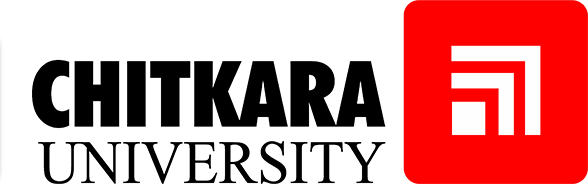
**Artificial Intelligence and Machine**

**Learning**

Project Report

Semester-IV (Batch-2022)

**Zillow Home Prediction**



**Supervised By: Submitted By:**

Mr.Rajeev Bhardwaj Tanish Dalal(2210990888)

**Department of Computer Science and Engineering**

**Chitkara University Institute of Engineering &Technology,**

**Chitkara University, Punjab**

**ABSTRACT**

In this project, we aim to predict the value of homes using a method called machine learning. We'll use data about different features of homes, like their location, size, and number of rooms, to train our model. The goal is to create a tool that can estimate how much a house is worth in the real estate market. By analyzing large amounts of data, our model will learn patterns and relationships to make accurate predictions. This project can help homeowners, buyers, and sellers make informed decisions about property investments.

In this project, we're diving into the world of real estate by predicting home values using machine learning techniques. By analyzing various aspects of homes, such as their location, size, and amenities, we aim to build a model that can estimate their market prices accurately. This involves collecting and processing large

amounts of data from sources like Zillow. Our model will learn from this data to identify patterns and relationships between different features and home prices. Ultimately, our goal is to create a valuable tool for homeowners, buyers, and sellers to make informed decisions about property transactions.

**INDEX**

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Title** | **Page Number(s)** |
| 1 | Introduction | 3-4 |
| 2 | Problem Statement | 5 |
| 3 | Software Requirements | 6 |
| 4 | Proposed Design | 7 |
| 5 | Code | 8-16 |
| 6 | Results | 16-20 |
| 7 | References | 21 |

# Introduction

House prediction using machine learning (ML) is a burgeoning field aimed at leveraging data analysis and predictive modeling to estimate the value of residential properties. With the advent of advanced algorithms and the availability of vast datasets, ML-based house prediction projects have gained traction in real estate markets worldwide. These projects typically involve the exploration of various features such as property size, location, amenities, and historical sales data to develop accurate models for predicting house prices.

* 1. **Background:**

The real estate market is a complex ecosystem influenced by numerous factors such as economic conditions, demographics, location, and property characteristics. Traditionally, predicting house prices has relied on manual analysis and expert judgment, which can be time-consuming and prone to biases. However, with the advent of machine learning (ML) and data-driven approaches, the landscape of house prediction has evolved significantly.

* 1. **Objectives:**

The primary objective of this house prediction project is to develop a machine learning (ML) model capable of accurately estimating house prices based on various property features and market dynamics. Specific objectives include:

1.Building a comprehensive dataset containing relevant information about past property sales, including factors such as property size, location, amenities, and sale prices.

2.Implementing ML algorithms, including but not limited to linear regression, decision trees, and ensemble methods, to analyze the dataset and develop predictive models for house prices.

3.Evaluating the performance of different ML algorithms using metrics such as mean absolute error, mean squared error, and R-squared to identify the most accurate model.

4.Incorporating advanced techniques such as feature engineering, data normalization, and model tuning to improve the accuracy and robustness of the predictive models.

**Significance:**

The significance of this house prediction project lies in its potential to transform the real estate industry by providing stakeholders with advanced tools for informed decision-making and optimizing property transactions. Key aspects of its significance include:

Improved Decision Making: By leveraging machine learning algorithms to analyze vast datasets and predict house prices accurately, this project empowers homeowners**, buyers, sellers, and** real estate professionals to make well-informed decisions regarding property investments, pricing strategies, and market trends.

Enhanced Efficiency: Traditional methods of house price prediction often rely on manual analysis and expert judgment, which can be time-consuming and subjective. By automating the prediction process through machine learning, this project streamlines the valuation process, saving time and resources for all parties involved.

Market Transparency: By providing accurate and transparent house price predictions based on objective data analysis, this project promotes market transparency and fairness, reducing information asymmetry and mitigating the risk of overpricing or underpricing properties.

# Problem Statement

In the realm of house prediction, the current methods for estimating property prices often lack accuracy and fail to account for the diverse range of factors influencing real estate values. Traditional approaches may rely heavily on manual assessment or simplistic models that overlook critical variables, leading to inaccurate predictions and suboptimal decision-making for stakeholders.

# Software Requirements

1. **Programming Language:** Python will be used as the primary programming language for developing the recommendation system due to its extensive libraries and frameworks for data processing and machine learning**.**
2. **Libraries:**

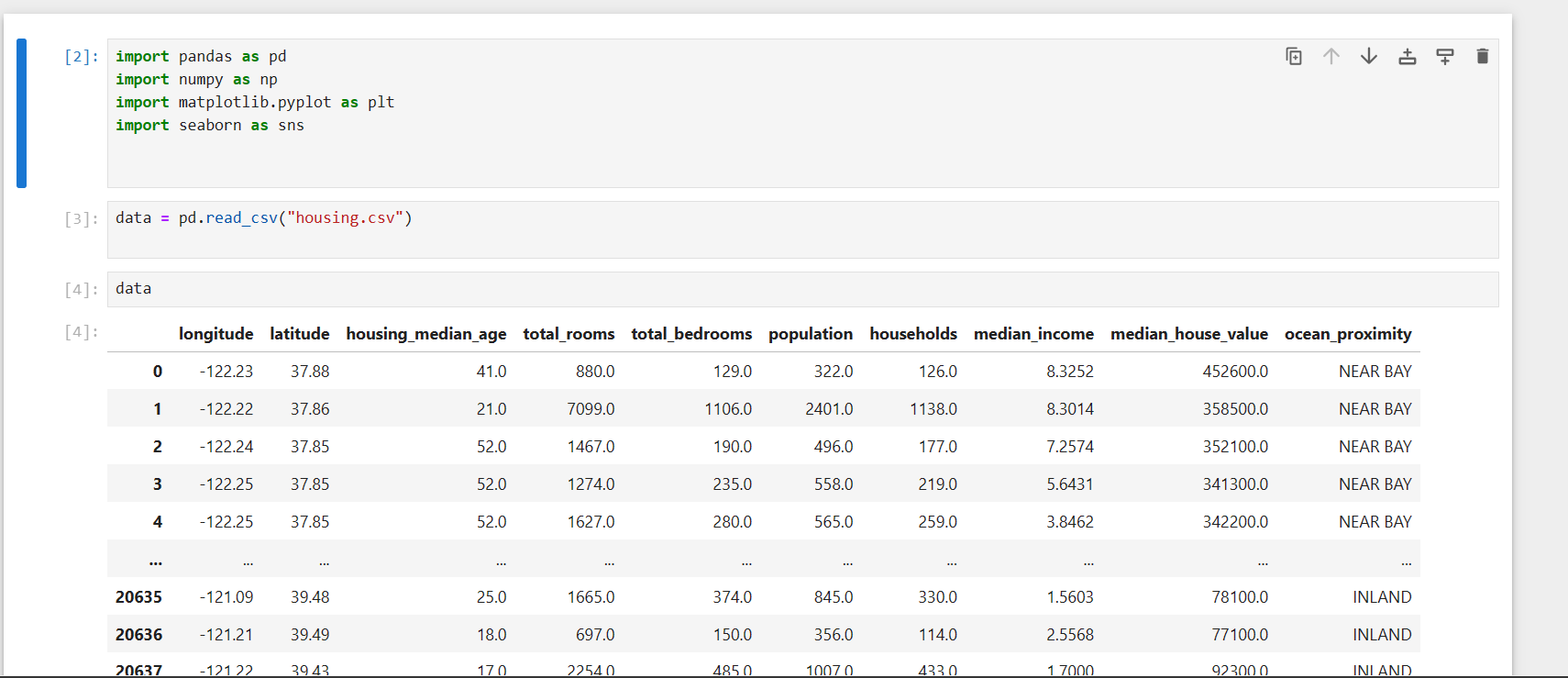
* **Pandas:** For data manipulation and analysis.
* **NumPy:** For numerical operations and array manipulation.
* **scikit-learn:** For machine learning algorithms and model evaluation.
* **Matplotlib and Seaborn:** For data visualization**.**

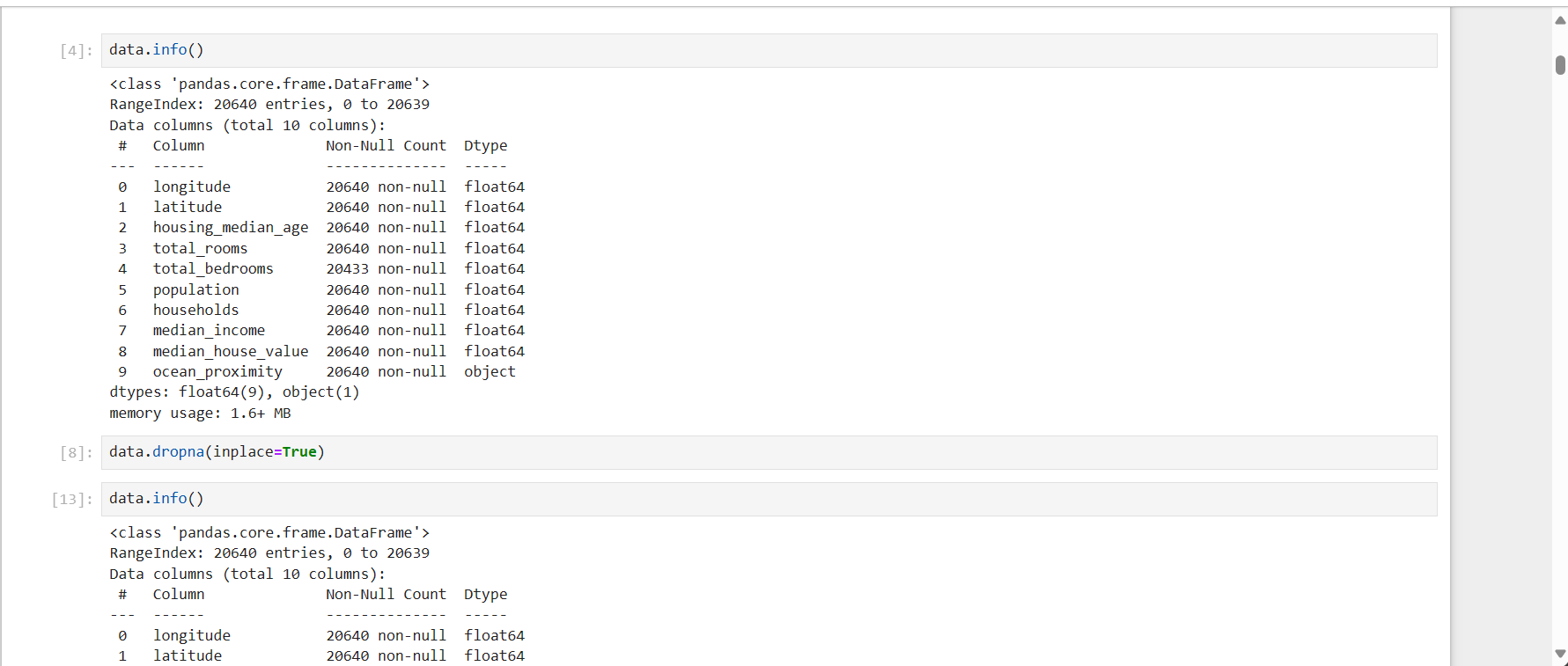
1. **Integrated Development Environment (IDE):** An IDE such as Jupyter Notebook will be used for coding, debugging, and testing the recommendation system.
2. **Version Control:** Git will be used for version control to manage changes to the source code and collaborate with team members.

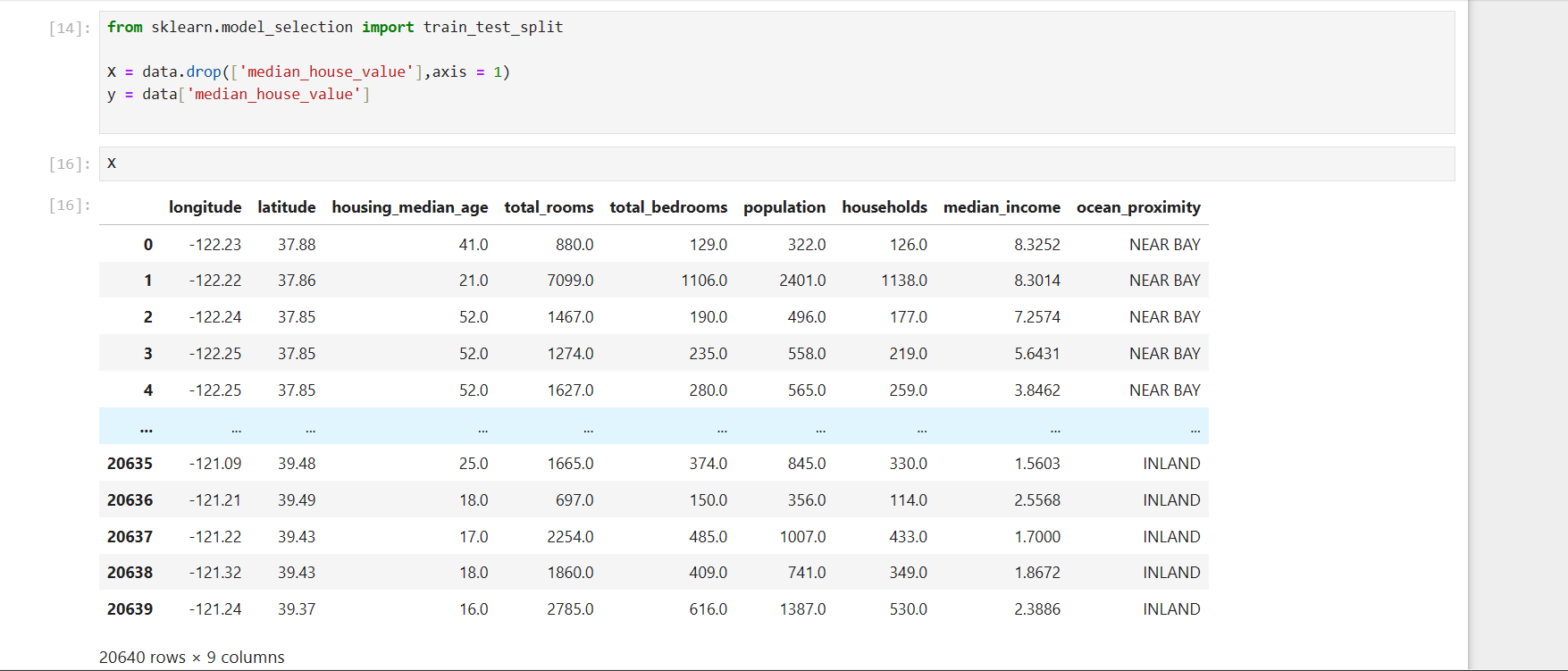
# Proposed Design

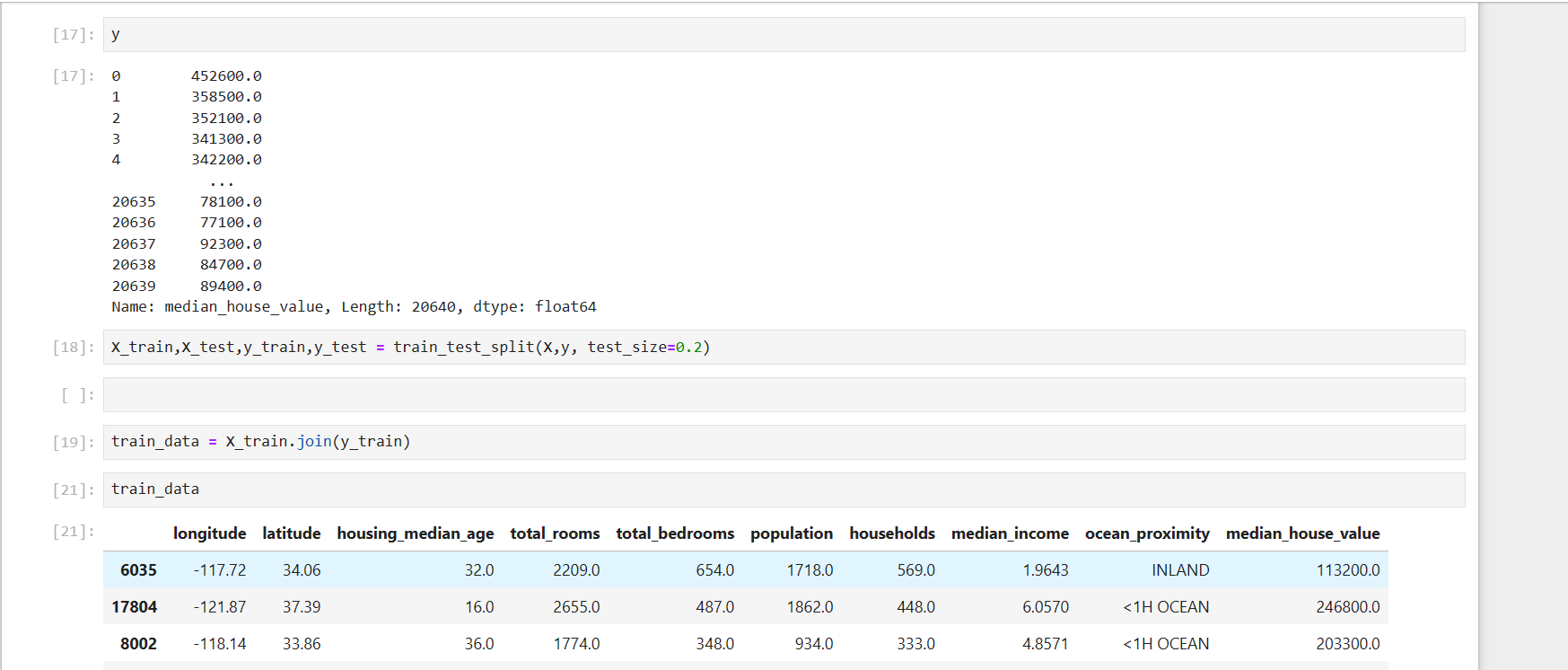
* **Importing Libraries**: Use Python libraries such as pandas, numpy, and scikit-learn for data manipulation, analysis, and machine learning.
* **Loading the Dataset**: Load a dataset containing information about songs, including genre, artist, and audio features**.**
* **Data Exploration:** Get the shape of the dataset to understand its size and displaying the first 5 rows of the dataset to inspect the data.
* **Data Preprocessing:** Visualize the number of duplicates in the 'track\_name' column and remove duplicates if necessary. Remove unwanted columns from the dataset. Check for null values in the dataset and visualize them for further analysis. Vectorize the 'genre' column to convert categorical data into numerical format for machine learning.

