Analysis of Coconut Production in Lumajang Regency, East Java, Indonesia

By Ahmad Kamiludin

Workflow

01 Business Understanding

02 Data Understanding

03 Data Preparation

04 Modelling

05 Evaluation

06 Deployment

Business Understanding

Lumajang Regency is one of the areas in the southern part of East Java Province. Lumajang Regency has considerable potential in the agricultural, plantation, and other sectors, although it has not been fully exploited optimally. In the plantation sector, the most considerable use of the area in Lumajang Regency in 2021 is the Coconut harvested area of 7,939 ha with a production yield of 8,259.78 tons. Then, followed by 3,962.5 hectares of robusta coffee plantations with a production of 1,312.29 tons. Based on these data, the potential for plantations in Lumajang Regency, especially Coconut, is vast.

Our Data Science team will help find related Attributes or Variables in the coconut production dataset, which can later be used as evaluation material for the Lumajang Regency Government, production efficiency, increasing production, and increasing people's income.

Purpose

Knowing and analyzing what variables are correlated or interconnected in the coconut production dataset in Lumajang Regency using the Correlation Matrix. The results of the Correlation Matrix will later be analyzed for evaluation or policy-making by the Lumajang Regency Government in the plantation sector, especially for coconut commodities.

Benefit

1. **Increase coconut production in Lumajang Regency both in quantity and quality.**
2. **Production efficiency.**
3. **Encouraging the development of coconut plants as one of the producers of industrial raw materials.**
4. **Increase people's income, especially in Lumajang Regency.**

Data Understanding

The Dataset used initially amounted to 4 files, namely datasets about Lumajang Regency plantations from 2018-to 2021. This Dataset is sourced from the official website, namely https://data.lumajangkab.go.id/. But we do not retrieve all the data contained in the Dataset. We only take data on plantation crops specifically for COCONUT commodities. Then, we will create a new excel file. This new Dataset has 4 rows and 8 columns.

Data Understanding

Commodity	Commodity Type.
Year	Year of Production.
Immature Plants	Plants that have not produced results because they are still young have never flowered/fruited or are not old enough to grow.
Mature Plants	Plants that are producing and or have already grown
Damaged Plants	Old plants are damaged and do not provide adequate results anymore; even though there are results, economically, they are no longer productive.
Total Land Area	Total Land Area (Ha).
Production	Total Coconut Commodity Production.
Productivity	A value that shows the average production yield per commodity per unit area of plantation crops in one reporting year.

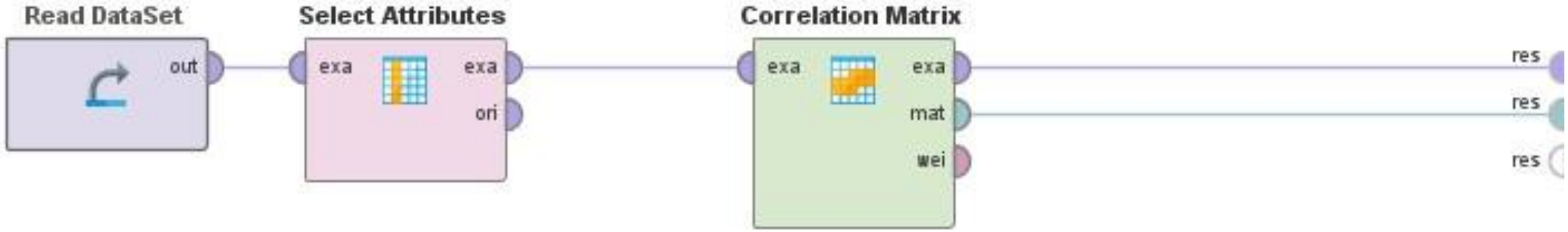
Data Preparation

The data used does not include all the columns in the dataset. Some variables that will not use in the filter first. Then the dataset used is a dataset that only contains variables:

1) **Immature Plants**
2) **Mature Plants**
3) **Damaged Plants**
4) **Total Land Area**
5) **Production**
6) **Productivity**

Modelling

The algorithm used to analyze this dataset is the Correlation Matrix. A Correlation Matrix is used to analyze and determine the relationship between variables in a dataset. Later, we can also find out what factors can affect coconut production.

