**Project Report: House Price Prediction**

**1. Introduction** The goal of this project is to predict house prices based on the size of the house using a simple linear regression model. The dataset used contains missing values, which were handled before training the model.

**2. Dataset Description**

* The dataset consists of two columns: Size (sq ft) and Price ($).
* Some values in the Size column were missing and needed to be filled before model training.

**3. Data Preprocessing**

* Loaded the dataset using pandas.
* Checked for missing values and identified that the Size column had NaN values.
* Replaced NaN values in the Size column with the **mean** of the available sizes.
* Verified that no missing values remained after imputation.

**4. Model Implementation**

* Used LinearRegression from sklearn.
* Defined the independent variable (Size) and dependent variable (Price).
* Trained the linear regression model using the cleaned dataset.

**5. Results & Findings**

* The model successfully fit the data, and a linear relationship was observed.
* The regression equation obtained was:

**Price = m \* Size + b**

where m is the slope (coefficient) and b is the intercept.

* Predictions for new house sizes could now be made using the trained model.

**6. Conclusion** This project demonstrated how to handle missing values and build a simple regression model to predict house prices. Future improvements could include using multiple features (e.g., number of rooms, location) for a more accurate prediction.

**7. Future Enhancements**

* Use more features to improve accuracy.
* Implement different regression models and compare performance.
* Visualize the data and model predictions using matplotlib or seaborn.

**8. Code Summary**

import pandas as pd

from sklearn.linear\_model import LinearRegression

# Load dataset

df = pd.read\_csv("house\_prices\_with\_nulls.csv")

# Handle missing values

df['Size'] = df['Size'].fillna(df['Size'].mean())

# Define model

reg = LinearRegression()

reg.fit(df[['Size']], df['Price'])

# Predict

reg.predict(size)

**End of Report**

**HAVE A GOOD DAY!**