**ASSIGNMENT 2 – DATA ANALYSIS & VISUALIZATION**

**Objective:**

Perform the Exploratory Data Analysis on the weather data.

**About the Dataset:**

This dataset contains daily weather observations from numerous Australian weather stations over a

11-Year period. This data set have 23 features and 100000 observations.

**Tasks to be carried out:**

1. Find the windy (Highest wind speed and frequent high-speed winds) places to build a windmill. Also find the sunny places to setup a solar field. Take values more than 75 percentile as minimum values for analysis.
2. Find the locations where humidity and temperature are more than 80 percentile value.
3. Find 3 places where rainfall, sunshine, humidity is not too high or too low[ Use logical threshold using percentile values]
4. Find the top 3 places where rainfall is high, but cloud cover is low and vice versa. Take cloud cover avg>5.
5. Which wind direction causes rains in which cities ?
6. Find top 5 coastal cities in which evaporation is higher than mean so that we can choose these places for salt preparation plants?
7. Suggest a strategy to a winter goods company to choose the location and months to promote their products.
8. What factors effect tomorrow rain column the most in weather dataset ?
9. Create a new column as rainfall level with values as no, low, medium, high, heavy as categories and do the bivariate analyses between rainfall level and wind direction   
    - no rain ------> 0 percentile  
    -low --------->0-25 percentile  
    -medium--------->25-50 percentile  
   -high--------->50-75 percentile  
   -heavy--------->75-100 percentile
10. Plot the correlation heatmap between the numerical variables.