

Solar power plant data analysis summary



Analysing performance data from a photovoltaic solar power plant in India

Context

- A renewable energy supplier owns solar power plants in India.
- Two of the power plants present power production anomalies affecting the total power generation and thus reducing business income.
- Prior to sending the technical team to fix the problem the data science team was required to analyse power plants' sensors data to make a first approximation to the root cause of failure of the equipment.
- What should the data science team do? How should they approach the problem?

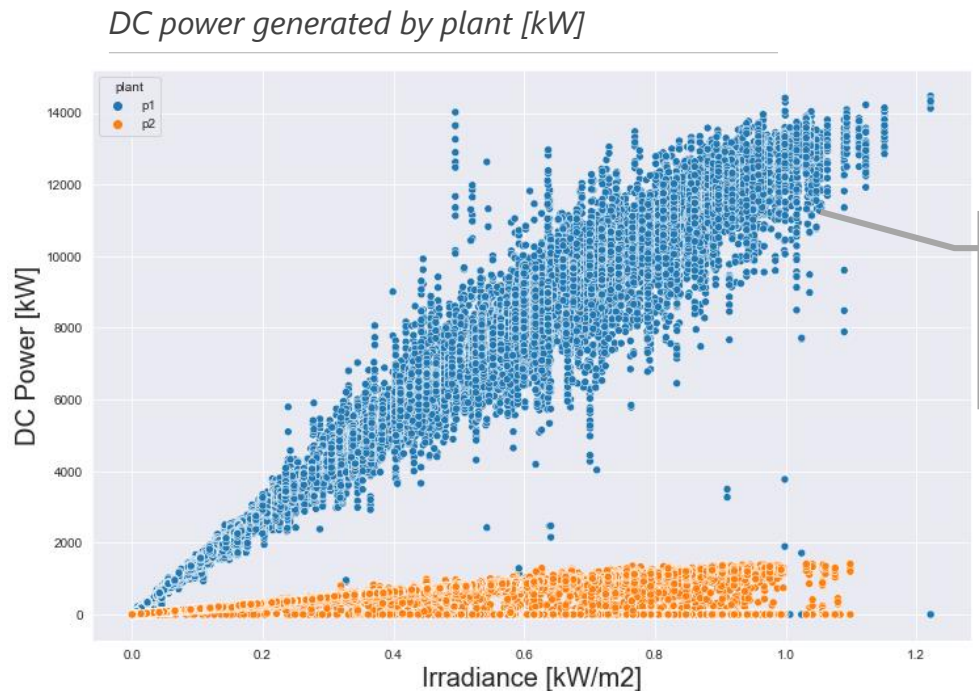
Methodology

1. Set the project objective
2. Establish the variables that affect the business objective
3. Define the KPIs of those variables
4. Enumerate and classify the relevant entities involved and the data that can be retrieved from them
5. Identify the first questions that need to be answered in order to start looking for significant insights into each of the variables affecting the business objective
6. Analyse the data and look for insights
7. Suggest an action plan

For the complete methodology description please refer to [this](#) link.

