Eas 503 Group 11 Final Project

NYS Solar Electric Programs Reported by NYSERDA

Abstract:

The NYSERDA has collected data on New York State Solar Electric Programs since December 2000. The New York Public Service Commission set a goal of achieving 6 gigawatts of distributed capacity solar by 2025. This project analyzes the NYSERDA Solar data, visualizes the production efficacy of solar projects with respect to the features provided, and determines that New York's solar energy projects since December 2000 will amount to at least 6 gigawatts of distributed solar energy at an estimated minimum.

Introduction:

The NY-Sun Initiative brings together and expands existing programs administered by the New York State Energy Research and Development Authority (NYSERDA), Long Island Power Authority (LIPA), and the New York Power Authority (NYPA), to ensure a coordinated and well-funded solar energy expansion plan (1). The NYSERDA has collected data on New York State Solar Electric Programs since December 2000 (2). The New York Public Service Commission extended the term of the program to support the state's goal of achieving 6 gigawatts of distributed solar by 2025 (3).

Data Description:

The "NYS Solar Electric Programs Reported by NYSERDA" dataset includes the following data points: Project number, city, county, state, zip code, sector, program type, solicitation, electric utility, purchase type, date application received, date completed, project status, contractor, primary inverter manufacturer, total inverter quantity, primary inverter model number, primary PV module manufacturer, total PV module quantity, primary PV module model number, project cost, incentive amount, total nameplate capacity, expected annual kilowatt-hour production, remote net metering, affordable solar, community distributed generation project and Green Jobs-Green New York participant (2). The feature TotNamePlateKWDC is measured in KW, the New York State goal of 6 GW solar is a function of KW/1,000,000 and is synonymous with the terms, "Nominal Power" and "Distributed Solar Capacity".

Data Schema:

TABLE1 (PROJECT LOCATION)		ProjectLocation	
Table	Columns	ColumnsNameDB	Column Type
T1	ProjectLocationID[Primary Key]	Auto(ProjectLocationID)	INTEGER
T1	Project Number[Foreign Key]	1(ProjectID)	TEXT
T1	City	3(City)	TEXT
T1	County	4(County)	TEXT
T1	Zip Code	6(Zip Code)	TEXT
T1	Location	30(Location)	TEXT

TABLE2 (INVERTER DETAILS)		InverterDetails	
Table	Columns	ColumnsNameDB	Column Type
T2	InverterDetailsID[Primary Key]	Auto(InverterDetailsID)	INTEGER
T2	Project Number[Foreign Key]	1(ProjectID)	TEXT
T2	Primary Inverter Manufacturer	16(PriInvetMft)	TEXT
T2	Total Inverter Quantity	18(TotalInvtQt)	INTEGER

TABLE3 (PV MODULE DETAILS)		PVModuleDetails	
T3	PVModuleDetailsID[Primary Key]	Auto(PVModuleDetailsID)	INTEGER
T3	Project Number[Foreign Key]	1(ProjectID)	TEXT
T3	Primary PV Module Manufacturer	19(PriPVMft)	TEXT
T3	Total PV Module Quantity	21(TotalPVQt)	INTEGER

TABLE4 (PROJECT TIMELINE)		ProjectTimeline	
T4	ProjectTimelineID[Primary Key]	Auto(ProjectTimelineID)	INTEGER
T4	Project Number[Foreign Key]	1(ProjectID)	TEXT
T4	Date Application Received	11(DtAppReceived)	TEXT
T4	Date Completed	12(DtCompleted)	TEXT
T4	Days Taken	TotalDays	REAL

TABLE5 (PROJECT COST)		ProjectCost	
T5	ProjectCostID[Primary Key]	Auto(ProjectCostID)	INTEGER
T5	Project Number[Foreign Key]	1(ProjectID)	TEXT
T5	Project Cost	22(ProjCost)	REAL
T5	$Incentive	23(Incentive)	REAL
T5	Remote Net Metering	26(RemNetMet)	TEXT
T5	Affordable Solar	27(AffSolar)	TEXT
T5	Green Jobs Green New York Participant	29(GreenCertified)	TEXT

TABLE6 (PROJECT)		Project	
T6	Project Number[Primary Key]	1(ProjectID)	TEXT
T6	Sector	7(Sector)	TEXT
T6	Program Type	8(Program Type)	TEXT
T6	Electric Utility	10(Electric Utility)	TEXT
T6	Purchase Type	11(PurchaseType)	TEXT
T6	Project Status	14(Project Status)	TEXT
T6	Contractor	15(Contractor)	TEXT

TABLE7 (PROJECT PRODUCTION)		ProjectProd	Column Type
T7	ProjectProdID[Primary Key]	Auto(ProjectProdID)	INTEGER
T7	Project Number(Common)[Primary Key]	1(ProjectID)	TEXT
T7	Total Nameplate kW DC	24(TotNamePlateKWDC	REAL
T7	Expected KWh Annual Production	25(ExpectKWhAnnProd	REAL
T7	Community Distributed Generation	28(CommDistGener)	TEXT

Used ProjectID as the primary key from the Projects table as a foreign key to all the other tables as the only unique identifier for each row in the database.

