

LAB 1: Develop a program to create **Histograms** for all numerical features and analyze the distribution of each feature. Generate **Box plots** for all Numerical Features and identify any Outliers. Use California Housing dataset.

Import necessary libraries

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.datasets import fetch_california_housing
```

Load the California Housing dataset

```
california = fetch_california_housing()
```

Convert to a Pandas DataFrame and include the target column

```
data = pd.DataFrame(california.data, columns=california.feature_names)
data["MedHouseValue"] = california.target # Add house price column
```

Display basic information about the dataset

```
print("Dataset Information:\n")
```

```
print(data.info())
```

```
# Print the first 10 rows of the dataset with house prices
```

```
print("\nFirst 10 Rows of the Dataset:\n")
```

```
print(data.head(10))
```

```
# Generate Histograms for all numerical features
```

```
plt.figure(figsize=(12, 8))
```

```
data.hist(bins=30, figsize=(12, 8), edgecolor='black', color='skyblue')
```

```
plt.suptitle("Histograms of All Numerical Features", fontsize=16)
```

```
plt.show()
```

```
# Generate Box Plots to identify outliers
```

```
plt.figure(figsize=(12, 8))
```

```
for i, col in enumerate(data.columns):
```

```
    plt.subplot(3, 3, i+1)
```

```
    sns.boxplot(y=data[col], color='lightcoral')
```

```
    plt.title(col)
```

```
    plt.ylabel('')
```

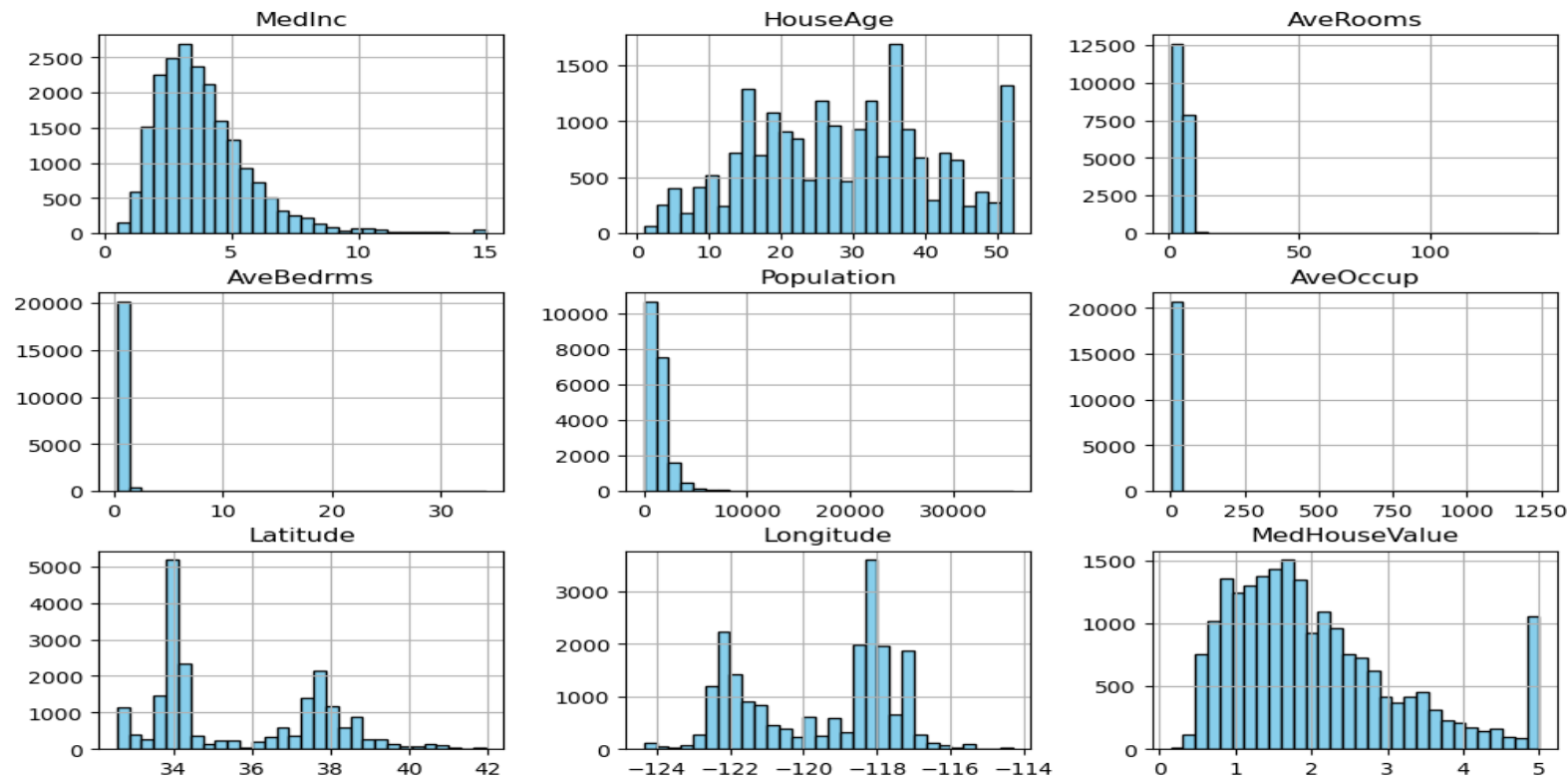
```
plt.tight_layout()
```

