MGT 6203 Group 19 Project Proposal

Team #: 19

Team Members:

1. Kirpa Singh; 903811445; I have a BA in Integrative Biology from UC Berkeley, and I currently work as a Database Associate at Gilead Sciences in Foster City, CA. I have been working in Biotech/Healthcare research for 2 years now. I'm in my second semester of OMSA and new to analytics.
2. Lyle Douglas Sweet; 903475010; I am a product manager focused on building monetization tools for media companies who create audio (radio, podcasts, streaming, etc.). I often work in SQL and occasionally write Python scripts for large updates via API.
3. Nai Ning Chi; 903817682; I have a MS in Horticultural Sciences from Texas A&M University and currently working as a Senior R&D Specialist at BASF in Davis, CA. I occasionally work in R and I've had experience running simulations on lab procedure improvements with the help of our statisticians. I joined the OMSA program for more structural training in analytics.
4. Karthik Balakrishna Madhur; 903214295; I currently work as a Structural Project Engineer with KPFF Consulting Engineers in Seattle, WA. I have been working for almost 5 years now, designing high-rise buildings on the west coast. This is my first semester in the OMSA program.
5. Chetna Chand Kewalramani; 903005593; I currently work as a Software Integration and Test engineer at Boeing near Seattle, WA. I have a BS Aerospace Engineering and have been working in Systems/Software Test for 5 years. I am in my second semester of OMSA and am new to analytics.

Project Title: Analyzing Solar Power Utilization and Adoption Trends

Background Information on chosen project topic:

Environmental concerns have been rising to the top of regional and global agendas. The McKinsey Global Energy Perspective 2022 report, global power consumption is projected to triple by 2050 as living standards and electrification grow [1]. Among the many different decarbonization efforts, electrification with renewable energies such as solar and wind power is the most economic and easiest to implement in most sectors. For solar power, it is predicted that the global solar PV (photovoltaic) capacity will grow 30 times by 2030. As countries aim for ambitious decarbonization targets, renewable energy, led by solar power is set to become the cornerstone of the worlds' power supply.

In the US, according to the solar market insight and research data from SEIA (Solar Energy Industries Association) [2], solar power has grown at a rate of 33% annually in the last decade. The cost of solar energy installation has dropped by more than over 60% over the past decade. Declining costs of solar technology and increasing demand for clean energy for both private and public sectors, with more than 135 gigawatts (GW) of solar capacity is installed as of 2021, enough to power 23 million homes in the U.S. and the number continues to grow. Identifying potential markets and customer demographics will help the solar industry target more customers and ensure the logistic and supply deployment to the markets.

Problem Statement (clear and concise statement explaining purpose of your analysis and investigation):

The purpose of this analysis is to gain insight into the interest in solar energy by examining its trends along with the primary factors affecting its adoption. We aim to use these insights to determine the possible future path of solar energy in the renewable energy industry.

State your Primary Research Question (RQ):

How has solar energy grown in recent years and what factors influence customer adoption of solar energy?

Add some possible Supporting Research Questions (2-4 RQs that support problem statement):

1. Is electricity cost a primary factor in the regions of solar energy utilization?
2. Is solar energy only accessible to those above a certain household income?
3. Difference in factors for leasing vs. buying

Business Justification:

Solar energy utilization in the U.S. has seen rapid growth over the past decade, as demonstrated by data from the U.S. Energy Information Administration [3]. A recent article by Indeed revealed that "solar consultant" is currently one of the fastest growing jobs in the U.S., having grown 135% in job share from 2020 to 2023 [4]. These data are all a testament to the fact the solar energy is a rapidly growing industry that should be further researched. Understanding solar markets in the U.S. is crucial as more money gets poured into growing industry. A large factor in understanding solar markets is understanding the target customer and their motivations, including income levels, and electricity costs, all of which we plan to delve into through the course of this project.

Data Sources (links, attachments, etc.):

1. U.S. Energy Information Administration – Net generation from solar by state
2. National Renewable Energy Laboratory – Understanding the Evolution of Customer Motivations and Adoption Barriers in Residential Solar Markets Survey Data
3. U.S. Energy Information Administration – Average retail price of electricity to ultimate customers

Data Description (describe each of your data sources, include screenshots of a few rows of data):

1. **Net generation from solar by state**: Provides data on net generation from solar in thousand megawatt hours at national and state level from January 2014-December 2022. It distinguishes between residential, commercial, and industrial sectors. We will be focusing mainly on the residential sector.

