

2. Develop a program to Compute the correlation matrix to understand the relationships between pairs of features. Visualize the correlation matrix using a heatmap to know which variables have strong positive/negative correlations. Create a pair plot to visualize pairwise relationships between features. Use California Housing dataset

```
# Import necessary libraries
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.datasets import fetch_california_housing

# Load the California Housing dataset
housing_data = fetch_california_housing(as_frame=True)
data = housing_data.frame

# Display the first few rows of the dataset
print(data.head())

# Compute the correlation matrix
correlation_matrix = data.corr() #The .corr() method computes the
correlation between each pair of features in the dataset.
#Strong positive correlations are closer to +1, and strong negative
correlations are closer to -1.

# Print the correlation matrix
print("\nCorrelation Matrix:")
print(correlation_matrix)

# Visualize the correlation matrix using a heatmap
plt.figure(figsize=(10, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', fmt='.2f',
linewidths=0.5)
plt.title("Correlation Matrix Heatmap")
plt.show()

# Create a pair plot for pairwise relationships
sns.pairplot(data, diag_kind='kde', plot_kws={'alpha': 0.7})
plt.show()
```

Output

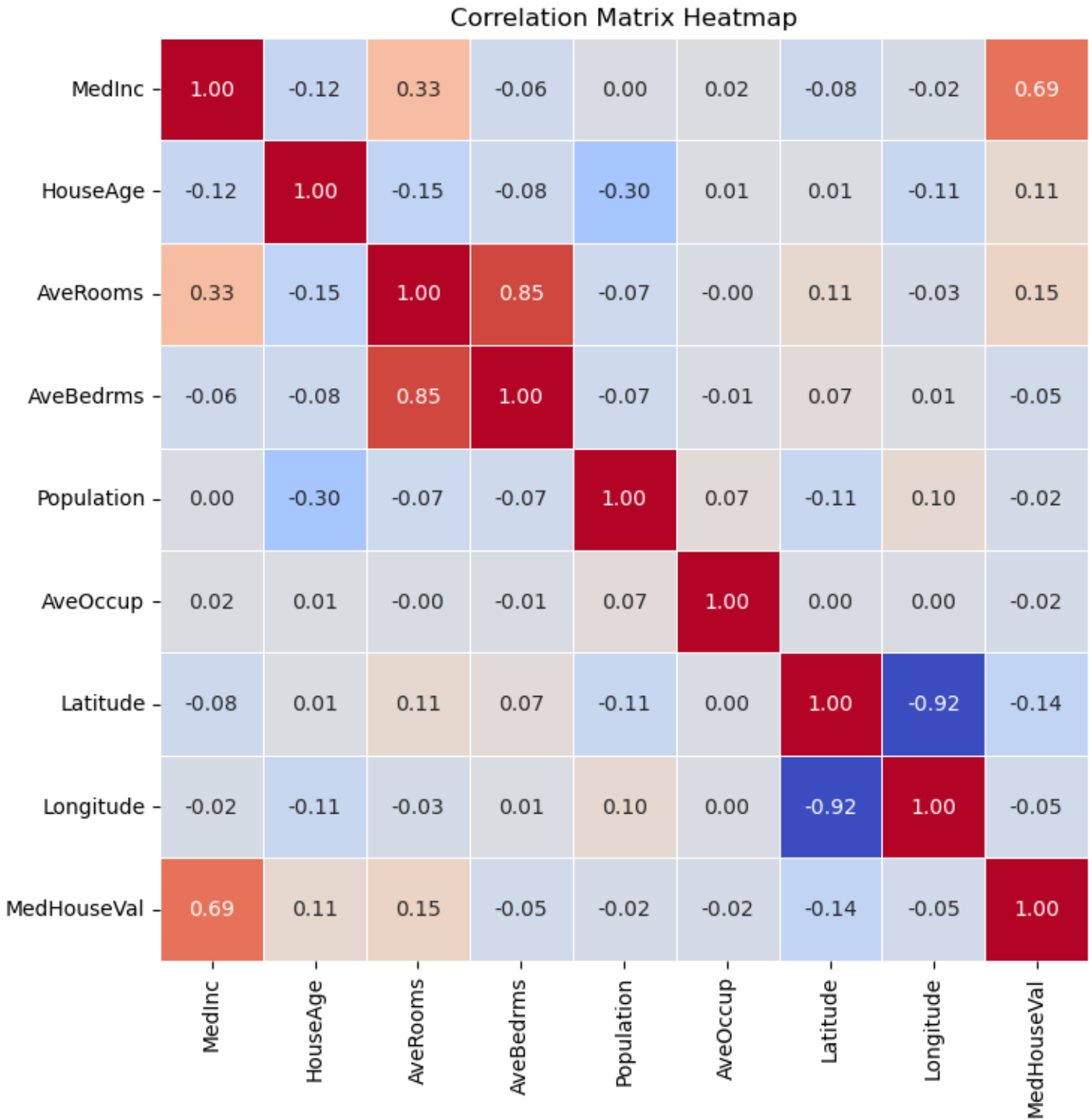
	MedInc	HouseAge	AveRooms	AveBedrms	Population	AveOccup	Latitude
0	8.3252	41.0	6.984127	1.023810	322.0	2.555556	37.88
1	8.3014	21.0	6.238137	0.971880	2401.0	2.109842	37.86
2	7.2574	52.0	8.288136	1.073446	496.0	2.802260	37.85
3	5.6431	52.0	5.817352	1.073059	558.0	2.547945	37.85
4	3.8462	52.0	6.281853	1.081081	565.0	2.181467	37.85

	Longitude	MedHouseVal
0	-122.23	4.526
1	-122.22	3.585
2	-122.24	3.521
3	-122.25	3.413
4	-122.25	3.422

Correlation Matrix:

	MedInc	HouseAge	AveRooms	AveBedrms	Population	AveOccup
MedInc	1.000000	-0.119034	0.326895	-0.062040	0.004834	0.018766
HouseAge	-0.119034	1.000000	-0.153277	-0.077747	-0.296244	0.013191
AveRooms	0.326895	-0.153277	1.000000	0.847621	-0.072213	-0.004852
AveBedrms	-0.062040	-0.077747	0.847621	1.000000	-0.066197	-0.006181
Population	0.004834	-0.296244	-0.072213	-0.066197	1.000000	0.069863
AveOccup	0.018766	0.013191	-0.004852	-0.006181	0.069863	1.000000
Latitude	-0.079809	0.011173	0.106389	0.069721	-0.108785	0.002366
Longitude	-0.015176	-0.108197	-0.027540	0.013344	0.099773	0.002476
MedHouseVal	0.688075	0.105623	0.151948	-0.046701	-0.024650	-0.023737

	Latitude	Longitude	MedHouseVal
MedInc	-0.079809	-0.015176	0.688075
HouseAge	0.011173	-0.108197	0.105623
AveRooms	0.106389	-0.027540	0.151948
AveBedrms	0.069721	0.013344	-0.046701
Population	-0.108785	0.099773	-0.024650
AveOccup	0.002366	0.002476	-0.023737
Latitude	1.000000	-0.924664	-0.144160
Longitude	-0.924664	1.000000	-0.045967
MedHouseVal	-0.144160	-0.045967	1.000000



C:\Users\HP\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning:
The figure layout has changed to tight
self._figure.tight_layout(*args, **kwargs)

