Task 3 – Data Visualization

By Mallela Preethi

0s

```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

9s

```
[3] from google.colab import files
uploaded = files.upload()
```

Choose Files householdtask3.csv

- **householdtask3.csv**(text/csv) - 5189 bytes, last modified: 6/3/2024 - 100% done

```
Saving householdtask3.csv to householdtask3.csv
```

Reading data

```
[4] df = pd.read_csv('householdtask3.csv')
```

```
display(df.head(10))
```

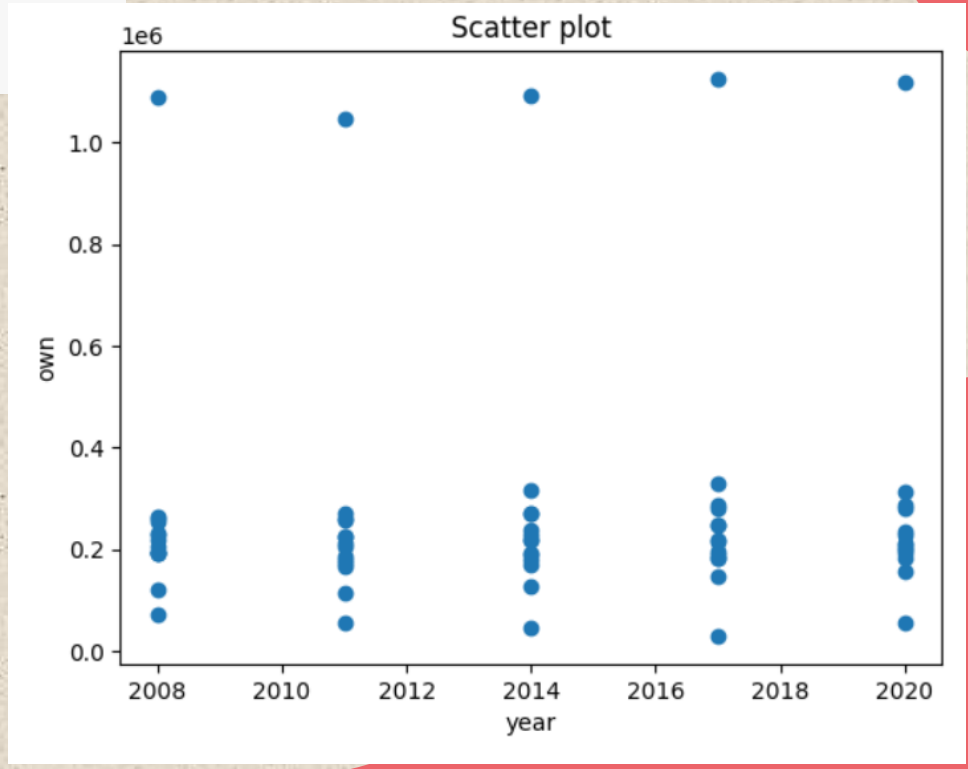
	year	tot_hhs	own	own_wm	own_prop	own_wm_prop	prop_hhs	age	size	income	expenditure	eqv_income	eqv_exp
0	2008	1560859	1087580	574406	69.7	36.8	100.0	35.9	2.7	46704	42394	26869	25132
1	2008	185965	71256	39405	38.3	21.2	11.9	29.9	2.6	23404	25270	14258	15824
2	2008	312376	191470	48424	61.3	15.5	20.0	40.0	2.3	16747	21145	13402	14408
3	2008	312333	196203	84171	62.8	26.9	20.0	34.7	2.8	31308	29855	18917	18266
4	2008	312240	217657	141318	69.7	45.3	20.0	31.5	3.0	49106	46561	26870	24672
5	2008	312336	229014	147658	73.3	47.3	20.0	35.3	2.6	61674	52776	36691	31958
6	2008	311574	253235	152835	81.3	49.1	20.0	39.3	2.5	96861	72822	55637	42932
7	2008	312761	194358	49448	62.1	15.8	20.0	38.7	2.5	23680	16413	15190	11015
8	2008	311973	206342	86390	66.1	27.7	20.0	36.1	2.7	34155	29085	20357	18121
9	2008	311840	194361	108065	62.3	34.7	20.0	33.0	2.8	49771	42662	27203	25132

