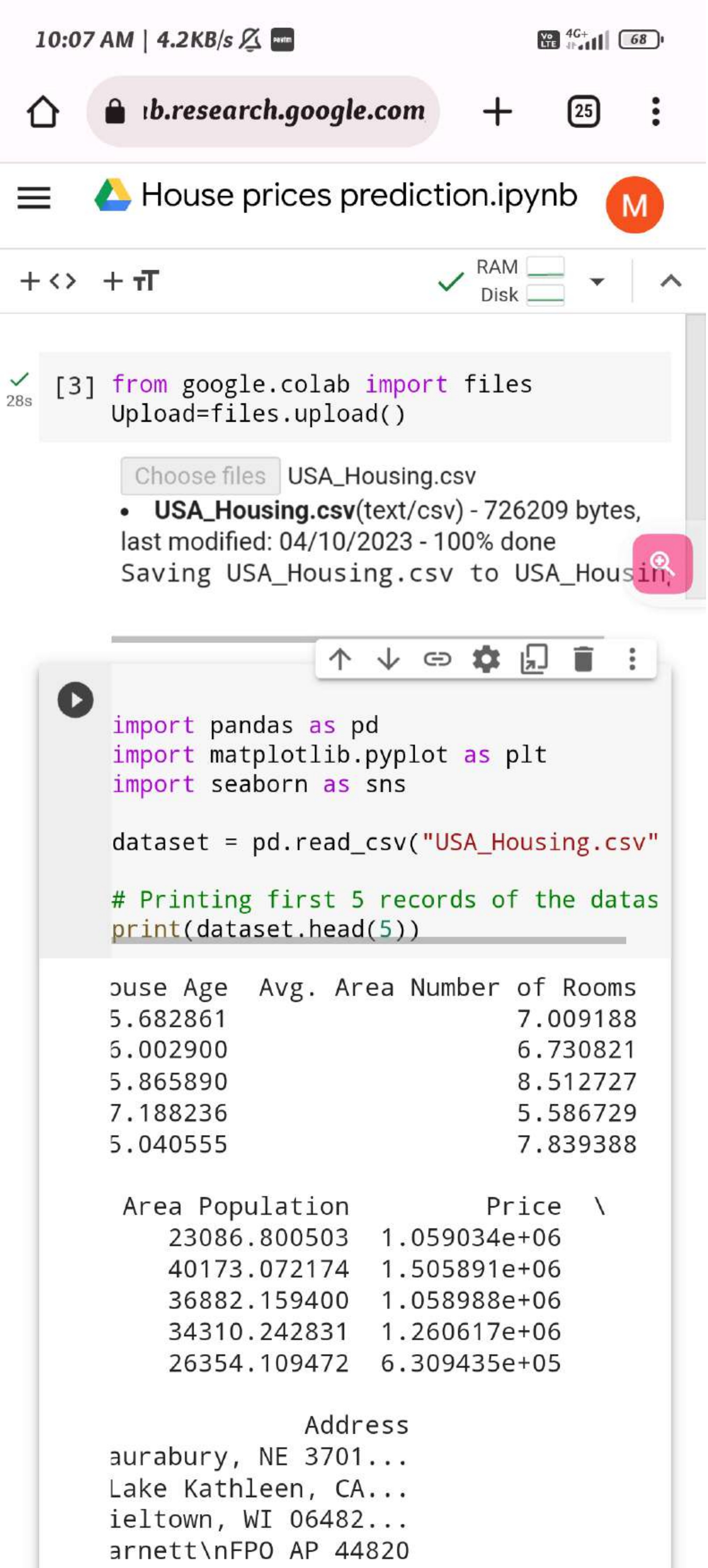


```
[3] from google.colab import files
    Upload=files.upload()
```

