**Weather Data and City Matching Project**

**Overview**

This project involves cleaning, formatting, and analyzing city data from various sources, performing fuzzy matching for city names, and retrieving weather data using an API. It includes several milestones that guide the data processing steps, from initial data loading to final data presentation and ethical considerations.

**Author**

Saron Yaya

**Initial work**

* [Portfolio Projects](https://github.com/Saron222/PortfolioProjects)

**Released on**

* GitHub

**My professional profile on LinkedIn**

* [My LinkedIn Profile](https://www.linkedin.com/in/saron-yaya/)

**Showcase**

**Data Cleaning and Formatting**

We start by loading and cleaning the city data from a CSV file and a web-scraped HTML file. This involves sorting, checking for duplicates, and handling missing data.

**Example:**

**Fuzzy Matching**

Using fuzzy matching algorithms, we ensure that city names entered by users are accurately matched to the correct entries in our dataset.

**Weather Data Integration**

We fetch real-time weather data for the cities using the OpenWeatherMap API and integrate this data with our city dataset.

**Technologies Used**

* **Python**: The primary programming language used for data processing and analysis.
* **Pandas**: For data manipulation and analysis.
* **BeautifulSoup**: For web scraping.
* **Requests**: For making HTTP requests to the weather API.
* **Fuzzywuzzy**: For fuzzy string matching.
* **Matplotlib**: For data visualization.
* **Jupyter Notebook**: For interactive development and presentation.

**Installation**

To set up this project on your local machine, follow these steps:

1. **Clone the Repository**: git clone [https:// https://github.com/Saron222/PortfolioProjects.git](https://github.com/saronyaya/DSC510.git)
2. **Prepare Data Files**:
   * Place the **worldcities.csv** file and **Cities by Latitude & Longitude.html** in the project directory.
   * Provide your OpenWeatherMap API key in a file.
3. **Run the Scripts**:
   * Execute the Jupyter Notebooks or Python scripts corresponding to each milestone to perform data processing and analysis.

**How to Run**

1. Open the Jupyter Notebooks in the project directory to follow along with the data processing steps.
2. Modify and run the Python scripts as needed to clean data, perform fuzzy matching, and fetch weather data.

**Contributing**

1. Fork the repository from [Saron222/PortfolioProjects](https://github.com/Saron222/PortfolioProjects/fork)
2. Create your feature branch: **git checkout -b feature/your-feature-name**
3. Commit your changes: **git commit -am 'Add your feature'**
4. Push to the branch: **git push origin feature/your-feature-name**
5. Create a new Pull Request in the original repository