```
df.shape
```

```
(70, 13)
```

Scatter Plot

```
plt.scatter(df['year'], df['own']) # year vs own column scatter plot
# Add title to scatter plot
plt.title('Scatter plot')
# Add x and y labels to plot
plt.xlabel('year')
plt.ylabel('own')
plt.show()
```

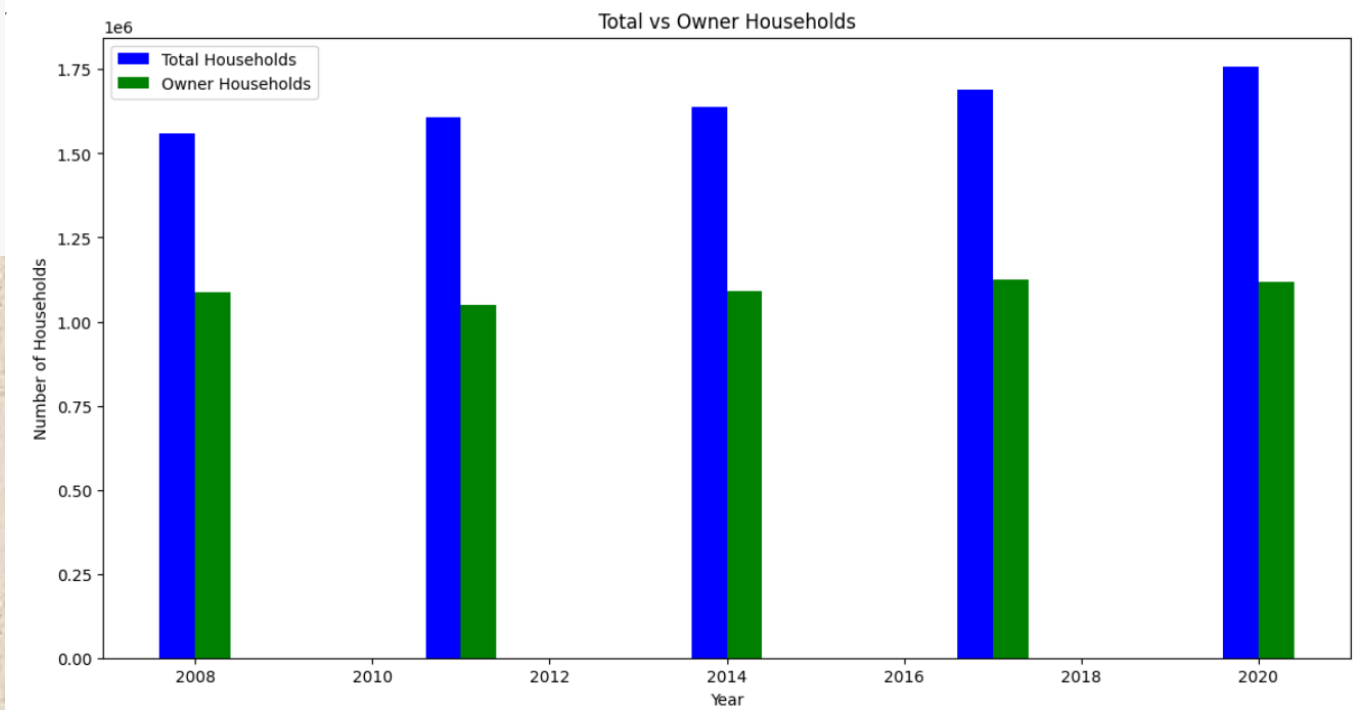


Line chart

```
[8]  # Plot line chart
     plt.figure(figsize=(10, 5))
     plt.plot(df['year'], df['income'], marker='o', label='Income', color='r')
     plt.plot(df['year'], df['expenditure'], marker='s', label='Expenditure', color='m')

     plt.xlabel('Year')
     plt.ylabel('Amount')

     plt.title('Income vs Expenditure')
     plt.legend()

     plt.show()
```

Bar Chart

```
plt.figure(figsize=(14, 7))

# Plotting total households and owner households with year as x-axis
plt.bar(df['year'] - 0.2, df['tot_hhs'], width=0.4, label='Total Households', color='b')
plt.bar(df['year'] + 0.2, df['own'], width=0.4, label='Owner Households', color='g')

# Adding x and y labels to the Bar chart
plt.xlabel('Year')
plt.ylabel('Number of Households')

# Add title to the chart
plt.title('Total vs Owner Households')

plt.legend() # Add legend

plt.show()
```



THANK YOU