Corellia Analytics

big data, data science

Descriptive Analysis on Solar Irradiance

Solar energy is one of the major renewable energy sources which can address many challenges faced by the world. The solar energy is self-sufficient and contributes to a sustainable future. It can be used to generate electricity as well as for heating purposes. Solar Irradiance is the rate at which the Earth receives solar energy. The solar irradiance figures are used to-

- **Plan the deployment of solar power systems**
- **Design energy-efficient buildings**
- **For research on climate modeling and weather forecasting**
- **Design spacecraft and for space travels.**

Objective

We tried to analyze solar irradiance across a few major cities in India and inform stakeholders on the trends that will enable them to utilize this natural energy source for the betterment of the environment and of the country.

Accomplishment

Changing seasons will have a direct impact on the solar irradiance. Therefore, we built a dashboard that helps users analyze the trends of solar irradiance in various cities across the seasons. We leveraged data cleaning and visualization techniques to achieve the same.

Technologies

We used the following technologies to create the dashboard:

Data Visualization

Data Cleaning

R Shiny

Data Sources

The SAM weather files for India combine the solar resource data and the weather data from two different sources.

"We need to bring sustainable energy to every corner of the globe with technologies like solar energy mini-grids, solar powered lights, and wind turbines."

Ban ki -moon

(South Korean Politician)

The solar resource data contains hourly global horizontal, direct normal, and diffuse horizontal data for entire India. This data is divided into approximately 10 km by 10 km square grid cells and is available on the India Solar Resource Maps website. The solar resource data were acquired from satellite imagery using a numerical model developed by the State University of New York (SUNY).

The weather data are from the Integrated Surface Database (ISD) maintained by the U.S. National Oceanic and Atmospheric Administration (NOAA). The weather data includes hourly values of air temperature, dew point temperature, atmospheric pressure, wind direction, and wind speed. For a description of the ISD, see the Integrated Surface Database web site.

Weather data from ISD are available for 87 stations whereas the hourly solar data are available for all locations in India. Table 1 depicts the selected ISD stations to represent a range of climate zones where solar projects may be of interest in India. Solar resource data was merged with ISD weather data to create the SAM weather files.

Selection criteria for the ISD stations

The following criteria were followed for selecting the stations from the 87 ISD stations:

- Availability of the measured data for a period of 2002 - 2008 to match the dates with SUNY solar radiation data.
- Data availability for at least 20,000 hourly observations over the measurement period.
- Availability of sufficient open land space to develop a large-scale solar power project. For example, New Delhi was excluded.
- Location away from the coast to avoid climate effects of large bodies of water. For example, Bombay was excluded.

Table 1

List of locations with Integrated Surface Database (ISD) station name, number and coordinates, SUNY grid cell coordinates, and years for which SAM weather files were prepared

Location Name	ISD Number	ISD Station		Center of SUNY Grid Cell		Available Years
		Lat. (oN)	Long. (oE)	Lat. (oN)	Long. (oE)	
Amritsar	420710	31.633	74.867	31.55	74.85	2002-2008
Jaipur/Sanganer	423480	26.817	75.8	26.95	75.85	2004-2008
Lucknow/Amausi	423690	26.75	80.833	26.85	80.85	2002-2008
Varanasi/Babatpur	424790	25.45	82.867	25.55	82.85	2002-2008
Ahmedabad	426470	23.067	72.633	23.25	72,75	2002-2008
Calcutta/Dum Dum	428090	22.65	88.45	22.55	88.45	2002-2008
Nagpur/Sonegaon	428670	21.1	79.05	21.25	79.05	2004-2008
Begumpet Airport	431280	17.45	78.467	17.35	78.65	2002-2008
Tiruchchirappalli	433440	10.767	78.717	10.75	78.75	2002-2008

For most locations shown in Table 1, SAM weather file is available for each year from 2002 through 2008. For Nagpur/Sonegaon and Jaipur/Sanganer, there was insufficient data for 2002 and 2003. Hence for those two locations, files are only available for 2004 through 2008.

The SAM weather files include hourly data for the following subset of elements in the standard format:

- Dry Bulb Temperature
- Dew Point Temperature
- Wet-bulb Temperature
- Percent Relative Humidity
- Wind Velocity
- Wind Direction
- Atmospheric Pressure
- Global Horizontal Radiation
- Direct Normal Radiation
- Diffuse Horizontal Radiation
- Latitude
- Longitude
- Site Elevation
- Time

For the remaining data elements not used by SAM, the value -9900 indicates a null value. SAM requires that each data element have value regardless of whether it is used during simulations.

Typical Meteorological Year and Single Year Data :

The SAM weather files for India do not contain typical meteorological year data. Each file contains the given location's hourly data for a single year.

A typical meteorological year file contains data for a single year that represents the solar resource and weather over a period of many years. To develop a meteorological year file, thirty years of data on climate patterns of the given locations would be necessary. With only six years of solar resource data, it was not possible to create typical meteorological years files.

Timestamp

The solar data in the SUNY data set are in Local Standard Time (LST) for GMT+5.5 hours and "time-shifted" so that each hourly value represents the average value from the previous hour.

The timestamp for the ISD weather data is on-the-hour GMT. The weather data time stamp was adjusted to match the solar data timestamp based on the GMT +5.5 time zone.

Missing Data

The solar data sets have no missing data. For each grid cell and year, the data file contains the hour-based solar radiation values of the entire year. The hour-based weather data sets from the ISD contain some missing data. Gaps in the weather data were filled using a method developed in the U.S. National Solar Radiation Database (NSRDB) 1991-2005 update.

The ISD data do not contain relative humidity values required by SAM. Those values were calculated based on the temperature and atmospheric pressure that are included in the data set, and using the method used for filling gaps in the NSRDB 1991-2005 update.

We acquired a total of 59 files of data for 9 cities, of which each file contains 58 features.

> *"There is a lot of willful incompetence in solar industry that is in the process of coming to light."* — Steven Magee (Author of Toxic Electricity)

Process

We first imported all the files into one data frame in r, where we selected few features like date, time, ghi, dni, dry_bulb. We collected a few features like the name of the location, ISD number, their latitude and longitude numbers, grid latitude and longitude numbers, year, etc from the document.

The Indian seasons are divided as follows :

Summer	-	Mar, Apr, May, Jun
Monsoon	-	Jul, Aug, Sep, Oct
Winter	-	Nov, Dec, Jan, Feb

The data is then grouped by month, season, and year. New columns were created by calculating average ghi, dni, dry_bulb and max ghi, dni, dry_bulb, keeping all the data in the new data frame.

An attractive dashboard was created where users were asked to enter details of the following :

1. Location name
2. Season
3. Year

The dashboard filters the data and plots the location name on the map, plot the graph of average dni, dry bulb, ghi.

> *"Renewable energy is an essential part of our strategy of decarbonization, decentralization, as well as digitalization of energy."* - Isabelle Kocher (CEO, Engie)

Visual insights

We have three visuals for our dashboard.

Image 1.1 Table

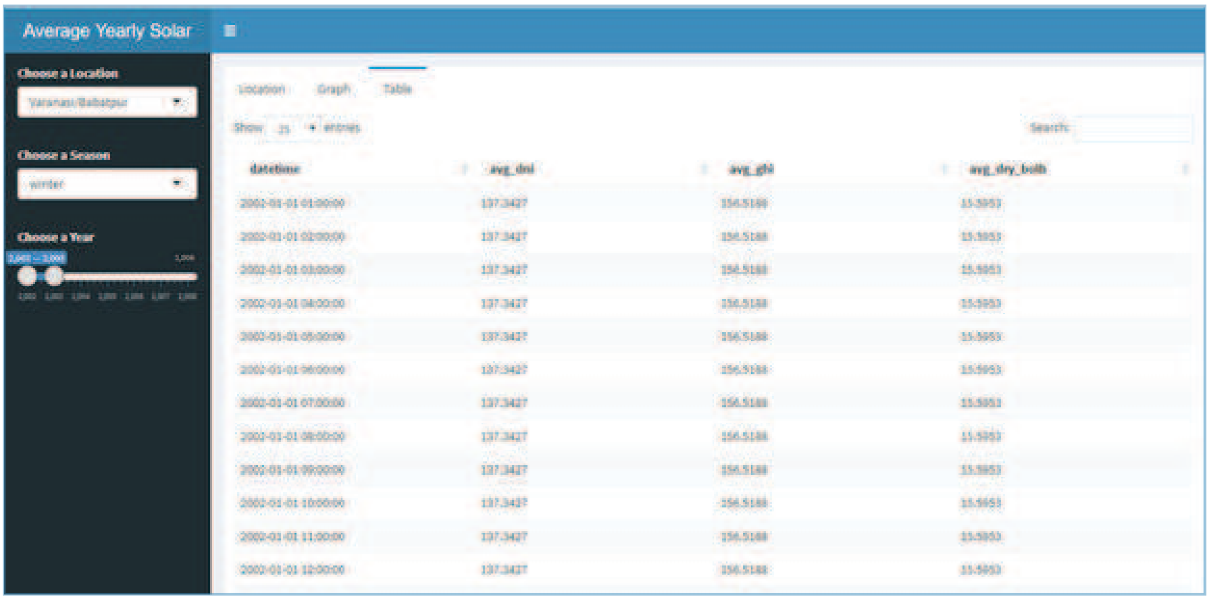
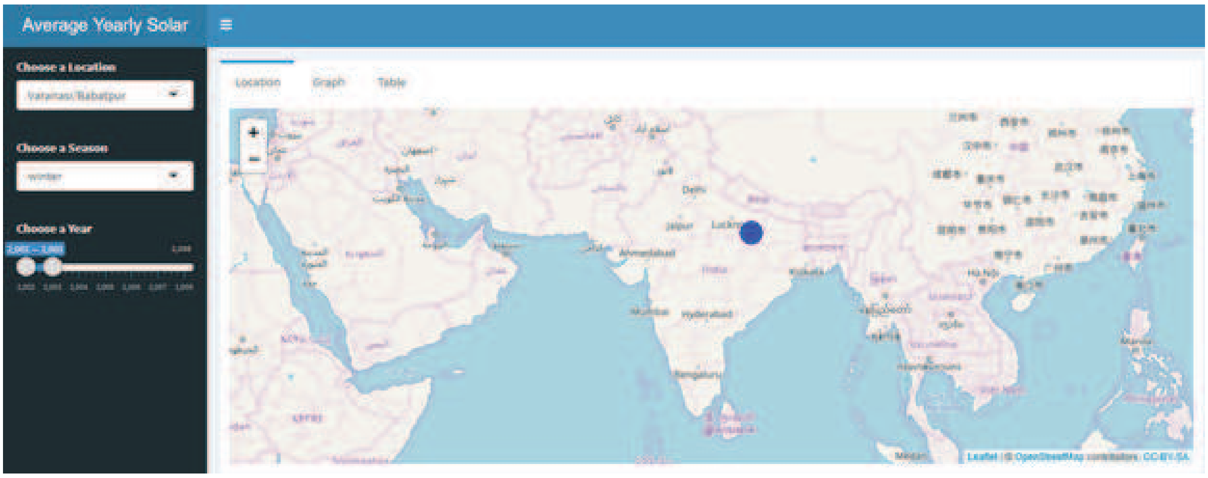


Image 1.2 Graph

Image 1.3 Map



Conclusion

With the data, we can analyze the trends of solar irradiance across the year in certain cities which helps potential industrialists, businesses, and enterprises to identify opportunities. Solar energy is renewable, abundant, and sustainable. Harnessing this energy doesn't cause any pollution and is environment-friendly. With the latest technological equipment like solar panels, solar energy plays a vital role in reducing electricity bills and reduces our dependence on non-renewable energy sources. Technological advancements in the solar power industry, like nanotechnology and quantum physics, have the capability to triple the electrical output of solar panels.

Corellia Analytics

big data, data science