GIS, 3D Mapping, Spatial Analysis and Data Science to analyze Solar feasibility

Reusable energy is on the rise, by combining the sciences of Geography, Data Science, Geographic Information System and Geospatial Analysis we can predict solar feasibility.

ABSTRACT

We have seen great increase in the use of renewable energy in recent years and the trend is here to stay. Data is continuously being updated, technology keeps on advancing and newer research and analysis keeps on adding to this growing area of interest. In this paper I have attempted to combine different fields of sciences to create a fuller model and help calculate the feasibility of solar energy. Detailed analysis will be added separate to the research paper in github/github pages.

INTRODUCTION

Aim of this paper is to provide basis and starting point for automated analysis engine that can work on address or region input and provide detailed analysis for solar feasibility. I have taken data and research that already exists and combined it to draw a better picture for solar potential. I will be comparing my result with the results achieved by researchers in the original papers and my data will be available on github.I have created a pointing system; the scale will evolve overtime. Initial proposal for categories

- Unit cost of electricity/Future unit cost
- Natural Solar potential: weather, length of days
- GIS Solar Potential Estimate
- Shadow analysis
- 3D/Urban Analysis
- Rooftop Area Analysis
- Energy demand/Future demand analysis
- Housing market value/Future value analysis
- Energy consumption growth

I have taken into consideration some very detailed papers and research recently published in all these categories to develop the scale. I am referring below the datasets and scientific research papers I am referred to develop the scaling system:

- Unit cost of electricity/Future unit cost
    - data.gov, US Census Bureau, USA.gov (1,2,3)
    - A linear regression pattern for electricity price forecasting in the Iberian electricity market (7)
- Natural Solar potential: weather, length of days
    - NOAA, The National Oceanic and Atmospheric Administration (9)
- GIS Solar Potential Estimate
    - ArcGIS (4)
- Shadow analysis
    - 3D Solar Potential in the Urban Environment: A Case Study in Lisbon (5)
- 3D/Urban Analysis
    - Applications of solar mapping in the urban environment (6)
- Rooftop Area Analysis
    - Applications of solar mapping in the urban environment (6)
- Energy demand /Future demand analysis
    - data.gov, US Census Bureau, USA.gov (1,2,3)
- Housing market value /Future value analysis
    - data.gov, US Census Bureau, USA.gov (1,2,3)
- Energy consumption growth analysis
    - data.gov, US Census Bureau, USA.gov (1,2,3)
    - Prediction of energy consumption: Variable regression or time series? A case in China (8)

ANALYSIS

Python scripts will be added to GITHUB and all the GIS maps will be available on ArcGIS and also WebApp created to view all the data in a single app with GIS maps. I will focus on areas from Harrisburg PA to Sayreville, NJ i.e from school to home. Following are the links:

Unit cost of electricity/Future unit cost

I used data from government organizations. data.gov, US Census Bureau, USA.gov (1,2,3), all of these have accurate historical data for USA. I have compared the regional cost and also plotted the data on GIS map using ArcGIS. For prediction of future cost to calculate if the model is sustainable I have used regression tools in Python scikit-learn (8) built on Numpy. And also Implemented and analyzed methods discussed in 'A linear regression pattern for electricity price forecasting in the Iberian electricity market' (7). Detailed analysis can be viewed on github pages.

First GIS map is average cost of electricity per customer by county, data is from 2016, following are the links to GIS map and Python script with analysis.

Python Annual Price analysis(cost dollars per KWh):

US dataset:
https://data.bls.gov/timeseries/APU000072610?amp%253bdata_tool=XGtable&output_view=data&include_graphs=true

PA Dataset:
https://data.bls.gov/timeseries/APUS12B72610?amp%253bdata_tool=XGtable&output_view=data&include_graphs=true

Python Code/Results:

In [22]:
```
import pandas as pd
```

In [23]:
```
US = pd.read_excel('US.xlsx', index_col=0, header=9)
```

In [24]:
```
PA = pd.read_excel('PA.xlsx', index_col=0, header=9)
```

In [25]:
```
print("US data")
```

```
US data
```

In [26]:
```
US
```

Out[26]:

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2010	0.124	0.123	0.125	0.126	0.127	0.132	0.133	0.133	0.132	0.127	0.125	0.125
2011	0.125	0.125	0.127	0.127	0.129	0.134	0.135	0.135	0.135	0.130	0.128	0.127
2012	0.128	0.128	0.127	0.127	0.129	0.135	0.133	0.133	0.133	0.128	0.127	0.127
2013	0.129	0.129	0.128	0.128	0.131	0.137	0.137	0.137	0.137	0.132	0.130	0.131
2014	0.134	0.134	0.135	0.131	0.136	0.143	0.143	0.143	0.141	0.136	0.134	0.135
2015	0.138	0.138	0.136	0.137	0.137	0.143	0.142	0.142	0.141	0.136	0.134	0.133
2016	0.134	0.134	0.134	0.134	0.133	0.138	0.139	0.139	0.139	0.134	0.131	0.133
2017	0.134	0.135	0.134	0.135	0.137	0.142	0.143	0.142	0.142	0.137	0.136	0.136
2018	0.135	0.135	0.135	0.134	0.136	0.139	0.139	0.139	0.138	0.136	0.134	0.135
2019	0.135	0.136	0.135	0.135	0.136	0.139	0.140	0.139	0.139	0.136	0.133	0.133
2020	0.134	0.134	0.134	0.133	0.134	0.137	0.137	0.137	0.137	0.135	0.136	NaN

In [27]:
```
print("PA data")
```

```
PA data
```

In [28]:
```
PA
```

Out[28]:

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2010	0.163	0.163	0.163	0.162	0.162	0.173	0.173	0.173	0.174	0.161	0.160	0.160
2011	0.163	0.165	0.164	0.163	0.163	0.167	0.172	0.173	0.172	0.169	0.169	0.169
2012	0.163	0.164	0.162	0.163	0.164	0.167	0.159	0.159	0.159	0.166	0.165	0.165
2013	0.156	0.157	0.157	0.163	0.163	0.159	0.159	0.159	0.164	0.161	0.162	0.164
2014	0.162	0.162	0.157	0.156	0.156	0.157	0.159	0.158	0.156	0.154	0.154	0.159
2015	0.159	0.160	0.156	0.157	0.156	0.160	0.159	0.159	0.158	0.155	0.155	0.155
2016	0.160	0.159	0.157	0.156	0.155	0.157	0.158	0.157	0.158	0.153	0.152	0.151
2017	0.152	0.151	0.150	0.151	0.152	0.153	0.153	0.152	0.152	0.145	0.148	0.150
2018	0.147	0.148	0.145	0.145	0.153	0.155	0.155	0.154	0.152	0.149	0.151	0.150
2019	0.118	0.155	0.155	0.155	0.156	0.156	0.156	0.155	0.154	0.152	0.152	0.153
2020	0.153	0.154	0.153	0.152	0.151	0.154	0.155	0.154	0.153	0.151	0.150	NaN

