Environmental Analysis on African Subcontinent

Team Members

1. Aman Khandelwal (2022201010)
2. Arun Das (2022201021)
3. Divyansh Negi (2022201014)
4. Nikhil Chawla (2022201045)
5. Nikhil Khemchandani (2022201042)
6. Piyush Singh (2022201032)

Expected Outcomes

1) Developing a website that presents fundamental statistical insights and visualizes environmental indicators through graphs and maps for diverse African nations.
2) Constructing a machine learning prediction model to address a specific environmental challenge.

Data and Methodology

We have managed to identify four major processes in the methodology for the outcomes.

1) Data Collection
   We have found various open environmental datasets for Africa and the countries in Africa. These dataset contain statistics for carbon emission, energy consumption, number of threatened species, etc over the span of a few years.

2) Data Cleaning and Processing
   Once we have the raw data, we need to clean it for processing by performing the following steps.
   - Handling Missing Values: Some of the values in the dataset are blank. We can try interpolating these values from the adjacent values or replace them by the mean depending on the dataset we are using.
   - Dealing with Outliers: We first have to understand why the outlier may be present in the dataset. If it is due to erroneous sensor readings or some other error, we can remove them.
   - Data Transformation: Once we have cleaned the data we need to find an appropriate representation for the machine learning model and the visualizations.
   - Feature Engineering: We can use techniques like dimensionality reduction to create features that make machine learning algorithms work better.

3) Exploratory Data Analysis (EDA)
   We can use the visual tools like bar charts, plots, etc to find some correlation between various parameters of the dataset. We can refine our Machine learning model by giving more weightage to parameters that are more correlated to our target parameter.

4) Data Modeling
   The final step is to use this data for our visualization and as input to our machine learning model. We can then tune the hyperparameters of the model and make better user interfaces for the visualizations.

Sample/Raw Data:

Here are some screenshots of the datasets that we have found. We have columns for the country in Africa and the year in which the reading was taken.

1) Threatened Species

T26	Threatened species					
Region/Country/Area		Year	Series	Value	Footnotes	Source
12	Algeria	2022	Threatene	105		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2004	Threatene	12		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2010	Threatene	21		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2015	Threatene	26		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2019	Threatene	40		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2020	Threatene	42		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2021	Threatene	42		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2022	Threatene	42		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2004	Threatene	2		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2010	Threatene	15		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2015	Threatene	17		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2019	Threatene	24		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2020	Threatene	28		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2021	Threatene	32		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2022	Threatene	33		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2004	Threatene	50		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2010	Threatene	105		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2015	Threatene	114		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.
12	Algeria	2019	Threatene	145		World Conservation Union (IUCN), Gland and Cambridge, IUCN Red List of Threatened Species publication, last accessed May 2022.

2) CO_2 Emission

T27	CO2 emission estimates					
Region/Country/Are		Year	Series	Value	Footnotes	Source
12	Algeria	1975	Emissions	13,692		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	1985	Emissions	42,446		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2005	Emissions	78,045		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2010	Emissions	96,452		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2015	Emissions	1,31,690		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2017	Emissions	1,31,701		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2018	Emissions	1,38,496		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2019	Emissions	1,43,586		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	1975	Emissions	0.8		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	1985	Emissions	1.9		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2005	Emissions	2.3		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2010	Emissions	2.7		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2015	Emissions	3.3		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June
12	Algeria	2017	Emissions	3.2		International Energy Agency, IEA World Energy Balances 2021 and 2006 IPCC Guidelines for Greenhouse Gas Inventories, last accessed June

Milestones

1) 06/09/23 : Relevant Data Collection .
2) 17/09/23 : Data Cleaning and Transformation(I).
3) 30/09/23 : Final Refined Data ,Understanding , Feature Generation , Transformation(II).
4) 17/10/23 : Modeling and Observation(I) , Visualization(II).
5) 01-05/11/23 : Modeling and Observation(II),Visualization(II),Documentation.

FrameWork Tools

1) React/JavaScript for frontend and backend tasks
2) EDA and ML modeling tasks using python.