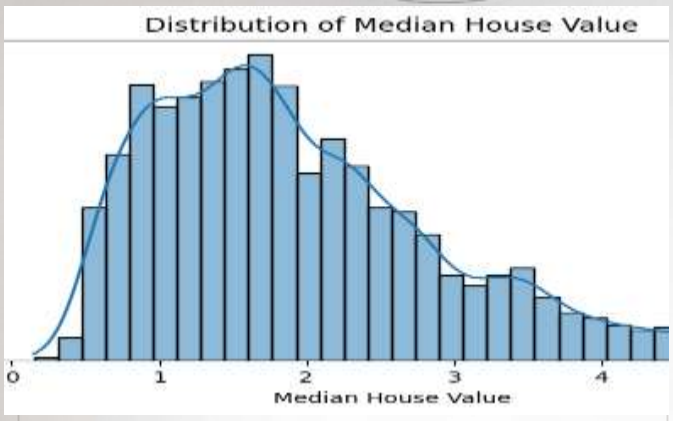
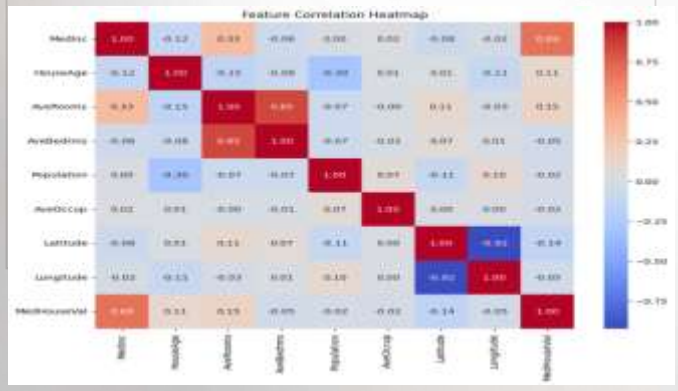
TITLE: HOUSE PRICE PREDICTION USING LINEAR REGRESSION

DATASET: CALIFORNIA HOUSING DATASET

OBJECTIVE: PREDICT MEDIAN HOUSE VALUES USING LINEAR REGRESSION

TOOLS: PYTHON, SCIKIT-LEARN, PANDAS, MATPLOTLIB, GOOGLE COLAB





DATASET & EDA

DATASET OVERVIEW:

Source: California Housing Dataset

Features: Income, house age, rooms, population, location

Target: Median House Value

EDA INSIGHTS:

No missing values found

Median Income has highest correlation with house prices

Geographic features influence pricing

MODEL & EVALUATION

EVALUATION METRICS:

- MAE: (0.5332001304956553)
- RMSE: (0.74558138301277)
- R^2 Score: (0.5757877060324)

INTERPRETATION:

- Model explains a significant portion of price variation
- Performs reasonably well for a baseline model

CONCLUSION&FUTURE IMPROVEMENTS

CONCLUSION:

- Linear Regression successfully predicts house prices
- Model performance can be improved with advanced algorithms
- Project demonstrates complete ML workflow

FUTURE IMPROVEMENTS:

- Try Ridge/Lasso regression
- Use Random Forest or XGBoost