PROJECT INSIGHTS

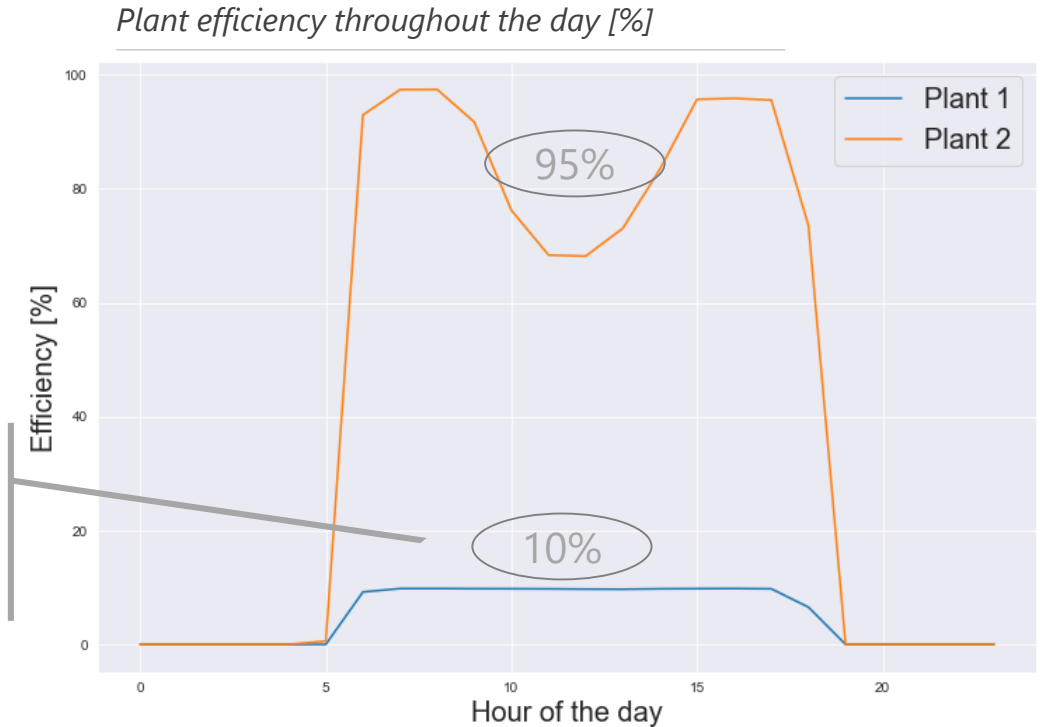
There are important data quality issues that suggest that the following statements are not conclusive



PENDING A
SECOND
REVIEW

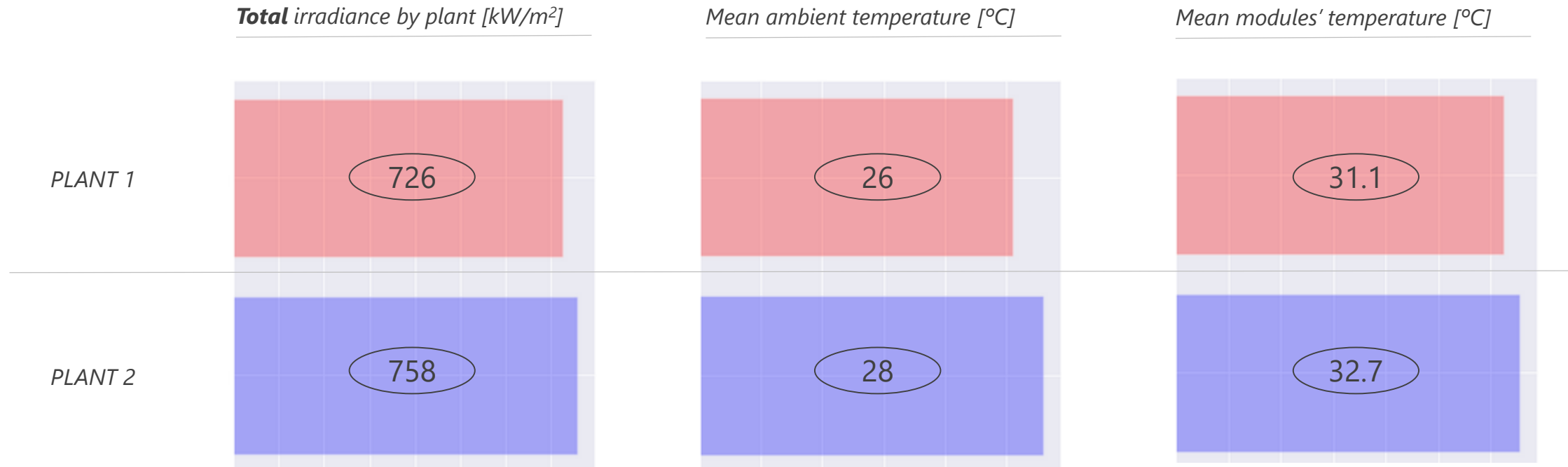
DC generation in
power plant 1 is
around 10 times
higher than in plant 2

