DATA CLEANING REPORT

DATA CLEANING PROCEDURE (Enumerate the steps to clean the data)	Data cleaning was performed using the Pandas library. First, missing values were detected using the isnull().sum() function, and these rows were removed from the dataset using the dropna() function. Then, duplicate rows were identified using duplicated().sum() and removed using the drop_duplicates() function.
DATA CLEANING SCRIPT(S) DIRECTORY URL (You will upload to somewhere, e.g. onedrive)	https://github.com/AGovem/ENG-346-project/blob/main/Crop_Yield_prediction_ai.py
CLEAN DATA URL (You will upload to somewhere, e.g. onedrive)	https://github.com/AGovem/ENG-346-project/blob/main/cleaned_data.txt
CLEANING TECHNIQUES (e.g. interpolation for column X of data file Y, median for column Z of data file D)	for missing values row deletion method was applied. for duplicate data removing duplicates method was applied.

NOTE: Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. Fixing may be done using various mathematical methods, such as interpolation, extrapolation, putting a fixed value as missing value, putting the mod/median/mean/max/min as missing value.