# ML Project: Predicting Selling Prices

## Table of Contents

- Requirements  
- Installation  
- Usage  
- Project Structure  
- Steps to Run

## Requirements

Ensure you have the following installed:

* Python 3.7 or higher
* Packages: pandas, numpy, scikit-learn, and matplotlib

You can install the necessary packages using the following command:

pip install -r requirements.txt

## Installation

1. Clone this repository:  
 git clone https://github.com/your-username/your-repository-name.git  
 cd your-repository-name  
2. Install dependencies as specified in requirements.txt.

## Usage

Run the main Python file (main.py) to execute the project step-by-step. Follow the instructions below to load, clean, train, and test the data.

## Project Structure

The project directory includes:

* main.py: Main script to load, clean, train, and test the dataset.
* dataset/: Folder containing the training and test datasets.
* requirements.txt: List of required packages.

## Steps to Run

1. Run the Python File: Begin by running main.py:

python main.py

1. Load the Dataset: From within the program, load the training dataset located in the dataset/ folder:

train\_data = pd.read\_csv('dataset/train.csv')

1. Clean the Training Data: The data needs to be cleaned before it can be used for training. This may include handling missing values, removing or imputing outliers, scaling features if necessary.

Train the Data: After cleaning, split the training data into input features (X) and target labels (y). Select the Linear Regression model:

1. from sklearn.linear\_model import LinearRegression  
   model = LinearRegression()

Train the model with the cleaned training data:

1. model.fit(X\_train, y\_train)

Predict Selling Prices on Test Data: Load the test data from dataset/ and use the trained model to predict selling prices.

1. test\_data = pd.read\_csv('dataset/test.csv')  
   predictions = model.predict(test\_data)

Output Predictions: The predicted selling prices can be displayed or saved to a file as follows:

1. import pandas as pd  
   pd.DataFrame(predictions, columns=['PredictedPrice']).to\_csv('predictions.csv', index=False)

## Conclusion

Following these steps, you can load, clean, and train a Linear Regression model on the training data and predict selling prices on the test data. For more details on each step, refer to main.py.