The Project Location table 1 gives us the details regarding the location where the project was set up such as the city, county, and the zip code.

The Inverter details table 2 and the Pv Modules table gives us information regarding the inverter manufacturer and quantity of inverters.

The PV Module details table 3 provides similar information to the inverter, including manufacturer and quantity of modules.

The Project Timeline table 4 contains when the project began and when it ends as well as the duration of the project in days.

The Project Cost table 5 provides the expenses for the project such in categories such as total Project Cost, Incentive (if any), whether it was Remote Net Metered or not, or if it was a part of the Affordable Solar or the Green Jobs programs.

The Project table 6 contains the details of the project such as the sector, program type, electric utility, purchase type, project status, and the appointed contractor.

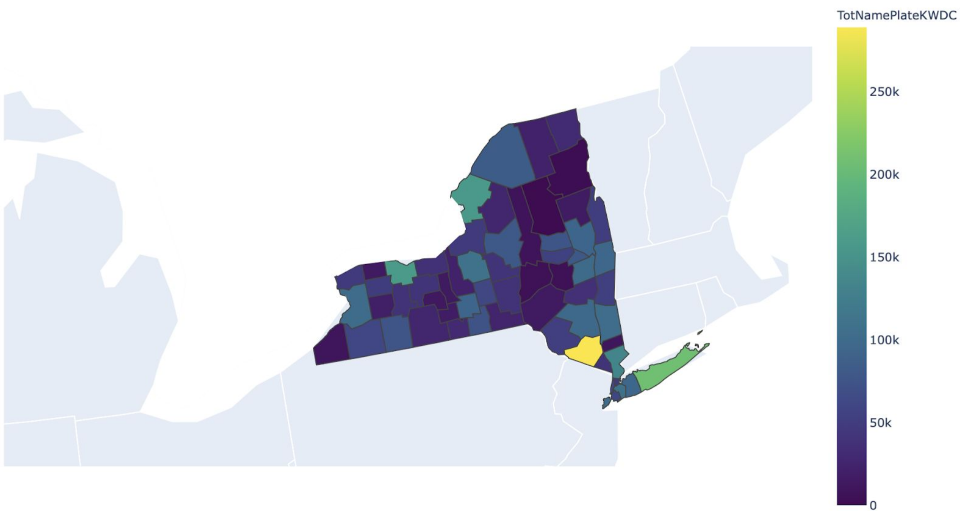
The Project Production table 7 contains the information correlating with the output we are most interested in measuring: TotalNamePlateKWDC, also known as the distributed solar capacity, Expected KWh Annual Production, and whether or not the production is participating in Community Distributed Generation.

Analysis:

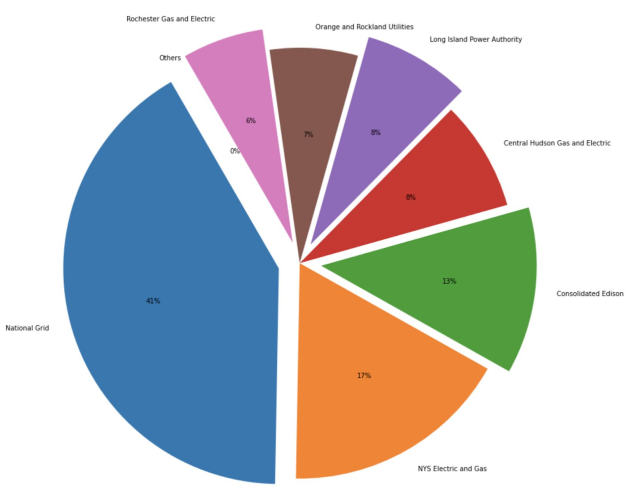
By analyzing the data, we want to answer any of the following questions: How does the data relate to an understanding of New York State solar? How is the distributed solar capacity relevant in relation to other features of every project? Most importantly, is NYS on track to reach the goal of 6 gigawatts of distributed solar by 2025?

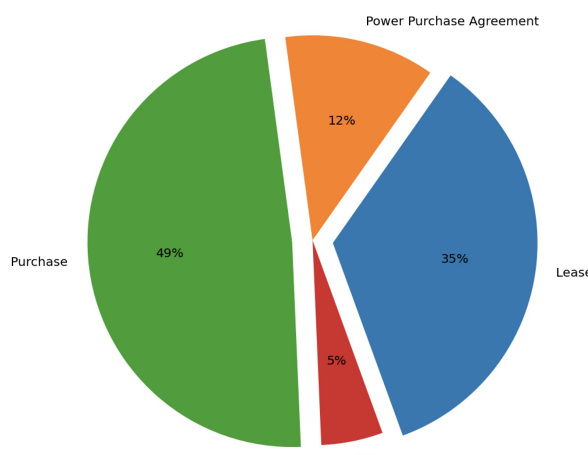
Choropleth Map - Taking a csv file of county codes throughout the state of NY to populate a choropleth library map. This map function takes a year, column, and table (ex: 2020, TotNamePlateKWDC, ProjectProd) and displays an interactive map of New York with the figures as they correspond to each county:

```
Enter year:2020
Enter column:TotNamePlateKWDC
Enter table:ProjectProd
```

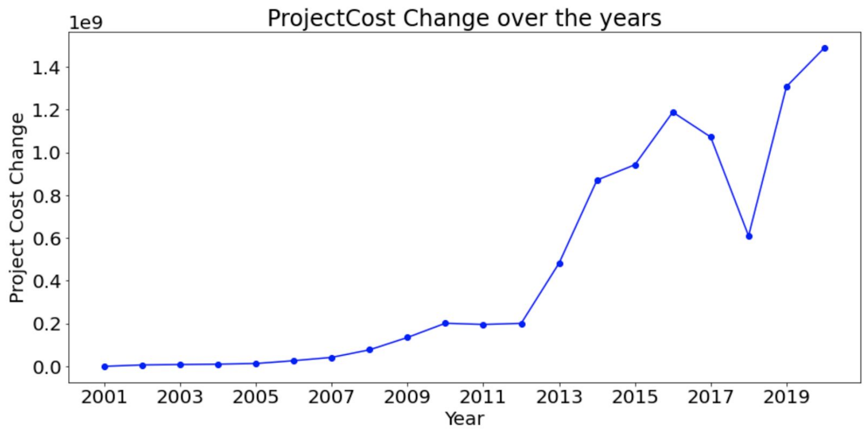


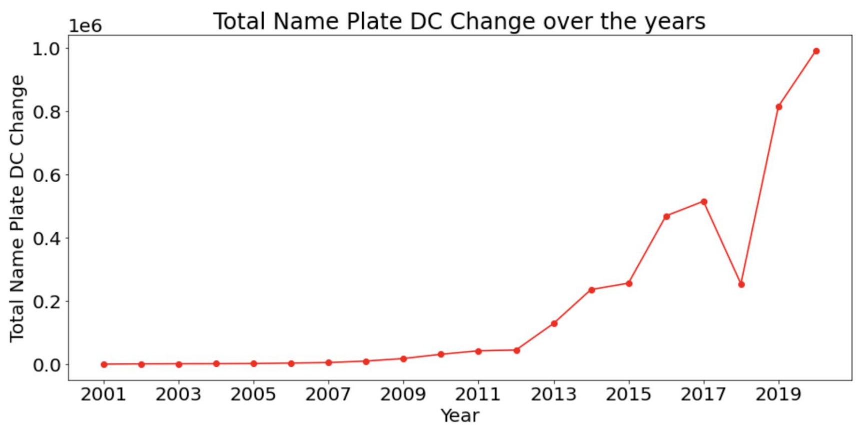
Pie Chart Distribution - The following charts visualize which electric utility providers are correlated with the most distributed solar capacity, and which type of purchase of solar project is used, respectively, as a percentage of the totals:



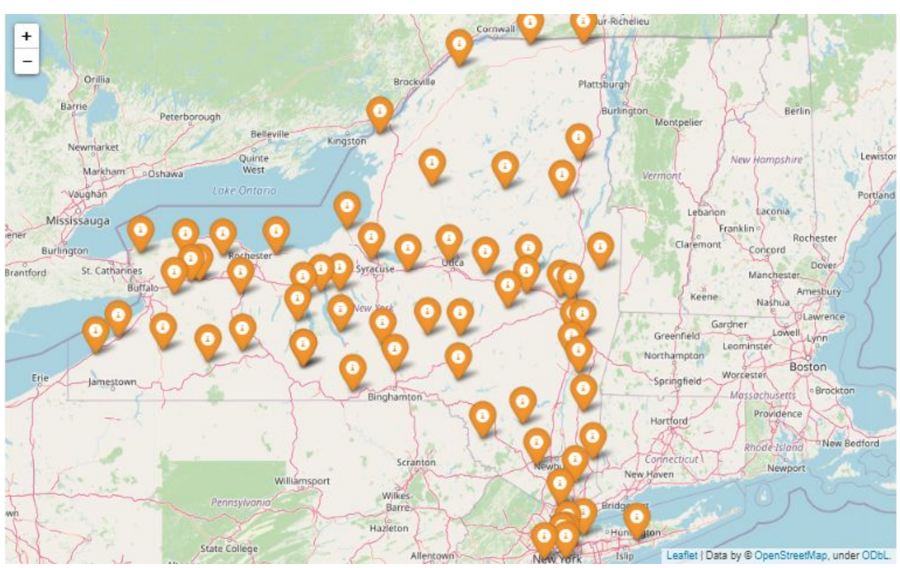


<u>Line Graph</u> - The following plots visualize the change of project costs and distributed solar capacity growth over the years, respectively:

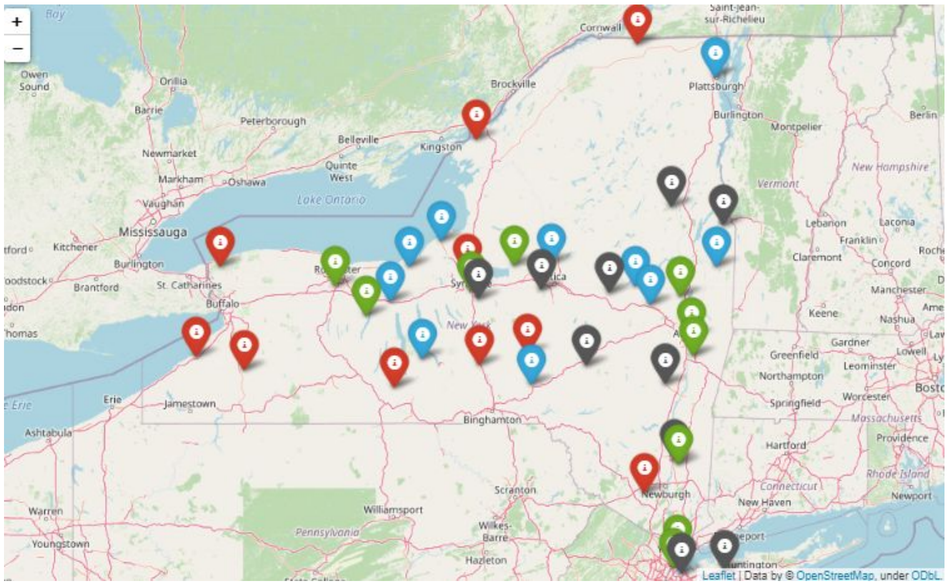




<u>Folium</u> - The following map illustrates the highest solar producing power station of each county in New York and:



The next map illustrates the top solar stations across New York State in terms of GreenCertified (green), Remote Net Metered (grey), Affordable Solar (blue), Community Distributed Generation (red) for each county:

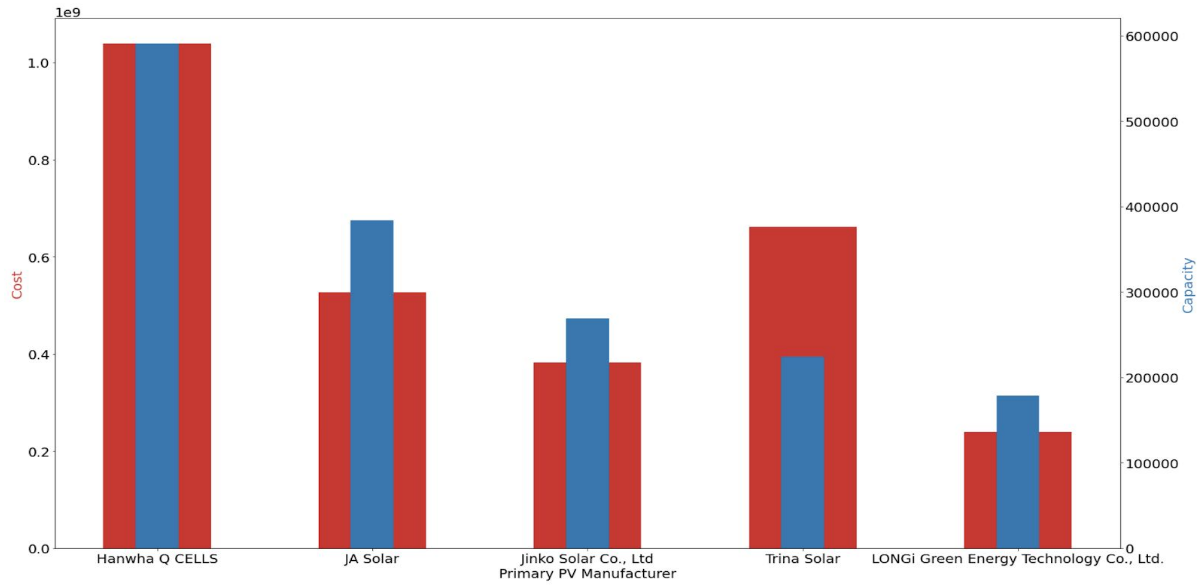


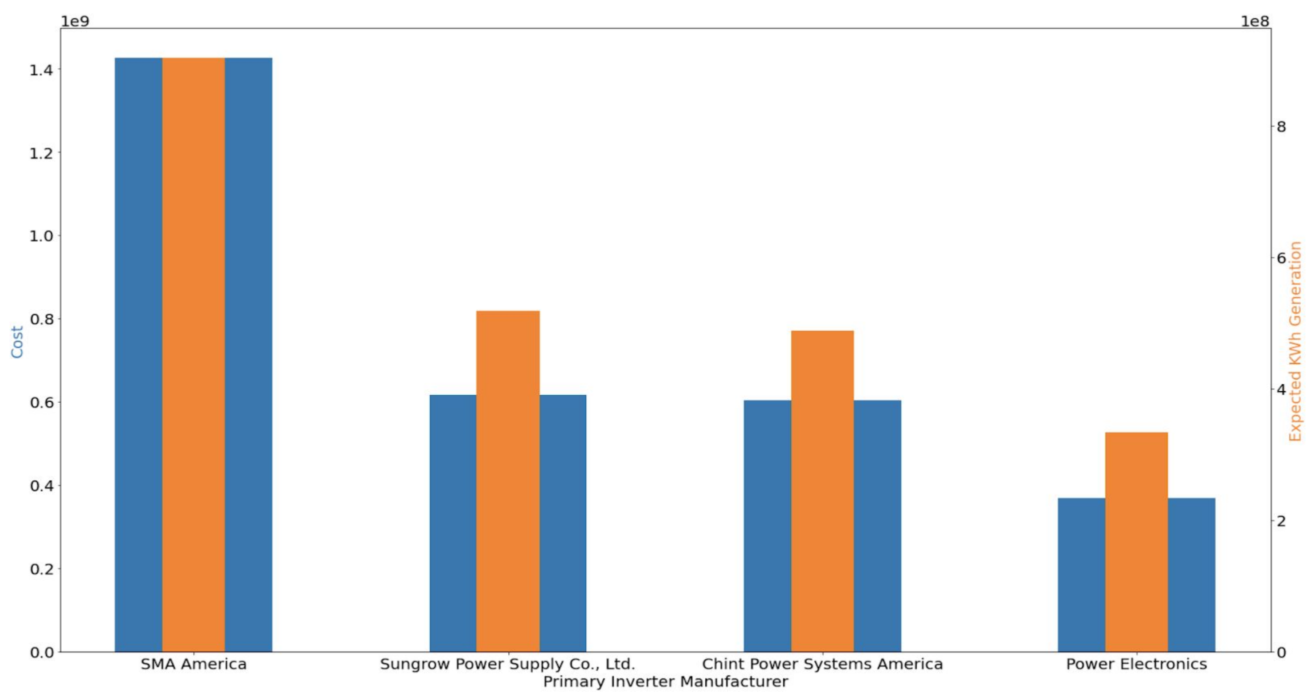
Top Manufacturers - The tables below demonstrate the highest producing manufacturers of PV Modules (left) and Inverters (right), followed. Important to note, the data contains some discrepancies in naming conventions for some manufacturers. SunPower and LG Electronics remain in the top three even so, but their quantities are dispersed by the discrepancies.

	PriPVMft	pvqt
0	SunPower	14351
1	Trina Solar	12411
2	LG Electronics Inc.	7806
3	Hanwha Q CELLS	7095
4	Canadian Solar	6669
5	LG Electronics Solar Cell Division	6215
6	REC Solar	4941
7	Sunpower	4218
8	SolarWorld	4195
9	Others	3935
10	LG Electronics	3507
11	Hanwha Q-Cells	3076
12	Trina Sola	2773
13	Yingli Energy (China)	2202
14	Hyundai Heavy Industries	2146

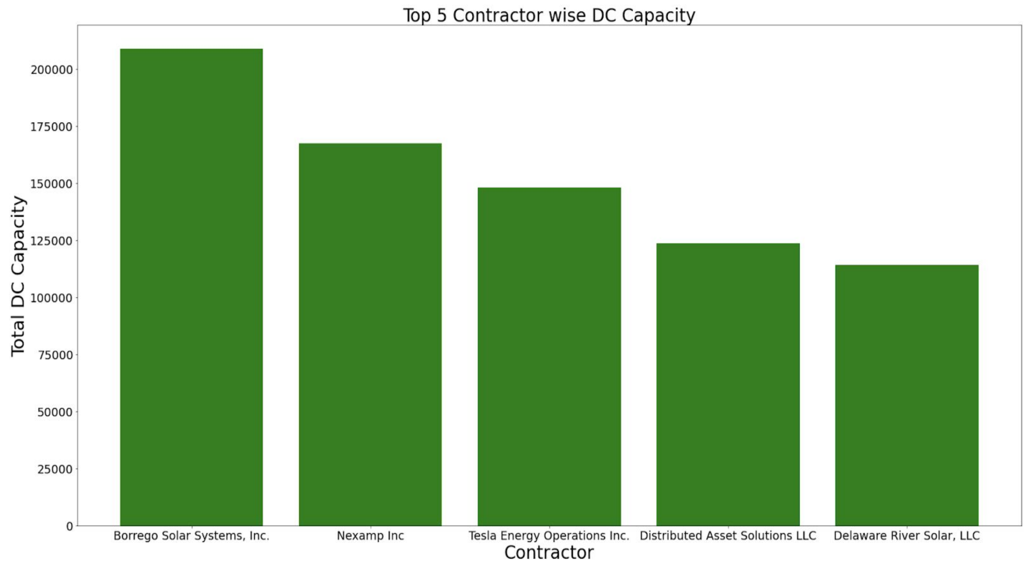
	PriInvetMft	invqt
0	SolarEdge Technologies	17474
1	SMA America	13406
2	Enphase Energy	12928
3	SolarEdge Technologies Ltd.	11387
4	SunPower	10859
5	Enphase Energy Inc.	10640
6	ABB	10441
7	Power-One	4606
8	Fronius USA	4256
9	Others	3987
10	Enphase En	3066
11	Delta Electronics	1489
12	SolarEdge	1224
13	SunPower C	948
14	PV Powered	630