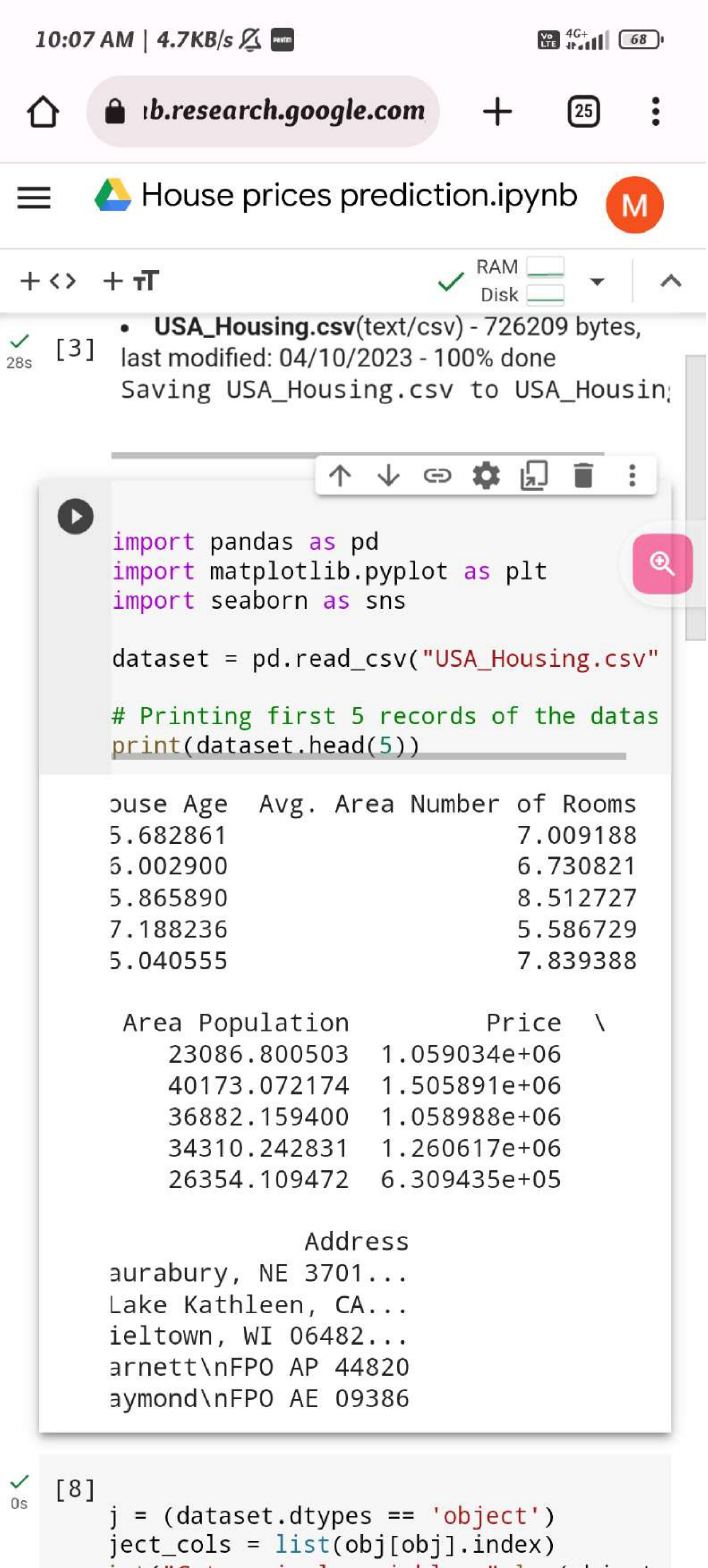
USA_Housing.csv

- **USA_Housing.csv**(text/csv) - 726209 bytes, last modified: 04/10/2023 - 100% done

```
Saving USA_Housing.csv to USA_Housin
```

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

dataset = pd.read_csv("USA_Housing.csv"

# Printing first 5 records of the datas
print(dataset.head(5))
```

```
ouse Age  Avg. Area Number of Rooms
5.682861                   7.009188
6.002900                   6.730821
5.865890                   8.512727
7.188236                   5.586729
5.040555                   7.839388