```
plt.suptitle("Box Plots of All Numerical Features", fontsize=16)
```

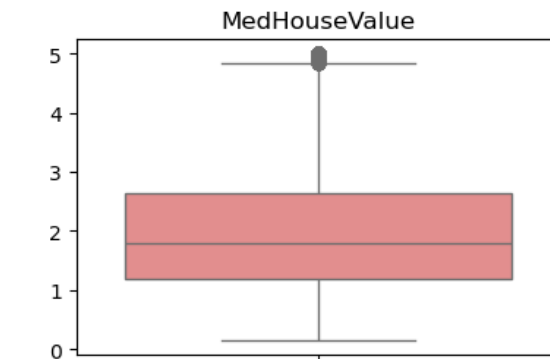
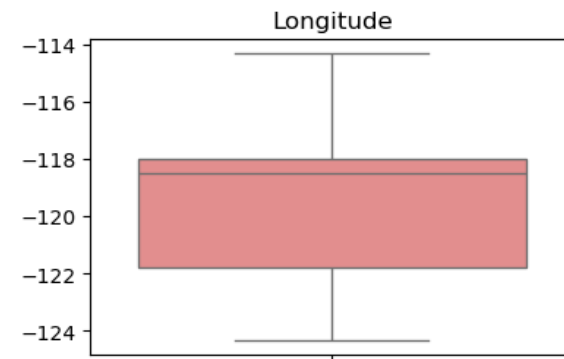
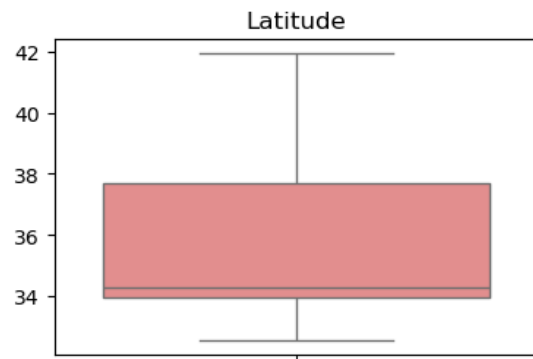
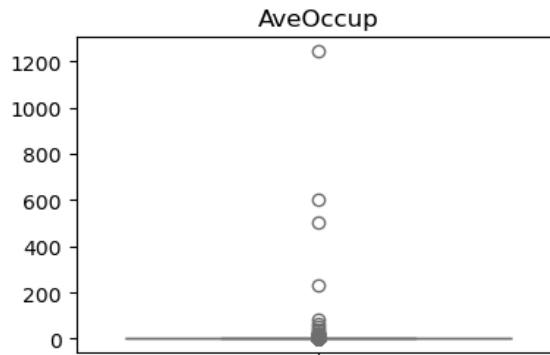
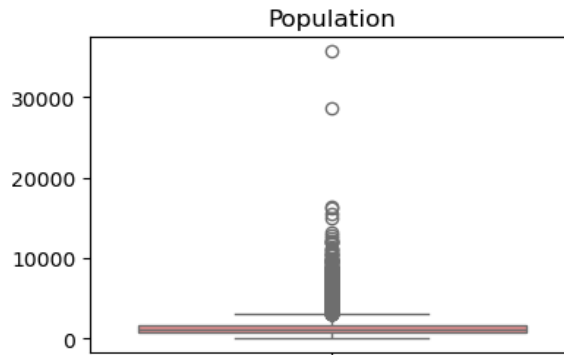
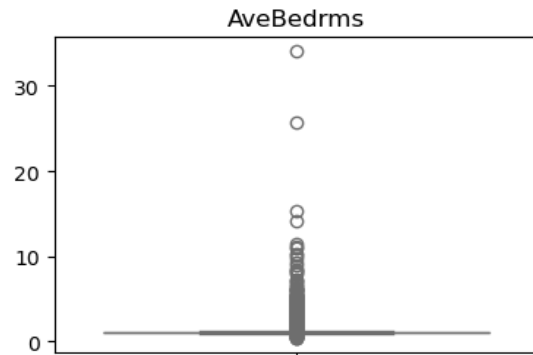
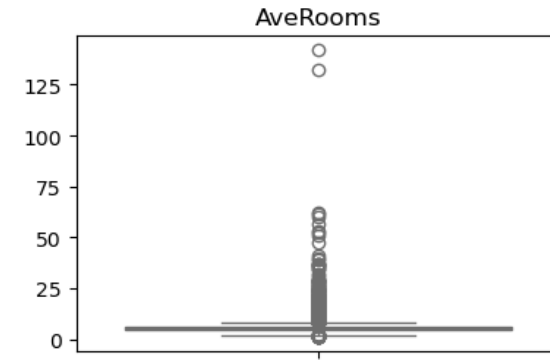
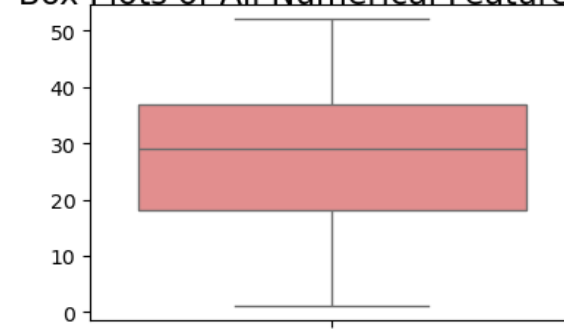
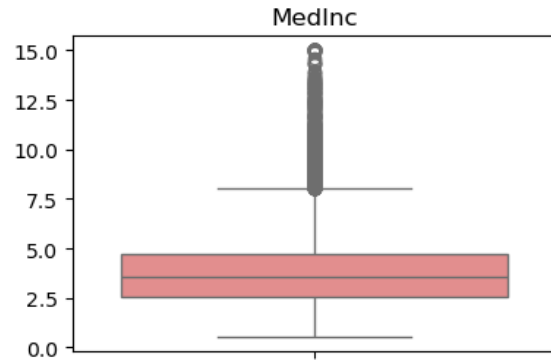
```
plt.show()
```

OUTPUT

Histograms of All Numerical Features



Box Plots of All Numerical Features



California Housing Dataset(Sample Records)

longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population	households	median_income	median_house_value
-122.23	37.88	41	880	129	322	126	8.3252	452600
-122.22	37.86	21	7099	1106	2401	1138	8.3014	358500
-122.24	37.85	52	1467	190	496	177	7.2574	352100
-122.25	37.85	52	1274	235	558	219	5.6431	341300
-122.25	37.85	52	1627	280	565	259	3.8462	342200
-122.25	37.85	52	919	213	413	193	4.0368	269700
-122.25	37.84	52	2535	489	1094	514	3.6591	299200
-122.25	37.84	52	3104	687	1157	647	3.12	241400
-122.26	37.84	42	2555	665	1206	595	2.0804	226700
-122.25	37.84	52	3549	707	1551	714	3.6912	261100
-122.26	37.85	52	2202	434	910	402	3.2031	281500
-122.26	37.85	52	3503	752	1504	734	3.2705	241800
-122.26	37.85	52	2491	474	1098	468	3.075	213500
-122.26	37.84	52	696	191	345	174	2.6736	191300
-122.26	37.85	52	2643	626	1212	620	1.9167	159200
-122.26	37.85	50	1120	283	697	264	2.125	140000
-122.27	37.85	52	1966	347	793	331	2.775	152500
-122.27	37.85	52	1228	293	648	303	2.1202	155500
-122.26	37.84	50	2239	455	990	419	1.9911	158700
-122.27	37.84	52	1503	298	690	275	2.6033	162900