Moch5 / karamoja_analysis

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

Watch 0

0 stars 0 forks 0 watching Branches Activity
Tags

Public repository

... 1 Branch 0 Tags Go to file t Go to file Add file Code

Moch5 intro edit f61b941 · 17 minutes ago

SHAPEFILES	Initial commit	3 days ago
TABLES	Data cleaning and Python Viz	2 days ago
viz_images	add karamoja image	2 hours ago
README.MD	display karamoja image	2 hours ago
karamoja_analysis.ipynb	intro edit	17 minutes ago
karamoja_dashboard.pdf	dashboard update	1 hour ago
karamoja_notebook.pdf	add notebook pdf	1 hour ago
karamoja_presentation.pdf	add ln link	1 hour ago

README

Karamoja Agricultural Production Analysis

python 3.8+ Jupyter Notebook pandas 1.3.0+ Matplotlib 3.5+ Seaborn 0.12+ Geospatial Shapefile

Karamoja is the most food-insecure region of Uganda. One of the main reasons is the low productivity of crops due to intense droughts as well as pest and disease outbreaks.

To help NGOs provide technical support and prioritize interventions, this project presents a comprehensive data analysis of agricultural production in Karamoja, examining sorghum and maize yields across 7 districts and 52 subcounties to assess food security patterns and inform targeted interventions.

Table of Contents

Project Overview

This analysis investigates agricultural productivity in Karamoja, Uganda's most food-insecure region, using district and subcounty-level data. By examining relationships between population, land use, and crop yields, this project provides actionable insights for food security interventions.

Core Business Questions

1. Which districts and subcounties have the highest and lowest agricultural productivity?
2. How do maize and sorghum yields compare across the region?
3. What is the relationship between population density and land use efficiency?
4. Which areas are most food secure based on per capita production?
5. Based on the above, what does the data inform

Primary Stakeholders

Stakeholder	Role
Uganda Ministry of Agriculture	Policy development and resource allocation
Food Security NGOs	Program targeting and intervention design
Local Government Officials	District and subcounty planning
Agricultural Extension Services	Field-level technical support

Data Architecture

Data Sources

- **Shapefiles**
  - District boundaries
  - Subcounty boundaries
  - Crop type maps
- **CSV Tables**
  - Production data
  - Population records
  - Agricultural metrics

Dataset Structure

We have two tables of data in our dataset, categorized by the administration structure:

District Level (`district_df`)

```
Shape: (7 rows, 11 columns)
Features: District name, population, area, crop yields, production totals

#### Subcounty Level (subcounty_df)

Shape: (52 rows, 13 columns)
Features: Subcounty name, district, population, yields, crop areas, production
```

Data Understanding

1. Data Loading & Inspection

```
# Load shapefiles for spatial analysis
district_shp = read_shp("SHAPEFILES/Uganda_Districts.shp")
subcounty_shp = read_shp("SHAPEFILES/Uganda_Subcounties.shp")

# Load production data
district_df = pd.read_csv("TABLES/Uganda_Karamoja_District_Crop_Yield_Population.csv")
subcounty_df = pd.read_csv("TABLES/Uganda_Karamoja_Subcounty_Crop_Yield_Population.csv")
```

Key Metrics:

- **7 Districts** with aggregated agricultural data
- **52 Subcounties** with detailed production information
- **Maize & Sorghum** yield analysis
- **Population dynamics** and land use efficiency

2. Feature Engineering

• Calculated **Per Capita Production** metrics to factor in the population in estimating food security levels.

• Derived **Land Use Efficiency** scores to establish which districts could be prioritized on agricultural support and training.

• Aggregated **district-level totals** for maize and sorghum to determine the combined output.