 Area Population         Price  \
    23086.800503  1.059034e+06
    40173.072174  1.505891e+06
    36882.159400  1.058988e+06
    34310.242831  1.260617e+06
    26354.109472  6.309435e+05

                 Address
aurabury, NE 3701...
Lake Kathleen, CA...
ieltown, WI 06482...
arnett\nFPO AP 44820
```



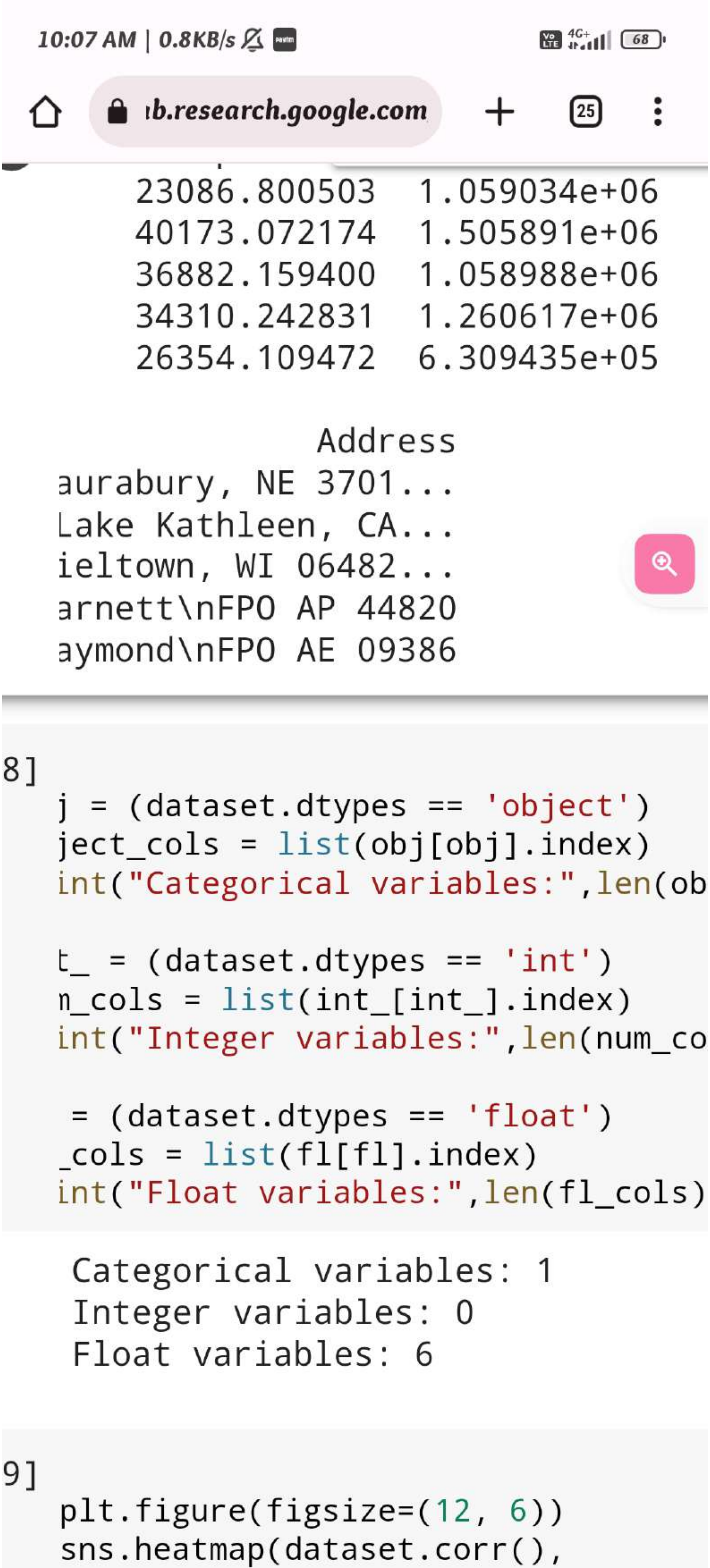
- **USA_Housing.csv**(text/csv) - 726209 bytes, last modified: 04/10/2023 - 100% done