While the efficiency
conversion is only
around 10%



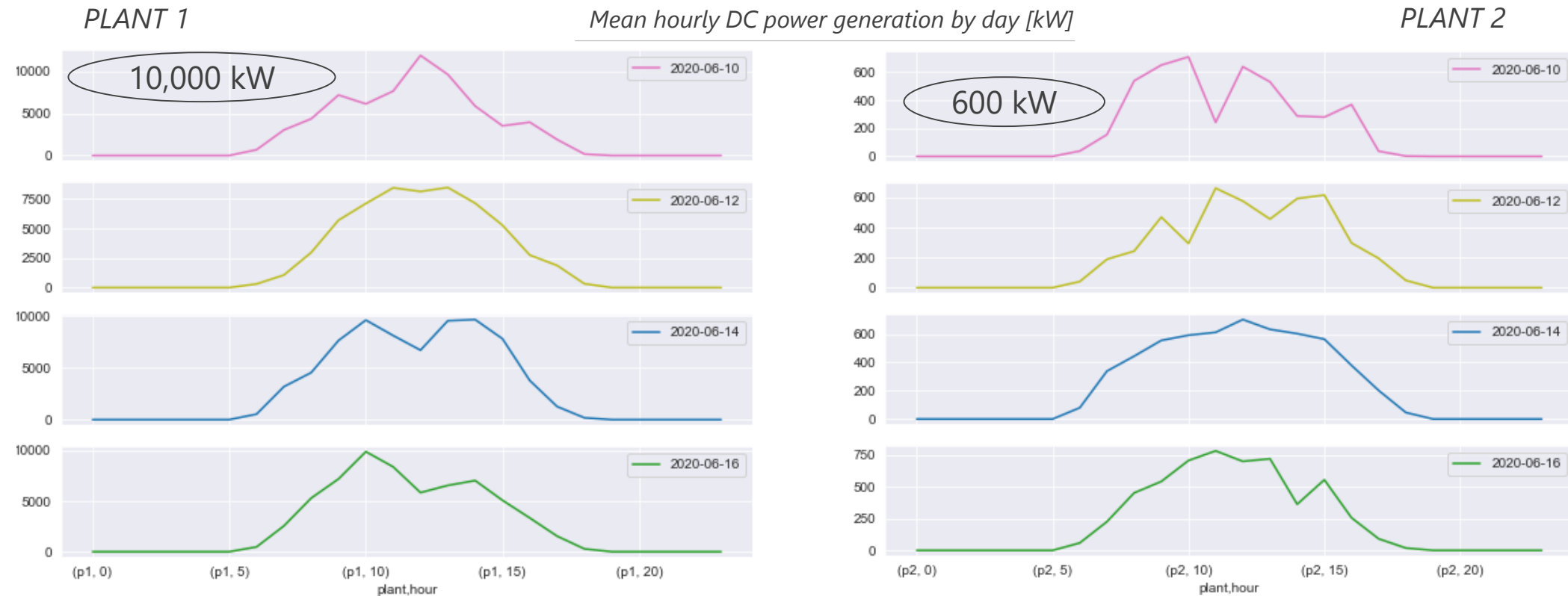
It has been required to review the data acquisition system to check if it has been a data quality issue.

Both power plants receive high, similar amount of irradiance, with no sign of power availability problems



The small difference found between both power plants suggests that they should have very similar DC power generation capabilities.

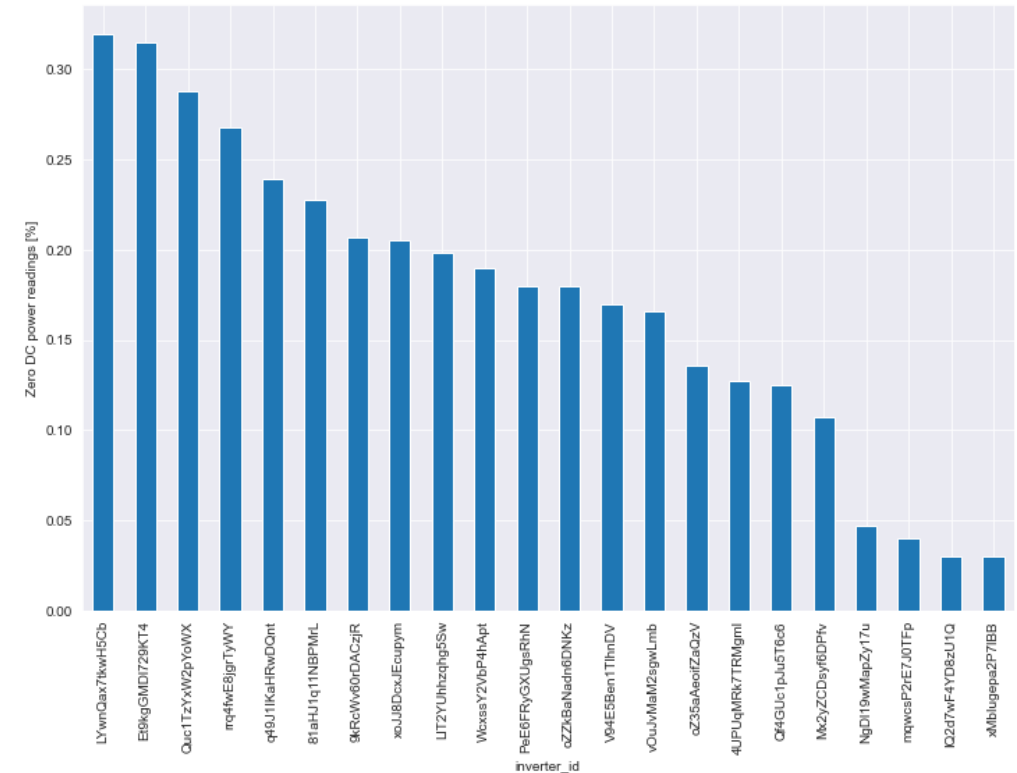
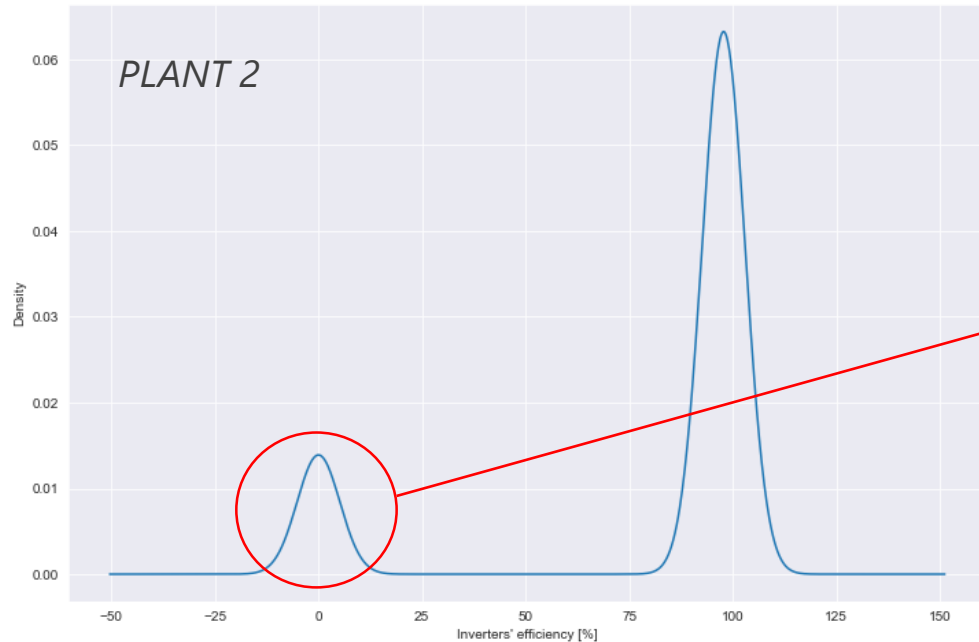
DC power generation in plant 1 is very high, approximately 10 MW of daily DC generation



Plant 1 and plant 2 have the same number of inverters but the DC power produced is around 10 times higher for plant 1. It is suspected that DC power generation in plant 1 is not measured correctly or is artificially scaled up.

Some inverters at plant 2 fail to convert DC power 30% of the time during irradiance peak hours

Mean inverters' efficiency [%] distribution between 09:00 and 15:00



It is highly suggested to check these inverters from plant 2 in the next maintenance check.

After carrying out the data analysis, the following action plan is suggested:

