

Learning Reflection Assignment:

Now you have successfully conquered the material. You tried hard to practice the questions, and you have a good result from the practice.

In this assignment, you should first show what is your findings with the dataset of your choice by using the package in the module. You then should summarize your learning experience briefly. You don't need to write lengthy answers, one or two paragraphs for one question will be fine. In terms of communication, providing some examples to support your arguments would be better.

You should prepare a PDF that contains both your practice/result and your answers to the questions.

Your learning reflection will be peer-reviewed by the cohort. Let's learn from each other's experiences and it is, once again, another type of learning.

Let's get started!

Regarding the data and your result:

Q1: Print your result of playing with the data set of your choice with the package in this module. Try to make good use of the data, and conduct clear and interesting communication with your audience.

Q2: What is the most difficult part of playing with the data? How did you work through the difficulty?

Q3: What is the most important takeaway from your result? Why?

Regarding your learning experience:

Q4: Do you have a better understanding and capability for data visualization? Why?

Q5: What is your most successful approach in learning this module?

```
In [19]: import matplotlib as mpl
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import sklearn
import seaborn as sns
import matplotlib.pyplot as plt
```

```
In [20]: housing = pd.read_csv("../housing/housing.csv")
```

```
In [21]: housing
```

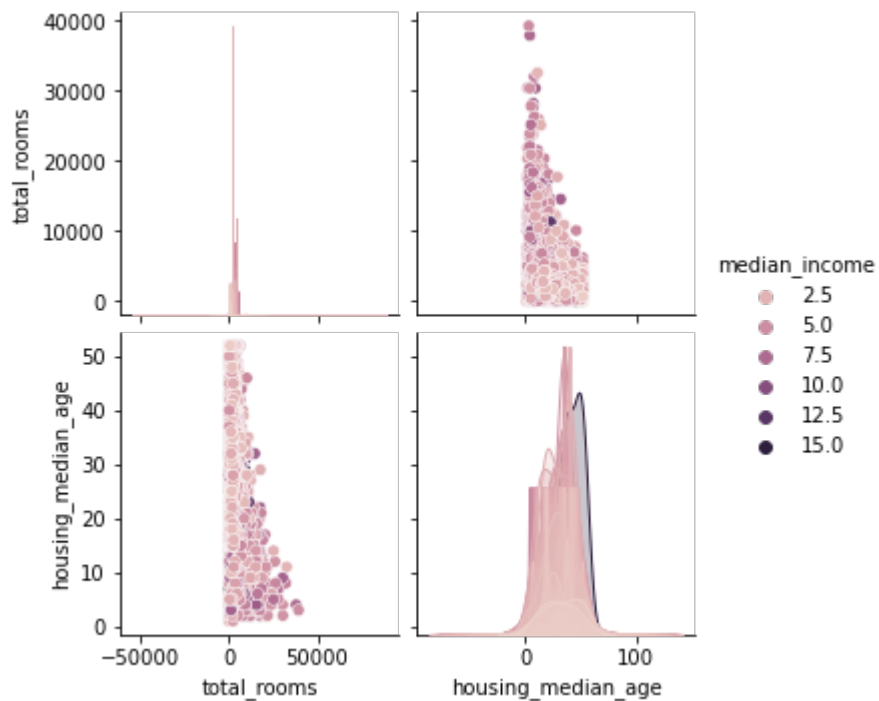
```
Out[21]:
```

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population	household_size
0	-122.23	37.88	41.0	880.0	129.0	322.0	2.58
1	-122.22	37.86	21.0	7099.0	1106.0	2401.0	2.18
2	-122.24	37.85	52.0	1467.0	190.0	496.0	2.48
3	-122.25	37.85	52.0	1274.0	235.0	558.0	2.46
4	-122.25	37.85	52.0	1627.0	280.0	565.0	2.59
...
20635	-121.09	39.48	25.0	1665.0	374.0	845.0	2.26
20636	-121.21	39.49	18.0	697.0	150.0	356.0	2.38
20637	-121.22	39.43	17.0	2254.0	485.0	1007.0	2.08
20638	-121.32	39.43	18.0	1860.0	409.0	741.0	2.49
20639	-121.24	39.37	16.0	2785.0	616.0	1387.0	2.29

20640 rows × 10 columns

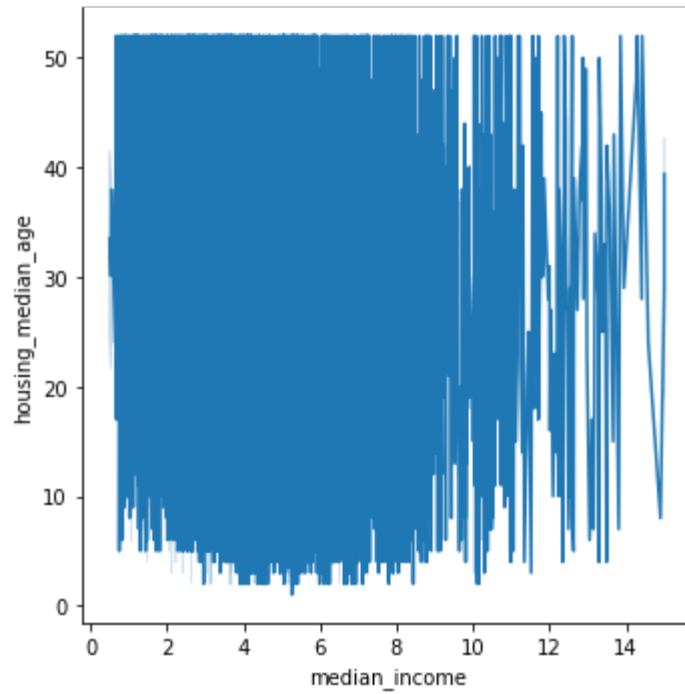
```
In [8]: pairplot = housing[["median_income", "total_rooms",  
                             "housing_median_age"]]  
sns.pairplot(pairplot, hue='median_income')
```

```
Out[8]: <seaborn.axisgrid.PairGrid at 0x7fb8c3154f10>
```



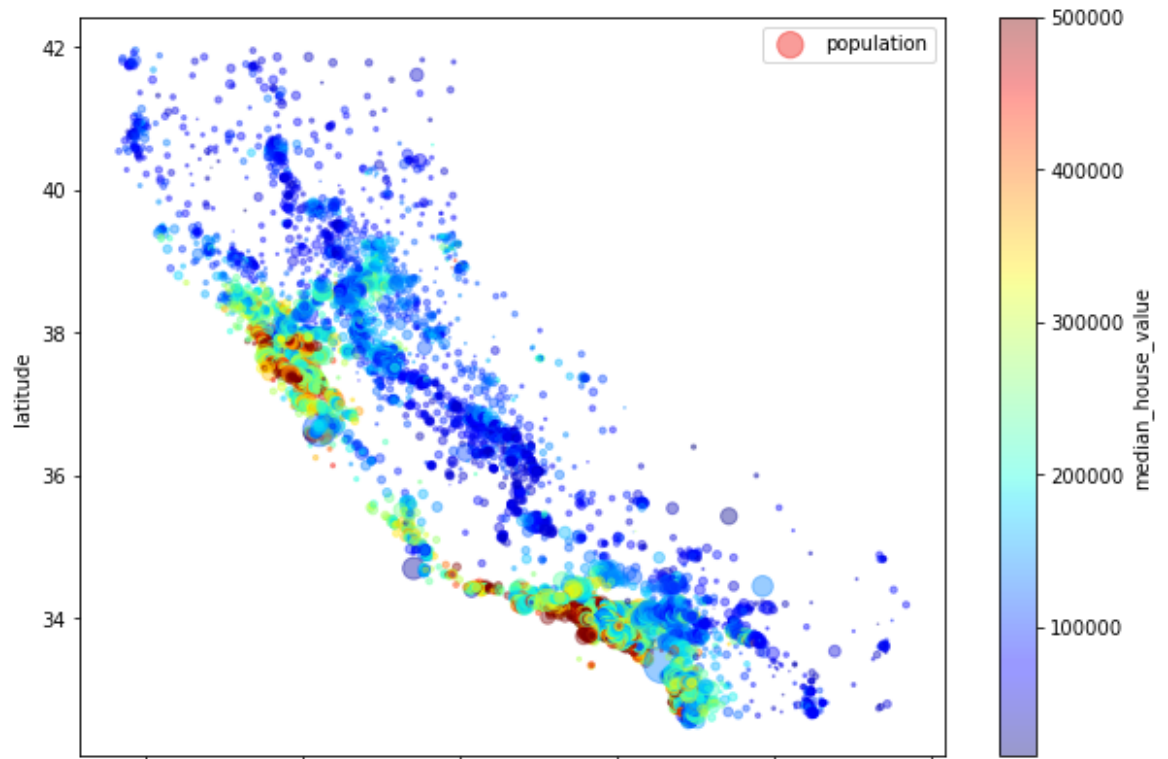
```
In [15]: sns.relplot(  
    data=housing, kind="line",  
    x=housing["median_income"], y=housing["housing_median_age"],  
    )
```

Out[15]: <seaborn.axisgrid.FacetGrid at 0x7fb8a695aa60>



```
In [23]: housing.plot(kind="scatter", x="longitude", y="latitude", alpha=0.4,  
s=housing["population"]/100, label="population", figsize=(10,7),  
c="median_house_value", cmap=plt.get_cmap("jet"), colorbar=True,  
)  
plt.legend()
```

Out[23]: <matplotlib.legend.Legend at 0x7fb8a7af8820>



Q1: see above

Q2: visualizing import information

Q3: price of a house is more expensive on coast

Q4: yes, seaborn is difficult but very powerful

Q5: practicing

In []: