Michelle Sullivan
Nicolas Tacos
Lorelei Wong

Motivation

Hypothesis:

- We are massively underutilizing solar power in the United States

Core questions:

- What does US solar power production look like?

Further questions:

- Is solar power as big of a trend in other states as it is in California? Do we love it because our energy is expensive?
- Are we underutilizing solar power?
- How much energy does the US use and how much of that is solar?
- Are states increasing solar production at a steady rate across the board or are some states doing better than others?

Motivation continued...

Why ask these questions?

- Importance of renewable energy
- Impact on our daily lives

Could we answer them?

- Sometimes!
- Finding exactly the right data sometimes got in our way

A Brief Summary of Findings

- Solar power is one of the least utilized energy sources in the US
- Coal makes up almost half of all energy production (1990-2019)
- California produces A LOT of solar power
- Energy prices in California are definitely increasing, as is solar power
- Other very sunny states are not producing at the same rate as California

Questions and Data

- How much energy is created by solar compared to other energy sources?
- What states could show potential for solar energy production?
  - Needed to understand potential sunshine hours for states and cities across the US
- How much solar power is the US producing by state ?
  - Last year
  - The past decade
- Does cost of energy affect a state's likelihood to produce more solar power?

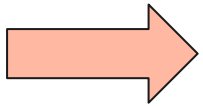
Data Clean-up and Exploration

- lambda functions are mini versions of a normal function without a return statement
- In the example above we used lambda to remove % symbol in our sunshine data in order to change the datatype to a float variable instead of an object

```
#strip the % column of the % symbol and convert to float
sunshine_df["Annual % Average Possible Sunshine"] = sunshine_df["Annual % Average Possible Sunshine"] \
                                                    .map(lambda x: x.rstrip('%'))
sunshine_df["Annual % Average Possible Sunshine"] = sunshine_df["Annual % Average Possible Sunshine"].astype(float)
sunshine_df
```

	CITY	STATE	ANNUAL % AVG POSSIBLE SUNSHINE	NO. OF YEARS OF DATA
0	YUMA	AZ	90%	42
1	REDDING	CA	88%	10

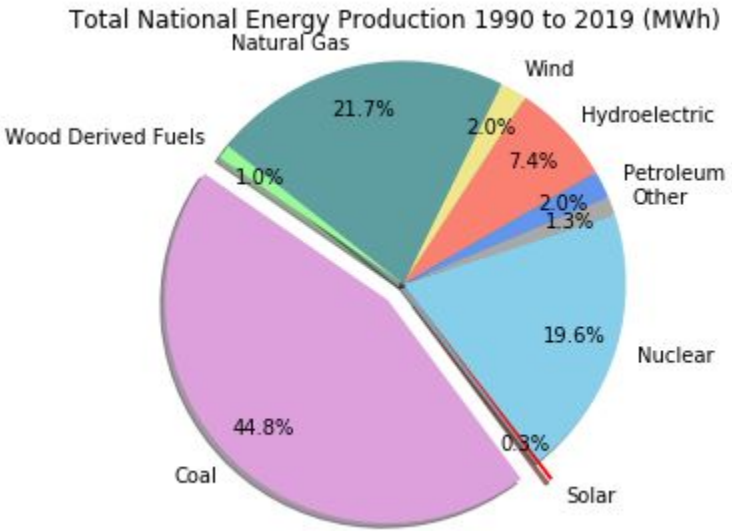
```
City                                 object
State                                object
Annual % Average Possible Sunshine   object
# of Data Years                      int64
dtype: object
```



	City	State	Annual % Average Possible Sunshine	# of Data Years
0	YUMA	AZ	90.0	42
1	REDDING	CA	88.0	10

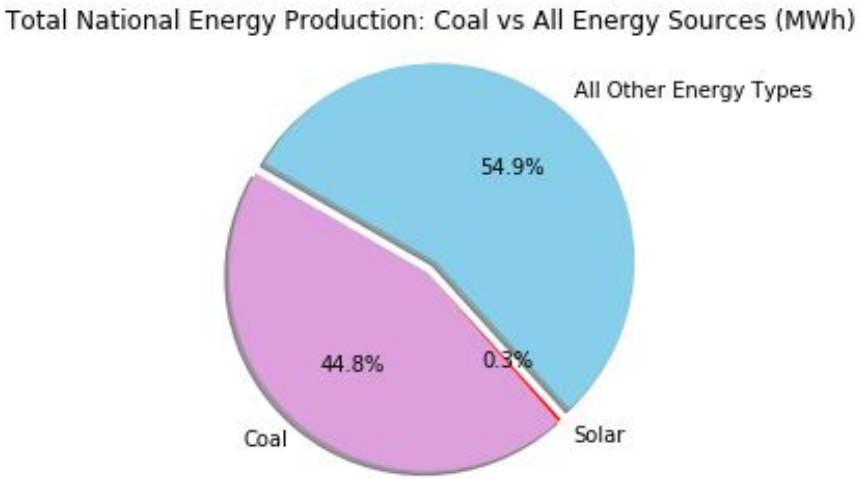
```
City                                 object
State                                object
Annual % Average Possible Sunshine   float64
# of Data Years                      int64
dtype: object
```

Data Analysis - National Energy Production



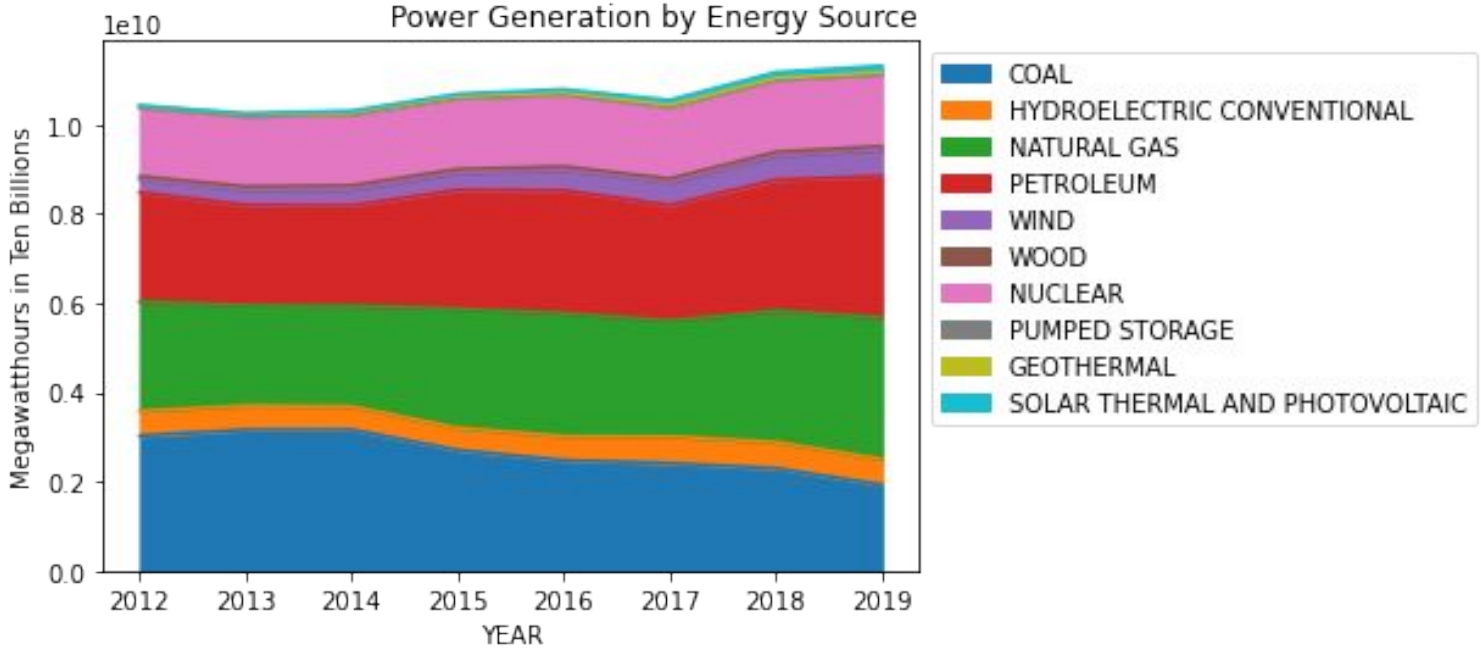
1. Solar power makes up for only 0.3% of energy production.
2. Largest contributors are coal, natural gas, and nuclear.
3. Other Category:
    a. Other biomass,
    b. Other gases,
    c. Pumped storage,
    d. Other (Not specified)

Data Analysis - National Energy Production



1. All other energy production sources combined equates to 54.9%.
2. Coal is by far the largest single contributor to energy production in the United states.
3. This visualization better grasps how underutilized solar energy.

Data Analysis - Energy Sources in the US



Data Analysis - Energy Sources in the US

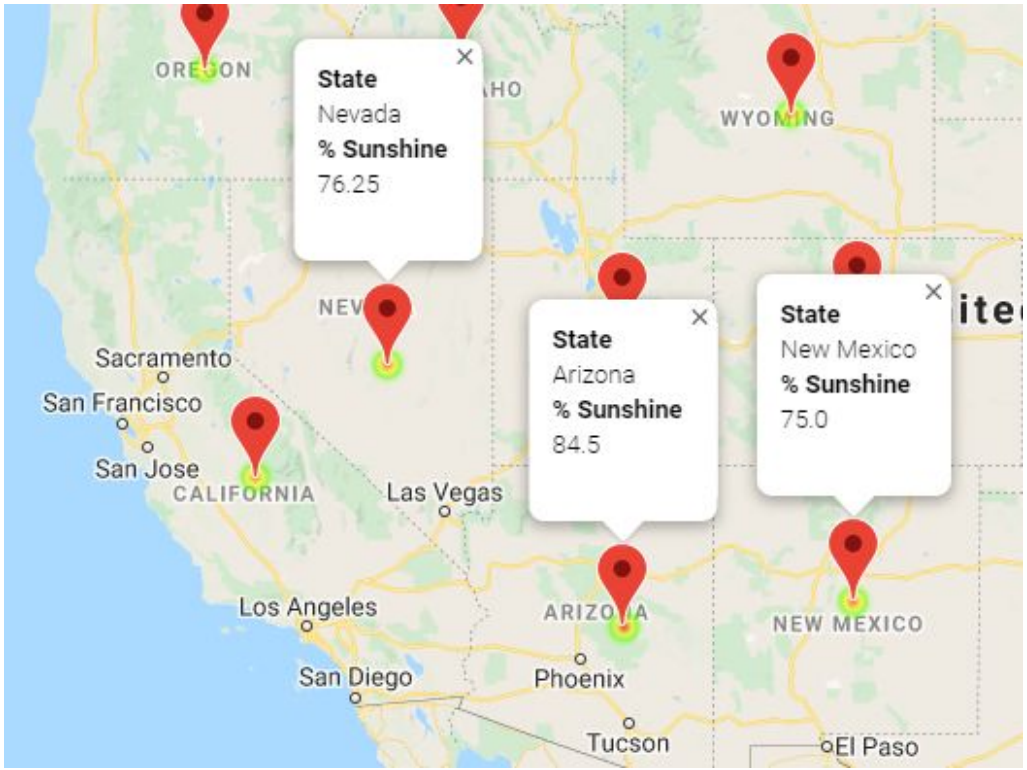
2012			2019			2012-2019		
1.	29.03%	Coal	1.	28.02%	Natural Gas	1.	24.76%	Coal
2.	23.51%	Petroleum	2.	28.02%	Petroleum	2.	24.65%	Petroleum
3.	23.51%	Natural Gas	3.	17.05%	Coal	3.	24.65%	Natural Gas
4.	14.75%	Nuclear	4.	14.30%	Nuclear	4.	14.92%	Nuclear
5.	5.29%	Hydroelectric	5.	5.21%	Wind	5.	5.14%	Hydroelectric
6.	2.70%	Wind	6.	5.08%	Hydroelectric	6.	4.05%	Wind
7.	0.72%	Wood	7.	1.27%	Solar	7.	0.76%	Wood
8.	0.29%	Geothermal	8.	0.68%	Wood	8.	0.66%	Solar
9.	0.09%	Pumped Storage	9.	0.27%	Geothermal	9.	0.29%	Geothermal
10.	0.08%	Solar	10.	0.09%	Pumped Storage	10.	0.11%	Pumped Storage

Data Analysis - Cities ≥ 70% Sunshine

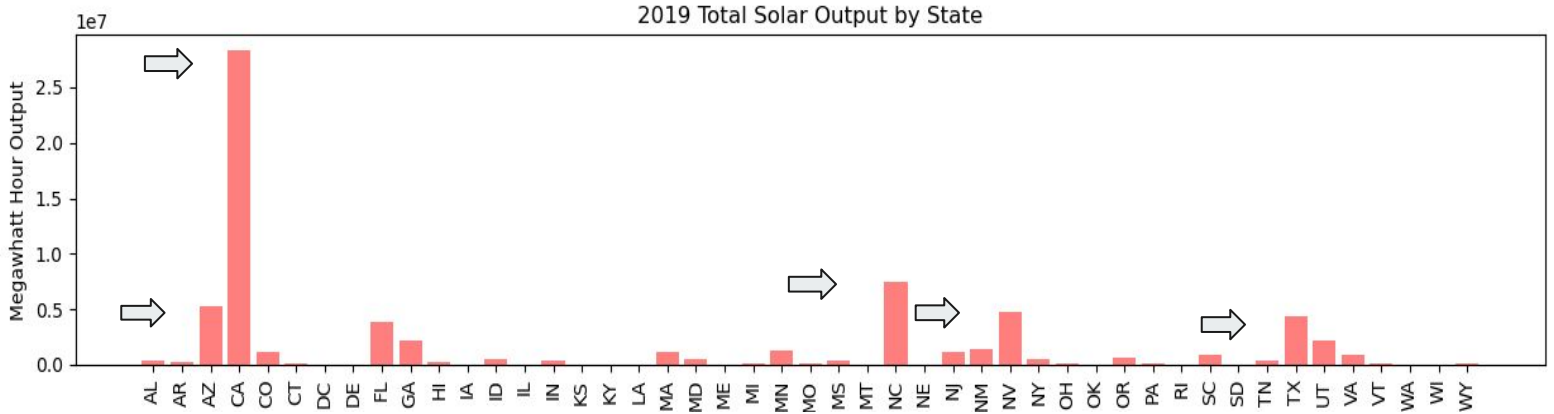
Data Analysis

Top State Avg % Sunshine:

1.	84.50%	Arizona
2.	76.25%	Nevada
3.	75.00%	New Mexico
4.	72.00%	Colorado
5.	71.86%	California
6.	68.00%	Utah
7.	66.83%	Florida
8.	65.75%	Kansas
9.	65.67%	Wyoming
10.	65.21%	Texas



Data Analysis - Solar Power by State



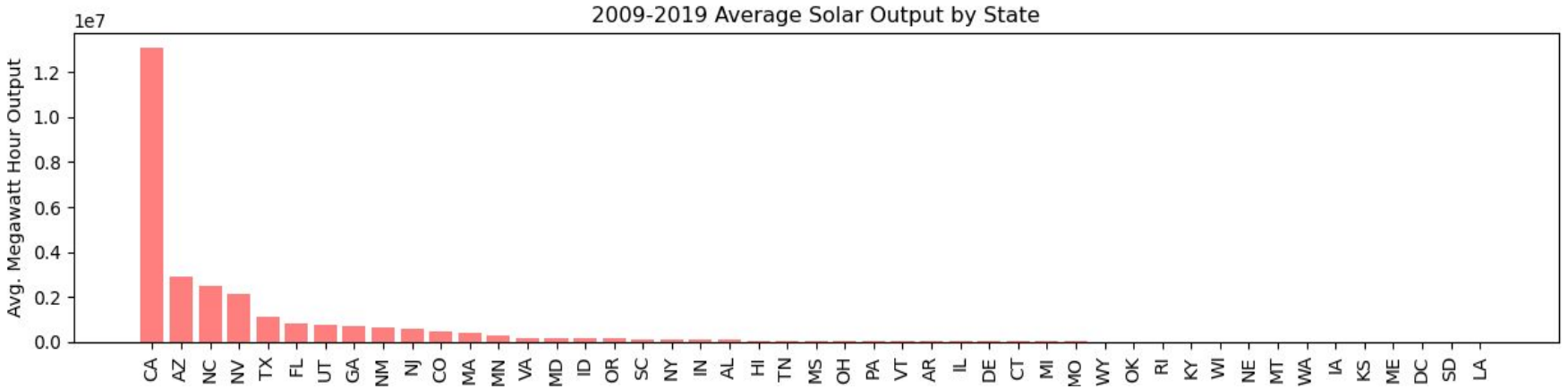
Top 5 States:

1. CA - 28,331,513
2. NC - 7,451,338
3. AZ - 5,278,019
4. NV - 4,810,511
5. TX - 4,365,125

Bottom 5 States:

1. LA - 1,506
2. SD - 1,829
3. ME - 7,050
4. KS - 11,323
5. IA - 15,436

Data Analysis - Solar Power by State



Top 5 States:

1. CA - 13,072,409
2. AZ - 2,890,447
3. NC - 2,471,711
4. NV - 2,137,167
5. TX - 1,149,194

Bottom 5 States:

1. LA - 544
2. SD - 592
3. ME - 2,451
4. KS - 2,823
5. IA - 3,188

Data Analysis - Is there a cost correlation?

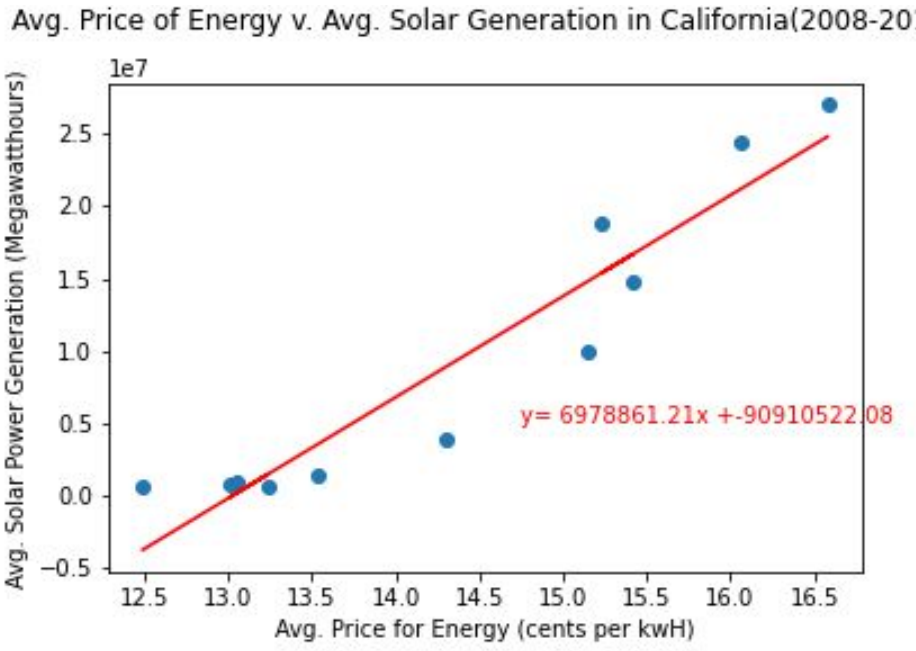
Hypothesis reminder:

- Californians like solar power because we pay so much for electricity and are trying to lower our costs.

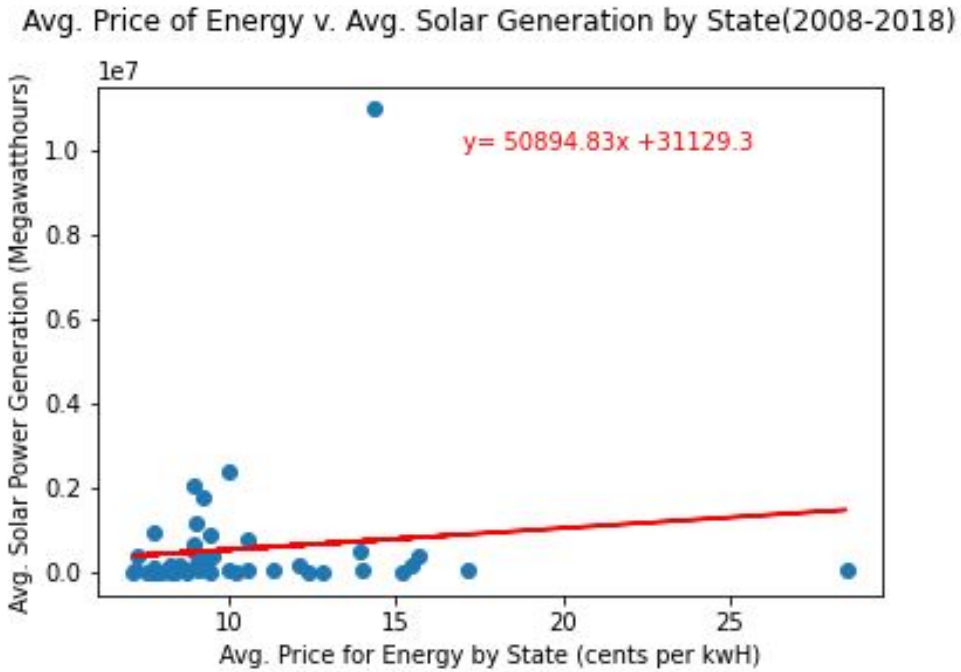
Is it true?

Let’s look at the data...

Data Analysis - Cost Correlation Cont...



Data Analysis - Cost Correlation Cont...



Discussion

- **Expected to see more** solar power usage throughout the US than we did.
- Our perception of solar power ubiquity is probably skewed by living in CA.
- Expected to see far more solar power creation in the Southwest US than we did. **North Carolina was a big surprise producer**!
- Solar power uptake seems partially based on sunshine availability but California's output figures suggest other **states are no where near their potential.**

Post Mortem

- Why has California solar production behaved this way?
- Do other states with high energy costs have similar correlations that we can't see in the ten year averaged all-state graph?
- To further investigate the potential solar power production for each state we would need to understand:
  - Potential sunshine hours/state
  - Solar power produced/sunshine hour
  - Capabilities and limitations of solar power
    - Space, maintenance, ROI, etc.
- Is solar the best option for power generation and distribution?
  - What would be the most efficient source of power for each state?

Questions?

Ask us questions!