

Data Science Project

DATA CENTRALIZATION & DASHBOARD DEVELOPMENT

M. HILAL PRAKOSO ✓

hilalprakoso63@gmail.com

M. HAFIZH NURSYAHID ✓

mhafizhnursyahid@gmail.com

M. FAUZAL DWIANSYAH ✓

anchairis10@gmail.com

WARIH MAHARANI ✓

wmaharani@telkomuniversity.ac.id

INTRODUCTIONS:

Issues:

1. Data management challenges in the coffee industry in Bandung
2. Monthly manual Excel records are time-consuming and error-prone
3. Need for efficient, accurate decision-making

Dataset: Bandung roastery's sales data, 843 rows



Solution:

1. Develop a data dashboard for centralized data management
2. Make use of data science to increase operational efficiency.
3. Implement ARIMA forecasting to predict trends.

METHODS:

Forecasting Method:

ARIMA (Autoregressive Integrated Moving Average)

$$\nabla^d X_t = \sum_{i=1}^q \beta_i \epsilon_{t-i} + \sum_{i=1}^k \alpha_i \nabla^d X_{t-i} + \epsilon_t, \quad (1)$$

$$\tilde{X}_t = \nabla^d \tilde{X}_t + \sum_{i=0}^{d-1} \nabla^i X_{t-1}. \quad (2)$$

For model evaluation, we employ **MAE**, **MSE**, and **AIC** for accuracy.



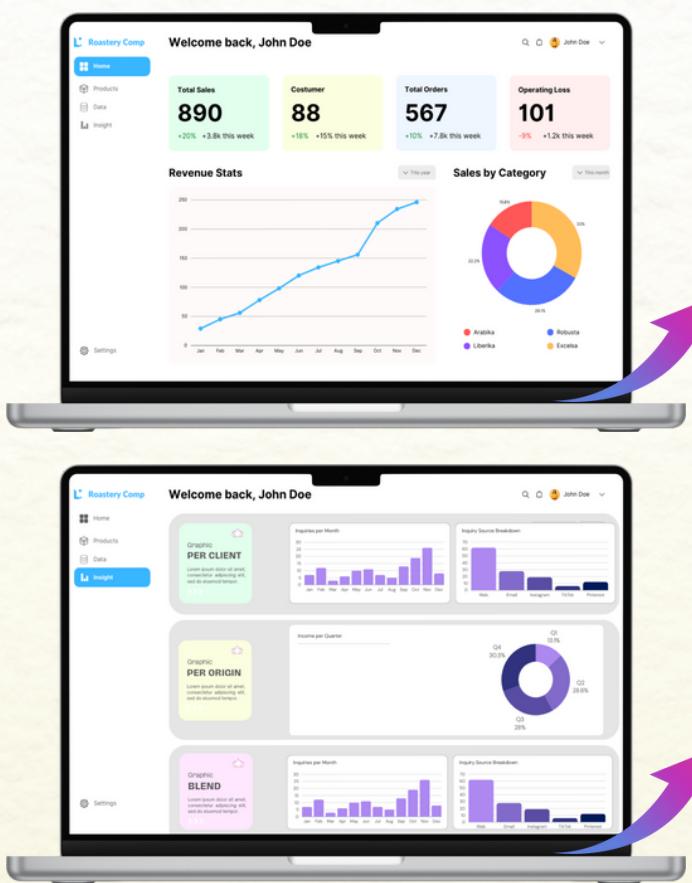
Using Streamlit, we transferred recording data and ARIMA forecasts to the dashboard.

REFERENCES:

- Carta, S., Medda, A., Pili, A., Reforgiato Recupero, D., & Saia, R. (2018). Forecasting e-commerce products prices by combining an autoregressive integrated moving average (ARIMA) model and google trends data. Future Internet, 11(1), 5.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
- Liu, C., Hoi, S. C., Zhao, P., & Sun, J. (2016, February). Online arima algorithms for time series prediction. In Proceedings of the AAAI conference on artificial intelligence (Vol. 30, No. 1).

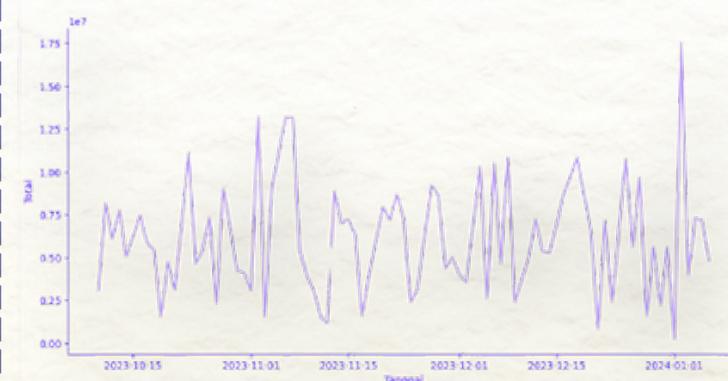
FINDINGS:

The adjacent image showcases the dashboard design for the **Home Page section**.



The accompanying image illustrates the dashboard design dedicated to the **ARIMA Prediction Analysis section**.

DISCUSSION:



The number of coffee beans roasted at the roastery saw an increase in **Q4**. Additionally, the data on the number of roasted coffee beans is normally distributed.

These fluctuations occurred in **Q1 & Q4**, each with distinct patterns.

Fluctuations > Production efficiency, profits, & customer satisfaction.

CONCLUSIONS:

In conclusion, processing the raw data into information are a critical step for effective decision-making, especially in this case for anticipating the future sales. We provide a centralized system and also with the dashboard to provide an overview of current sales situation, allowing the decision-makers to respond swiftly and accurately based on data.