```
Saving USA_Housing.csv to USA_Housin
```

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

dataset = pd.read_csv("USA_Housing.csv"

# Printing first 5 records of the datas
print(dataset.head(5))
```

```
ouse Age  Avg. Area Number of Rooms
5.682861                   7.009188
6.002900                   6.730821
5.865890                   8.512727
7.188236                   5.586729
5.040555                   7.839388

 Area Population         Price  \
    23086.800503  1.059034e+06
    40173.072174  1.505891e+06
    36882.159400  1.058988e+06
    34310.242831  1.260617e+06
    26354.109472  6.309435e+05

            Address
aurabury, NE 3701...
Lake Kathleen, CA...
ieltown, WI 06482...
arnett\nFPO AP 44820
aymond\nFPO AE 09386
```

```
j = (dataset.dtypes == 'object')
ject_cols = list(obj[obj].index)
```

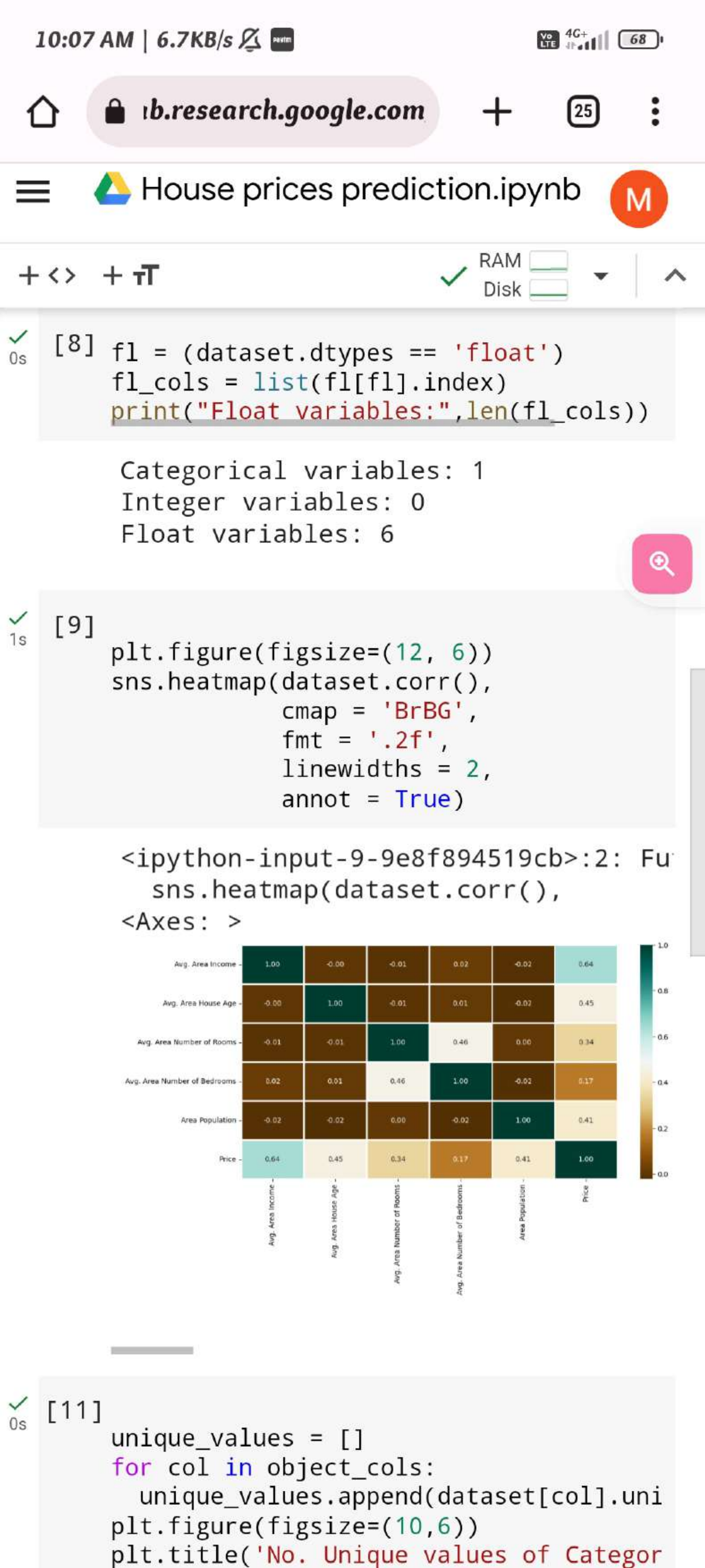


```
      23086.800503  1.059034e+06
      40173.072174  1.505891e+06
      36882.159400  1.058988e+06
      34310.242831  1.260617e+06
      26354.109472  6.309435e+05

                 Address
aurabury, NE 3701...
Lake Kathleen, CA...
ieltown, WI 06482...
arnett\nFPO AP 44820
aymond\nFPO AE 09386
```

```
8]
j = (dataset.dtypes == 'object')
ject_cols = list(obj[obj].index)
int("Categorical variables:",len(ob

t_ = (dataset.dtypes == 'int')
n_cols = list(int_[int_].index)
int("Integer variables:",len(num_co

 = (dataset.dtypes == 'float')
_cols = list(fl[fl].index)
int("Float variables:",len(fl_cols)
```

```
Categorical variables: 1
Integer variables: 0
Float variables: 6
```

```
9]
plt.figure(figsize=(12, 6))
sns.heatmap(dataset.corr(),
```



