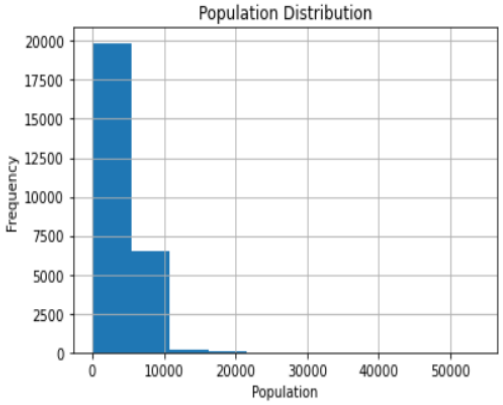
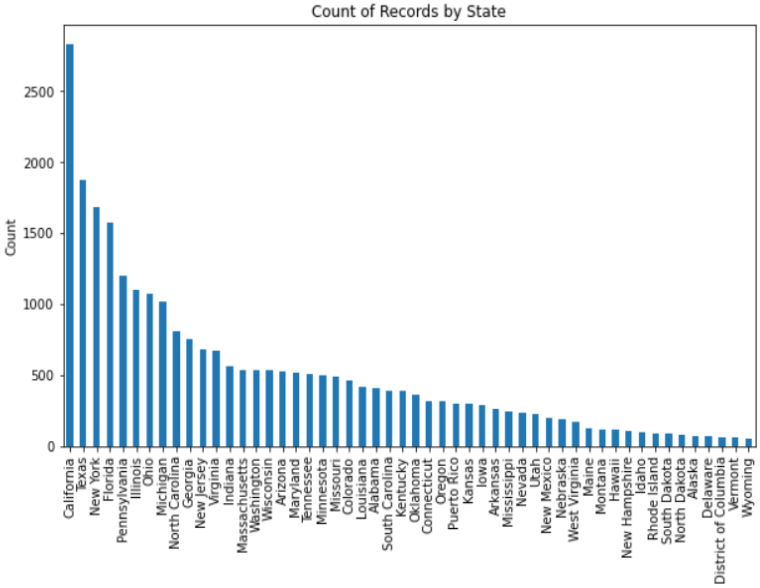
Real Estate Project Screenshots

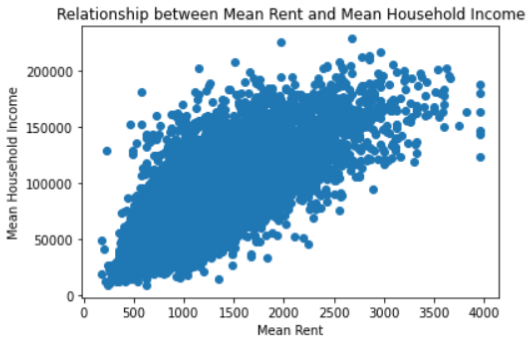
In [105]:
```
# Population vs Distribution

train_data['pop'].hist()
plt.xlabel('Population')
plt.ylabel('Frequency')
plt.title('Population Distribution')
plt.show()
```