Further, in terms of top manufacturers, the charts below represent those that have the highest cumulative solar capacity output followed by the manufacturers with the highest expected annual solar output with their cumulative cost:

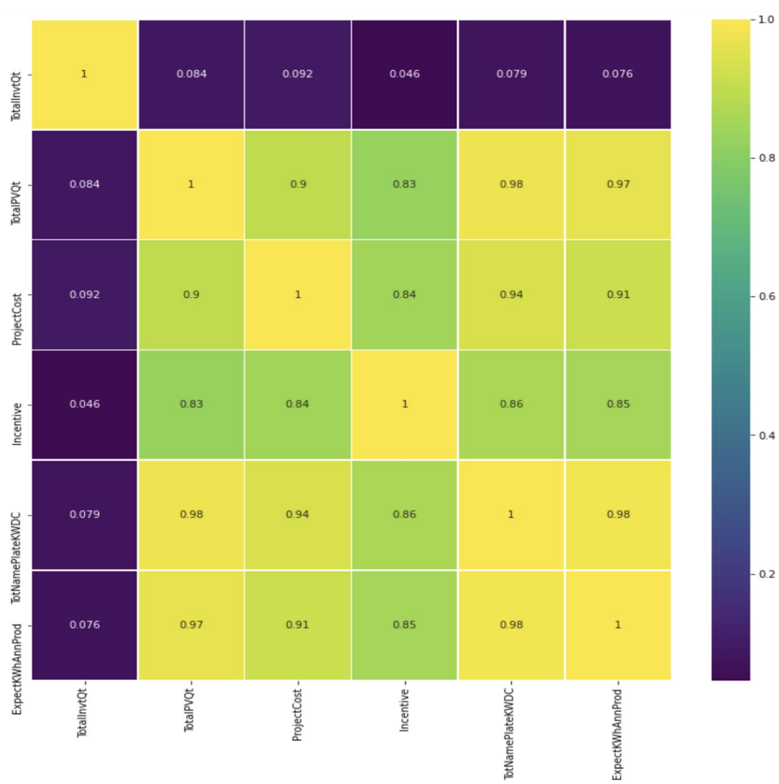




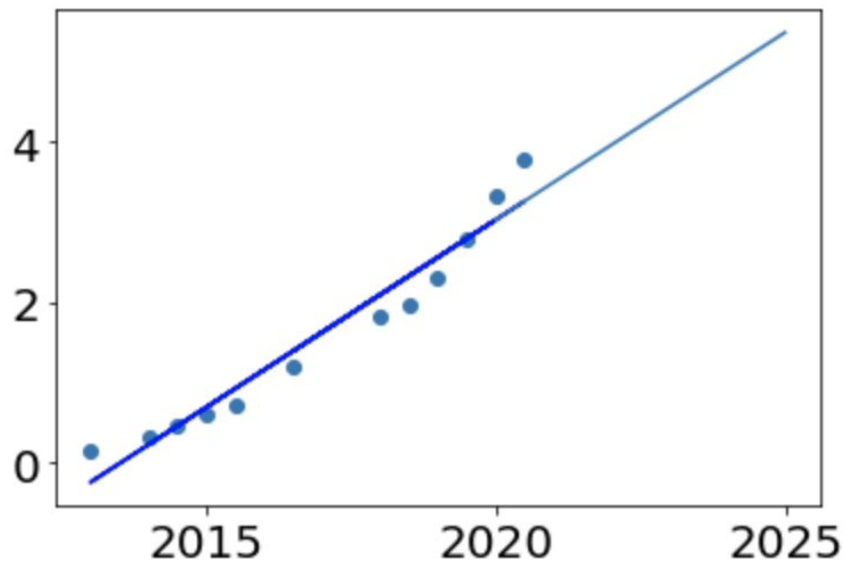
Top Contractors - The following graph represents the Solar Capacity Contractors with the greatest total production in New York.



