Quiz 4

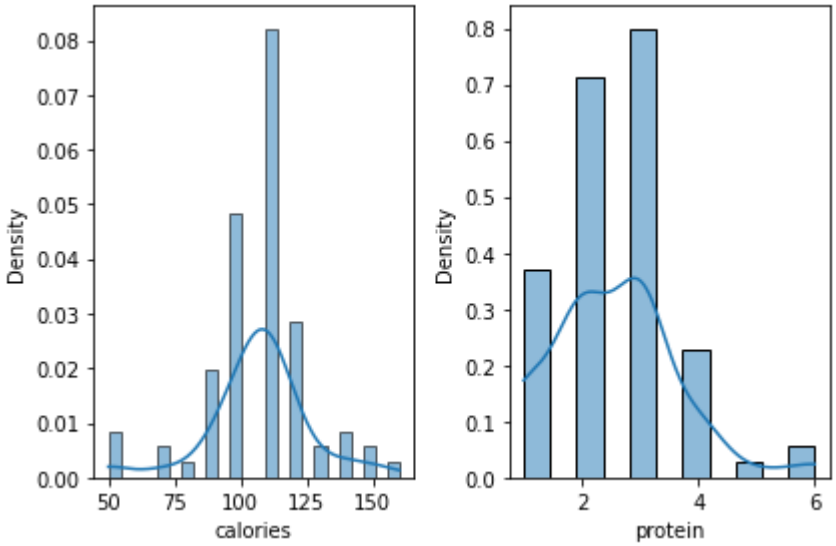
Loan Pham and Brandan Owens

In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
df = pd.read_csv('../dataFiles/cereal.csv')
df.head()
```

Out[1]:

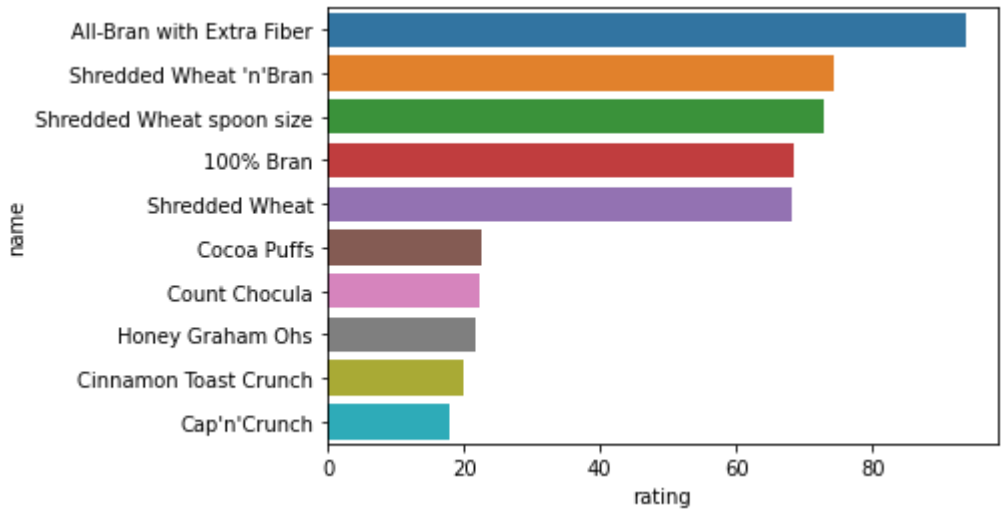
	name	mfr	type	calories	protein	fat	sodium	fiber	carbo	sugars	potass	vitamins	shelf	we
0	100% Bran	N	C	70	4	1	130	10.0	5.0	6	280	25	3	
1	100% Natural Bran	Q	C	120	3	5	15	2.0	8.0	8	135	0	3	
2	All-Bran	K	C	70	4	1	260	9.0	7.0	5	320	25	3	
3	All-Bran with Extra Fiber	K	C	50	4	0	140	14.0	8.0	0	330	25	3	
4	Almond Delight	R	C	110	2	2	200	1.0	14.0	8	-1	25	3	

In [2]:
```
#q.1.a
fig, ax = plt.subplots(1,2)
ax1 = sns.histplot(df, x='calories', stat='density', kde=True, ax=ax[0])
ax2 = sns.histplot(df, x='protein', stat='density', kde=True, ax=ax[1])
plt.tight_layout()
```



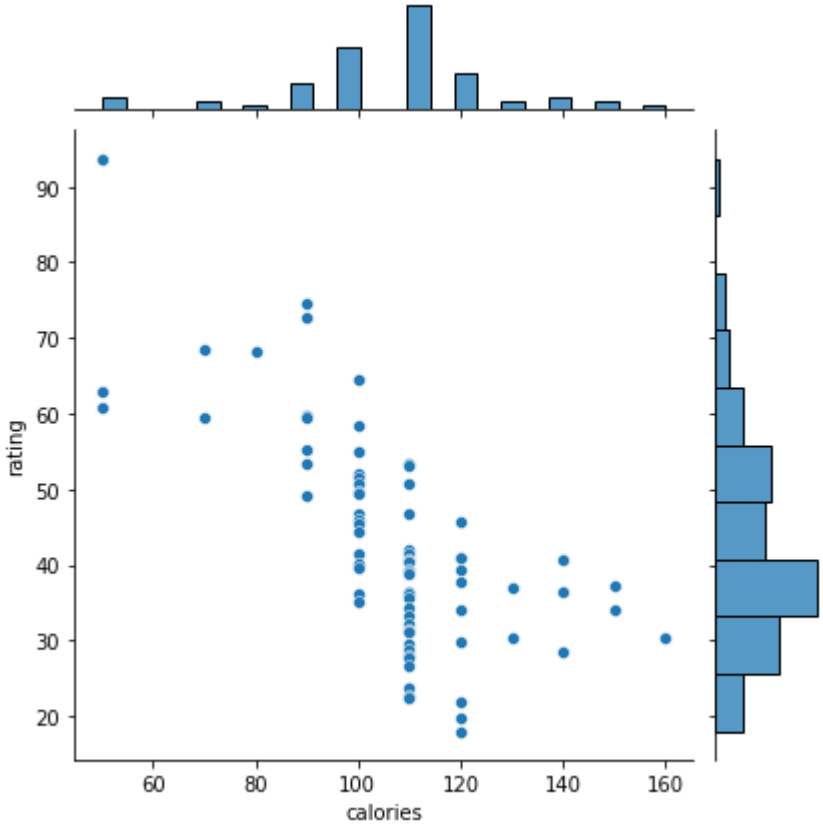
In [3]:
```
#q.1.b
df.sort_values(by='rating', ascending=False, inplace = True)
tops_bottoms = df.head().append(df.tail())
sns.barplot(x="rating", y="name", data=tops_bottoms)
```

Out[3]: <AxesSubplot:xlabel='rating', ylabel='name'>



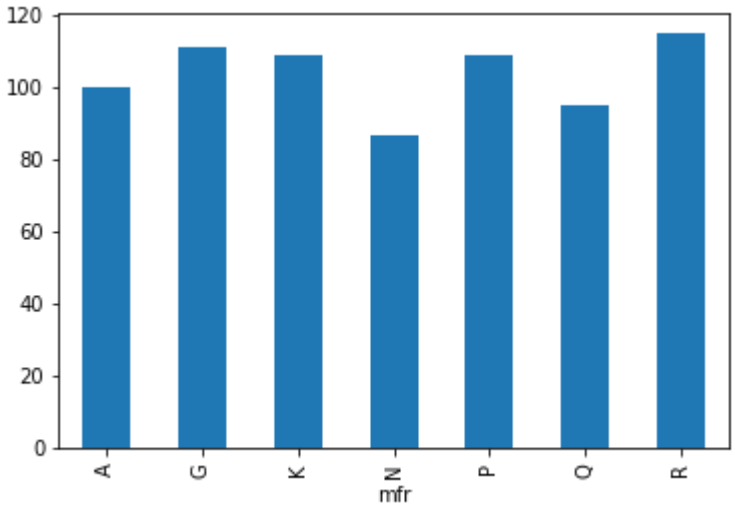
In [4]:
```
#q.1.c
sns.jointplot(data=df, x='calories', y='rating')
```

Out[4]: <seaborn.axisgrid.JointGrid at 0x1bff91227c0>



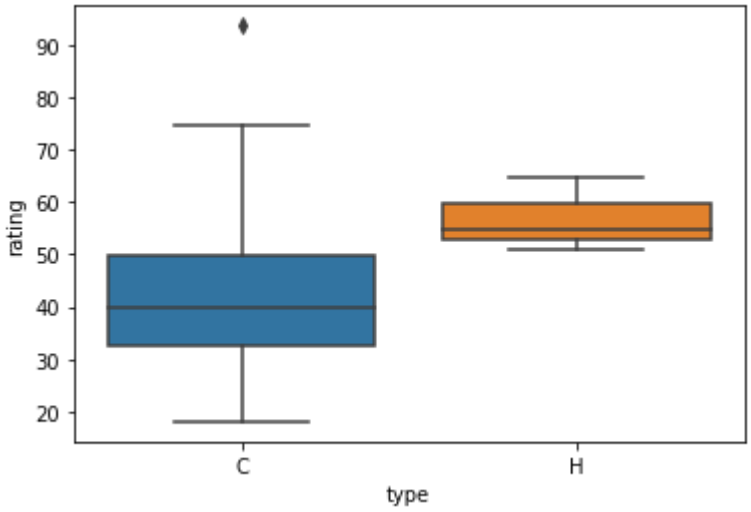
In [5]:
```
#q.1.d
mfr_cal = df.groupby('mfr')['calories'].mean().reset_index()
mfr_cal.plot.bar(x='mfr', y='calories', legend=False)
```

Out[5]: <AxesSubplot:xlabel='mfr'>



In [6]:
```
#q.1.e
sns.boxplot(x='type', y='rating', data=df)
```

Out[6]: <AxesSubplot:xlabel='type', ylabel='rating'>


90
80
70
60
50
40
30
20
rating
C
H
type