[8]

```
fl = (dataset.dtypes == 'float')
fl_cols = list(fl[fl].index)
print("Float variables:",len(fl_cols))
```

```
Categorical variables: 1
Integer variables: 0
Float variables: 6
```

[9]

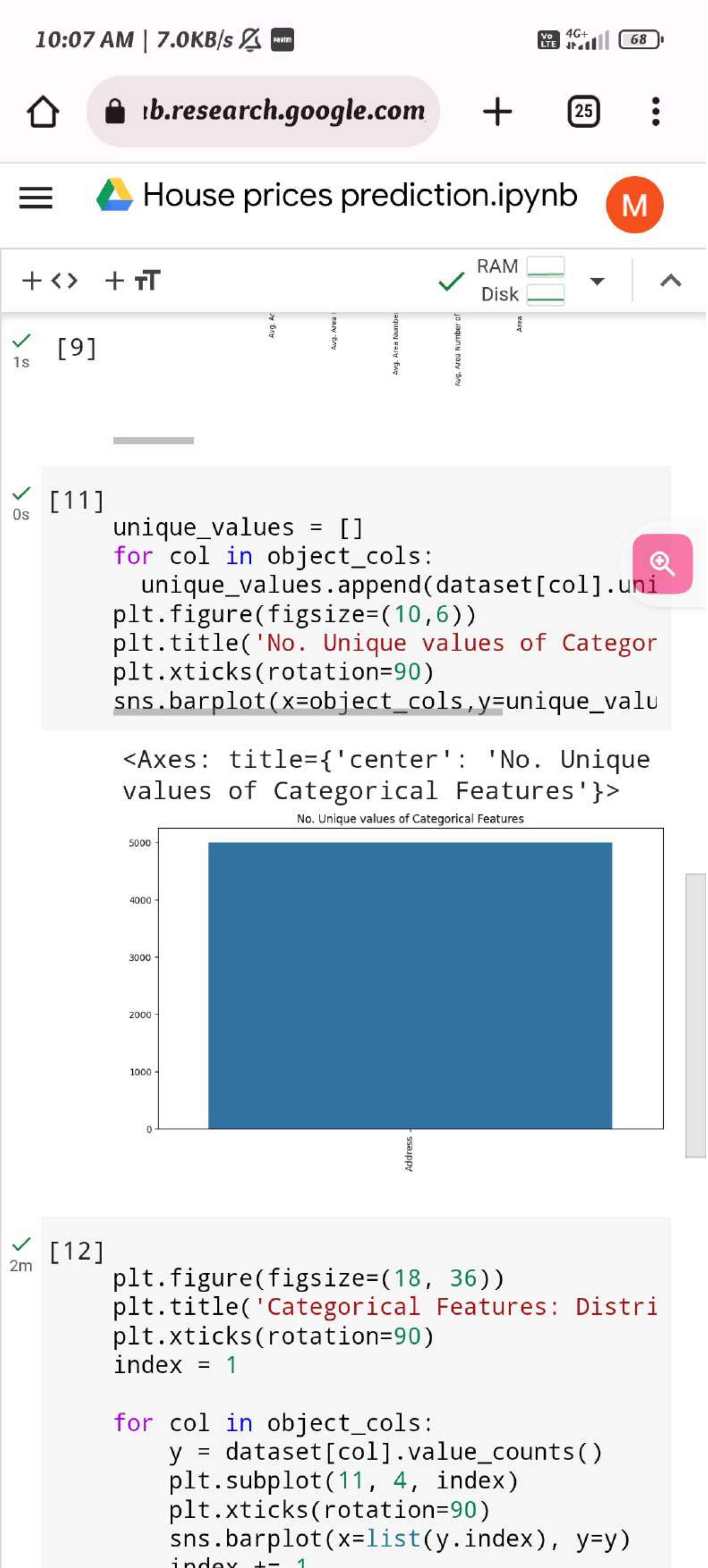
```
plt.figure(figsize=(12, 6))
sns.heatmap(dataset.corr(),
            cmap = 'BrBG',
            fmt = '.2f',
            linewidths = 2,
            annot = True)
```

```
<ipython-input-9-9e8f894519cb>:2: Fu
  sns.heatmap(dataset.corr(),
<Axes: >
```

	Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price
Avg. Area Income	1.00	-0.00	-0.01	0.02	-0.02	0.64
Avg. Area House Age	-0.00	1.00	-0.01	0.01	-0.02	0.45
Avg. Area Number of Rooms	-0.01	-0.01	1.00	0.46	0.00	0.34
Avg. Area Number of Bedrooms	0.02	0.01	0.46	1.00	-0.02	0.17
Area Population	-0.02	-0.02	0.00	-0.02	1.00	0.41
Price	0.64	0.45	0.34	0.17	0.41	1.00

[11]

```
unique_values = []
for col in object_cols:
  unique_values.append(dataset[col].uni
plt.figure(figsize=(10,6))
plt.title('No. Unique values of Categor
```



1s [9]