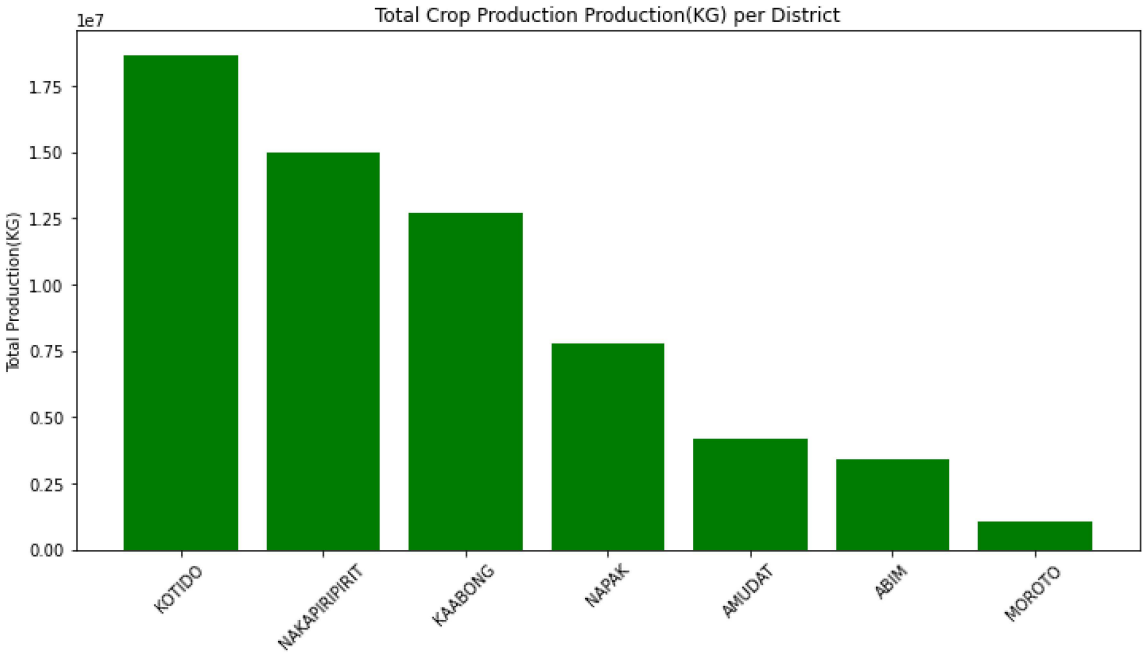
3. Data Integration

Merged subcounty and district data to create a combined enriched master dataset:

```
master_df = pd.merge(subcounty_df, district_df,
                     left_on='DISTRICT_NAME', right_on='DNAME')
```

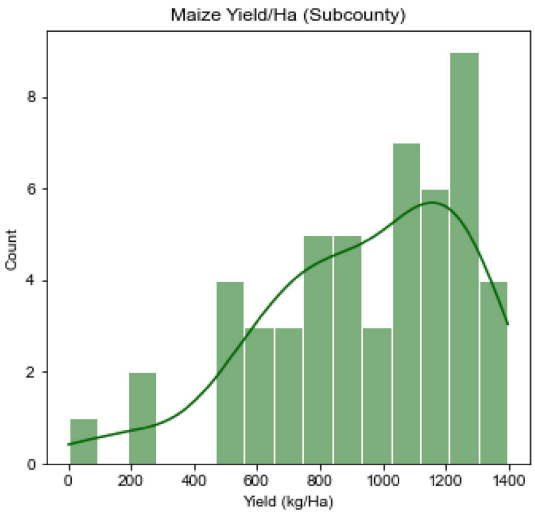
Exploratory Data Analysis

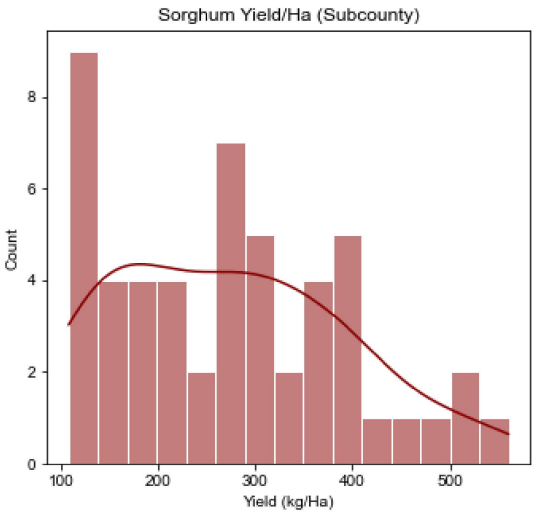
District Production Comparison



Kotido leads production at 18.6 million kg, while **Moroto** underperforms at 1 million kg, indicating significant regional disparities.

Crop Yield Distributions





Maize yields show a **positive** distribution (1000–1200 kg/ha peak), while sorghum shows a **negative** distribution (100–300 kg/ha peak), indicating that maize is more established across the region.