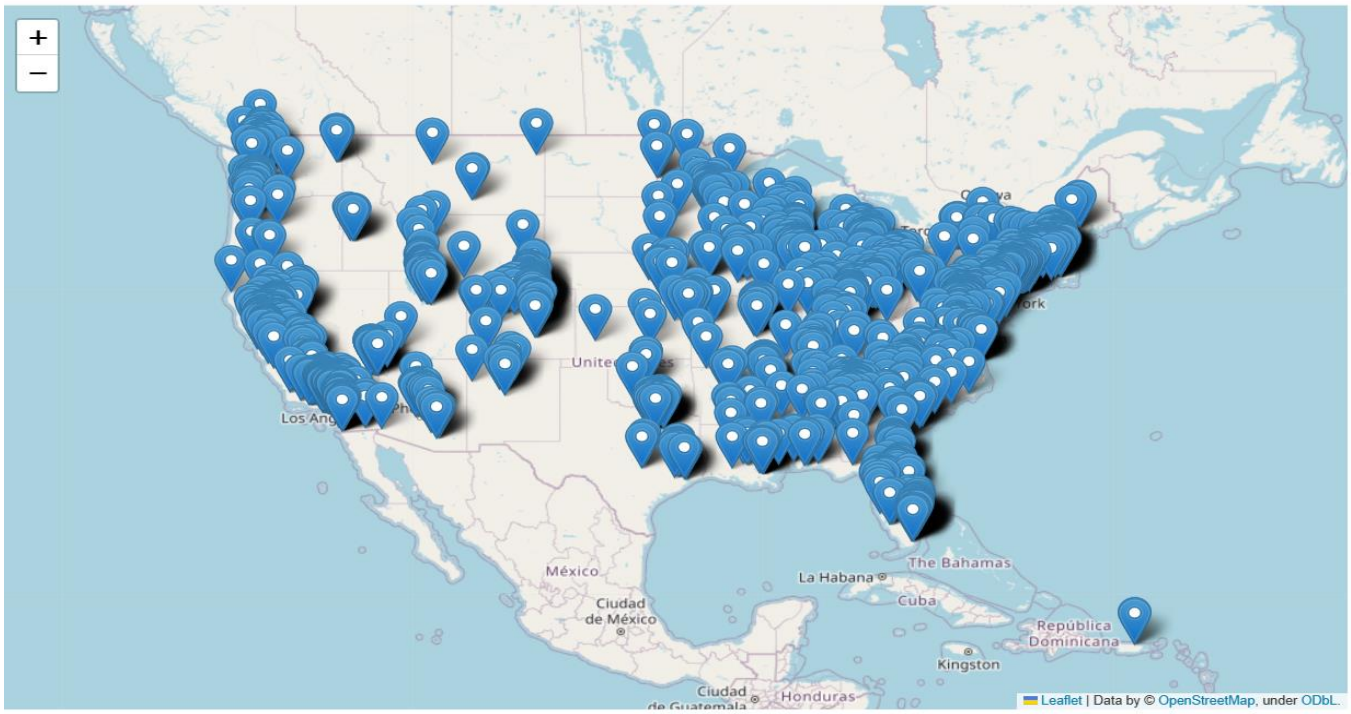
In [106]:
```
state_counts = train_data['state'].value_counts()
plt.figure(figsize=(10, 6))
state_counts.plot(kind='bar')
plt.xlabel('State')
plt.ylabel('Count')
plt.title('Count of Records by State')
plt.show()
```



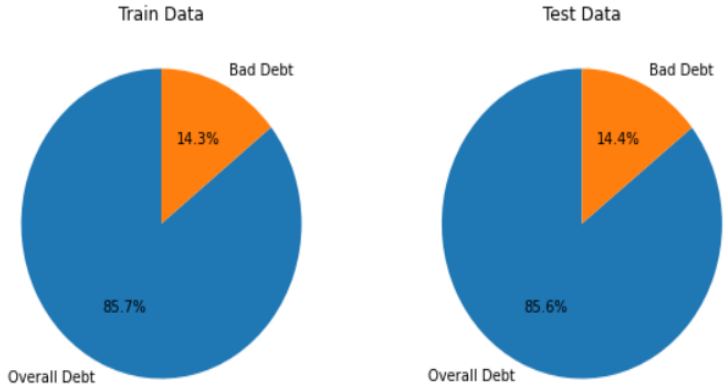
In [107]:
```
plt.scatter(train_data['rent_mean'], train_data['hi_mean'])
plt.xlabel('Mean Rent')
plt.ylabel('Mean Household Income')
plt.title('Relationship between Mean Rent and Mean Household Income')
plt.show()
```




```
import folium

# Filter the data
filtered_data = train_data[(train_data['second_mortgage'] <= 0.50) & (train_data['pct_own'] > 0.10)]
top_2500_loc = filtered_data.nlargest(2500, 'second_mortgage')

# Create a map
map = folium.Map(location=[37.0902, -95.7129], zoom_start=4)  # Set the initial map location

# Add markers for each location
for index, row in top_2500_loc.iterrows():
    lat, lng = row['lat'], row['lng']
    folium.Marker(location=[lat, lng]).add_to(map)

# Display the map
map
```