In [29]:
```
print("US annual averages")
```

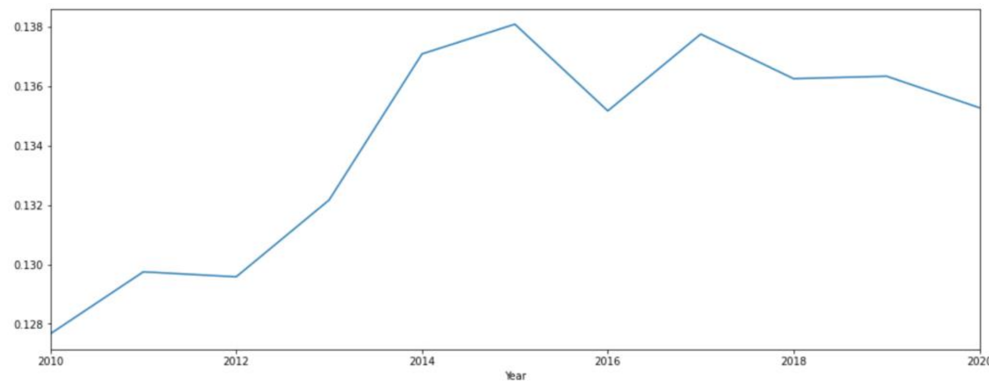
```
US annual averages
```

In [30]:
```
US_mean = US.mean(axis = 1)
US_mean
```

Out[30]:
```
Year
2010    0.127667
2011    0.129750
2012    0.129583
2013    0.132167
2014    0.137083
2015    0.138083
2016    0.135167
2017    0.137750
2018    0.136250
2019    0.136333
2020    0.135273
dtype: float64
```

In [31]:
```
US_mean_plot = US_mean.plot(kind = 'line', figsize=(16,6))
US_mean_plot
```

Out[31]:
```
<AxesSubplot:xlabel='Year'>
```



In [32]:
```
print("US monthly averages")
```

```
US monthly averages
```

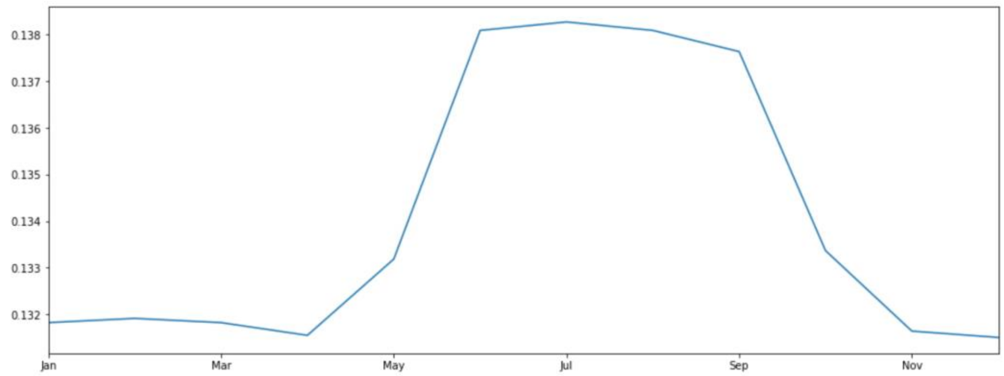
In [33]:
```
US_mean_mon = US.mean(axis = 0)
US_mean_mon
```

Out[33]:
```
Jan    0.131818
Feb    0.131909
Mar    0.131818
Apr    0.131545
May    0.133182
Jun    0.138091
Jul    0.138273
Aug    0.138091
Sep    0.137636
Oct    0.133364
Nov    0.131636
Dec    0.131500
dtype: float64
```

In [34]:
```
US_mean_mon_plot = US_mean_mon.plot(kind = 'line', figsize=(16,6))
US_mean_mon_plot
```

```
C:\Users\User\Anaconda3\lib\site-packages\pandas\plotting\_matplotlib\core.py:1182: UserWarning: FixedFormatt
er should only be used together with FixedLocator
  ax.set_xticklabels(xticklabels)
```

Out[34]: <AxesSubplot:>



In [35]:
```
print("PA annual averages")
```

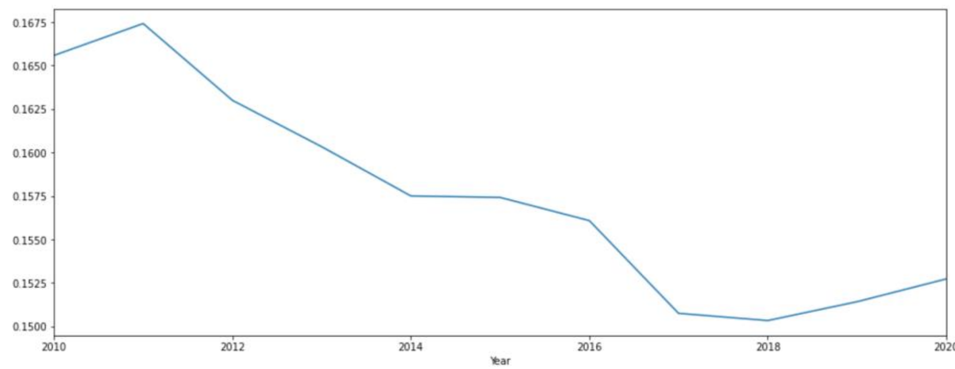
```
PA annual averages
```

In [36]:
```
PA_mean = PA.mean(axis = 1)
PA_mean
```

Out[36]:
```
Year
2010    0.165583
2011    0.167417
2012    0.163000
2013    0.160333
2014    0.157500
2015    0.157417
2016    0.156083
2017    0.150750
2018    0.150333
2019    0.151417
2020    0.152727
dtype: float64
```

In [37]:
```
PA_mean_plot = PA_mean.plot(kind = 'line', figsize=(16,6))
PA_mean_plot
```

Out[37]: <AxesSubplot:xlabel='Year'>



In [38]:
```
print("PA monthly averages")
```

```
PA monthly averages
```

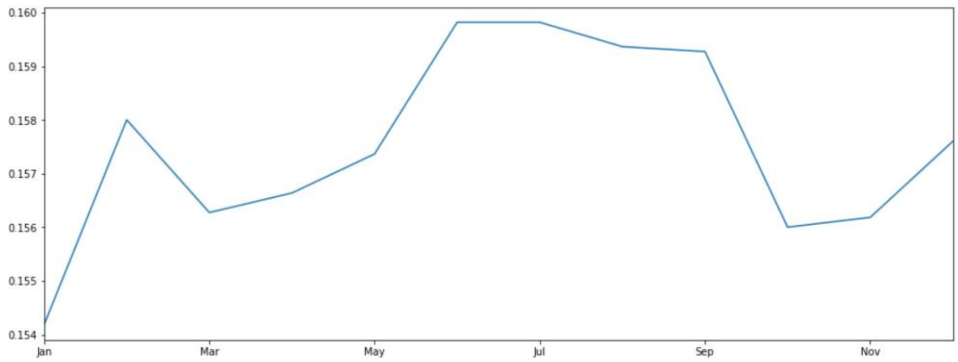
In [39]:
```
PA_mean_mon = PA.mean(axis = 0)
PA_mean_mon
```

Out[39]:
```
Jan    0.154182
Feb    0.158000
Mar    0.156273
Apr    0.156636
May    0.157364
Jun    0.159818
Jul    0.159818
Aug    0.159364
Sep    0.159273
Oct    0.156000
Nov    0.156182
Dec    0.157600
dtype: float64
```

In [40]:
```
PA_mean_mon_plot = PA_mean_mon.plot(kind = 'line', figsize=(16,6))
PA_mean_mon_plot
```

```
C:\Users\User\Anaconda3\lib\site-packages\pandas\plotting\_matplotlib\core.py:1182: UserWarning: FixedFormatt
er should only be used together with FixedLocator
  ax.set_xticklabels(xticklabels)
```

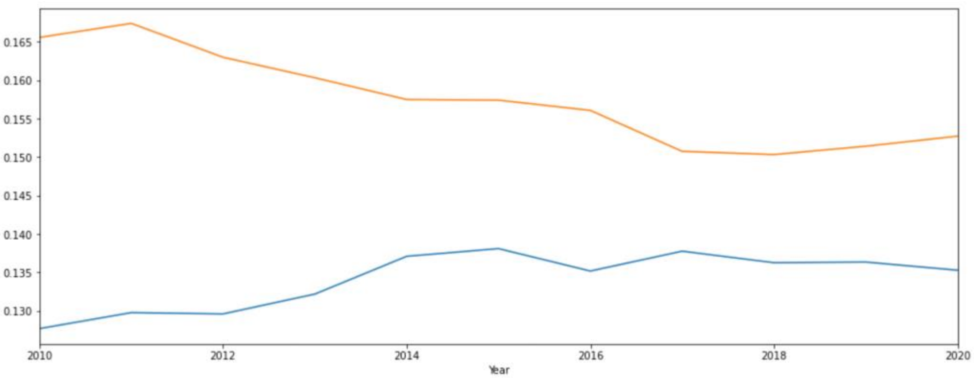
Out[40]: <AxesSubplot:>



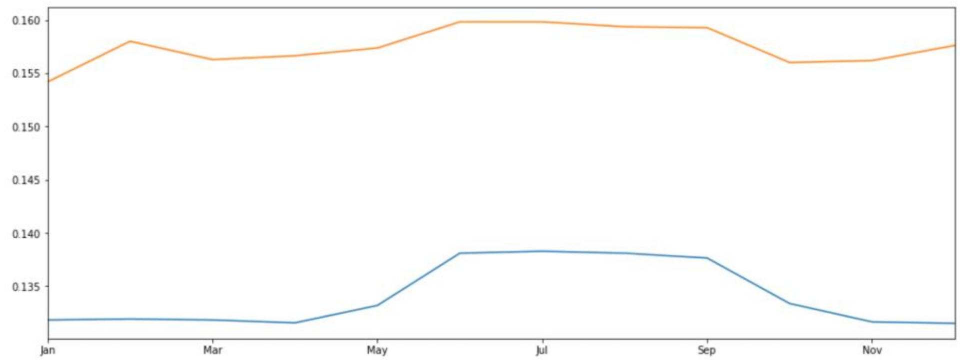
print("Comparing PA prices to US average prices")

In [41]:
```
ax = US_mean.plot()

PA_mean.plot(ax=ax, figsize=(16,6))
```

Out[41]: <AxesSubplot:xlabel='Year'>



```
In [42]: ax = US_mean_mon.plot()

         PA_mean_mon.plot(ax=ax, figsize=(16,6))

C:\Users\User\Anaconda3\lib\site-packages\pandas\plotting\_matplotlib\core.py:1182: UserWarning: FixedFormatt
er should only be used together with FixedLocator
  ax.set_xticklabels(xticklabels)

Out[42]: <AxesSubplot:>
```



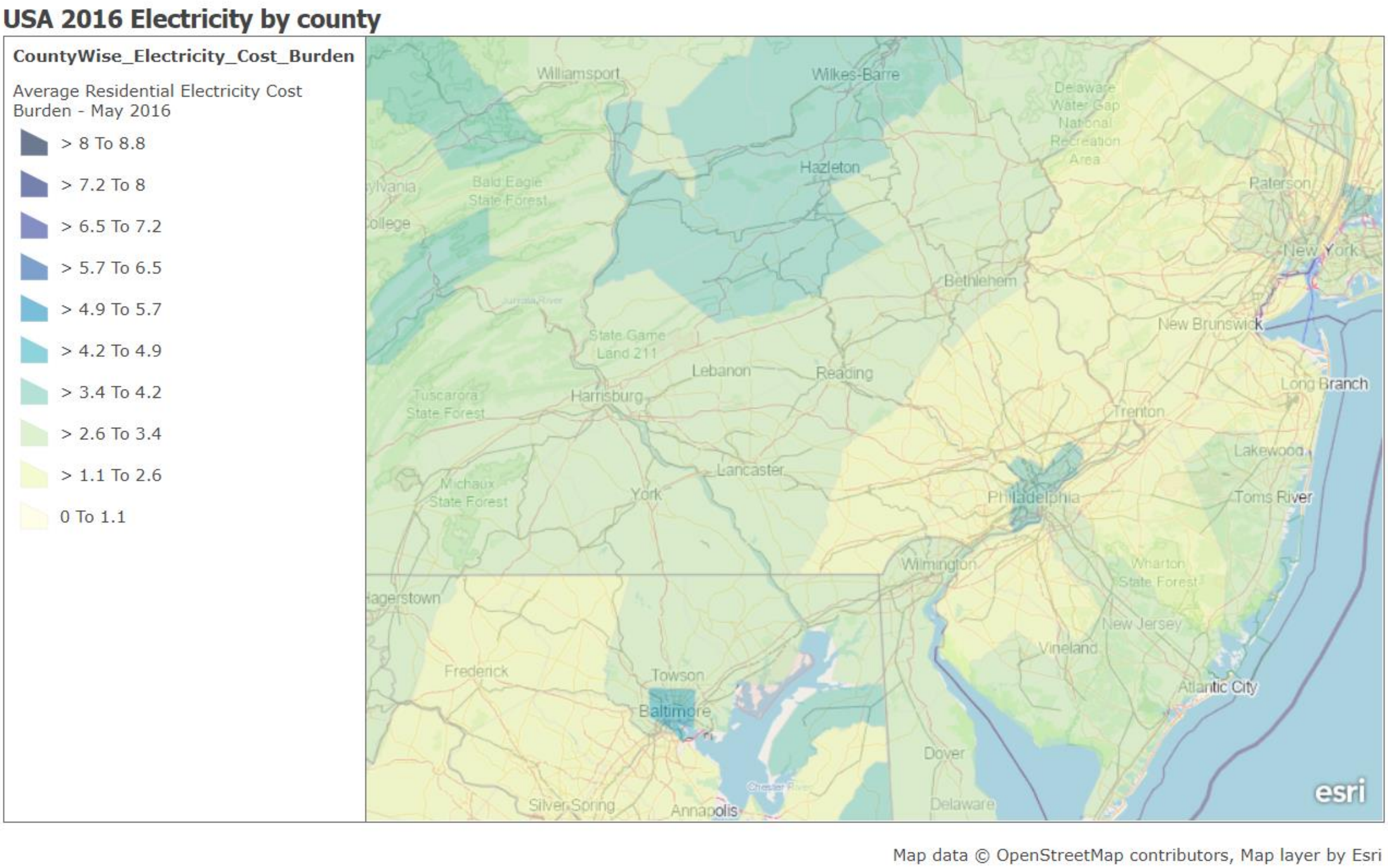

```
In [ ]:

In [ ]:
```

GIS Mapping:

GIS Dynamic MAP:
https://www.arcgis.com/apps/View/index.html?appid=4e9976871cd24a53b8b0b8e1477b0a6b

RAW DataSet: https://www.arcgis.com/home/item.html?id=c670ccbaef1c42d7a58057cd461578ec

Below are the static images of analysis and GIS map.



The ‘Cost Burden’ map does show the stark difference between cities and suburbs, cities look like a good place for solarization but we will look further into it as the cost of land and area of houses are very different from the suburbs.

Natural Solar potential: weather, length of days

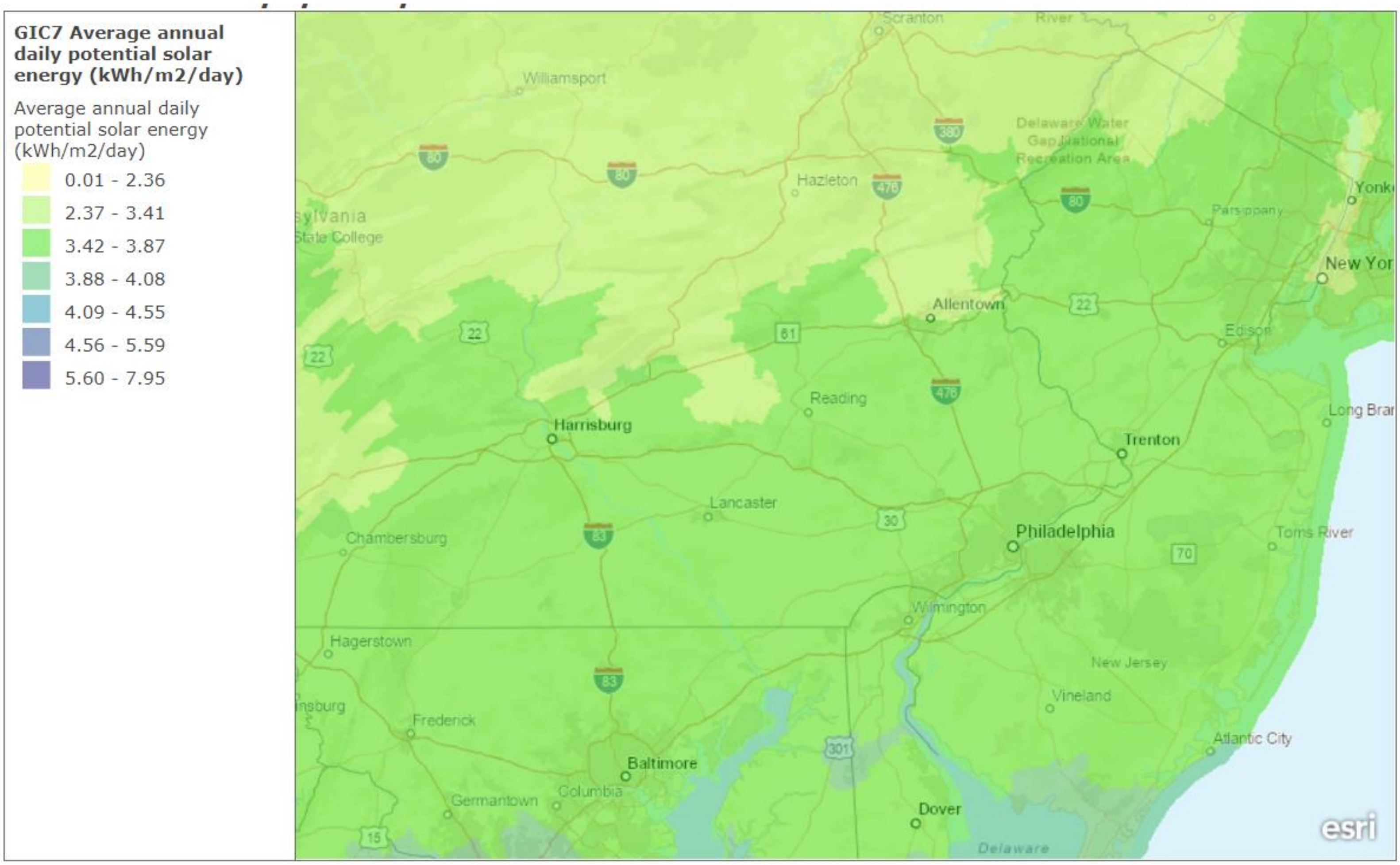
NOAA The National Oceanic and Atmospheric Administration (9), scientific agency within the United States Department of Commerce has extensive datasets. Regression analysis is not necessary for this data set, so I have utilized simpler mean and mode for analysis of data. I have attempted to use regression for future prediction but weather models are not easy to predict. I have also used ArcGIS for analyzing and comparing regional disparities.

GIS Mapping:

Below is the GIS map with average daily annual solar potential

GIS Dynamic App:
https://www.arcgis.com/apps/View/index.html?appid=f6b412741bd049719e306190e5e0a074

Static Map:



EnviroAtlas | Esri, HERE, Garmin, FAO, NOAA, USGS, EPA, NPS

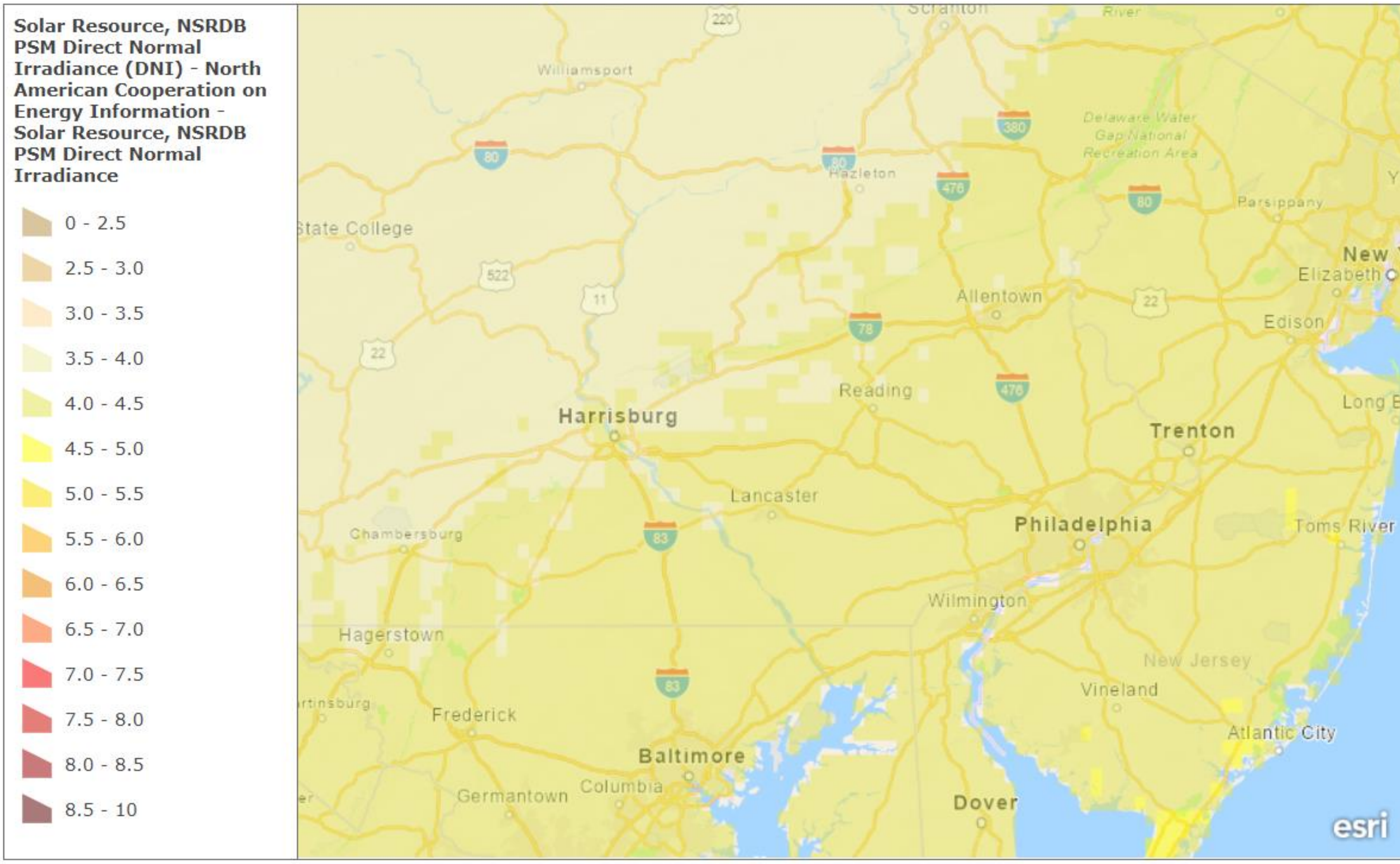
GIS Solar Potential Estimate

Solar potential estimation can be done using ArcGIS with available raster data. There are a lot of raster data and layers available that can be used for solar potential. I have used a couple of them and used ESRIs lessons to implement explorable GIS map to view and explore solar potential of different regions.

GIS Mapping:

GIS App for solar estimate:
https://www.arcgis.com/apps/View/index.html?appid=a56b5b1dec7142e3a21db3e8243adca3

Static Map for solar Estimate:



National Renewable Energy Laboratory ("NREL"), Alliance for Sustainable Energy, LLC, U.S. Department of Energy ("DOE"). This GIS data was developed by the National Renewable Energy Laboratory ("NREL"), which is operated by the Alliance for Sustainable Energy, LLC for the U.S. Department of Energy ("DOE"). The user is granted the right, without any fee or cost, to use, copy, modify, alter, enhance and distribute this data for any purpose whatsoever, provided that this entire notice appears in all copies of the data. Further, the user of this data agrees to credit NREL in any publications or software that incorporate or use the data. Access to and use of the GIS data shall further impose the following obligations on the User. The names DOE/NREL may not be used in any advertising or publicity to endorse or promote any product or commercial entity using or incorporating the GIS data unless specific written authorization is obtained from DOE/NREL. The User also understands that DOE/NREL shall not be obligated to provide updates, support, consulting,

Solar Potential Utilization:

GIS Map app:
https://www.arcgis.com/apps/View/index.html?appid=98fbe0f78ef74674b364bb75b7559a82

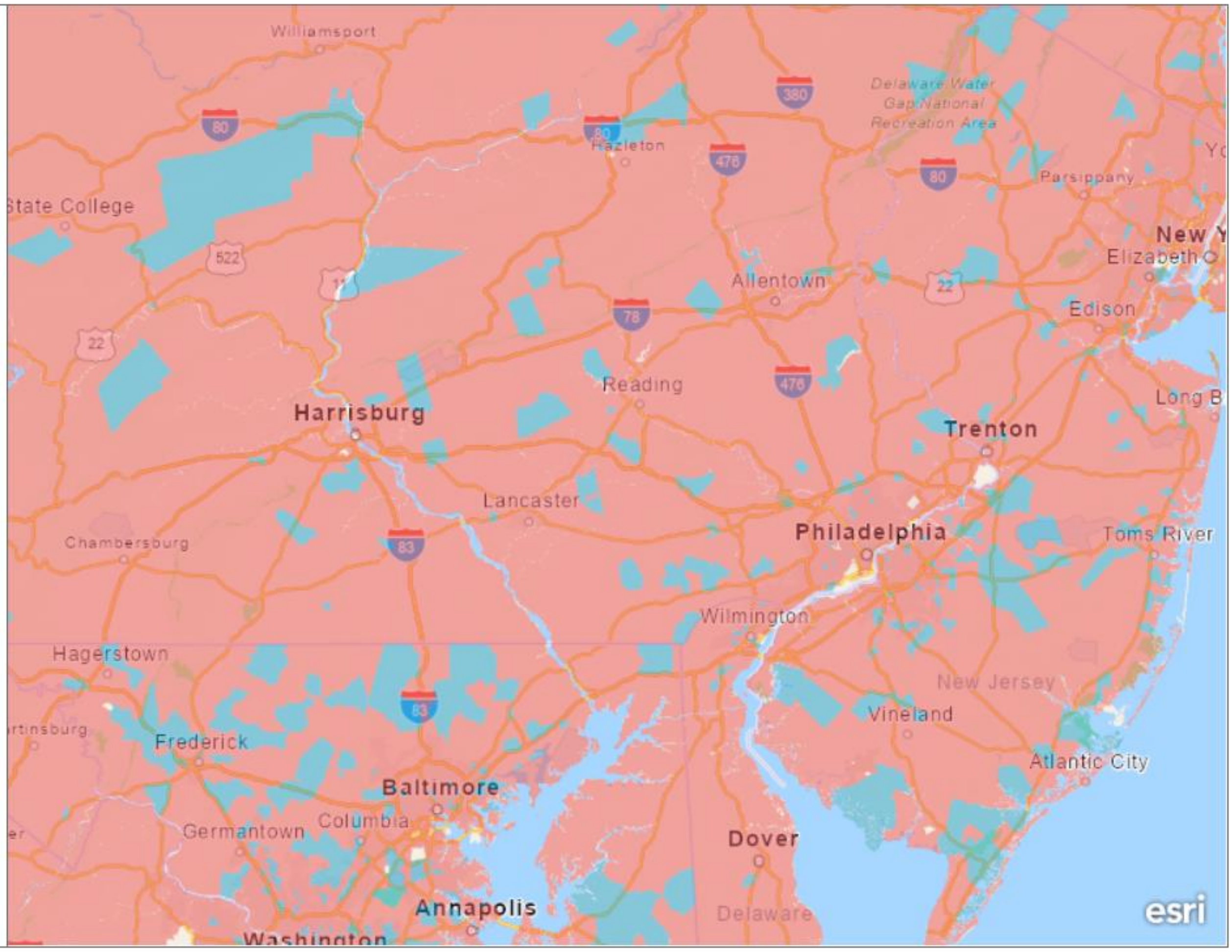
GIS Static Map

Solar Utilization

Untapped Solar Potential

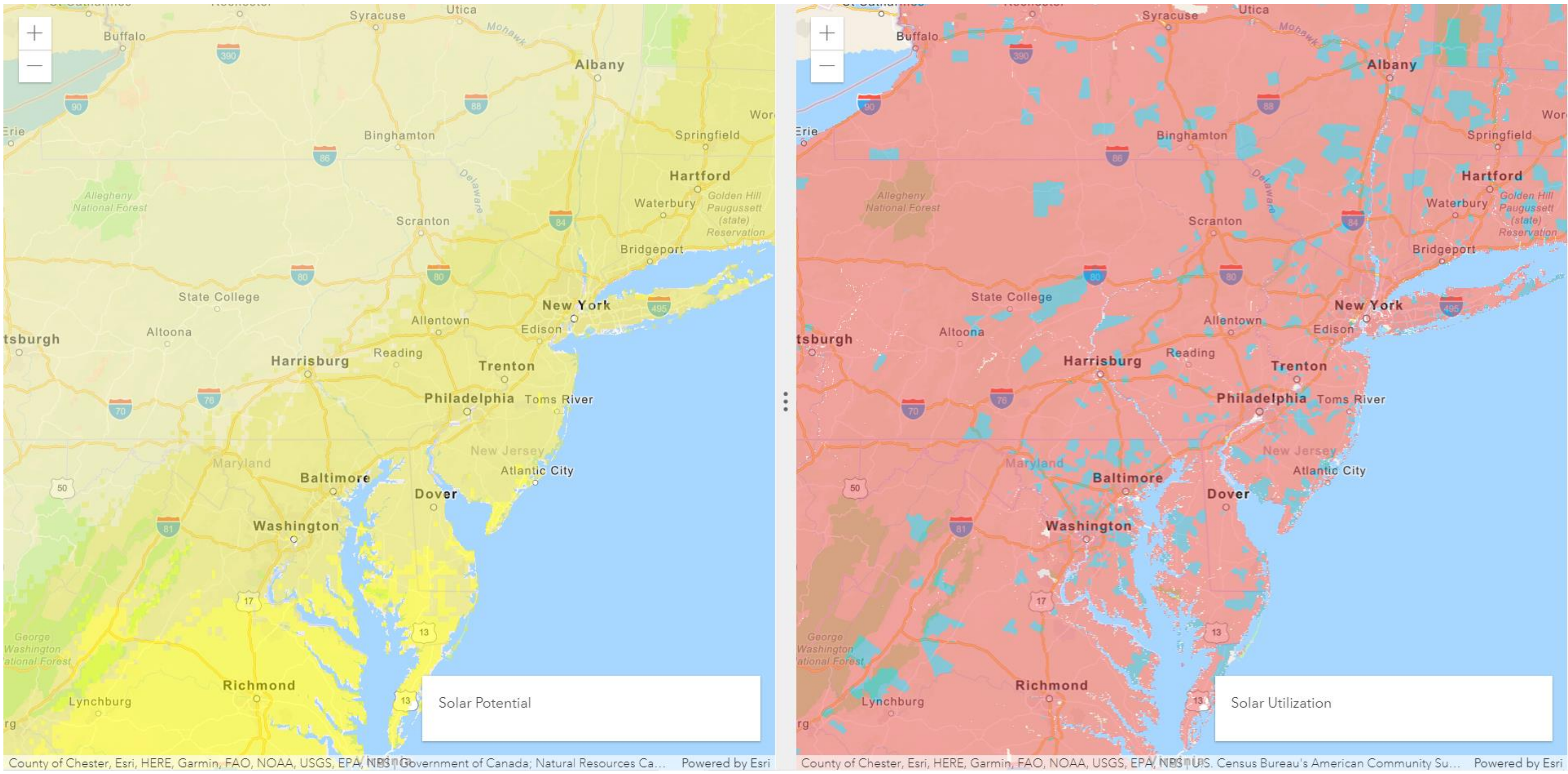
Tract Status

- This area has low solar potential
- Households here are already using solar



National Renewable Energy Laboratory ("NREL"), Alliance for Sustainable Energy, LLC, U.S. Department of Energy ("DOE"). This GIS data was developed by the National Renewable Energy Laboratory ("NREL"), which is operated by the Alliance for Sustainable Energy, LLC for the U.S. Department of Energy ("DOE"). The user is granted the right, without any fee or cost, to use, copy, modify, alter, enhance and distribute this data for any purpose whatsoever, provided that this entire notice appears in all copies of the data. Further, the user of this data agrees to credit NREL in any publications or software that incorporate or use the data. Access to and use of the GIS data shall further impose the following obligations on the User. The names DOE/NREL may not be used in any advertising or publicity to endorse or promote any product or commercial entity using or incorporating the GIS data unless specific written authorization is obtained from DOE/NREL. The User also understands that DOE/NREL shall not be obligated to provide updates, support, consulting,

Comparing Solar Potential to Solar utilization:
https://www.arcgis.com/apps/Compare/index.html?appid=6d9b06e4a8ce458ea059e945ce373dbb



Shadow analysis

For shadow analysis, I analyzed ‘3D Solar Potential in the Urban Environment: A Case Study in Lisbon’, shadow analysis has not improved dramatically in recent years, as light and how it interacts with buildings and contours has not changed, my application of the shadow analysis is on github.

3D/Urban Analysis

I analyzed ‘Applications of solar mapping in the urban environment’ for this section. It is a very detailed subject; the research is comprehensive and accurate. I have recreated the research by utilizing different geographic location and added GIS mapping to the analysis.

Rooftop Area Analysis

This is an extension of the 3D analysis, with rooftop analysis I also added GIS mapping and regression to create mappable data layers and raster data. GIS maps and layers are on ArcGIS.

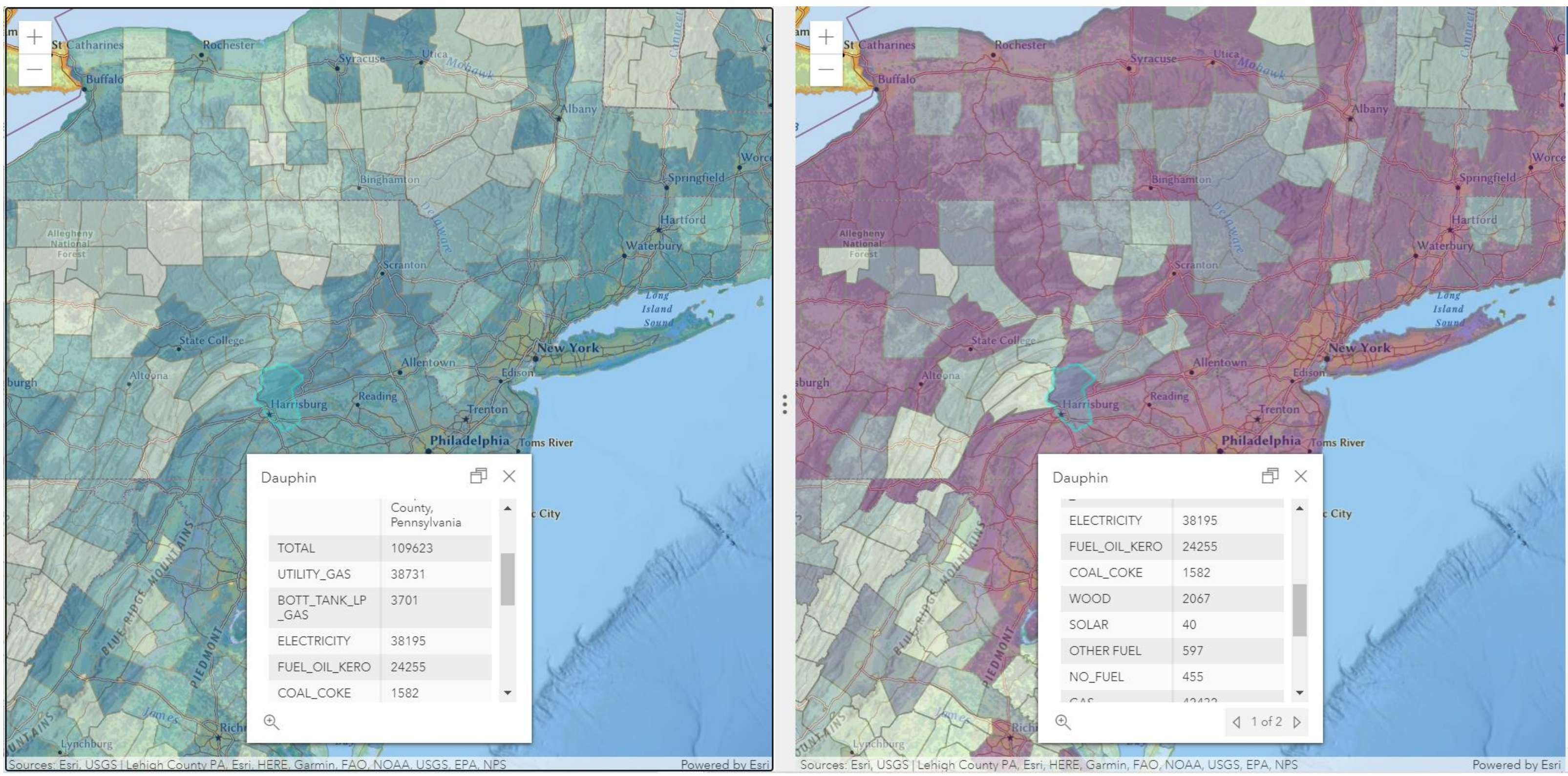
Energy demand/Future demand analysis

Extensive historical energy data is available from data.gov, US Census Bureau, USA.gov. For prediction of future cost to calculate if the model is sustainable I have used regression tools in Python scikit-learn (8) built on Numpy. And also Implemented and analyzed methods discussed in 'Prediction of energy consumption: Variable regression or time series? A case in China'. I utilized multiple regression techniques with different data sources.

Dynamic Web app to compare side by side Electricity vs Natural gas Consumption by county:

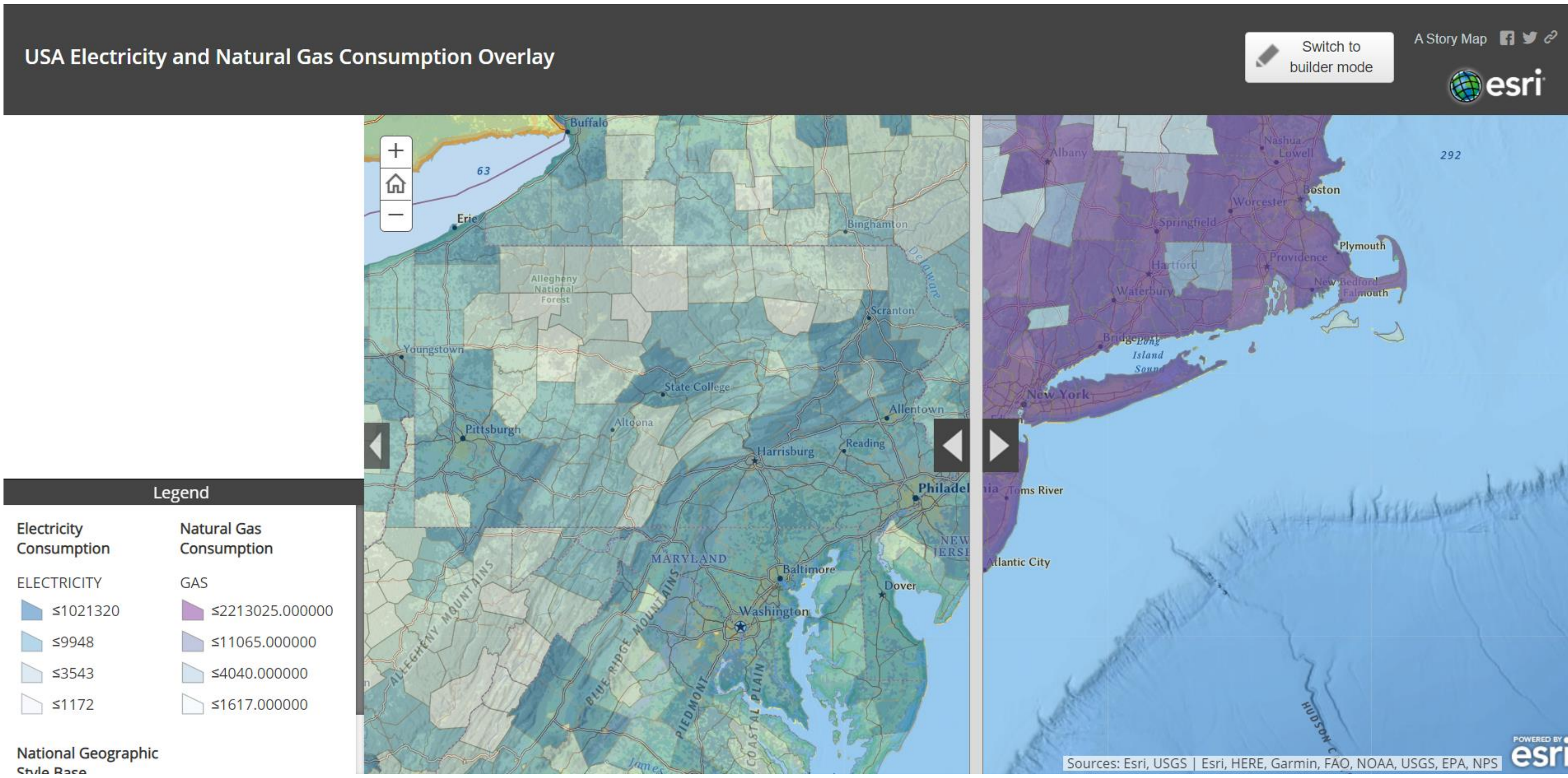
https://www.arcgis.com/apps/Compare/index.html?appid=a12cd39ed58f460eb181ffb14d2738cb

