T5 Project Proposal
Exploratory Data Analysis (EDA)

- **Background:**

The General Authority of Meteorology and Environmental Protection received requests from Ministry of Energy, and they asked to provide them with the three most areas with clear and sunny weather for most of the year, as well as the three highest areas in the average wind speed to set up projects in solar and wind energy.

- **Questions:**
  1-What are the areas where the average clear weather is high?
  2-What are the regions where the average wind speed is the highest and
     most regular?

- **Data Description:**
  o We will use hourly historical weather data for all Saudi Arabia cities from 2017 to 2019 which was collected from Kaggle website the dataset can be found from this link: https://www.kaggle.com/esraamadi/saudi-arabia-weather-history

  o **Columns:**(index, region, date, time, year, month, day, hour, minute, weather, temp,wind, humidity, barometer, visibility)
  o **Data type:** float64(1), int64(9), object(5)
  o **Data size (no. of rows and no of columns):** (249023, 15)

- ➢ A modification has been made to the dataset: ==unnamed column renamed to index and city column renamed to region==.

- **Tools:**
  o Program: Python.
  o Libraries: Pandas, Matplotlib, Numpy.

- **MVP Goal:**
  The MVP goal of this project it is to determine the best suitable place to set up projects in renewable energy (solar energy and wind energy) based on the mentioned dataset.