description	Jan-14	Feb-14	Mar-14	Apr-14	May-14	Jun-14	Jul-14	Aug-14	Sep-14	Oct-14	Nov-14	Dec-14	Jan-15	Feb-15	Mar-15	Apr-15	May-15	Jun-15	Jul-15	Aug-15	Sep-15	Oct-15	Nov-15	Dec-15	Jan-16
United States : all sectors	1375	1499	2224	2476	2842	3024	2936	3019	2879	2682	2171	1798	1902	2299	3206	3643	3898	3966	4114	4156	3547	3107	2712	2484	2465
United States : residential	263	277	382	421	468	478	502	503	472	445	373	363	340	375	536	609	676	693	741	746	679	618	515	471	520
Connecticut : all sectors	7	8	11	12	13	14	15	15	14	12	9	10	10	11	16	19	21	21	25	24	22	18	14	14	17
Connecticut : residential	3	3	4	4	5	5	6	6	5	5	4	4	4	5	7	8	10	10	11	11	10	9	7	7	8
Maine : all sectors	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	1	1	2
Maine : residential	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

2. **Understanding the Evolution of Customer Motivations and Adoption Barriers in Residential Solar Markets Survey Data**: There are 3 different surveys - "Adopter", "General Population", and "Considerer" survey. All survey responses combined provide data for understanding how those who do not have rooftop PV (photovoltaic) differ from those who do. Survey questions include state of residence, home type, household income, age, education level, purchasing habits, etc.

CASE_ID	GPS_NAC_ADOPTER	SURVEY_SOURCE	HOME	STATE	AGE_BINNED	EDUC_BINNED	F1	F2	TRIG1	TRIG3	TRIG2	TRIG4	N1	N2	N3
1	1	1	1	1	4	4	1	0	0	1	0	0	0	0	0
2	1	1	1	4	3	4	1	1	0	1	1	0	0	1	0
3	1	1	1	2	4	1	1	1	0	1	1	0	0	0	0
4	1	1	1	4	4	2	0	1	1	1	1	0	0	0	0

3. **Average retail price of electricity to ultimate customers:** This dataset provides data on retail prices of electricity in cents per kilowatt-hour from January 2001-December 2022 in residential, commercial, industrial, and transportation sectors. We will be focusing mainly on the residential sector.

description	Jan 2001	Feb 2001	Mar 2001	Apr 2001	May 2001	Jun 2001	Jul 2001	Aug 2001	Sep 2001	Oct 2001	Nov 2001	Dec 2001	Jan 2002	Feb 2002	Mar 2002	Apr 2002	May 2002	Jun 2002	Jul 2002
United States : all sectors	6.75	6.87	7.01	7.02	7.17	7.58	7.88	7.84	7.62	7.43	7.02	7.03	6.95	6.97	6.95	6.95	7.11	7.45	7.68
United States : residential	7.73	8.04	8.32	8.46	8.83	9.07	9.03	9.01	8.92	8.84	8.47	8.29	8.07	8.19	8.17	8.37	8.64	8.73	8.82

Key Variables:

Dependent – Have they adopted solar energy?

Independent- Household incomes*, retail price of electricity*, state of residence, motivation for adoption, availability of buying and leasing options as recorded in survey data

* Indicates that the team hypothesizes it will be an important factor

Planned Approach:

We plan to use logistic regression using survey factors (ex: household income) on the dependent variable (solar adoption). The variables in the survey dataset (household income, state of residence) have already been transformed into categorical variables. We plan to train this model using 70% of the data and test on the remaining 30%. We would determine the quality of the model by comparing using AIC and R^2.

We plan to do regression using time series data on monthly residential generation of solar (according to the EIA) and monthly average retail price of electricity (also using EIA data) to identify trends in solar adoption. We will be regressing retail price on generation of solar.

Anticipated Conclusions/Hypothesis

We predict that higher income households will be correlated with higher solar adoption rates. We also predict that increased retail price of electricity will be strongly correlated with increases in net residential generation of solar.

What business decisions will be impacted by the results of your analysis? What could be some benefits?

Our analyses will help us better understand solar markets by understanding the thought process behind solar adoption. The regression models will give us insight into some of the factors that affect solar adoption, which can help companies when identifying areas for growth.

Project Timeline/Mention key dates you hope to achieve certain milestones by:

Date	Submission Deadlines	Milestones
March 12	Project Proposal	March 18 – Complete draft of slides for Proposal Presentation
March 24	Project Proposal Presentation	March 25 – Finish cleaning/organizing datasets
April 2	Progress report	April 9 – Have all analyses completed April 12 – Complete draft of Final presentation slides
April 16	Final Report	
April 19	Final Presentation Video	