



Learning Portfolio: Update

Victoria Zabel, SE 2023S437253

```

241         expensive
17398        affordable
36608        affordable
44731         adequate
18104    very expensive
...
11284        expensive
44732         adequate
38158         adequate
860         expensive
15795    very expensive

```

```
[36139 rows x 10 columns]
```

By having turned cut into a numerical variable, we are now able to analyze its effect on other variables.

```
30]: data.corr()
```

```
30]:
```

	carat	cut	depth	table	x	y	z
carat	1.000000	-0.139938	0.033351	0.175629	0.975212	0.948318	0.945965
cut	-0.139938	1.000000	-0.169529	-0.409105	-0.131307	-0.125194	-0.148888
depth	0.033351	-0.169529	1.000000	-0.297566	-0.020945	-0.026081	0.098007
table	0.175629	-0.409105	-0.297566	1.000000	0.190703	0.178906	0.144665
x	0.975212	-0.131307	-0.020945	0.190703	1.000000	0.970684	0.962937
y	0.948318	-0.125194	-0.026081	0.178906	0.970684	1.000000	0.940992
z	0.945965	-0.148888	0.098007	0.144665	0.962937	0.940992	1.000000

The following variables were transformed.

```
34]: from sklearn.compose import ColumnTransformer
      from sklearn.pipeline import make_pipeline
      from sklearn.preprocessing import FunctionTransformer
      from sklearn.preprocessing import StandardScaler
      import numpy as np
      from sklearn.preprocessing import OneHotEncoder

      default_num_pipeline = make_pipeline(StandardScaler())

```



+ Code + Text

✓ RAM
Disk

✓ [54] iris.drop('sepal_width', axis=1, inplace=True)

✓ [55] iris.drop('species', axis=1, inplace=True)

✓ [56] iris.head(10)

1 to 10 of 10 entries

Index	petal_length	petal_width
0	1.4	0.2
1	1.4	0.2
2	1.3	0.2
3	1.5	0.2
4	1.4	0.2
5	1.7	0.4
6	1.4	0.3
7	1.5	0.2
8	1.4	0.2
9	1.5	0.1

Show per page✓ [78] ilength = iris["petal_length"]
#ilength

iwidth = iris["petal_width"]
#iwidth