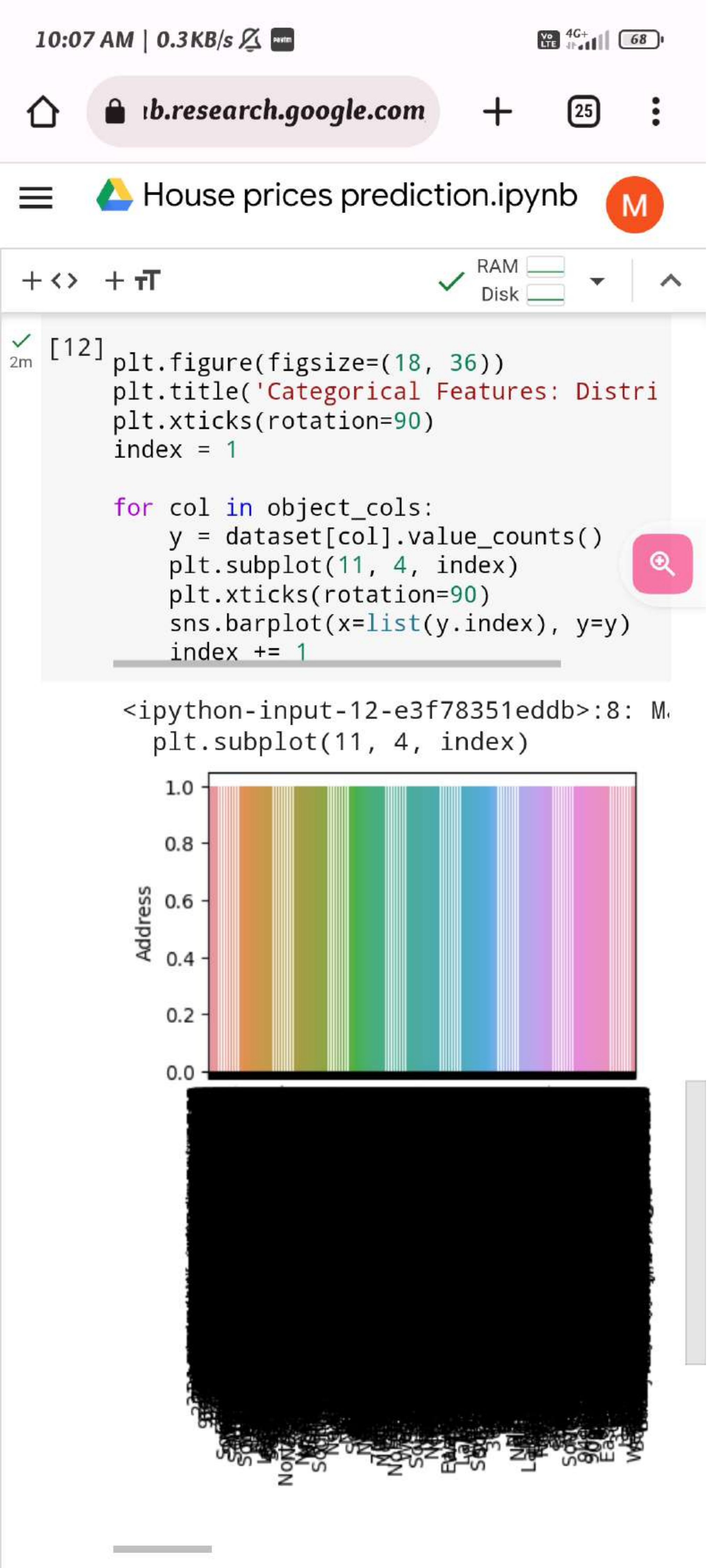
0s [11]

```
unique_values = []
for col in object_cols:
  unique_values.append(dataset[col].uni
plt.figure(figsize=(10,6))
plt.title('No. Unique values of Categor
plt.xticks(rotation=90)
sns.barplot(x=object_cols,y=unique_valu
```

```
<Axes: title={'center': 'No. Unique
values of Categorical Features'}>
```

