Assignment No. 1
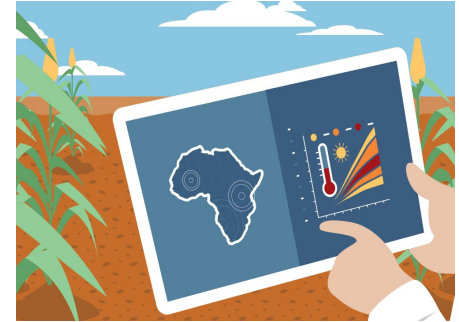
Climate Stakeholders Quest

Ihssene Brahimi

In this assignment, you are asked to finish the following **Quests** :

- **Quest 1 : Exploratory Data Analysis**
- Quest 2 : Forecasting the Future
- Quest 3 : Climate Stakeholders, Time to Act !

- Follow all instructions in the [Notebook](#).
- There are tasks that need to be implemented by code and there are others (e.g. Reflections) that need to be answered by text (Add markdown cells or write them in a separate file).
- When you finish the notebook, summarize all your important reflections and output results (visualizations, stats, model performance, insights, etc.) in this presentation to share with the other AI Forge stakeholders.

Important: Before starting, you should select a stakeholder role, then based on it, you will decide what mission you can solve with the provided dataset, what patterns are more important, etc.

Use Cases

Following are potential challenges or use cases to solve with the dataset :

- **Climate Analysis:** Study long-term climate patterns and variations in different regions.
- **Weather Prediction:** Build models for weather forecasting based on historical data.
- **Environmental Impact:** Analyze air quality and its correlation with various weather parameters.
- **Tourism Planning:** Use weather data to help travelers plan their trips more effectively.
- **Geographical Patterns:** Explore how weather conditions differ across countries and continents.

How would you Use this Data if you were
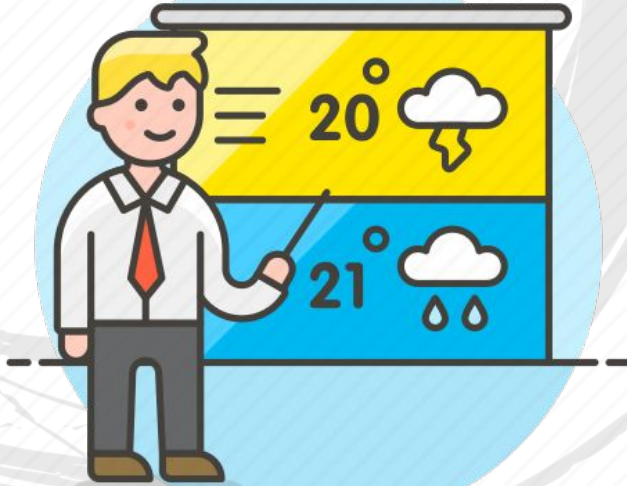
1

A **meteorologist** in Algeria who wants to **validate the accuracy of forecasts** specifically for **their city's weather station**

2

A **Climate Researcher** studying **regional climate trends** in East Africa

3

A **Farmer** in Nairobi

4

A **Local Government Policymaker** in South Africa

Quest 1
Exploratory Data Analysis

- What is your role ?
- Based on your role and [the dataset](#) that you have, what challenge can you solve ?
- What insights do you get from the different types of EDA: Univariate Analysis, Bivariate Analysis, Multivariate Analysis ?
- How would you react if you encounter many missing values that are crucial for your analysis ? What strategy will you pick to overcome that ? Should you remove them or not ?

- What Features are the most important to you ?
- Which preprocessing steps are most helpful for you ?
- How does removing outliers (if Any) affect the distribution precipitation, temperature, etc. (Select features according to what matters to you most) ?
- What insights can you get from the correlation matrix (heatmap) ?
- What challenge can you solve based on your EDA findings ?

References

Dataset:

- https://www.kaggle.com/datasets/nelgiriyewithana/global-weather-repository?resource=download&select=GlobalWeatherRepository.csv

EDA:

- https://www.ibm.com/think/topics/exploratory-data-analysis
- https://www.machinelearningplus.com/machine-learning/how-to-detect-outliers-using-iqr-and-boxplots/
- https://online.stat.psu.edu/stat200/lesson/2/2.2/2.2.10
- https://online.stat.psu.edu/stat200/lesson/3/3.1
- https://online.stat.psu.edu/stat200/lesson/2/2.2/2.2.2

AI Forge Helper Notebooks (Seaborn, Pandas, etc.):

- https://github.com/AI-Forge-Program-Algeria/AI-Forge-Course-2025/tree/main/1_courses/1_beginner_track/2_machine_learning_libraries