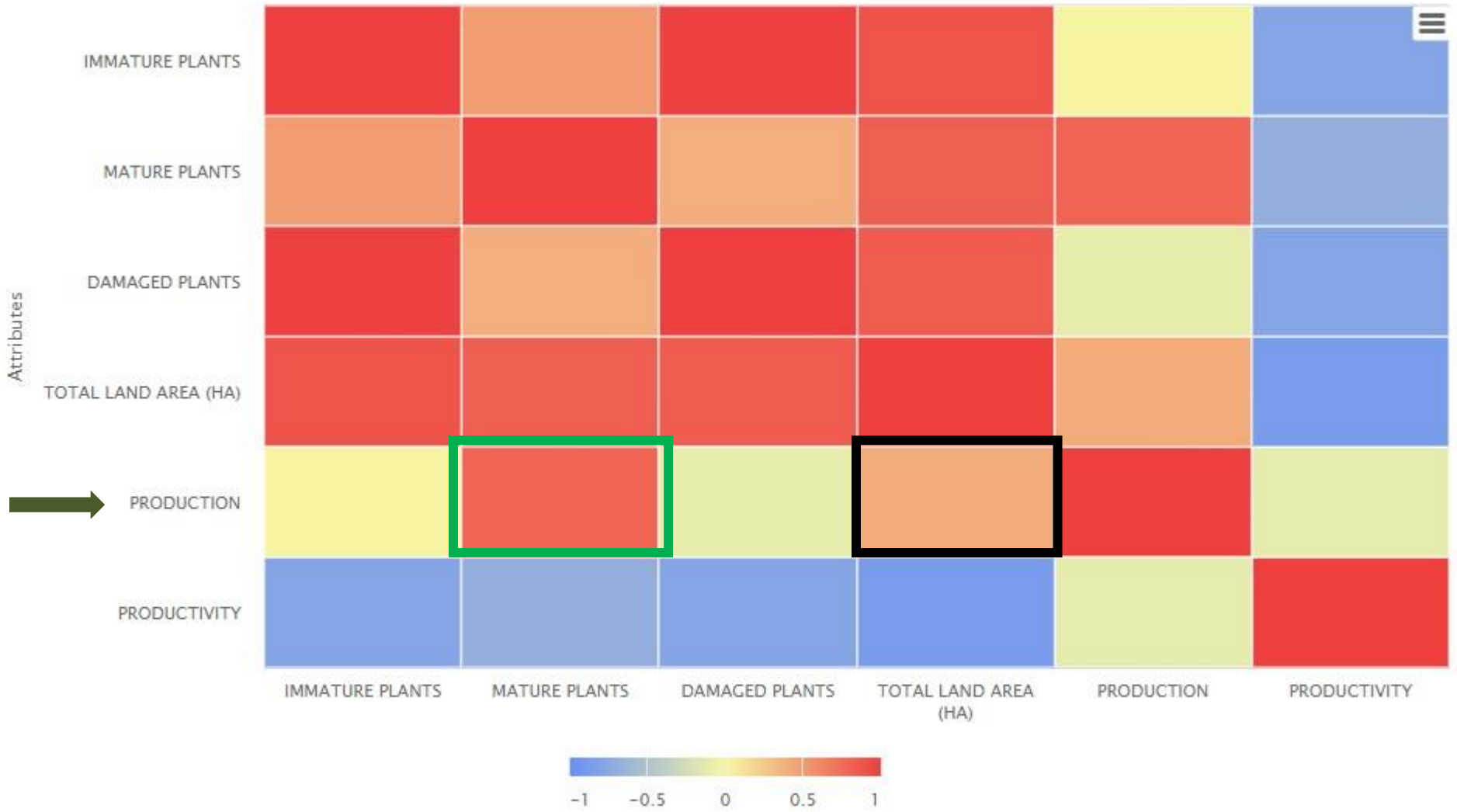


Correlation Matrix

Attributes	IMMATURE PLANTS	MATURE PLANTS	DAMAGED PLANTS	TOTAL LAND AREA ...	PRODUCTION	PRODUCTIVITY
IMMATURE PLANTS	1	0.490	0.993	0.897	0.019	-0.792
MATURE PLANTS	0.490	1	0.402	0.825	0.802	-0.689
DAMAGED PLANTS	0.993	0.402	1	0.849	-0.098	-0.787
TOTAL LAND AREA (HA)	0.897	0.825	0.849	1	0.410	-0.874
PRODUCTION	0.019	0.802	-0.098	0.410	1	-0.119
PRODUCTIVITY	-0.792	-0.689	-0.787	-0.874	-0.119	1

- **The higher the value (the thicker the purple), the higher the correlation level.**

Matrix Visualizations

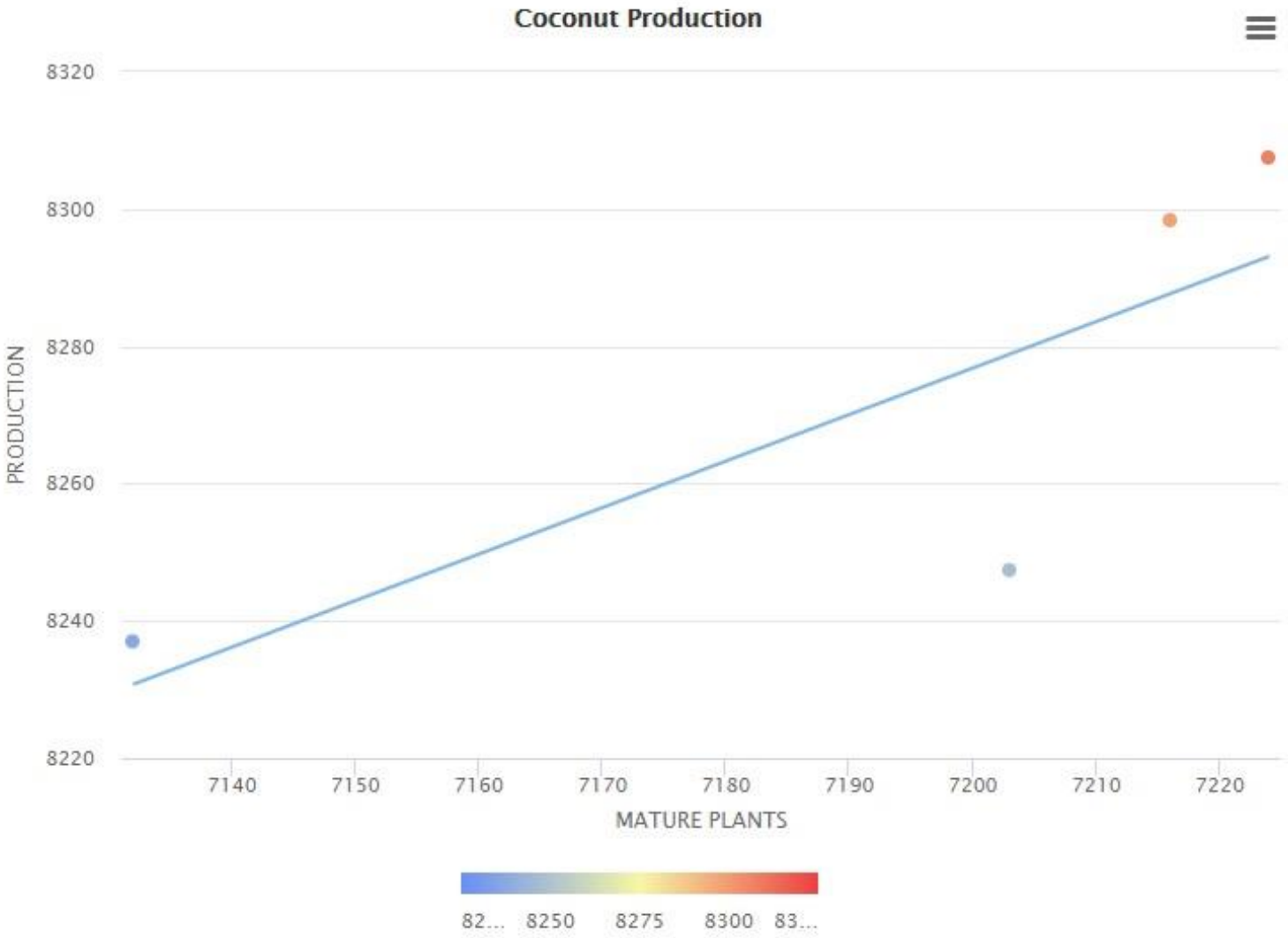


Evaluation

Attributes	IMMATURE PLANTS	MATURE PLANTS	DAMAGED PLANTS	TOTAL LAND AREA ...	PRODUCTION	PRODUCTIVITY
IMMATURE PLANTS	1	0.490	0.993	0.897	0.019	-0.792
MATURE PLANTS	0.490	1	0.402	0.825	0.802	-0.689
DAMAGED PLANTS	0.993	0.402	1	0.849	-0.098	-0.787
TOTAL LAND AREA (HA)	0.897	0.825	0.849	1	0.410	-0.874
PRODUCTION	0.019	0.802	-0.098	0.410	1	-0.119
PRODUCTIVITY	-0.792	-0.689	-0.787	-0.874	-0.119	1

