**REPORT**

* **Importing Libraries**:
* import pandas for data manipulation, matplotlib for visualizations, and numpy for numerical operations.
* **Loading the Dataset**:
* df = pd.read\_csv('housing\_prices.csv'): Loads the housing data from a CSV file into a pandas DataFrame for analysis.
* **Dataset Overview**:
* df.info() shows information about the dataset (column names, data types, and missing values).
* **Handling Missing Data**:
* df.isnull().sum() checks for missing values in the dataset.
* df['price'].fillna(df['price'].mean()) fills missing values in the 'price' column with the mean of the column.
* **Removing Duplicates**:
* df.drop\_duplicates() removes any duplicate rows to ensure clean and unique data.
* **Calculating Mean Price**:
* df['price'].mean() calculates and prints the average house price in the dataset.
* **Filtering Data (More Than 3 Bedrooms)**:
* filtered\_data = df[df['bedrooms'] > 3] filters the dataset to include only houses with more than 3 bedrooms.
* avg\_price\_by\_bedrooms = df.groupby('bedrooms')['price'].mean() groups the data by the number of bedrooms and calculates the average price for each group.
* **Bar Chart (Average Price by Bedrooms)**:
* avg\_price\_by\_bedrooms.plot(kind='bar') plots a bar chart showing the average price of houses for each number of bedrooms.