AlvimPedro / **ppge_machine_learning**   Public

<> **Code**   ⊙ Issues   ⑂ Pull requests   ▷ Actions   ⊞ Projects   ▯ Wiki   ⊘ Security   ⩘ Insights   ⚙ Settings

Beta   Try the new code view

⑂ main ▾   **ppge_machine_learning** / lista1 / **ex4.ipynb**   Go to file   ···

AlvimPedro Lista 1   Latest commit `77a5ac8` 5 minutes ago   ⏱ **History**

⩘ **1** contributor

282 lines (282 sloc) | 7.21 KB   <>   ▯   Raw   Blame   ✎ ▾   ⧉   🗑

In [38]:
```python
import pandas as pd
import math

df = pd.read_csv('housing.csv')
df.head()
```

Out[38]:

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population	households	median_income	median_house_value
0	-122.23	37.88	41.0	880.0	129.0	322.0	126.0	8.3252	452600.0
1	-122.22	37.86	21.0	7099.0	1106.0	2401.0	1138.0	8.3014	358500.0
2	-122.24	37.85	52.0	1467.0	190.0	496.0	177.0	7.2574	352100.0
3	-122.25	37.85	52.0	1274.0	235.0	558.0	219.0	5.6431	341300.0
4	-122.25	37.85	52.0	1627.0	280.0	565.0	259.0	3.8462	342200.0

Choosing `median_income` and `median_house_value`

Mean

In [39]:
```python
m1 = df['median_income'].mean()
m2 = df['median_house_value'].mean()
```

Variance

In [40]:
```python
v1 = df['median_income'].var()
v2 = df['median_house_value'].var()
```

Standard Deviation

In [41]:
```python
s1 = math.sqrt(v1)
s2 = math.sqrt(v2)
```

Covariance

In [42]:
```python
cov = ((df['median_income'] - m1)*(df['median_house_value'] - m2)).sum() / len(df)
```

Correlation

In [44]:
```python
corr = cov / (s1*s2)
```

In [52]:
```python
print(f'A correlação entre media de renda do bairro e a media de preço das casas do bairro é de: {corr:.4f}'
```

A correlação entre media de renda do bairro e a media de preço das casas do bairro é de: 0.6880

In [ ]:

Give feedback