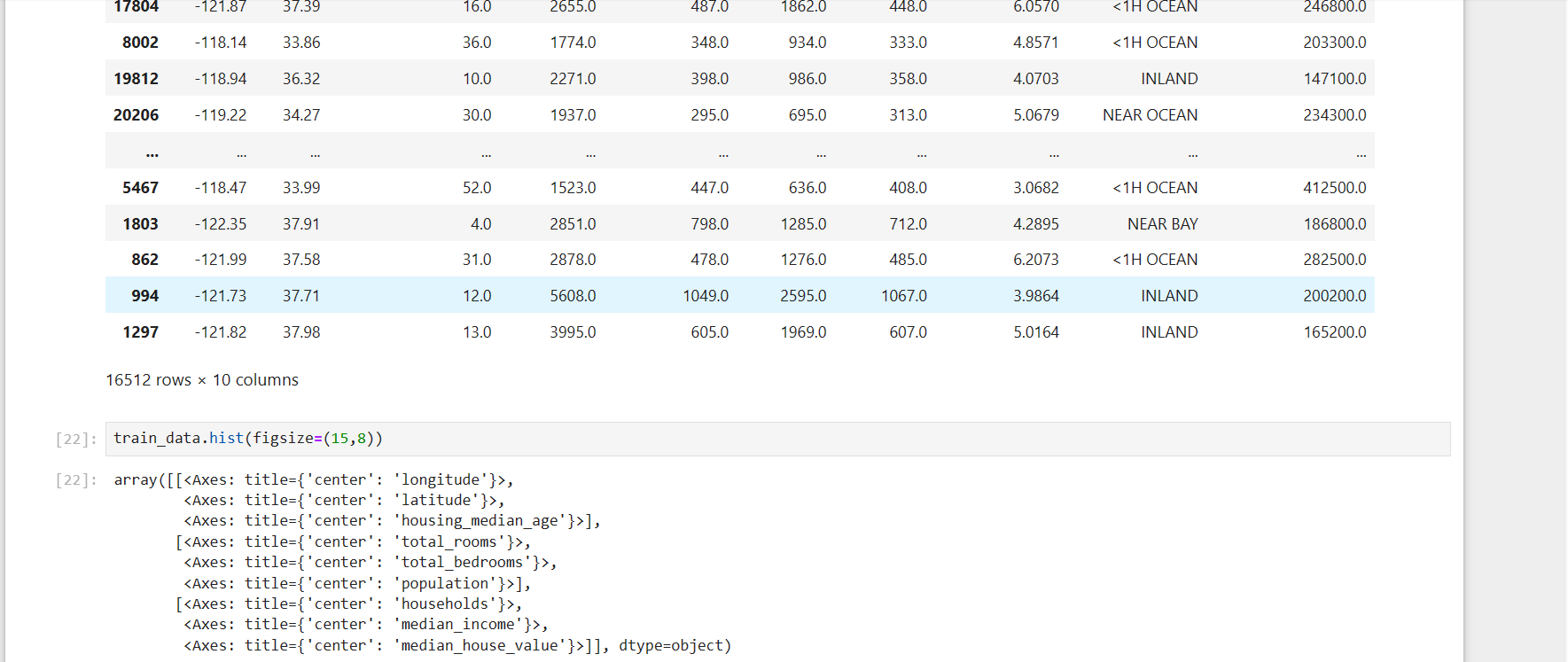
# Code

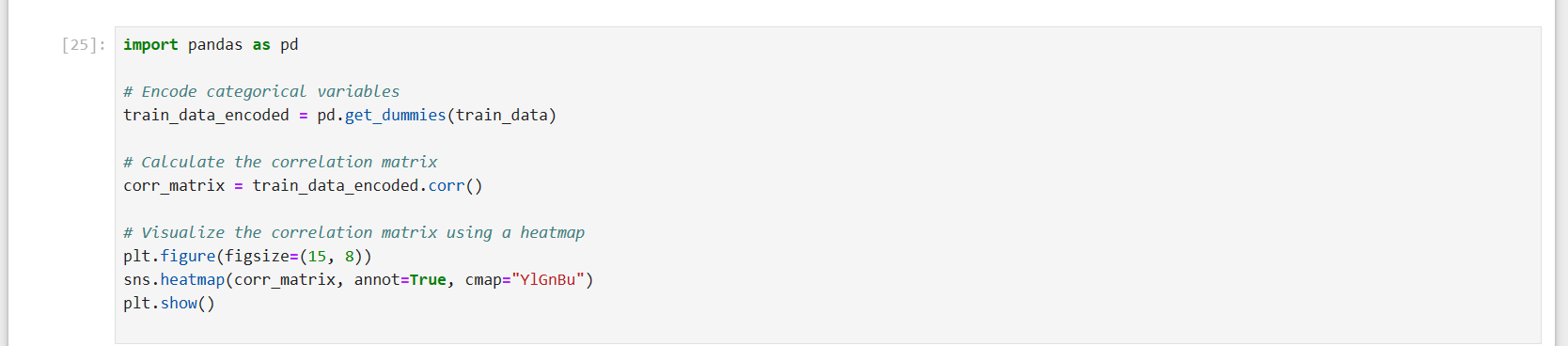




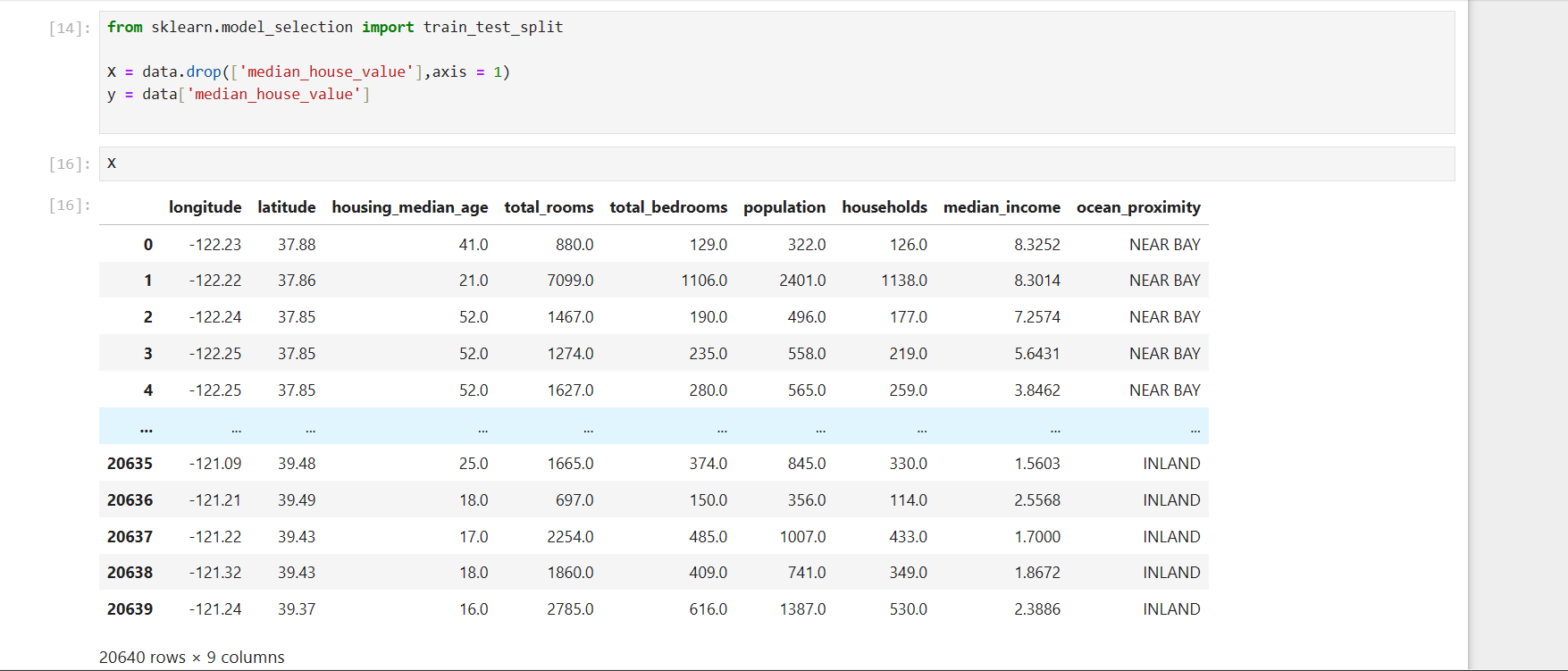


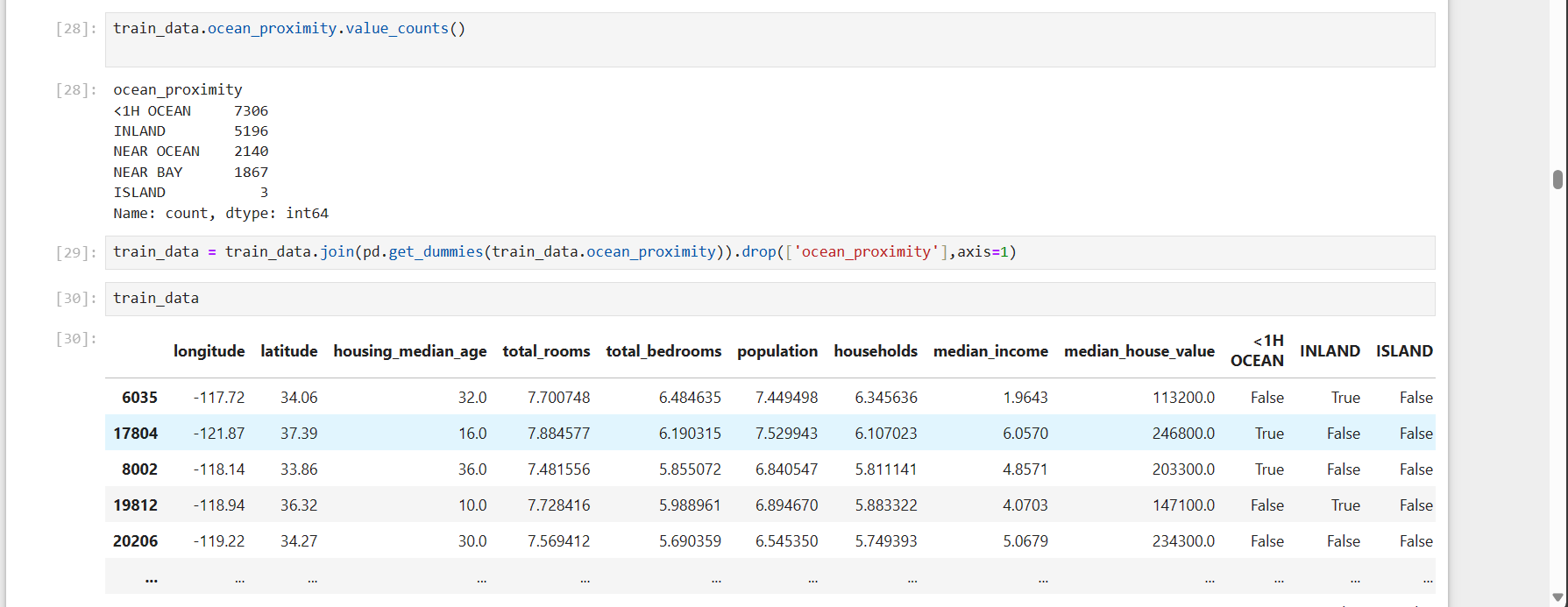


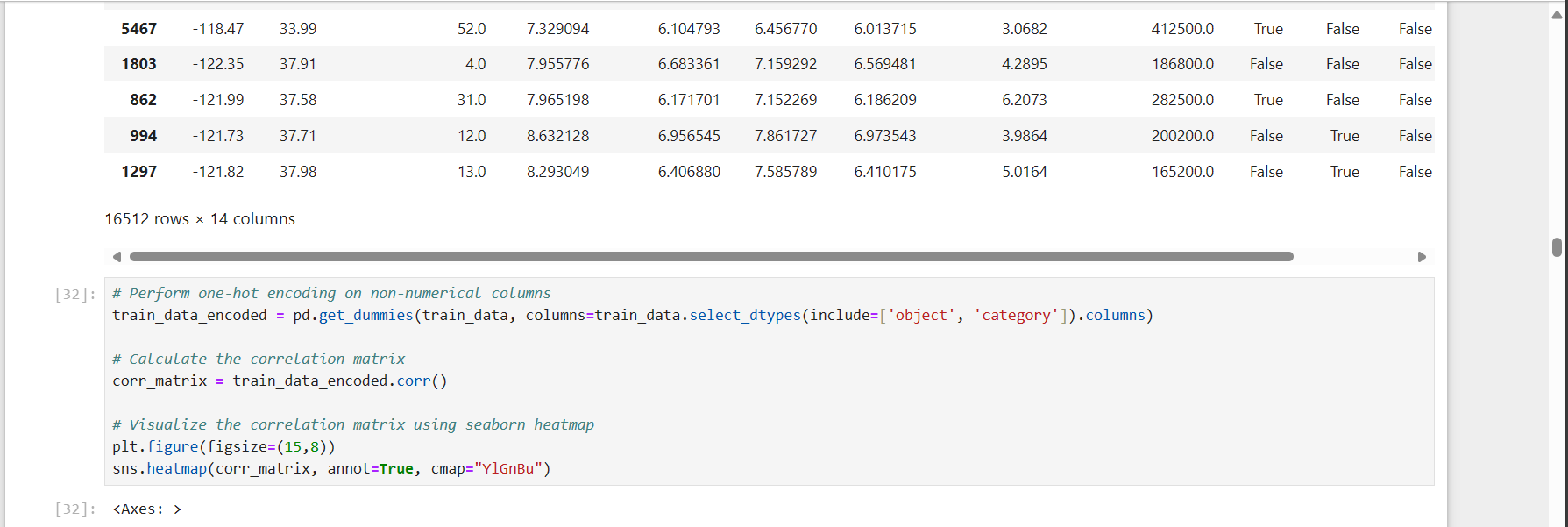










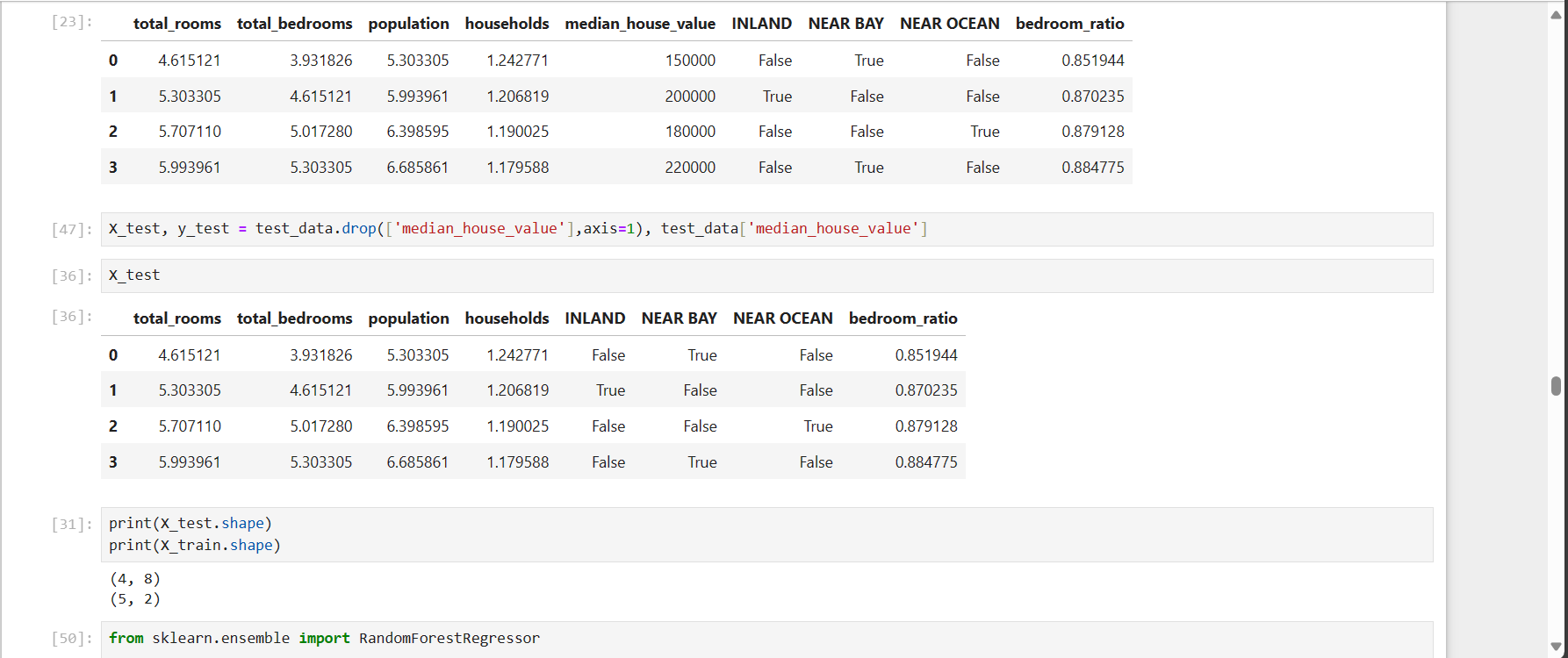


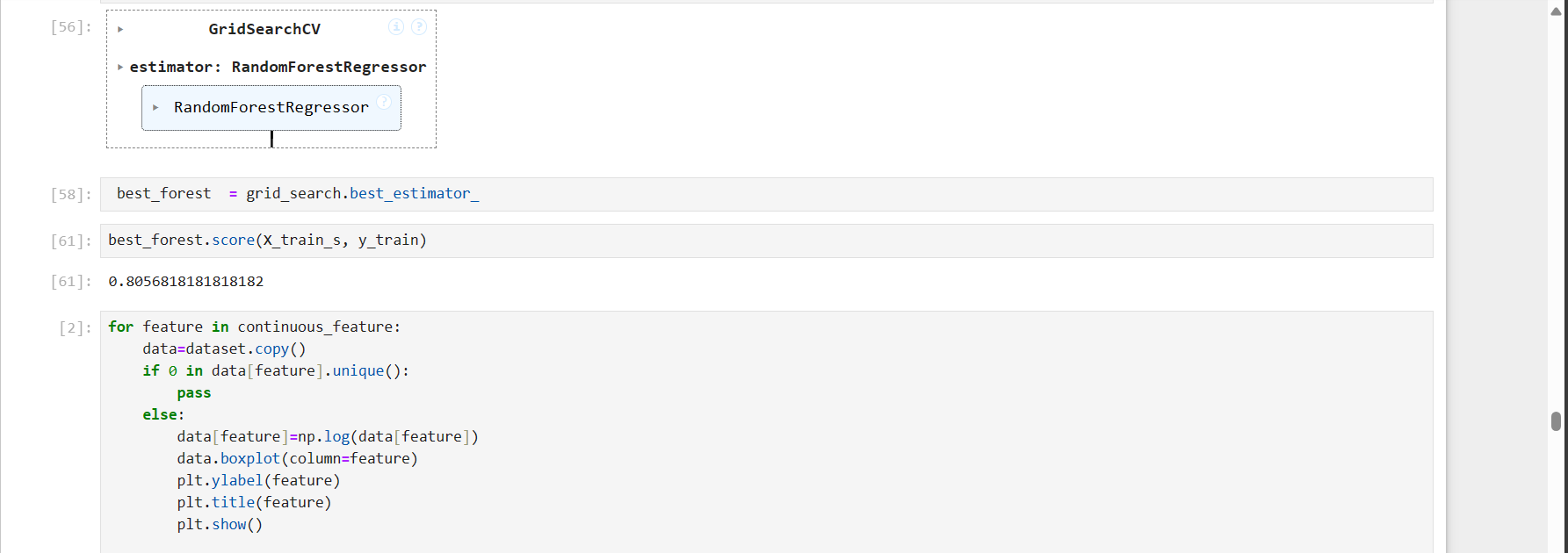


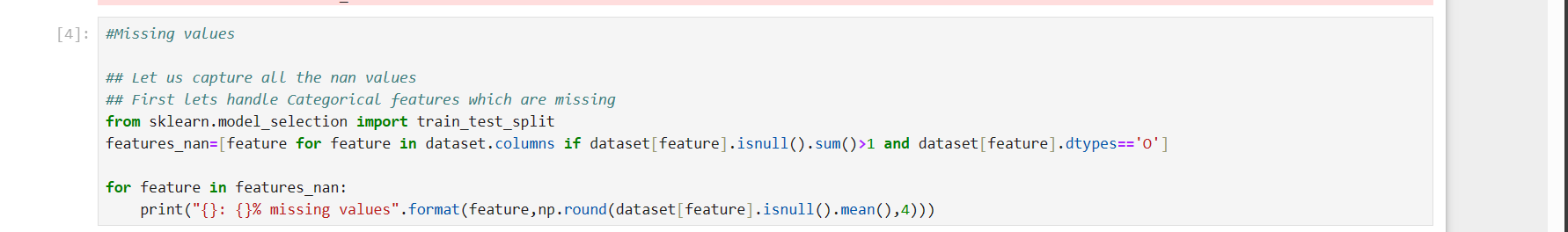




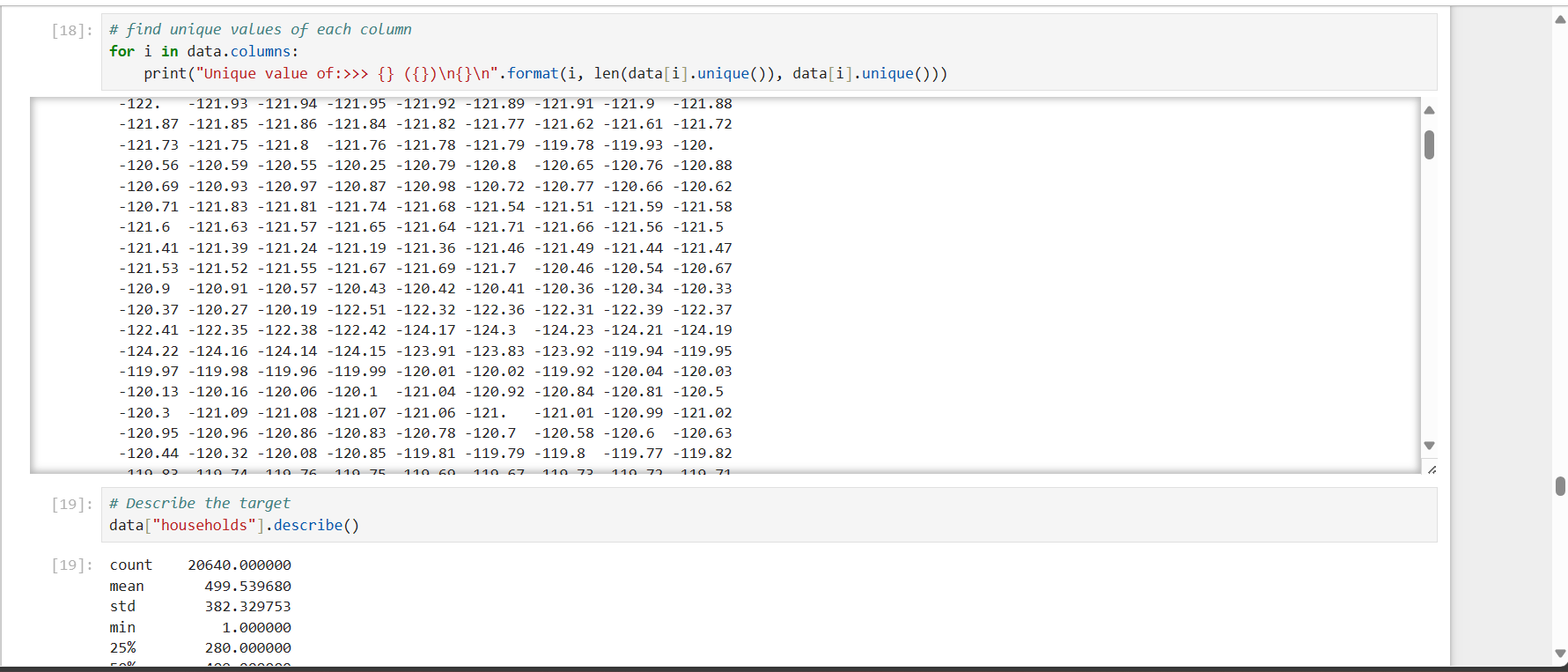


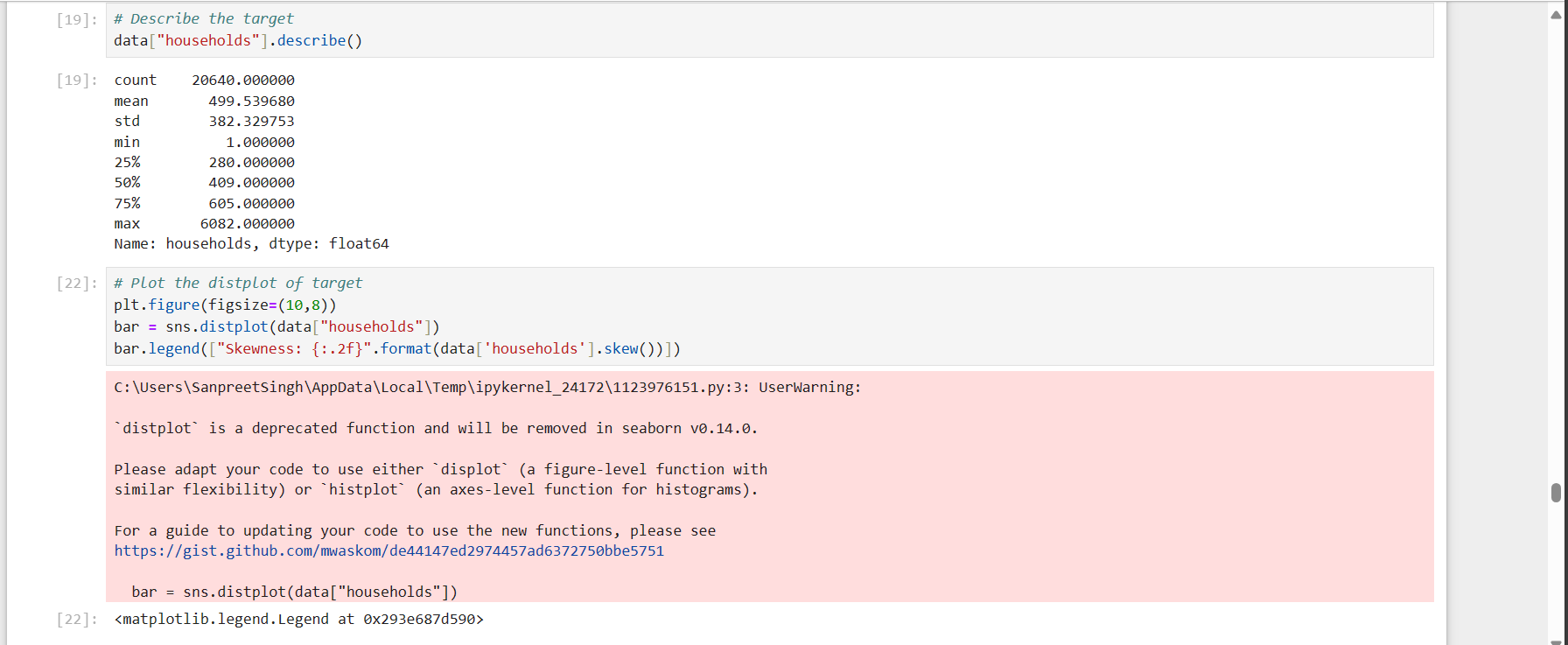


