README Content

Project Title

California Dreaming: An In-Depth Analysis of 1990 Housing Prices

Brief Description of the Dataset

This project analyzes the California Housing Prices dataset, derived from the 1990 census. It includes data on median house prices and other housing statistics for various districts in California. The dataset features variables such as longitude, latitude, housing median age, total rooms, total bedrooms, population, households, median income, median house value, and ocean proximity.

Summary of Analysis Steps

Data Loading and Initial Exploration: Loaded the dataset and examined its structure and basic statistics.

Data Cleaning and Preprocessing: Handled missing values in the 'total_bedrooms' column.

Exploratory Data Analysis (EDA): Conducted a thorough exploration of the dataset through summary statistics and visualizations.

Feature Engineering: Created new features, including rooms per household, bedrooms per room, population per household, income category, and age category, to gain deeper insights.

Visualizations: Generated histograms, bar charts, a correlation heatmap, scatter plots, and box plots to visualize the relationships and distributions in the data.

Folder Structure and Contents Description

- Data: Contains the original and modified datasets.
- Notebooks: Includes Jupyter notebooks for data cleaning, EDA, feature engineering, and visualizations.
- Visualizations: Stores all the visual outputs of the analysis.
    - Histograms
    - Bar Charts
    - Correlation Heatmap
    - Scatter Plots
    - Box Plots

- Reports: Holds the detailed analysis report and other documentation.

Instructions for Replicating the Analysis

To replicate this analysis:

Download the dataset and place it in the 'Data' folder.
Run the Jupyter notebooks in the 'Notebooks' folder in sequential order.
Visualizations generated in the notebooks can be found in the 'Visualizations' folder.

Contact Information or Credits

For any inquiries or further information, please contact Mohamad Chokr at mohamad00122@gmail.com.