---3---

Time Series Challenge

## **Coffee Sales Time Series Analysis**

#### ***Overview of the Task***

This data science challenge focuses on analyzing and forecasting coffee sales using a time series dataset. The dataset includes various features related to sales transactions, and the goal is to develop models that can accurately predict future sales, identify trends, and optimize sales strategies. Participants will need to preprocess the data, apply time series analysis techniques, and deploy a forecasting solution.

#### ***Problem Statement***

This dataset contains detailed records of coffee sales from a vending machine. Your task is to build a time series forecasting model to predict future coffee sales. You will need to handle data preprocessing, feature engineering, and model selection to generate accurate forecasts.

#### ***Objective***

1. **Data Exploration and Preprocessing**: Clean and preprocess the data to ensure it is suitable for analysis. This includes handling missing values, encoding categorical features, and creating any necessary additional features.
2. **Feature Engineering**: Derive meaningful features from the raw data. This may include extracting time-based features from the 'date' and 'datetime', aggregating sales data by coffee type or payment method, and creating lag features or rolling statistics.
3. **Model Building and Evaluation**: Develop and evaluate various time series forecasting models. Explore methods like ARIMA, SARIMA, Prophet, or LSTM networks. Compare the models using appropriate evaluation metrics.
4. **Forecasting**: Generate and validate forecasts for future coffee sales.
5. **Deployment**: Implement a deployment strategy to make the forecasting model accessible.

#### ***Data Description***

The dataset consists of the following features:

* **date**: The date of the sales transaction in YYYY-MM-DD format.
* **datetime**: The precise timestamp of the sales transaction, including the time of day.
* **cash\_type**: Indicator of the payment method used for the transaction (e.g., cash, credit card).
* **card**: Information related to card transactions, which might include card type or masked card number.
* **money**: The amount of money received in the transaction.
* **coffee\_name**: The name of the coffee product sold.

#### ***Evaluation Criteria***

|  |  |
| --- | --- |
| **Data Collection, Connectivity, and Cleaning** | **30%** |
| **Logic and Results** | **40%** |
| **Presentation** | **5%** |
| **Code Quality** | **5%** |
| **Deployment / Running App** | **20%** |

Good luck with your analysis and forecasting! We look forward to innovative solutions and detailed insights from your work.