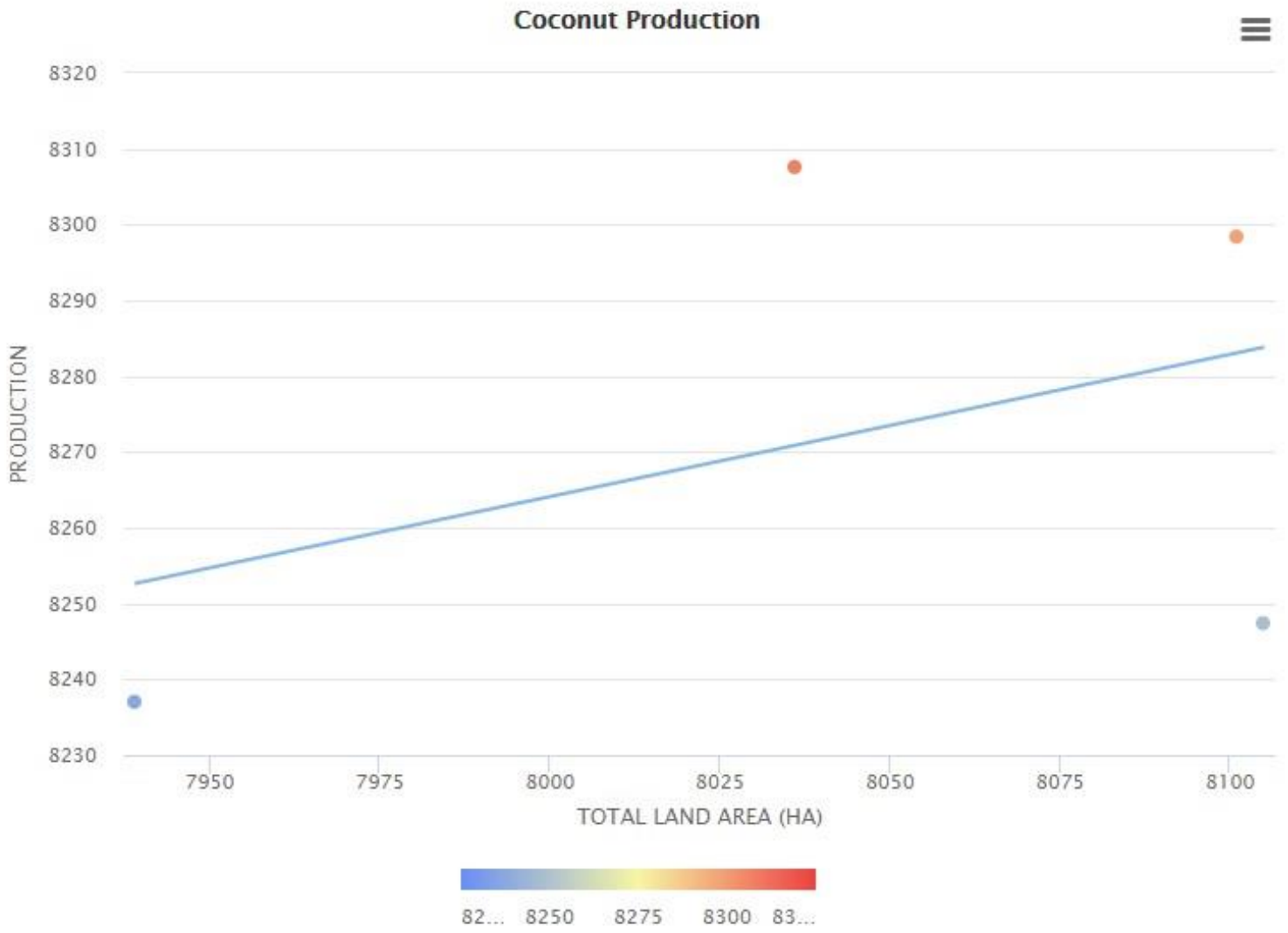
- **The attribute (factor) that has the most significant effect (positive relationship) on coconut production is the mature plant.**
- **Attribute (factor) that has a positive relationship but is pretty low in coconut production is Total Land Area (Ha).**
- **Other attributes (factors) have no relationship or correlation with coconut production.**

Evaluation



- **The graph shows that the mature plants correlate with coconut production.**

Evaluation



- **The graph shows that Total Land Area (ha) positively correlates with Production, but there are also anomalies. Sometimes the Total Land Area is high, but the Production is low.**

Deployment

1. **Focus on the care of Mature Plants, such as Plant sanitation, weed control, fertilization, and others.**
2. **Then, holding special training on coconut commodity maintenance with speakers who are experts in their fields.**
3. **Determine and prepare labor and time for plant maintenance.**
4. **Re-evaluate if there is an expansion of land for coconut production because the amount of land area has a low impact on coconut production.**

Thank You

By Ahmad Kamiludin