Introduction

Generating solar power relies on certain conditions like weather, number of sensors and how the data is gathered through an inverter. In this project we will be studying two different power plants and how faulty equipment can affect the data readings.

The data is sorted into two files with seven columns each, one containing the data for the power that is generated, the other for the weather.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| DATE\_TIME  Date given in date hour and minute | PLANT\_ID  Describes the plant | SOURCE\_KEY  Inverter meaning the one that reads the data for the generator and the sensor for the weatherdata | DC\_POWER  Amount of direct current generated | AC\_POWER  Amount of alternating current generated | DAILY\_YIELD  Daily production of generated power | TOTAL\_YIELD  Total production generated across the dataset. |

We can observe that there is no use for the PLANT\_ID as we are only dealing with two power plants. The So