Robust yield prediction of various farm processing units

Problem Description:

A new fast food chain is seeing rapid expansion over the past couple of years. They are now trying to optimize their supply chain to ensure that there are no shortages of ingredients. For this, they've tasked their data science team to come up with a model that could predict the output of each food processing farm over the next few years. These predictions could further increase the efficiency of their current supply chain management systems.

In this competition you are expected to build a machine learning model(s) that could predict the output of the food processing farms for the next year.

About Data:

There are 5 datasets along with a sample submission file provided to you in this competition. The datasets are named as follows:

- ***'train_data.csv':***
  - **date:** The timestamp at which the yield of the food processing farm was measured
  - **farm_id:** The farm identifier that recognizes the farm food processing plant
  - **ingredient_type:** The type of ingredient being produced
  - **yield:** The yield of the plant in tonnes

- ***'farm_data.csv':***
  - **farm_id:** The farm identifier that recognizes the farm food processing plant
  - **founding_year:** They year when the operations commenced on the farm and food processing plant.
  - **num_processing_plants:** The number of processing plants present on the farm
  - **farm_area:** The area of the farm in square meters
  - **farming_company:** The company that owns the farms
  - **deidentified_location:** The location at which the farm is present

- ***'train_weather.csv':***
  - *For each location where the farms are present, the weather data is also provided by timestamp*

- ***'test_data.csv' and 'test_weather.csv' are also provided to you***

Objectives:

In this hackathon, you are expected to:

1. Explore the data and engineer new features
2. Predict the yield for each farm during the given timestamps
3. Given the forecasted demand for the next few months for a particular ingredient, device a strategy to source it

Creating a strategy to source 'ing_w' ingredient type

Below is the forecasted demand for the next year for 'ing_w', using the model you built, create a sourcing strategy for 'ing_w' such that the following conditions hold:

- You should source the ingredient at least a month before
- The ingredient has a long shelf life, capable of being used for at least 8 months once stocked

month	demand
1	90430481
2	75344853
3	91917783
4	82591524
5	87419669
6	297500000
7	198900000
8	97367031
9	91049286
10	91755300
11	86533266
12	88514610