Python2 Assignment Report: Plotting the Grow Dataset

Objective

The goal of this assignment was to read and clean the GrowLocations dataset, fix any errors in the data, and plot the sensor locations on a UK map using Python. This involved handling invalid latitude and longitude values, correcting column labels, cleaning serial numbers and verifying that the data was within the defined bounding box.

Implementation

1. Loading and Cleaning Data:

- The dataset was loaded into a pandas DataFrame using pd.read_csv.
- Column names were verified and corrected (swapped 'Latitude' and 'Longitude' labels).
- Invalid rows were removed by filtering latitude and longitude values that did not fall within the UK bounding box:
    - **Longitude Range:** [-10.592, 1.6848]
    - **Latitude Range:** [50.681, 57.985]
- Serial numbers were cleaned to retain only the main serial identifier.

2. Visualisation:

- A UK map image was loaded using the PIL library (Image.open) and plotted as the background using matplotlib.
- Sensor locations were represented as red points on the map to clearly indicate their positions.
- Interactive tooltips were implemented to display detailed sensor information when a data point was hovered over.

3. Key Features:

- Interactive tooltips dynamically update to show relevant information.
- The plot includes a title, labels, and a legend for clarity.

Output

The final output clearly plots the sensor locations on the UK map. The use of tooltips improves interactivity, making it easy to explore individual sensor details.

Tools and Libraries

- **pandas** for data manipulation.
- **matplotlib** for plotting.
- **Pillow (PIL)** for image handling.

Conclusion

The code successfully meets the objectives of data cleansing, correcting errors, and visual representation. The interactive map offers an intuitive method for analysing sensor locations throughout the UK.