ADS-507 – Initial Final Team Project Proposal

Fill out this form and submit it by the end of Module 3 in Blackboard.

Team Number: **2**

Team Leader/Representative: **Aaron Carr**

Full names of team members:

1. **Aaron Carr**
2. **Lai Ieng Chan**
3. **Kyle Esteban Dalope**

Title of your Production-Ready Data Pipeline project: **Pipeline for Analysis of Climate Change Data in Relation to UN Climate Change Committee Global Policy Recommendation Goals**

Short description of your project and objectives:

The project is intended to bring in three datasets in relation to human population and climate related factors (emissions and temperature).

To that aim the team have been tasked by the 28th UN Conference of Parties (COP28) organizing committee to collect data from publicly available source systems with the goal of establishing a database system and pipeline that can be used for climate change impact analyses to inform recommendations for changes in global climate action policies.

Specifically, the team will develop a tracking system for countries' temperature, emissions, and population change data so that policy recommendations can be given to the responsible agencies for each country regarding carbon emission targets and their association to global temperature trends to predict whether the agreed upon goals of not exceeding the 1.5 degrees Celsius are met.

Name of your selected datasets:

1. **World Population Dataset**
2. **Temperature change Dataset**

3. **Emissions by Country Dataset**

Description of your selected datasets (data source, format, size of dataset, etc.):

1. World Population Dataset
  - **Source: Kaggle (https://www.kaggle.com/datasets/iamsouravbanerjee/world-population-dataset)**
  - **Format: .csv**
  - **Size:**
    - **234 instances**
    - **17 columns**

2. Temperature change Dataset
  - **Source: Kaggle (https://www.kaggle.com/datasets/sevgisarac/temperature-change)**
  - **Format: .csv**
  - **Size:**
    - **9656 instances**
    - **66 columns**

3. Emissions by Country Dataset
  - **Source: Kaggle (https://www.kaggle.com/datasets/thedevastator/global-fossil-co2-emissions-by-country-2002-2022)**
  - **Format: .csv**
  - **Size:**
    - **63105 instances**
    - **11 columns**

Please provide the link for your GitHub repository here:
https://github.com/amcarr-ds/ads507_data_engineering

How many times have your members met in the last two weeks? **Twice**

What was the agreed upon method of communication? Are you using any teamwork project management software, such as Trello or Asana? **Slack**

Comments/Roadblocks: **The team has discussed and agreed upon a weekly meeting based on our schedules and to consistently communicate with another via Slack and upload the corresponding data/documents through a shared Google folder for visibility.**