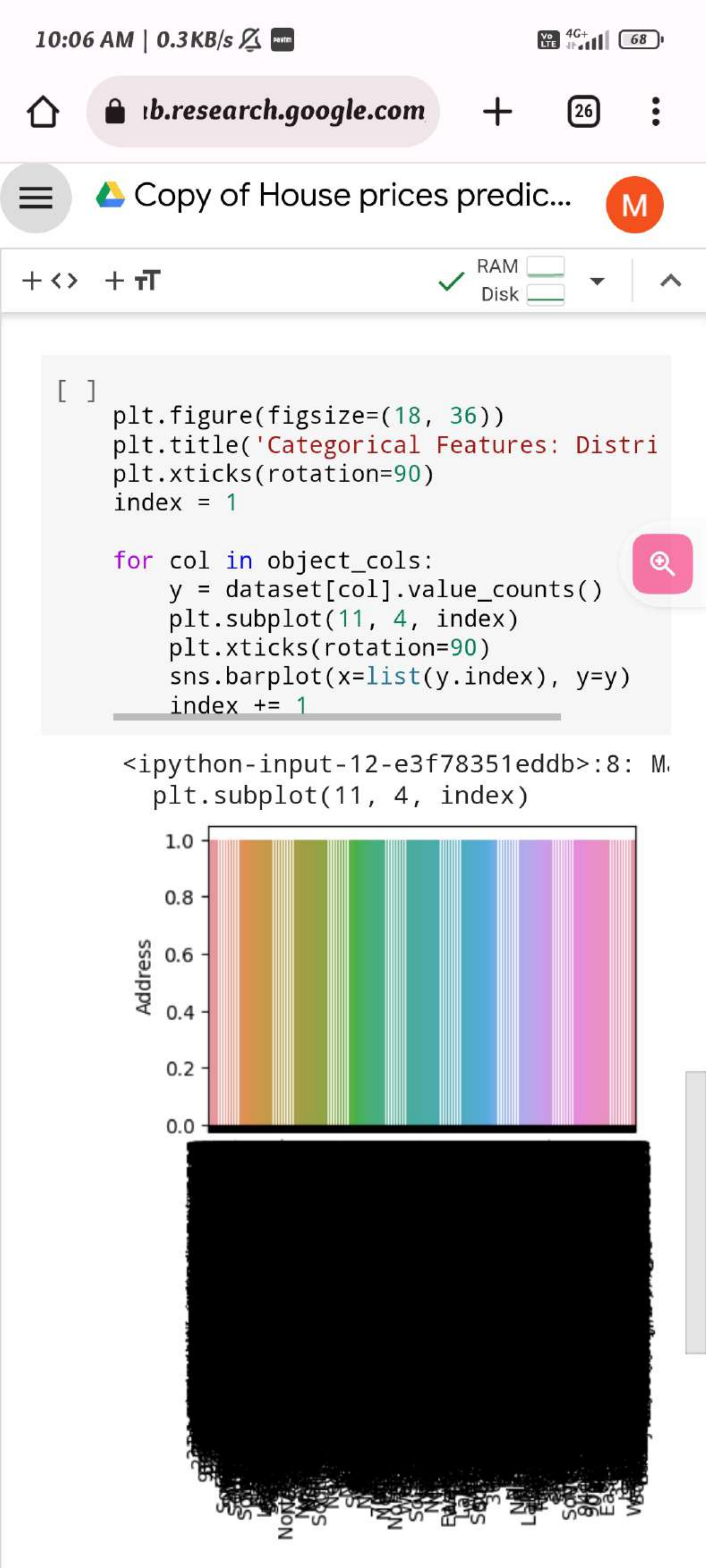
2m [12]

```
plt.figure(figsize=(18, 36))
plt.title('Categorical Features: Distri
plt.xticks(rotation=90)
index = 1

for col in object_cols:
    y = dataset[col].value_counts()
    plt.subplot(11, 4, index)
    plt.xticks(rotation=90)
    sns.barplot(x=list(y.index), y=y)
    index += 1
```



```
[12]
plt.figure(figsize=(18, 36))
plt.title('Categorical Features: Distri
plt.xticks(rotation=90)
index = 1

for col in object_cols:
    y = dataset[col].value_counts()
    plt.subplot(11, 4, index)
    plt.xticks(rotation=90)
    sns.barplot(x=list(y.index), y=y)
    index += 1
```

```
<ipython-input-12-e3f78351eddb>:8: M
  plt.subplot(11, 4, index)
```



```
[ ]
    plt.figure(figsize=(18, 36))
    plt.title('Categorical Features: Distri
    plt.xticks(rotation=90)
    index = 1

    for col in object_cols:
        y = dataset[col].value_counts()
        plt.subplot(11, 4, index)
        plt.xticks(rotation=90)
        sns.barplot(x=list(y.index), y=y)
        index += 1
```

```
<ipython-input-12-e3f78351eddb>:8: M
  plt.subplot(11, 4, index)
```