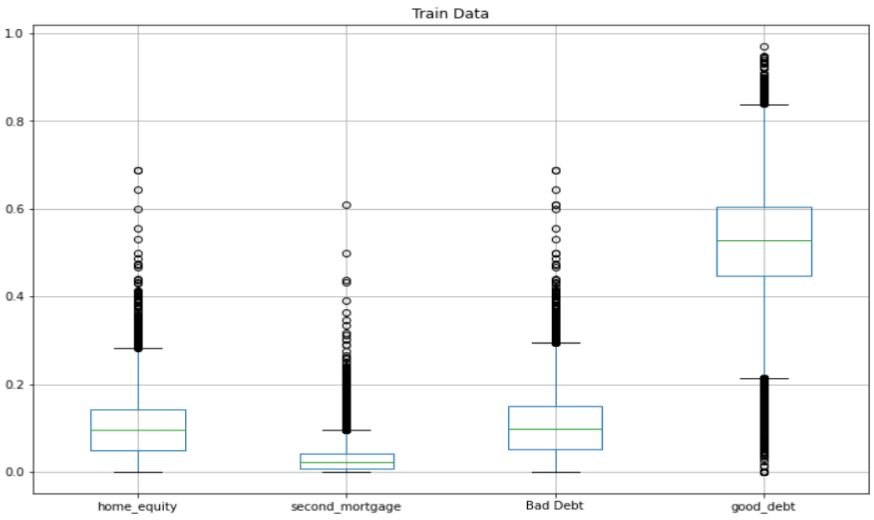


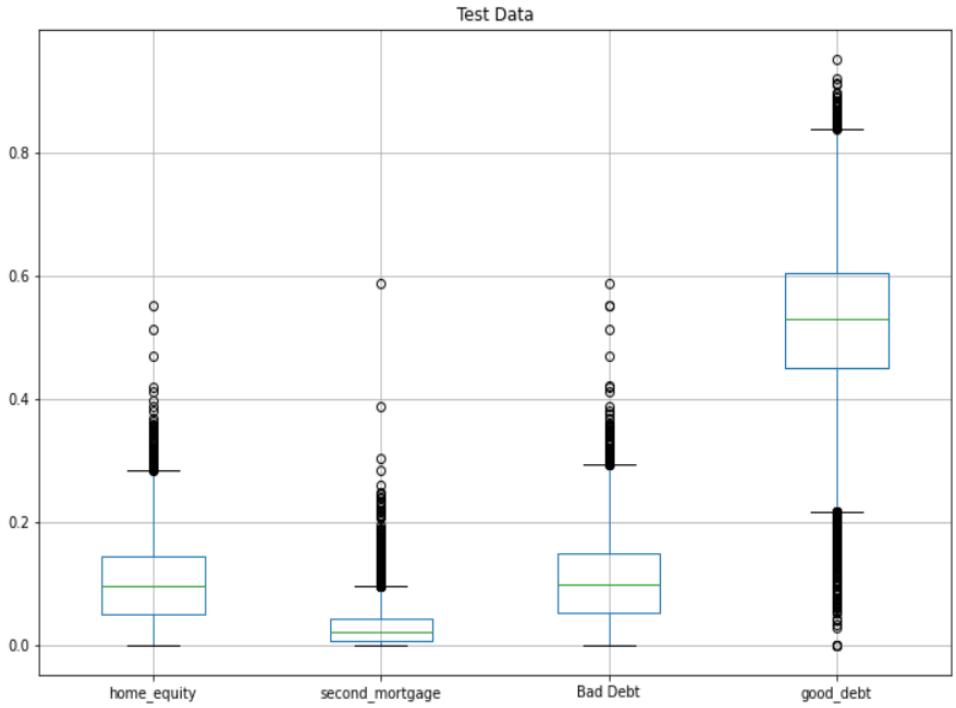

```
# Calculate overall debt
train_overall_debt = train_data['debt'].sum()
test_overall_debt = test_data['debt'].sum()

# Calculate bad debt
train_bad_debt = train_data['Bad Debt'].sum()
test_bad_debt = test_data['Bad Debt'].sum()

# Create data for pie charts
labels = ['Overall Debt', 'Bad Debt']
train_values = [train_overall_debt, train_bad_debt]
test_values = [test_overall_debt, test_bad_debt]

# Create subplots for train_data and test_data
fig, axes = plt.subplots(1, 2, figsize=(10, 5))

# Plot pie chart for train_data
axes[0].pie(train_values, labels=labels, autopct='%1.1f%%', startangle=90)
axes[0].set_title('Train Data')

# Plot pie chart for test_data
axes[1].pie(test_values, labels=labels, autopct='%1.1f%%', startangle=90)
axes[1].set_title('Test Data')

# Set aspect ratio to be equal so that pie charts are circular
axes[0].set_aspect('equal')
axes[1].set_aspect('equal')

# Display the pie charts
plt.show()
```

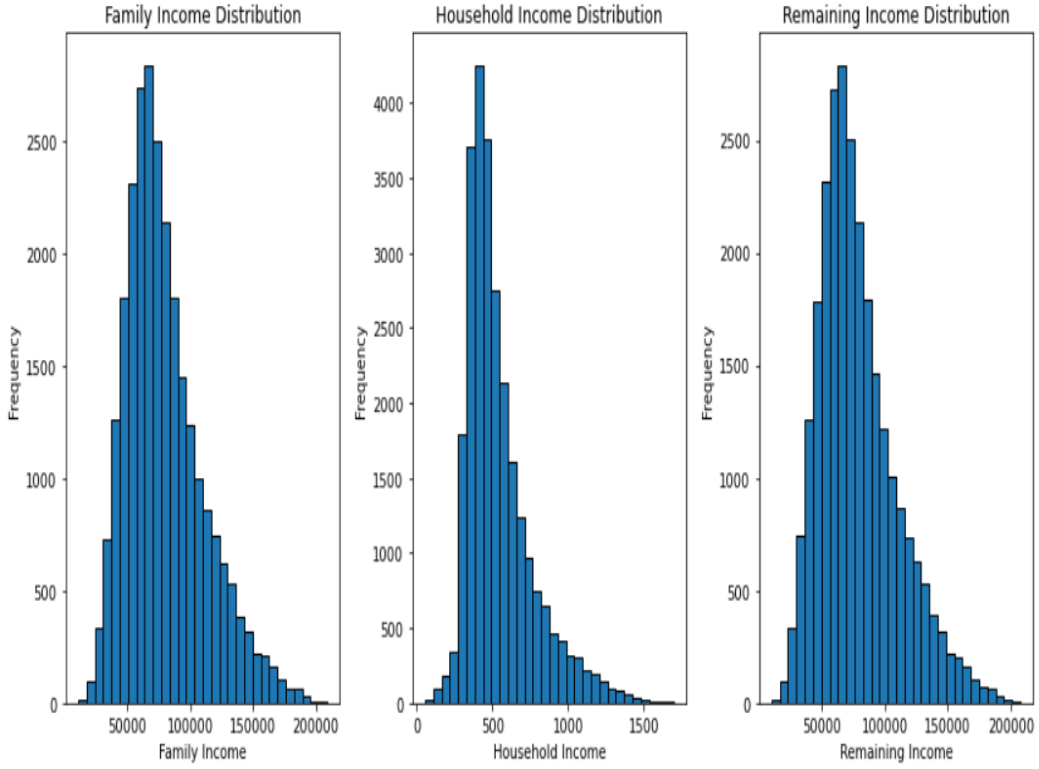


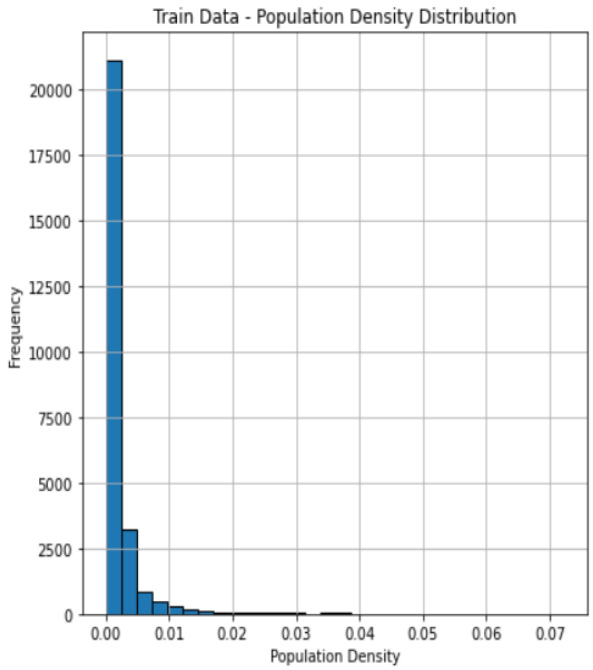

```
# Calculate good_debt as debt minus bad_debt for train dataset
train_data['good_debt'] = train_data['debt'] - train_data['Bad Debt']

# Calculate good_debt as debt minus bad_debt for test dataset
test_data['good_debt'] = test_data['debt'] - test_data['Bad Debt']
```

