Introduction

The California Solar Statistics https://www.californiasolarstatistics.ca.gov), is a California state agency that provides to the general public data summaries on a variety of solar initiatives.

This study looked at the interconnected solar installation data set, available here (https://www.californiasolarstatistics.ca.gov/data_downloads/) that contains all interconnected roof top projects, with a single row for each interconnection.

The goal of this study is to evetnually provide a tool that can be used to identify effective policies and installation approaches, across California regions, to improve the realization of solar energy opportunities.

The objectives of this Phase I application are:

- 1) Familarization with the data set and identification of initial trends in time.
- 2) Compare installations across regions on California. Solar policies and incentives are different across California. Are there differences between regions?
- 3) Set a framework for future development of the tool.

Solar Installs Application

In this application,

- 1) The interconnected solar installation data set is read in
- 2) Some minor data cleanup is performed
- 3) Data is summarized by California County
- 4) A Shiny app presents the data , allowing the user to select a California county of interest and observe, for that county, the number of solar projects in a Quarter, the amount of MegaWatts delivered by those projects, the average time of delivery of a solar project and the installers responsible for the projects.
- 5) The user can also specify residential, commercial or both sectors.

Example Use of Data

Following is the amount of Kilowatts delivered by solar projects by county for the top 8 counties in California. Note Southern California is much higher than other counties. To be answered with future studies, is this a result of more irradiance, higher energy needs (for example, pools) and/or policies in the reqion.

```
SolarData <- read.csv("C:/2015 R Folder/Solar/NEM_CurrentlyInterconnectedDataset_2015-10-31 - Brief and
SolarData$System.Size.DC       <- as.numeric(SolarData$System.Size.DC)
sd <- data.frame(aggregate(SolarData$System.Size.DC ,by=list(SolarData$Service.County), FUN=sum, na.rm
colnames(sd)<- c('County', 'KiloWatts');sd <- sd[order (sd$KiloWatts,decreasing = TRUE), ];sd[1:10,]
```

```
##            County  KiloWatts
## 36      SAN DIEGO 77452518.2
## 28         ORANGE  3331354.5
## 17    Los Angeles   316175.0
## 31      Riverside   272582.7
## 8          Fresno   190664.9
## 34 San Bernardino   188613.1
## 42    Santa Clara   183582.1
```

Future Work

The next steps for the application:

- 1) Merge in population data for each country so that the data can be normalized based on population across regions. A better estimate of solar project penetration would then be available.
- 2) Merge in solar irradiance data. Each county has different weather patterns and therefore different amount of available solar energy.
- 3) Bring in data regarding policies regarding rebates and other incentives for each reqion.
- 4) Factor in consumer behavior and demographics. For example, there are probably more home pools in Southern California which may increase the demand for cheaper energy.

Application References

The current version of the application is here (https://github.com/beaisis/DataProducts).

Supporting code and documentation for the application is on gitHub (https://beaisis.shinyapps.io/Solar/)