Assignment 5

Objective:
The objective of this assignment is to analyze the Rural Municipality Yield Data from 1938 to 2021 and identify any trends, patterns or insights that can be derived from the data. This assignment is designed to provide students with hands-on experience in data science, including data cleaning, exploration, visualization, feature selection anmachine learning model deployment.

Data:
1. **Yield Data**

The Rural Municipality Yield Data contains the yield data for crops grown in rural municipalities from 1938 to 2021. The data includes the following fields:

Year: The year for which the yield data was collected.
Municipality: The name of the rural municipality.
Crops: The type of crop grown.
Yield (bu/acre): The yield of the crop in kilograms per hectare.
Instructions:
2. **GIS Data**

Clean and prepare the data. This may include removing missing values, handling outliers, and transforming the data if necessary.

Explore the data by generating summary statistics, creating visualizations, and identifying trends, patterns, and outliers.

Perform statistical analysis on the data, such as regression analysis and supervised or unsupervised ML(**One Supervised and One Unsupervised Required)** to understand the relationship between different variables and the yield of crops.

Develop insights and conclusions based on your analysis. For example, you might identify which crops have the highest yields, which rural municipalities have the most productive farmland, or which years had the best yields overall.

Note: For this assignment you do not have to write a report or create presentation slides, but a Jupyter Notebook is required.