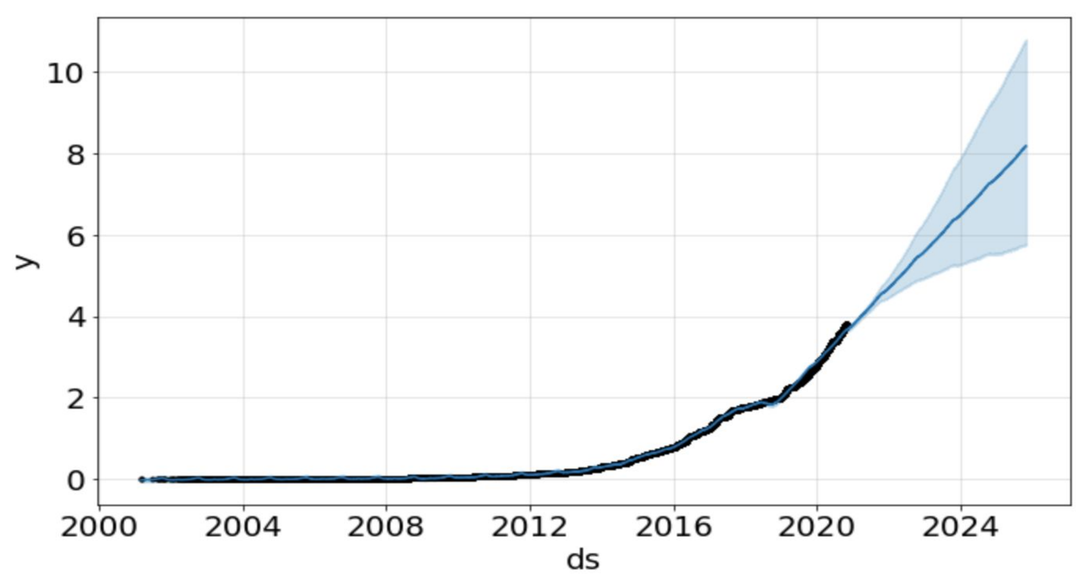
Correlation Matrix - The matrix-like figure below shows the interdependence of each variable with one another. There is clear evidence that the Total NamePlate KW DC is highly proportional to the Total PhotoVoltaic Quantity, Project Cost, Incentive and others.

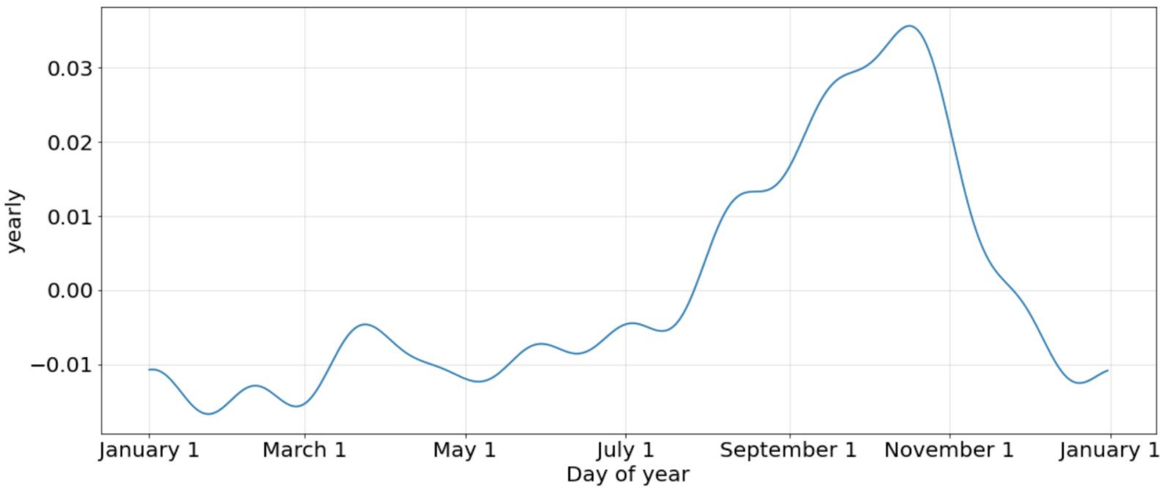


<u>Linear Regression</u> - The first intuition toward predicting the total distributed solar capacity in New York is to apply a linear regression by every year. However, due to the fact that solar capacity hasn't ramped up until after 2011, the results were highly inaccurate. Adjusting for only years after 2011 only reduced the data and produced an unfavorable model which is elaborated in the results. The points are the real totals up to 2020, the dark blue line is the fit, and the lighter line is the prediction:



<u>fbProphet</u> - The following graph is the result of our solution to prediction. Using Facebook's powerful, Open Source fbProphet library, we were able to forecast the projected potential of New York State's total solar capacity five years into the future (2025). fbProphet also provides insight into trends with respect to solar capacity such as what months the projects were beginning in (the succeeding graph):





0.03
0.02
0.01
0.00
−0.01
yearly
January 1
March 1
May 1
July 1
September 1
November 1
January 1
Day of year

Results:

Location:

- Inspecting the choropleth map, the most dense distributed solar producing counties are Orange county, Suffolk Jefferson and Monroe.

Purchase Type:

- The payment of solar systems is majorly purchased (whether it's through loans or not) at 49%. However, leases compete at 35% of the purchase type.