Static Map:



Dynamic Web app to compare overlap Electricity vs Natural gas Consumption by county:

https://www.arcgis.com/apps/StorytellingSwipe/index.html?appid=4442bfcf06e2432f8967028881be1b54



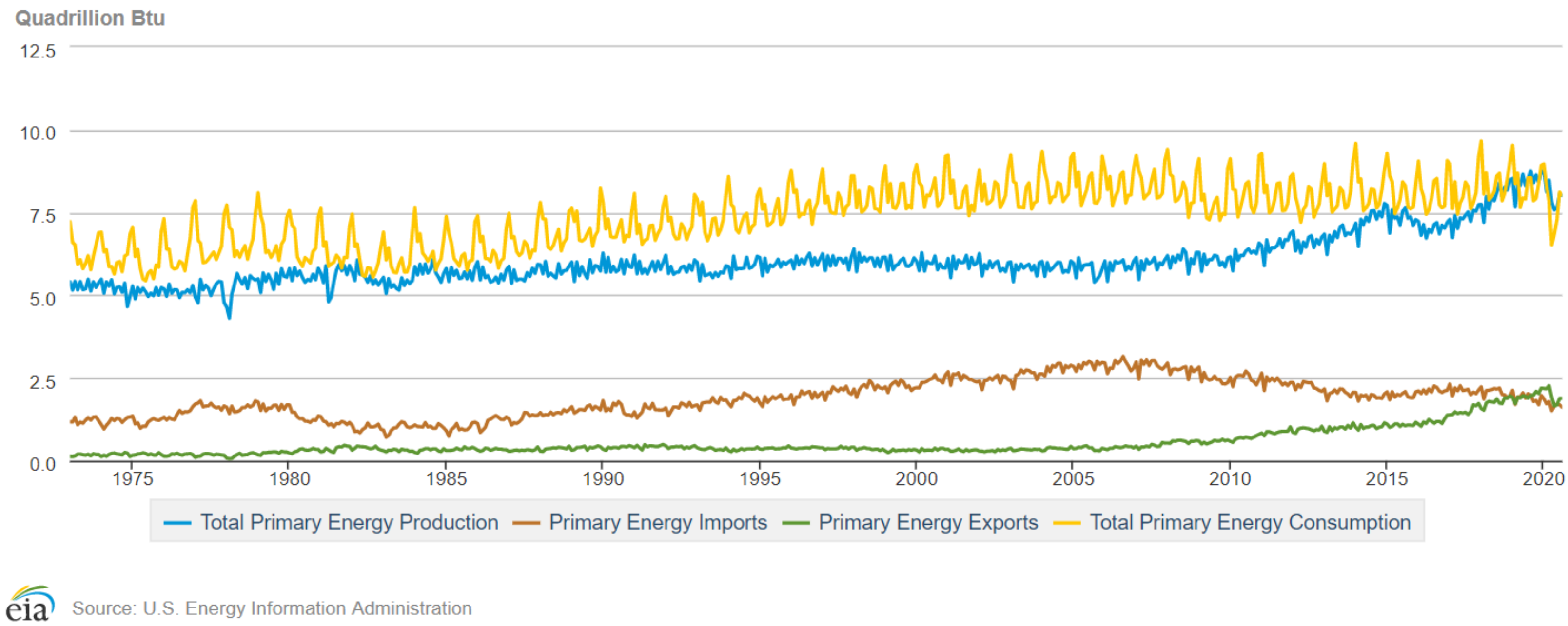
Housing market value/Future value analysis

I used GIS mapping and regression analysis for historical housing market. I have used simple regression techniques as this is not a major influence on solar energy

Energy consumption growth

I used data sets from data.gov, US Census Bureau and USA.gov. On historical data I applied regression analysis for predicting future energy utilization to analyze profitability and feasibility of solar energy for the region. Details of the analysis are on github and github pages.

Below is the energy utilization and production data from EIA(US Energy Information Administration) Chart is taken directly from the EIA website



REFERENCES

1. Usdatagov @usdatagov, & Usdatagov. (2020, September 14). Data.gov. Retrieved December 11, 2020, from https://www.data.gov/
2. Bureau, U. (n.d.). Census.gov. Retrieved December 11, 2020, from https://www.census.gov/
3. Official Guide to Government Information and Services: USAGov. (n.d.). Retrieved December 11, 2020, from https://www.usa.gov/
4. ArcGIS Online. (n.d.). Retrieved December 11, 2020, from https://www.arcgis.com/
5. Brito MC, Redweik P, Catita C, Freitas S, Santos M. 3D Solar Potential in the Urban Environment: A Case Study in Lisbon. Energies. 2019; 12(18):3457.
6. T. Santos, N. Gomes, S. Freire, M.C. Brito, L. Santos, J.A. Tenedório, Applications of solar mapping in the urban environment, Applied Geography, Volume 51, 2014, Pages 48-57, ISSN 0143-6228, https://doi.org/10.1016/j.apgeog.2014.03.008 (http://www.sciencedirect.com/science/article/pii/S0143622814000587)
7. Ramos, Jenice & Ferreira, Ângela & Fernandes, Paula. (2019). A linear regression pattern for electricity price forecasting in the Iberian electricity market. Revista Facultad de Ingeniería Universidad de Antioquia. 10.17533/udea.redin.20190522.
8. Li, Yan. (2019). Prediction of energy consumption: Variable regression or time series? A case in China. Energy Science & Engineering. 7. 10.1002/ese3.439.
9. National Oceanic and Atmospheric Administration. (n.d.). Retrieved December 12, 2020, from https://www.noaa.gov/