

Prediction and statistical analysis of Land Prices

Problem Statement: The project involves scraping land price data from a static website focused on Nagpur city. Utilizing Python libraries like BeautifulSoup and pandas, the collected data will be stored in a CSV file and then analyzed and visualized using matplotlib. Further, machine learning algorithms from scikit-learn will be employed to build a predictive model for land prices based on features such as location and area.

Softwares used: Jupyter, Python, Python libraries(NumPy,Pandas,Seaborn,Requets, BeautifulSoup,SciKit,MatPlotLib,JobLib,CSV), Html, CSS,etc.

Process:

Step 1: **Data Generation:** Generated data for various land prices of Nagpur city for analysis purpose.

Step 2: **Static Webpage Creation:** Created a static webpage to showcase the areas/lands in Nagpur city and added that generated data to the webpage.

Step 3: **Web Scrapping:** scrapped the data from the webpage and stored it in a CSV file. This step demonstrates that if the land price data is already available in some other webpage than we can scrap it and our model can analyze that.

Step 4: **Model Training:** Train a linear regression model on housing data, make predictions, evaluate using Mean Squared Error, visualize results, and save the model along with actual vs. predicted prices.

Step 5: **Prediction:** Load the trained Linear Regression model, read new land price data, preprocess it, ensure column consistency with the training data, make predictions, and save the dataset with predicted prices to 'ml_generated_price.csv'.

Step 6: **Displaying the predicted results:** Displayed the generated results according to the location provided by the user.

Step 7: **Analysis:** Used python libraries such as MatPlotLib and Seaborn for further analysis and graph generation for the ML generated dataset.