Executive Summary

By using data from varied sources (The World Bank, The Humanitarian Data Exchange, and Meteo Blue), a global understanding of relationships between potential climate change factors that could explain some key variables such as Temperature and Precipitation could be drawn for Africa as a whole and Kenya more specifically.

The project undertook as my capstone project at Springboard will be investigating the relationship between energy and mining consumption in Nairobi, Kenya and how it affects climate change. Analysis will be briefly global (Africa) before focusing specifically on Kenya.

Even though most of the project was centered around Nairobi, two other African cities (Algiers and Cape Town) were added to initially compare whether some trends in Nairobi's temperature could also be correlated with other parts in Africa. However, temperature variation is challenging to rely on to extract an obvious trend as temperature varies throughout seasons.

The top three possible causes and consequences that saw the biggest increase from 1985 to 2019 were then estimated. The biggest cause was the foreign direct investment (by percentage of GDP), followed closely by the C02 emissions, while the precipitation dominated the top possible leading consequence of climate change in Nairobi during this time.

Beyond relationship trends, correlation matrix offered further insights:

- CO2 seems to be more connected with the Total population growth and Methane emissions.
- Precipitation seems to be related to Ores & metal exports, agricultural land, CO2 emissions, foreign direct investments, Methane & Nitrous Oxide emissions, population in urban agglomerations, the total population, the total cloud cover, and the soil moisture.
- For the temperature variation, we could note some positive relationships with the following: Ores & metals exports, CO2 emissions, total population growth, the urban population, the sunshine duration, and the soil temperature

Eventually, using machine learning, a forecast prediction was achieved for the precipitation and temperature for Nairobi from 2019 to 2029 using data dating as far back as 1985. The predictions, using the eight most possible dominant causes and consequences features of climate change estimated a 0.5% and 3.5% increase in temperature and precipitation respectively by 2030. Those values were in line with predictions from the World Bank.

1. Introduction

In June 2021, I have embarked in an online Data Science bootcamp journey from Springboard that combines programming, mathematics, and technology to analyze data and identify insights.

Data has transformed every industry, and the ability to recognize trends and insights is a highly desirable skill. My goal, at the end of this 6-month journey will be to solidify my skills and tools to leverage statistics and programming to make predictions, optimize outcomes and help guide business decisions.

Founded in 2013, the San Francisco-based Online e-learning website offers several curriculums such as: Data Science & Analytics, Design, Coding and Cyber Security and has boasted about having 2500+ graduates to date.

Each student gets 550+ hours of hands-on curriculum, with 1:1 industry expert mentor oversight. Skills that are sought to be mastered are: Python, SQL, data analysis, data visualization, hypothesis testing and machine learning.

The Data Science track that I have been following is centered around the Data Science Method (DSM), the series of steps data scientists follow to complete projects successfully. Half-way into the curriculum, I have been learning the ins and outs of each of these steps and applying my knowledge to case studies and my capstone projects, one of them being used as my internship.

The project undertook as my capstone project will be investigating the relationship between energy and mining consumption in Nairobi, Kenya and how it affects climate change. Analysis will be briefly global (Africa) before focusing specifically on Kenya.

The six steps that this project followed are:

- Context & Problem Identification
- Data Wrangling
- Exploratory Data Analysis (EDA)
- Pre-Processing and Training Data Development
- Modeling

2. Context

Climate risks pose serious threats to Kenya's sustainable development goals. With the largest economy in East Africa and a population of 48.5 million, Kenya serves as the regions financial, trade and communications hub.

The country's economy is largely dependent on rainfed agriculture and tourism, each susceptible to climate variability and change and extreme weather events. Increasing interseason variability, increasing temperatures, heavy rainfall events and sea level rise lead to severe crop and livestock losses, famine and displacement. High population growth in urban areas is leading to expanding informal settlements, which are at risk from water scarcity, flooding and heat.

Most of the country's coast is low-lying, with coastal plains, islands, beaches, wetlands and estuaries at risk from sea level rise. A sea level rise of 30cm is estimated to threaten 17 percent (4,600 hectares) of Mombasa with inundation. Models estimate that by 2030 climate variability and extremes will lead to losses equivalent to 2.6 percent of GDP annually.

Kenya's geography is dominated by arid and semi-arid plains, with a temperate highland plateau (reaching over 5,000 m) in the center, and a hotter, wetter climate along the coast and the shores of Lake Victoria. Two-thirds of the country receive less than 500 mm (19.6 in) of rainfall per year; coastal and highland areas receive annual averages upwards of 1,100 mm (43 in) and 2,000 mm (78 in), respectively.

Kenya has two rainy seasons: "long rains" from March to June (about 70 percent of total annual rainfall); and "short rains" from October to December. In the west and along the coast, additional significant rainfall occurs outside of these two rainy seasons. Temperatures range from an average of 18°C (64 DegF) in high elevation areas like Nairobi to 26°C (78.8 DegF) in coastal areas such as Mombasa.

Located on the equator, Kenya experiences little seasonal temperature variation.

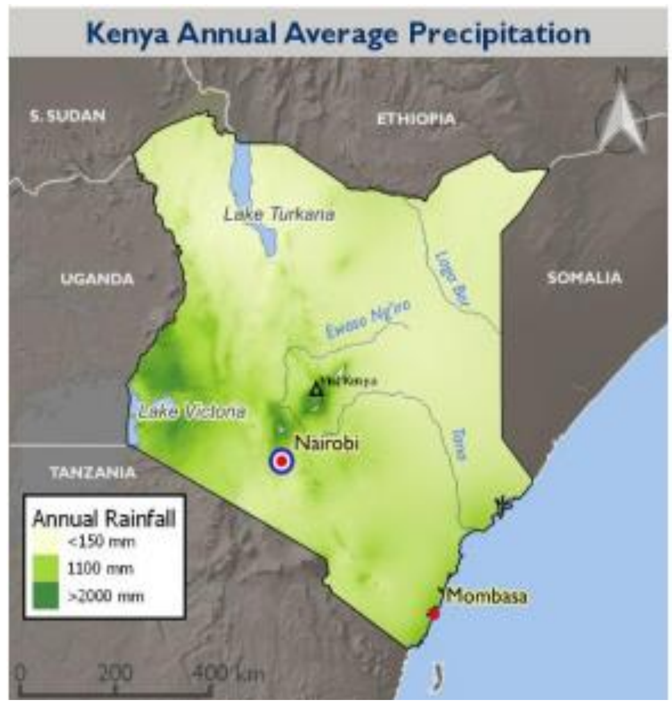

Fig.1 Map of African Countries (Wikipedia), left. Kenya Annual Average Precipitation in 2019 (World Bank)

3. Problem Identification

Stemming from different sources from the World Bank Open Data (***WBOD***), the Humanitarian Data Exchange (***HDE***) and Meteo Blue (meteorological service created at the University of Basel, Switzerland in 2006), several climate-related data were collected to extract insights into possible correlations between causes and consequences of climate change in Kenya. 20+ years of daily data were analyzed and, using machine learning, a prediction of main climate consequences was simulated up until 2030.