Correlation Analysis

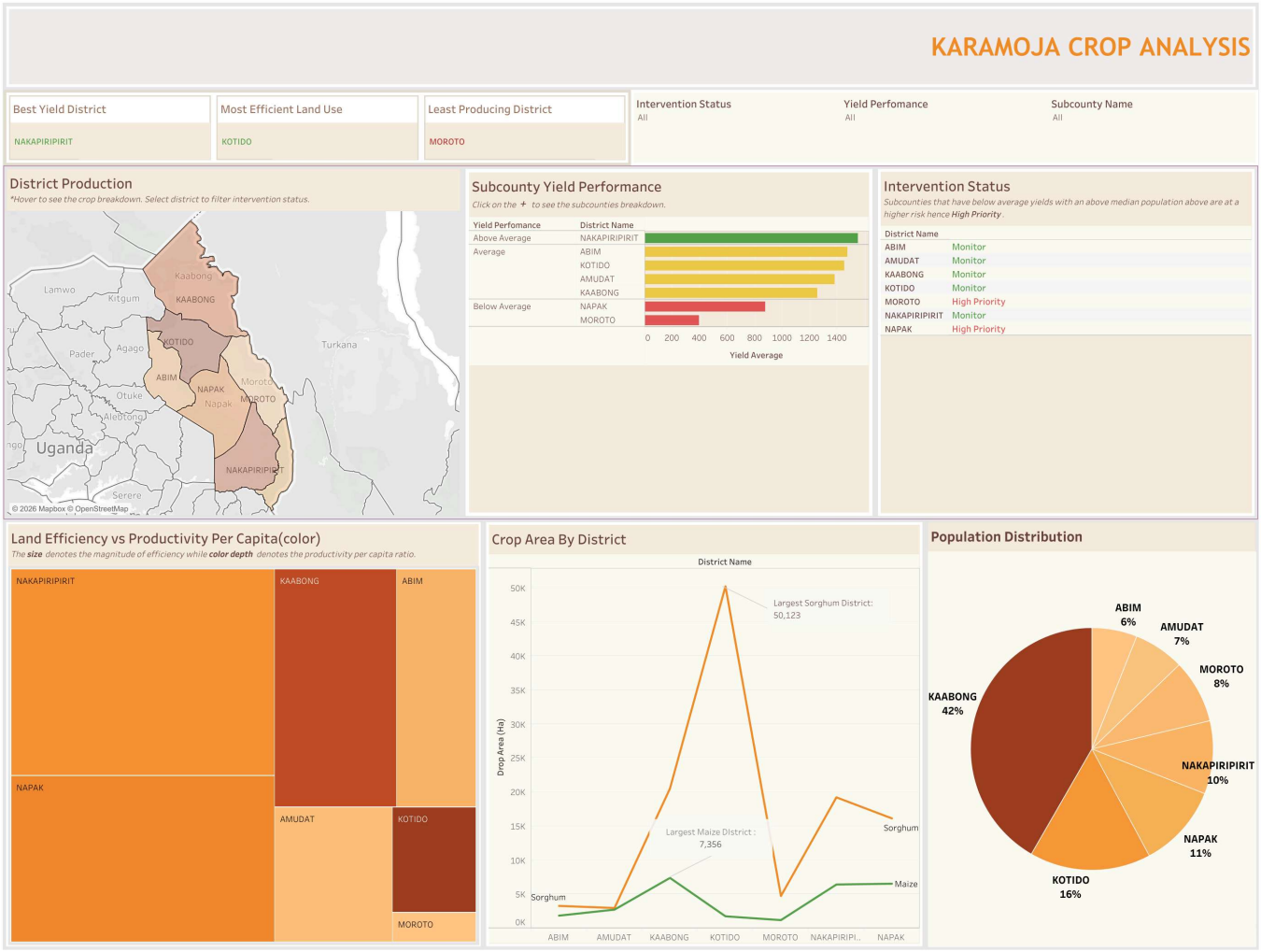
Strong correlation between crop yields suggests shared environmental factors. Weak population–land use correlation indicates the need for agricultural education.

Tableau Dashboard

This dataset was further visualized in Tableau for detailed interactive analysis. This further reinforced our findings as below:

1. Moroto district should be a top priority in the as the intervention status places all its districts under high priority level.
2. Sorghum is always cultivated on a bigger land mass than maize. Whether this is cultural or climatic will inform on what farm inputs/ agricultural training to support the communities with.
3. The most land-use efficient district, Nakapiripirit, has quite a low per capita productivity while the inverse is true for Kotido which has poor land use efficiency but excellent per capita productivity.

Link:
https://public.tableau.com/app/profile/mochama7964/viz/KaramojaCropAnalysis_17716932392280/KaramojaDashboard?publish=yes



Key Findings

1. Dramatic Production Disparities

Kotido district leads with 18.6 million kg in total production, while Moroto produces only 1 million kg, representing an 18-fold difference and signaling an urgent need for targeted support.

2. Crop Performance Patterns

- Maize significantly outperforms sorghum, with average yields of 940 kg/ha versus 274 kg/ha.
- Maize shows strong, consistent production, with most subcounties achieving 1000–1200 kg/ha.
- Sorghum underperforms, with most subcounties producing only 100–300 kg/ha. This is despite the area of sorghum always being significantly higher than that of maize
- These patterns confirm maize as the more established crop, while sorghum presents clear opportunities for improvement.

3. Correlation Insights

- Strong correlation (0.48) between maize and sorghum yields
- Weak correlation (0.1) between population and land use efficiency
- Negligible relationship between maize and sorghum cultivation areas

Recommendations

- Prioritize intensive agricultural interventions in Moroto district, which shows extreme underperformance.
- Implement sorghum improvement programs across all subcounties to address yield gaps.
- Expand agricultural training initiatives region-wide due to weak population–land use correlation as well as poor sorghum yields per hectare.
- Establish ongoing per capita production monitoring to identify food-insecure subcounties.

Future Development

Enrich our dataset with time factor so we can analyze progress over the years from the intervention and possibly perform

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

- **Jupyter Notebook** 100.0%