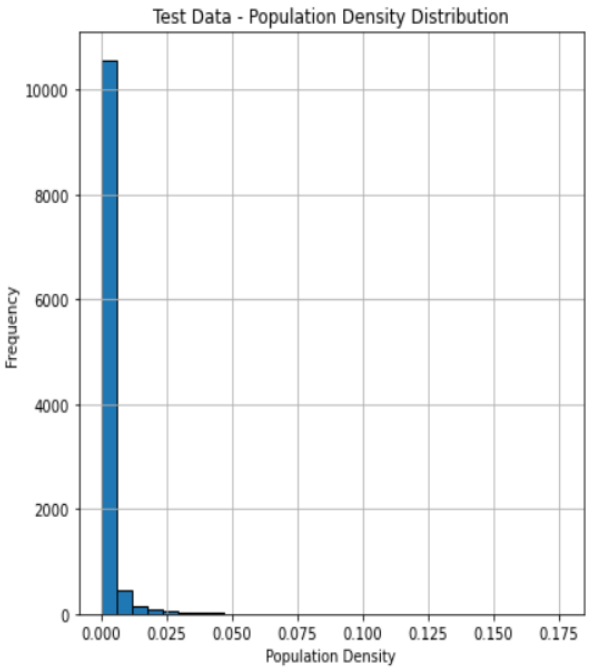
```
# Subset the required variables from train_data
train_cities = train_data[['home_equity', 'second_mortgage', 'Bad Debt', 'good_debt']]

# Subset the required variables from test_data
test_cities = test_data[['home_equity', 'second_mortgage', 'Bad Debt', 'good_debt']]

# Plot the box and whisker plot for train_data
train_cities.plot.box(figsize=(12, 8), grid=True)
plt.title('Train Data')
plt.show()

# Plot the box and whisker plot for test_data
test_cities.plot.box(figsize=(12, 8), grid=True)
plt.title('Test Data')
plt.show()
```

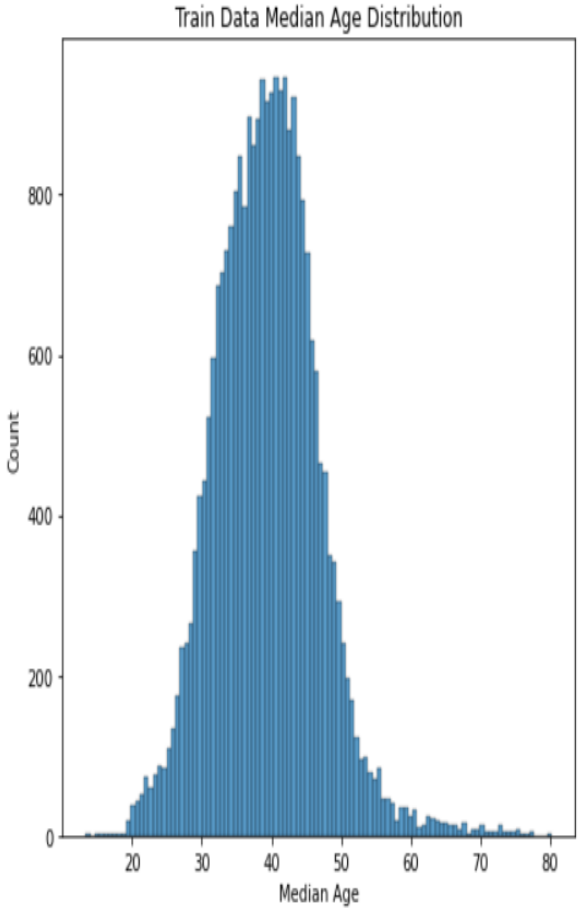


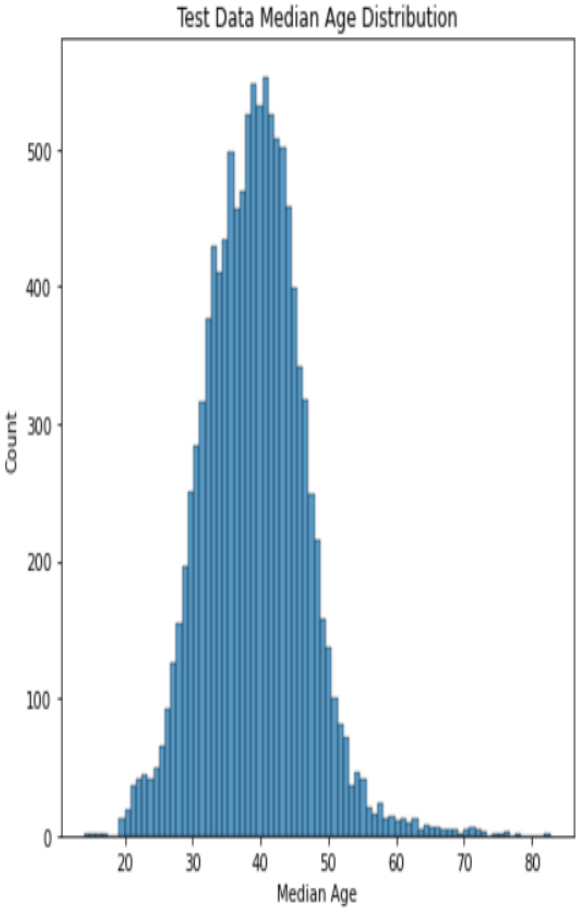

Test Data
0.8
0.6
0.4
0.2
0.0
home_equity
second_mortgage
Bad Debt
good_debt

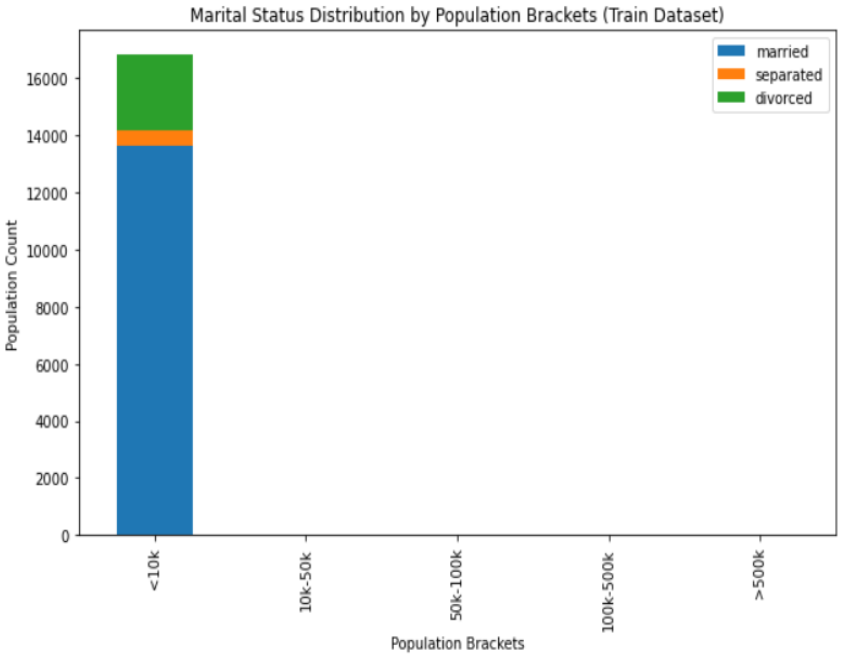

Family Income Distribution
Frequency
Family Income
50000
100000
150000
200000
Household Income Distribution
Household Income
0
500
1000
1500
Remaining Income Distribution
Remaining Income

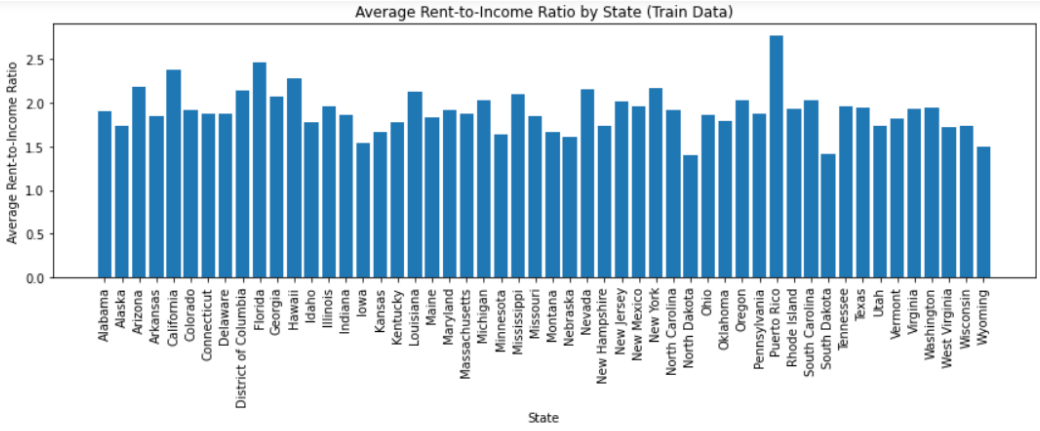

Train Data - Population Density Distribution
Frequency
Population Density
0
2500
5000
7500
10000
12500
15000
17500
20000
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.07

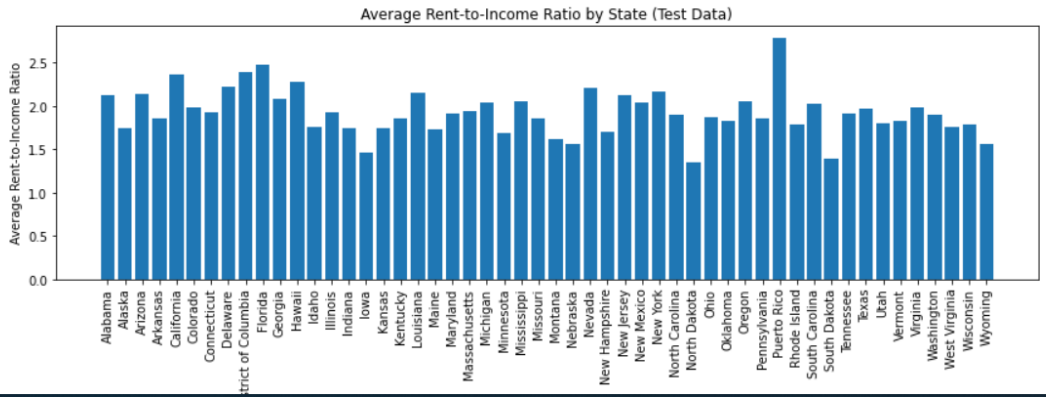

Test Data - Population Density Distribution
Frequency
Population Density
0
2000
4000
6000
8000
10000
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175

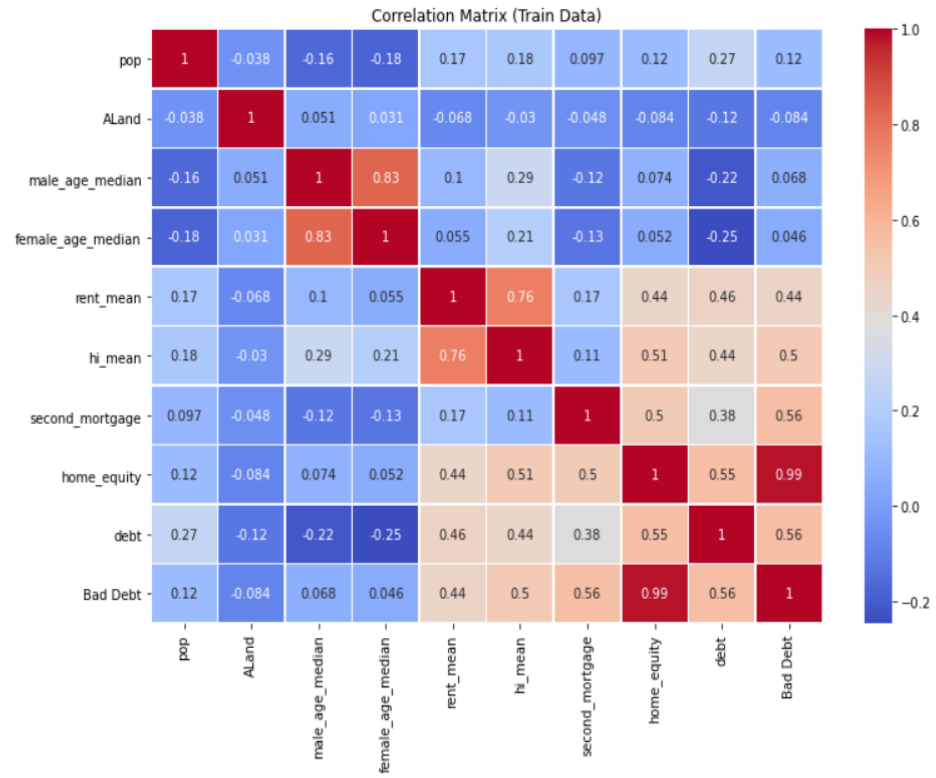

Train Data Median Age Distribution
Count
Median Age
0
200
400
600
800
20
30
40
50
60
70
80