Providers and Manufacturers:

- As the supplier of 41% of electric utilities, National Grid comes out on top of NYS Electric and Gas at 17%, and Con Edison at 13%.
- Industry giant, Tesla, broke its way into the top 5 solar contractors in New York just short of 3 years in business in the state.

Project Cost:

- The year on year increase or decrease of ProjectCost and TotalNamePlateKWDC visualises the efficacy or productivity of the solar power plants over the years. Both features are highly correlated and show the growth of the solar sector in the last decade. There was a slight dip in the years 2017-2018 that signifies a decrease in production. The data analyses used don't provide insight into why such a decrease occurred, but this opens up a door for further research.

Solar Stations:

- The folium maps of power plants throughout New York show that each county is contributing in some proportion. Combining these figures' results with choropleth we can see the pin point location of the highest producing solar plants in counties such as Orange, Suffolk, Monroe, and Jefferson which may help in setting up new projects, determining if these location are oversaturated with power, and if New York State is better off prioritizing a wider spread of farms instead of clustering in particular areas. This also helps in reconciliation of RemoteNetMetered credits and Community Distributed Solar usage. The second map gives the extrapolation of the first map segregated on the specific parameter to get a better understanding of each poweplant's requirements and market segment.

Forecasting Distributed Solar Capacity:

- In terms of the predictive model of linear regression, it was a good segue into employing the power of the many available Python libraries for Data Science. The general linear regression model's inaccuracy generated an unfavorable model with a testing error of 17%, and a test prediction on 2020 of 3.0 gigawatts of solar capacity, while we know the true total up until now is 3.82 gigawatts. This inaccurate model projected the total for 2025 to be 5.36 gigawatts. One shining piece of information to be salvaged from this attempt is that solar capacity hasn't seen much increase until after 2011. Upon some research, this happens to historically coincide with the Power New York Act of 2011, which directed NYSERDA to conduct a study on increasing generation from photovoltaic devices in New York (6).
- Applying the fbProphet model yielded a far more reliable prediction range, with 6 gigawatts of distributed solar capacity in fact being the minimum potential of New York State, 8 gigawatts being the expected projection, and a maximum distributed solar capacity of 10 gigawatts.

Conclusion:

- New York State has an increasingly popular solar program, allowing for providers and manufacturers to compete while providing a worthwhile investment in sustainable energy.
- The NYSERDA provides grants for financing and leasing options so that purchasing and installing solar is more feasible through the Affordable Solar program (4). Leasing may be a popular secondary option because New York State's tax credit for purchasing solar also applies for leases which allows for up to $5,000 in tax breaks. This makes leasing more affordable, because it can cut 25% of what you would be paying, and allow one to take a shorter lease and therefore a lower interest rate (5).
- Understanding New York's goals and how they are met is a greater challenge than simply running numbers. For a full picture, we would need an inside understanding on how the budget is allocated for the NYSERDA and the tax incentive rates are formed. With that said, it is safe to assume that when New York hits its goal, it is not expected to overperform and reach heights such as 8 gigawatts as when it isn't planned. The safest conclusion to be made from our forecast is that New York is more than capable of reaching its goal of 6 gigawatts by 2025.
- It is also interesting to note that applying the fbProphet model provided a trend graph of the most active months: Projects with the highest production of distributed nominal power begin in the fall between September and November. This is likely so that there is enough time to complete each project before the lucrative sunny Summer months roll around again. Countries such as the United States, which lie in the middle latitudes, receive more solar energy in the summer not only because days are longer, but also because the sun is nearly overhead (7).

Future Research Directions:

This project provided valuable insight on an industry that hasn't seen the *light of day* until 2011. Future research directions include but are not limited to:

- Assessing impact on solar distribution by newcomer contractors in the industry such as Tesla, which only began production in Buffalo, NY in 2017, yet the entity already made its way into the top 5 contractors of distributed solar by 2020.
- Understanding the mysterious dip from 2017-2018 in distributed solar production growth.
- Calculating where New York can optimize its power distribution and what features of a solar project contribute to the greatest output over time.

References:

[1] Governor Cuomo Announces NY-Sun Initiative:

https://www.governor.ny.gov/news/governor-cuomo-announces-comprehensive-ny-sun-initiative-expand-solar-development-new-york

[2] The NYSERDA dataset:
https://data.ny.gov/Energy-Environment/Solar-Electric-Programs-Reported-by-NYSERDA-Beginn/3x8r-34rs

[3] The extension of the program to 2025:

https://enerknol.com/new-york-extends-ny-sun-incentive-program-with-nearly-600-million-funding/

[4] Affordable Solar Program

https://www.nyserda.ny.gov/all-programs/programs/ny-sun/solar-for-your-home/paying-for-solar/incentives-and-financing

[5] Solar Energy Tax Credit

https://www.tax.ny.gov/pit/credits/solar_energy_system_equipment_credit.htm

[6] New York Solar Study:

https://www.nyserda.ny.gov/About/Publications/Solar-Study

[7] Solar Energy production is best in the Summer:

https://www.energy.gov/eere/solar/solar-radiation-basics