This data has been gathered at two solar power plants in India over a 34-day period. It has two pairs of files - each pair has one power generation dataset and one sensor readings dataset. The power generation datasets are gathered at the inverter level - each inverter has multiple lines of solar panels attached to it. The sensor data is gathered at a plant level - single array of sensors optimally placed at the plant.

There are a few areas of concern at the solar power plant -

1. Can we predict the power generation for next couple of days? - this allows for better grid management
2. Can we identify the need for panel cleaning/maintenance?
3. Can we identify faulty or suboptimally performing equipment?

Task Dated 8-1-2021

From Plant\_Generation Data:

1. Out of total yield power give the list of total\_yield power which is in between 90% to 100% along with Date\_time, Plant Id, Source\_Key and Total\_yield
2. Like the above task(1) give the results between 60% to 80%, 40% to 60% and 20% to 40%
3. Give the top 10% of the data of the sum of DC\_Power, AC\_power,Daily\_Yield along with Plant ID and Source Key