Test Data Median Age Distribution
Count
Median Age
0
100
200
300
400
500
20
30
40
50
60
70
80

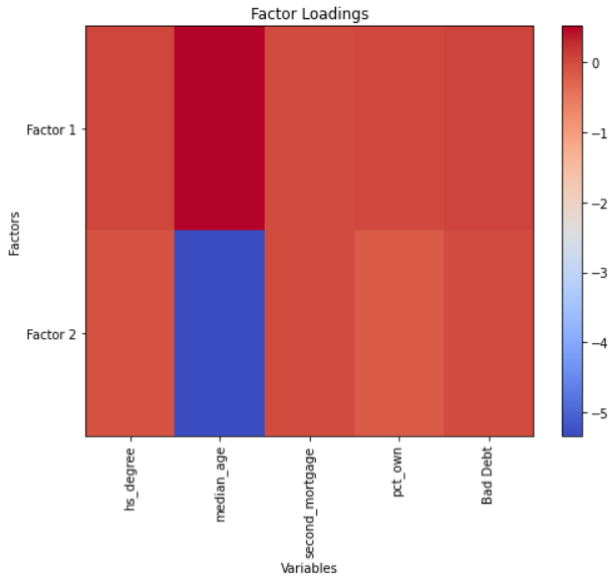

Marital Status Distribution by Population Brackets (Train Dataset)
married
separated
divorced
Population Count
0
2000
4000
6000
8000
10000
12000
14000
16000
<10k
10k-50k
50k-100k
100k-500k
>500k
Population Brackets


Average Rent-to-Income Ratio by State (Train Data)
Average Rent-to-Income Ratio
0.0
0.5
1.0
1.5
2.0
2.5
Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
District of Columbia
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
New Mexico
New York
North Carolina
North Dakota
Ohio
Oklahoma
Oregon
Pennsylvania
Puerto Rico
Rhode Island
South Carolina
South Dakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
West Virginia
Wisconsin
Wyoming
State


Average Rent-to-Income Ratio by State (Test Data)
Average Rent-to-Income Ratio
0.0
0.5
1.0
1.5
2.0
2.5
Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
New Mexico
New York
North Carolina
North Dakota
Ohio
Oklahoma
Oregon
Pennsylvania
Puerto Rico
Rhode Island
South Carolina
South Dakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
West Virginia
Wisconsin
Wyoming



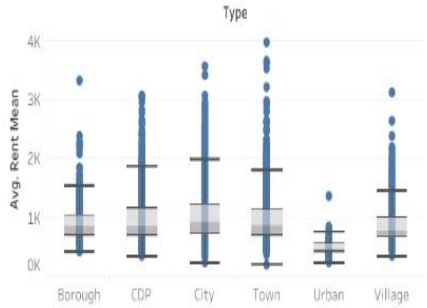

```
# Plot the loadings
plt.figure(figsize=(8, 6))
plt.imshow(loadings, cmap='coolwarm', aspect='auto')
plt.colorbar()
plt.xticks(range(len(columns_for_factor_analysis)), columns_for_factor_analysis, rotation=90)
plt.yticks(range(len(fa.components_)), ['Factor 1', 'Factor 2'])
plt.xlabel('Variables')
plt.ylabel('Factors')
plt.title('Factor Loadings')
plt.show()
```



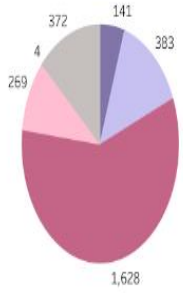
[86]:

Real Estate Data Reporting

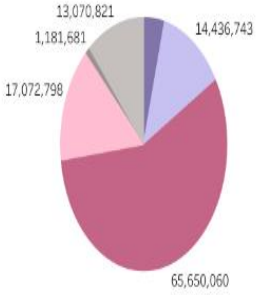
Box plot of distribution of average rent by type of place



Pie charts to show overall debt and bad debt



Pie chart to show the population distribution



Top 2500 Location

Heat map for correlation matrix