

[California Housing] Project Overview

Dataset

California Housing Data from 1990 census

Purpose

Data Analysis to get insights about California housing

Prediction Model for California housing price

Skills used

Python: pandas, numpy matplotlib, seaborn

Data Analysis

[FINDING 1] Coastal Area has higher house value

[FINDING 2] Coastal Area has higher house value

Prediction Model

Model Type Linear Regression

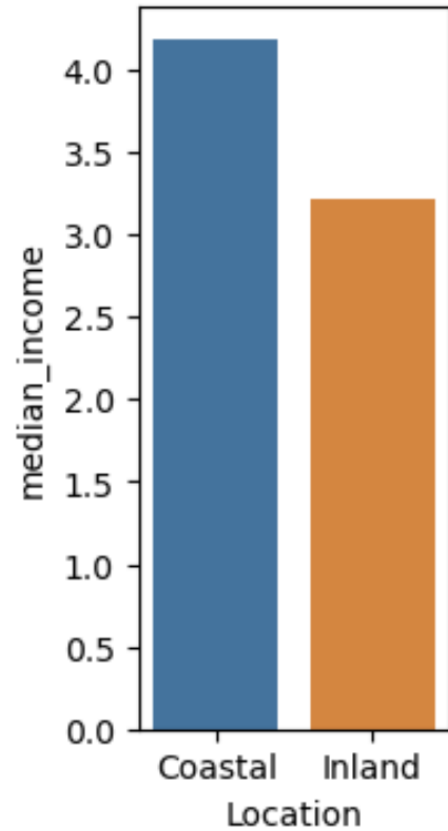
Outcome

MSE : 21212 / **R2Score** : 80% of accuracy

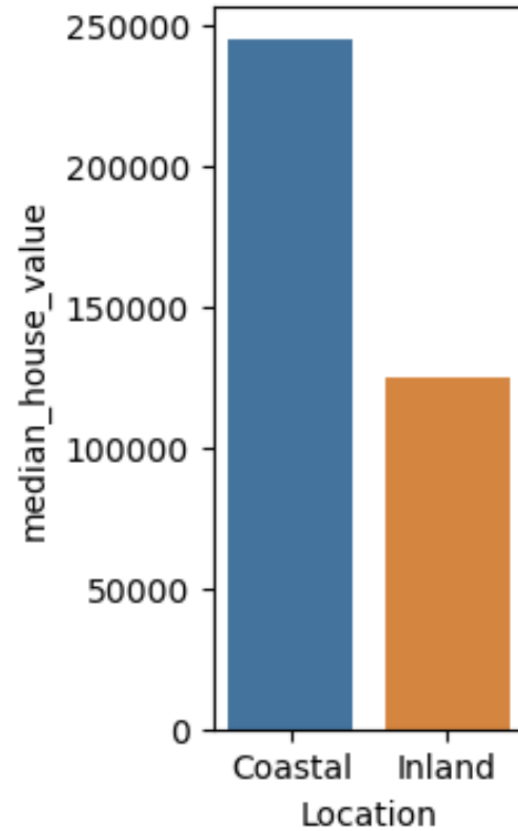
DATA ANALYSIS

[KEY FINDING 1] Coastal area has higher housing value

Income (Coastal vs Inland)



House Value (Coastal vs Inland)



The original Dataset has 5 categories for ocean Categories