Adding new columns to the data frame

November 6, 2020

1 Adding new columns to the data frame

[1]:
```
import pandas as pd
import seaborn as sns
```

[2]:
```
diamonds_url = "https://raw.githubusercontent.com/TrainingByPackt/
↪Interactive-Data-Visualization-with-Python/master/datasets/diamonds.csv"
```

[3]:
```
diamonds_df = pd.read_csv(diamonds_url)
```

[4]:
```
diamonds_df['price_per_carat'] = diamonds_df['price']/diamonds_df['carat']
```

[5]:
```
diamonds_df.head()
```

[5]:

	carat	cut	color	clarity	depth	table	price	x	y	z	\
0	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43	
1	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31	
2	0.23	Good	E	VS1	56.9	65.0	327	4.05	4.07	2.31	
3	0.29	Premium	I	VS2	62.4	58.0	334	4.20	4.23	2.63	
4	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75	

	price_per_carat
0	1417.391304
1	1552.380952
2	1421.739130
3	1151.724138
4	1080.645161

2 Conditional addition of columns

[6]:
```
import numpy as np

diamonds_df['price_per_carat_is_high'] = np.
↪where(diamonds_df['price_per_carat']>3500,1,0)
```

[7]:
```
diamonds_df.head()
```

[7]:

	carat	cut	color	clarity	depth	table	price	x	y	z	price_per_carat	price_per_carat_is_high
0	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43	1417.391304	0
1	0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31	1552.380952	0
2	0.23	Good	E	VS1	56.9	65.0	327	4.05	4.07	2.31	1421.739130	0
3	0.29	Premium	I	VS2	62.4	58.0	334	4.20	4.23	2.63	1151.724138	0
4	0.31	Good	J	SI2	63.3	58.0	335	4.34	4.35	2.75	1080.645161	0

[ ]: