**Climate Connect**

Solar Power Generation from the plant.

**Objectives:** To Predict the Power Generation from the Forecast Data.

**Introduction:** Design a profitable solar power plant by analyzing the radiation and metrological data. You will benefit in the operation by reduced maintenance costs and high efficiency.

The sun radiation on earth surface combines **Direct normal irradiation (DNI)** and **Diffuse Horizontal Irradiation (DHI)** .Both are linked in the formula for **Global Horizontal Irradiation (GHI).**

**GHI +DHI+DNI-**

**GTI** is total amount of direct and diffuse irradiation received from above by a tilted surface.

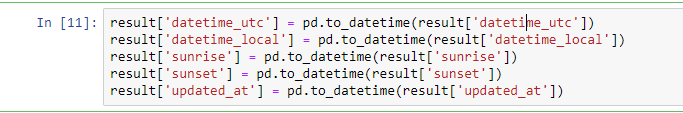
From the Given Data **GHI** in known to you and how much Power is generated with help of GHI.

From the [Research](https://www.ammonit.com/en/wind-solar-wissen/solarmessung#top) the Parameter that influence the Solar Energy Production **Air-temperature, wind-speed, wind-direction, humidity etc.** details which contains in Weather Actuals file.

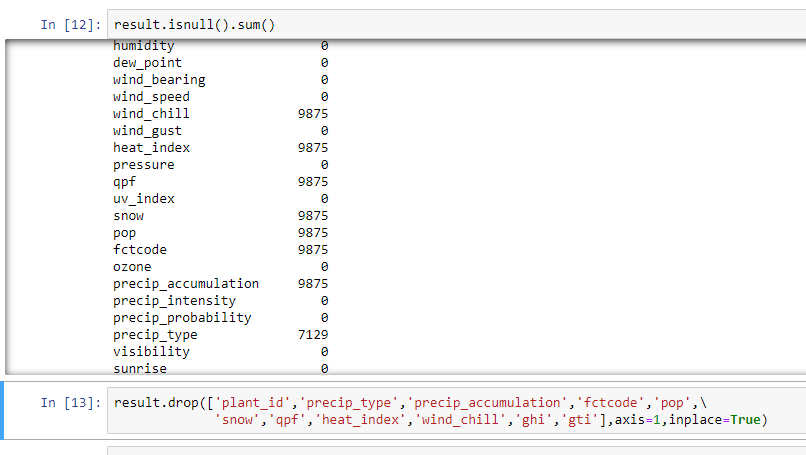
**Approach:**

**Feature Engineering**

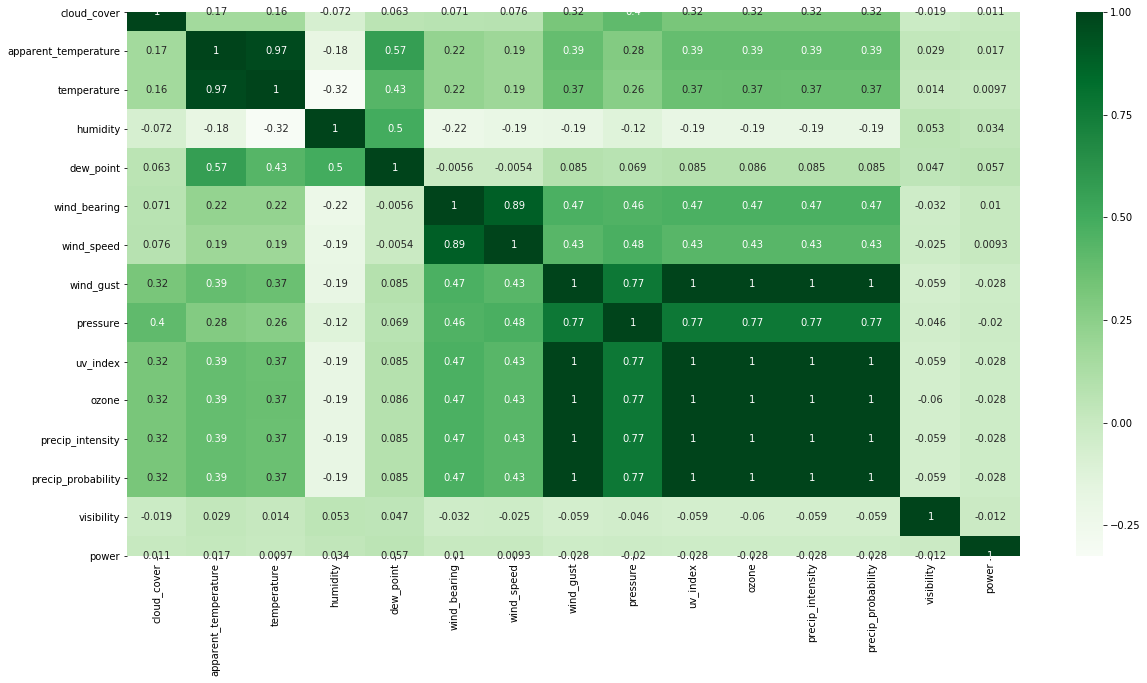
1. Converting variables into proper **date time** format.



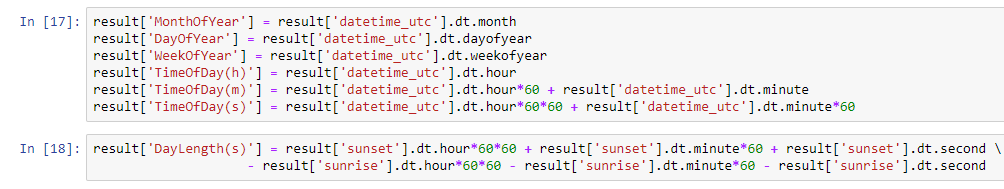
2. Identifying **Null values** and **Drop** the variables.



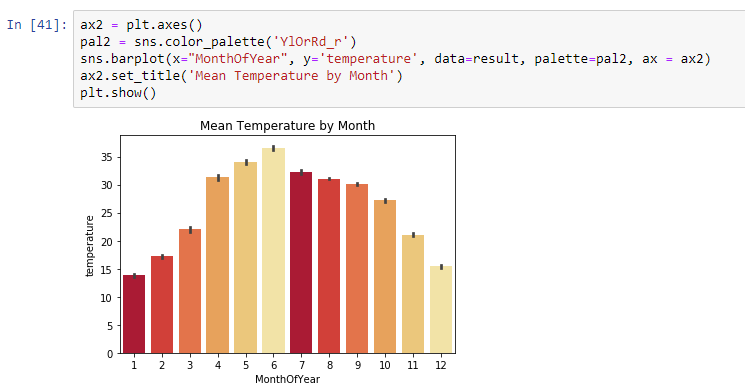
3. **Heat map** describing the correlation between the variables and drop the variables highly correlated with each other.



4. Calculating the time duration between **Sunrise** and **Sunset**.

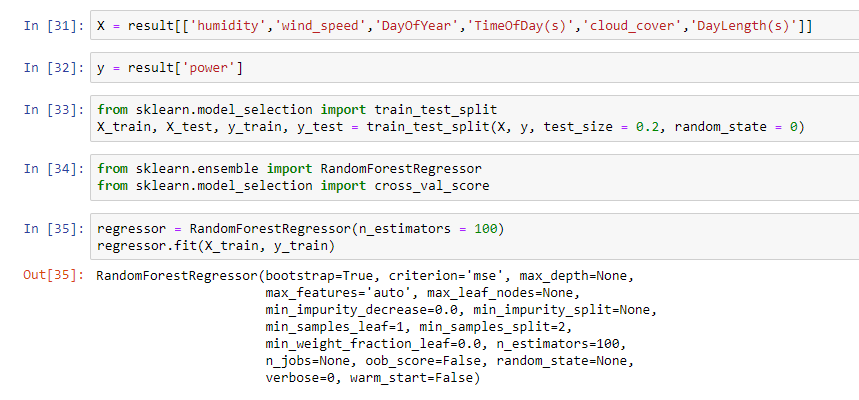


5. visualizing the **temperature** vary in month. It describes in **month of 4, 5, 6(summer)** keeps on increasing and decreases **after summer over**.

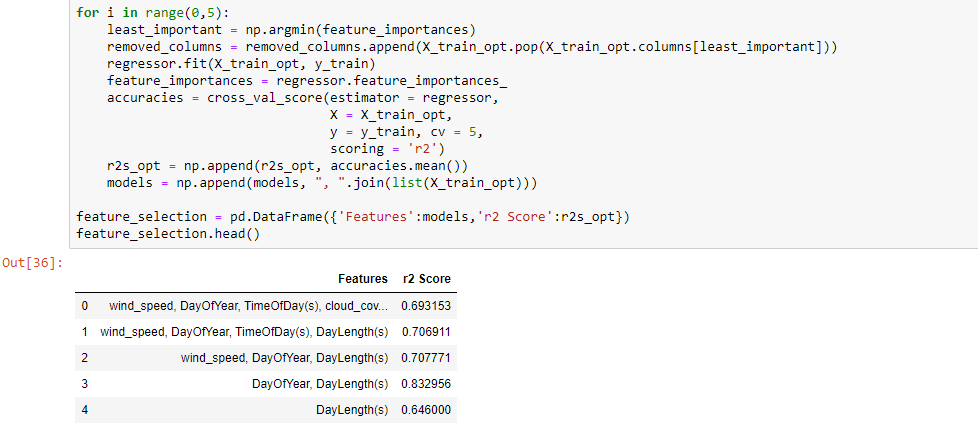


**Model Selection:**

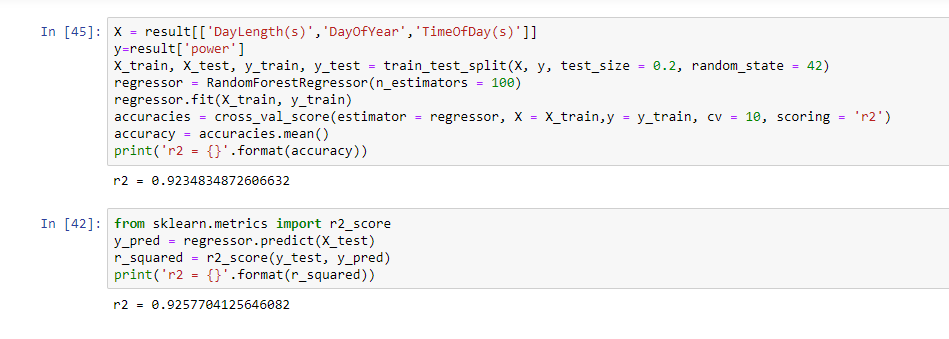
1]. Select the suitable variables for prediction and split the data and applying **RandomForestRegressor (RFR)** Algorithm.



2. Applying feature Importance we get r2 score. From this we can Say **DayOfYear and DayLength(s)** are highly correlated with target variable with **r2 score** of **0.832956.**



3. Finally the Best variables which predict the power supply are **DayOfLength(s), DayOfYear, and TimeOfDay** with **r2 score** of **0.9234** even it performs better on test data with r2 score more than train data with r2 score of **0.92577** on test data as well.



**Conclusion:**

From the Given Data DayLength in seconds and DayOfYear and TimeOfDay in seconds gives best predictive model for prediction of Solar Power Generation.