**BATTLE OF THE SOLAR POWER PLANTS**

#### **Introduction:**

Pakistan' peak electricity demand is estimated at 25,000 MW while its total power generation capacity is 28,000 MWs. Pakistan meets two-thirds of its energy requirement from fuel oil and natural gas. The share of renewables in Pakistan's energy mix is around 5% only. In November 2019 the Government of Pakistan unveiled its new Renewable Energy Development Policy. The policy aims to increase the share of renewables in Pakistan's energy mix by 30%. Around 8,000MW cheap renewable clean and green energy will be added to the system by 2025 while it will be increased to 20,000MW by 2030. In order to meet the goals stated in this policy, Pakistan will have to invest heavily in Solar and Wind energy plants. This will not be possible without Utility-Scale Solar Projects. In this study we will try to determine the best places in Pakistan for setting up a solar Independent Power Plant (Solar IPP) which will provide better electricity output at optimized CAPEX.

#### **Problem:**

Like with all major utility scale projects, transmission cost and transmission losses are a major bottleneck in a project's feasibility. Pakistan's regions which receive very high Direct Normal Irradiation (DNI) of above 6 kWh/m2/day are very distant from the national power grid. Which means developing a solar IPP in one of those regions will have huge transmission costs. In order to overcome this problem I have decided to choose sites which are nearest to existing power stations. Another major factor in developing a Solar IPP is the cost of Land. The nearer the IPP site is to a city or a major population center the higher will be the cost of land and there will also be difficulties in acquiring a single plot of land for the IPP.

#### **Data:**

The World Bank Group has worked extensively on Pakistan's electricity transmission grid data. A GEOJSON file listing all nodes in the national grid is available at their website <https://energydata.info/dataset/pakistan-electricity-transmission-network-2017> I have used that data in this study. There are a lot of un-named entries in the data set which I cleaned to get the locations of the grid stations. Then I by using the Foursuare API I explored venues nearby the grid stations. The purpose was to explore how many venues are there within a range of 15 kms of the Grid Station and how what was the mean distance of those venues from the Grid locations. Greater the number of venues near a grid and smaller mean distance of those venues from the Grid Location would mean that the Grid station is in an urban area and land development costs would be much higher. Finally I use the Global Solar Atlas website <https://globalsolaratlas.info/map?c=30.637912,68.994141,5&r=PAK> to get the Specific photovoltaic power output (PVOUT) values of the grid locations. This results in a dataframe with coordinates of all grid locations, their nearby venues count, mean distances from the venues and the Specific photovoltaic power output (PVOUT) values for each location. Now we can run the K-means clustering algorithm to cluster the grid nodes and label them to find out the most appropriate grid nodes for the development of a Solar IPP.