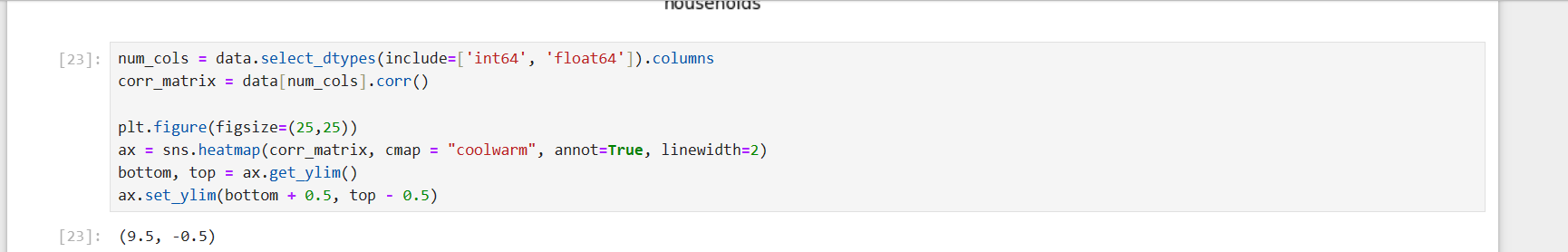
 



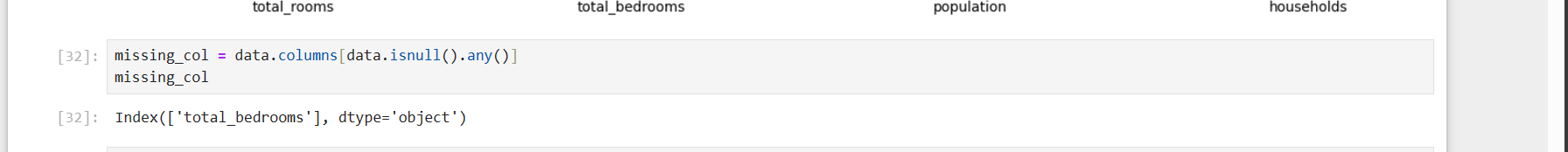




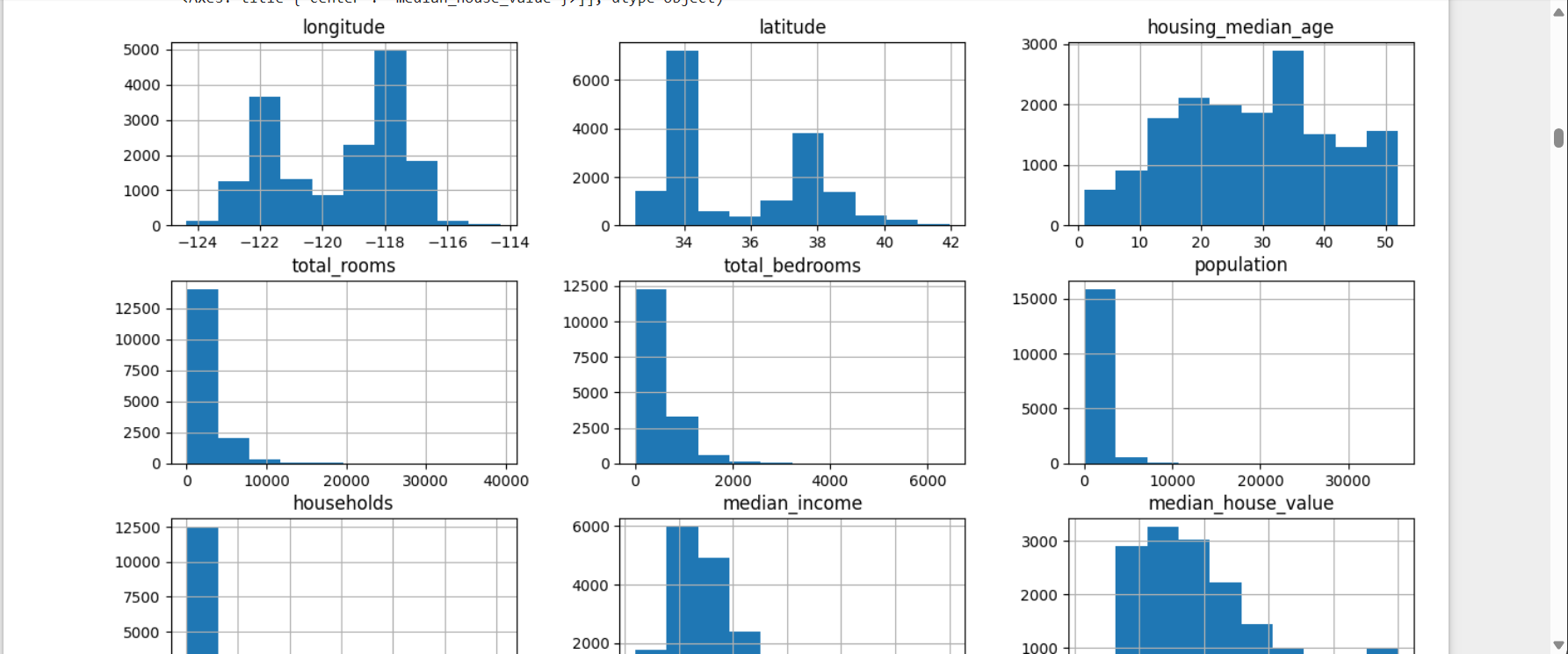


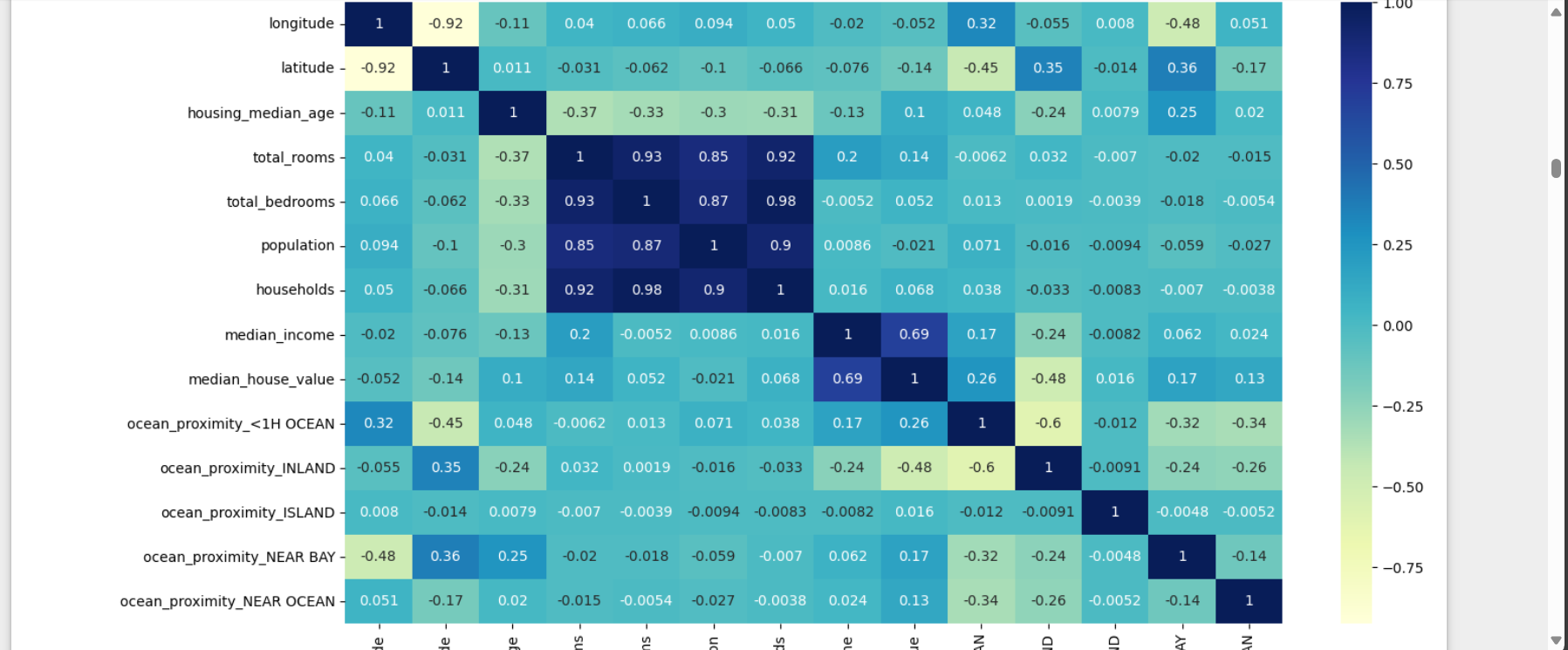


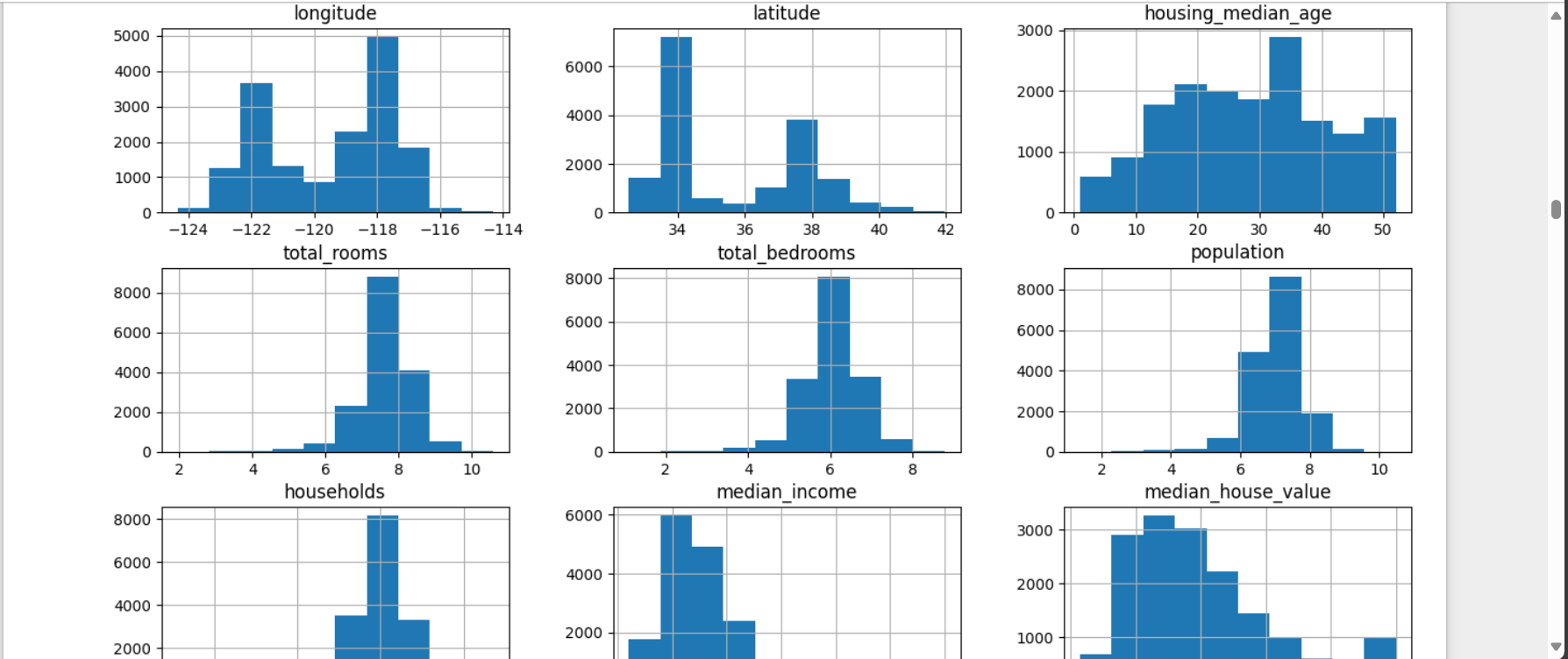


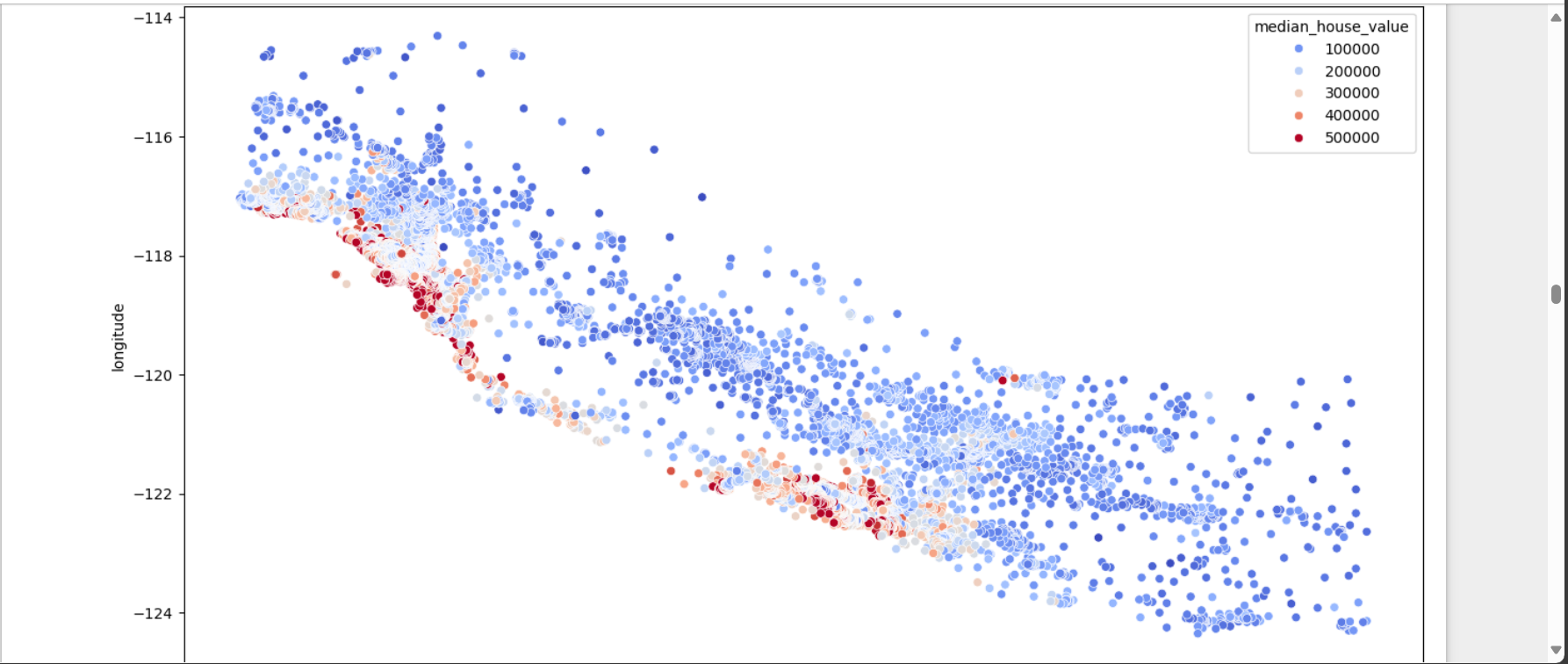
 

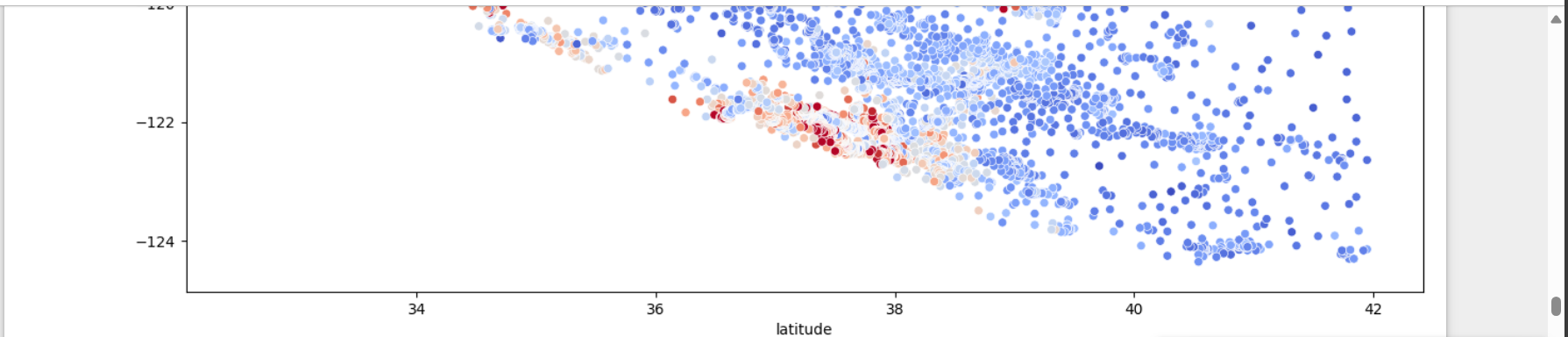
# Results



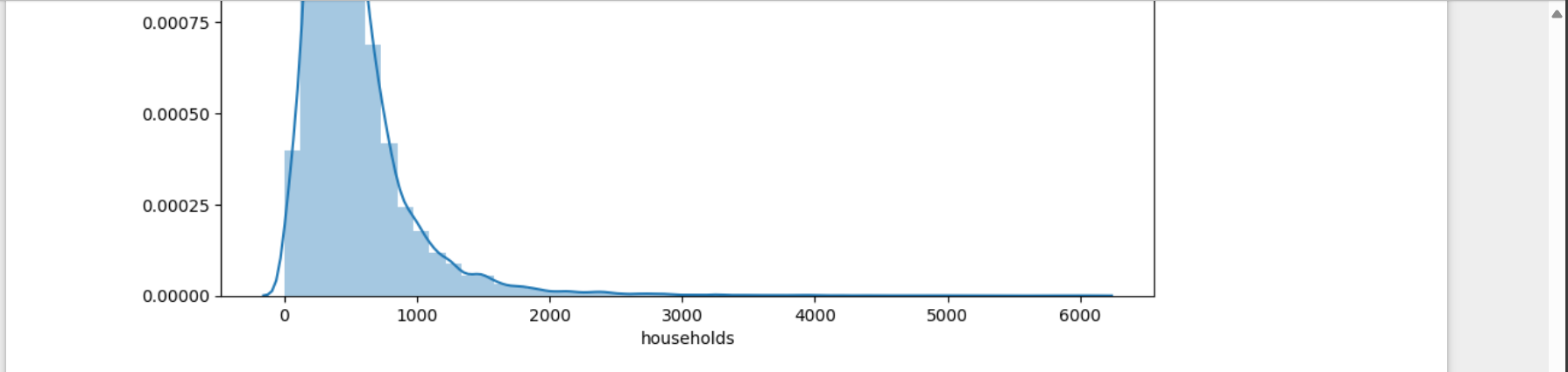


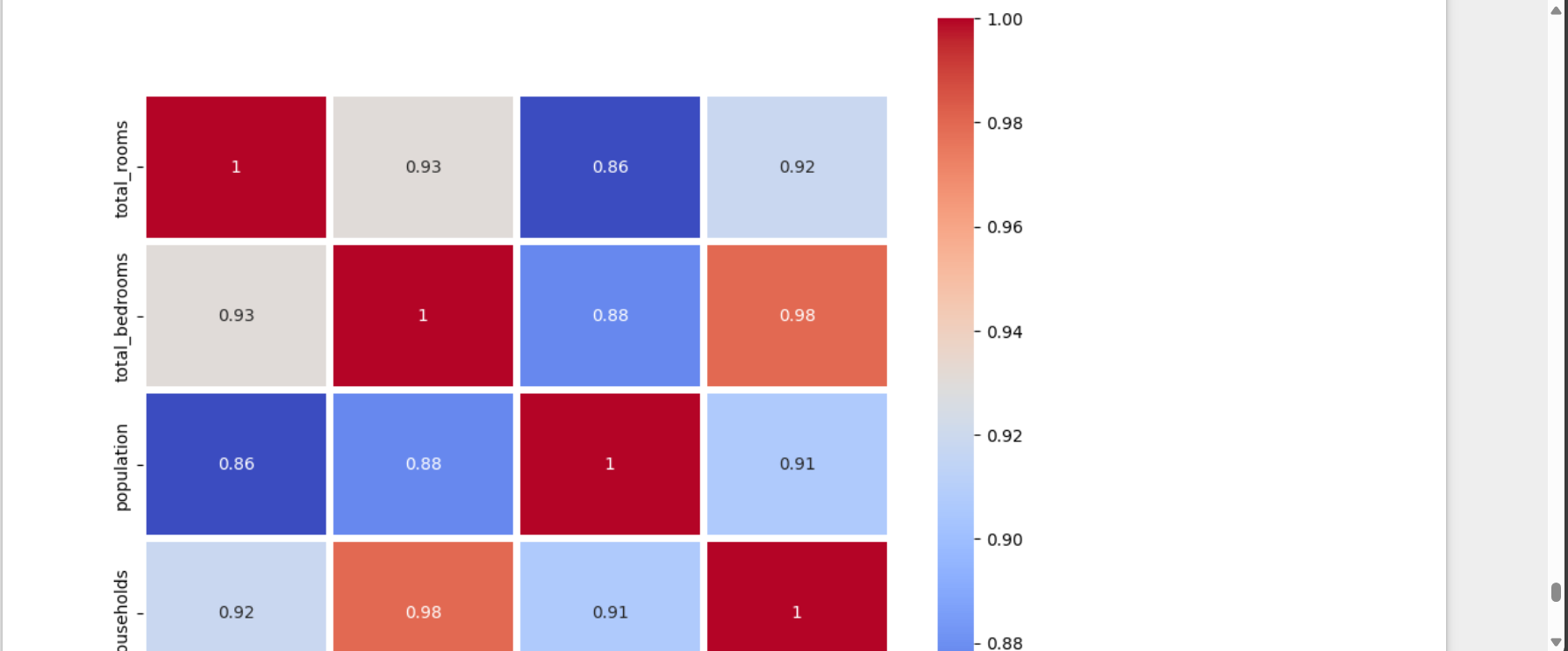


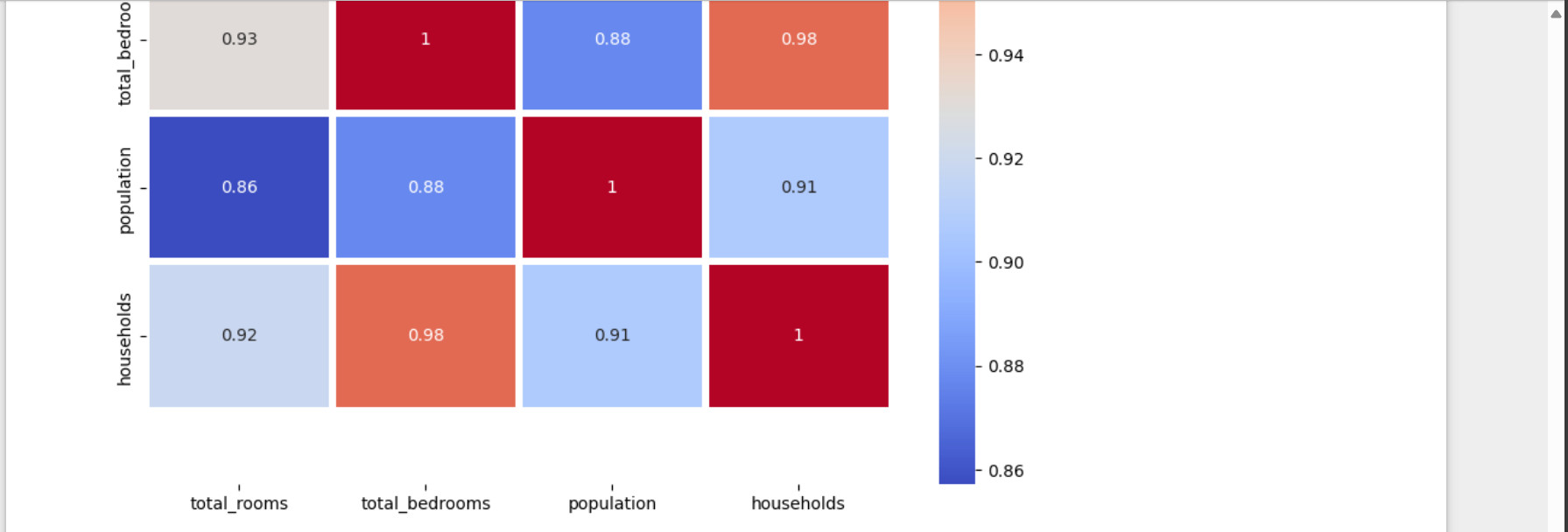


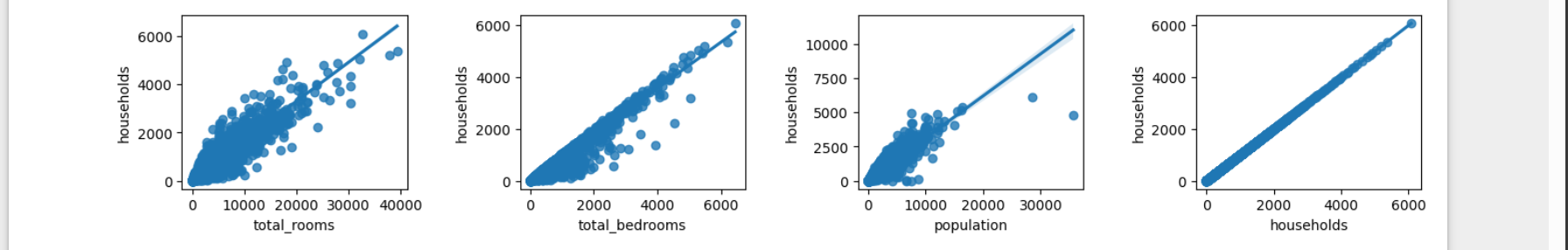












# References

* Scikit-learn: <https://scikit-learn.org/stable/>
* Matplotlib:  [https://matplotlib.org/](%20https://matplotlib.org/)
* Seaborn: <https://seaborn.pydata.org/>
* Tkinter: <https://docs.python.org/3/library/tkinter.html>
* W3Schools: <https://www.w3schools.com/python/python_ml_getting_started.asp>
* GeeksForGeeks: <https://www.geeksforgeeks.org/machine-learning/>