Quiz 4

Team: Phil Kim, Jeremiah Snow, Ryan Schuck

Q1 - cereal.csv

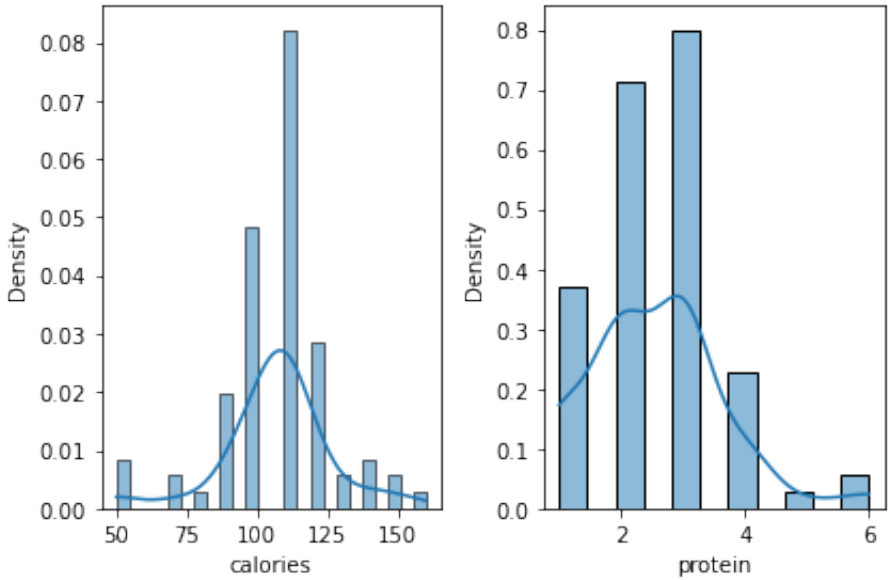
In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

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

Out[1]:

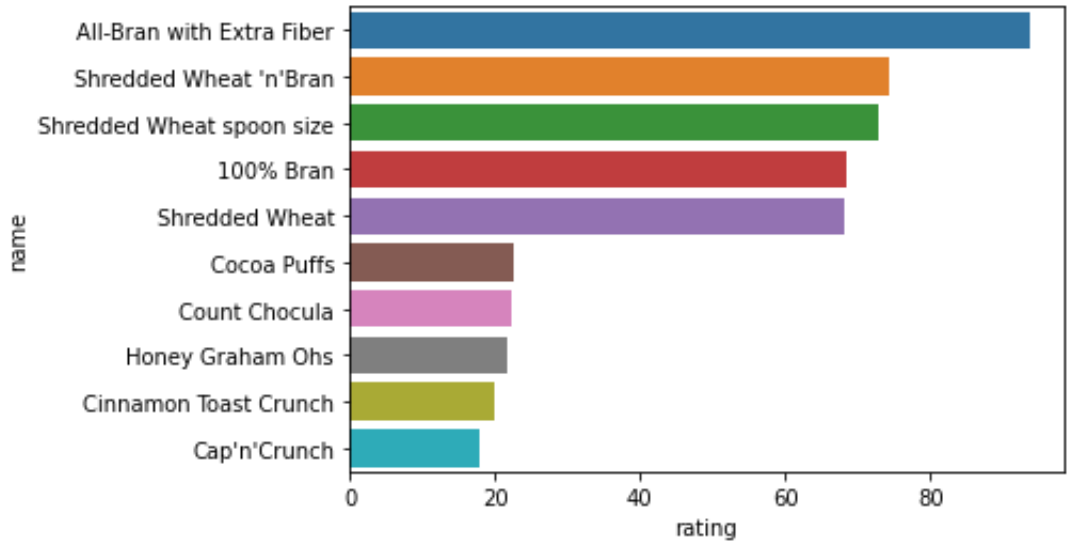
	name	mfr	type	calories	protein	fat	sodium	fiber	carbo	sugars	potass	vitamins	shelf	weight	cups	rating
0	100% Bran	N	C	70	4	1	130	10.0	5.0	6	280	25	3	1.0	0.33	68.402973
1	100% Natural Bran	Q	C	120	3	5	15	2.0	8.0	8	135	0	3	1.0	1.00	33.983679
2	All-Bran	K	C	70	4	1	260	9.0	7.0	5	320	25	3	1.0	0.33	59.425505
3	All-Bran with Extra Fiber	K	C	50	4	0	140	14.0	8.0	0	330	25	3	1.0	0.50	93.704912
4	Almond Delight	R	C	110	2	2	200	1.0	14.0	8	-1	25	3	1.0	0.75	34.384843

In [2]:
```
# (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()
```



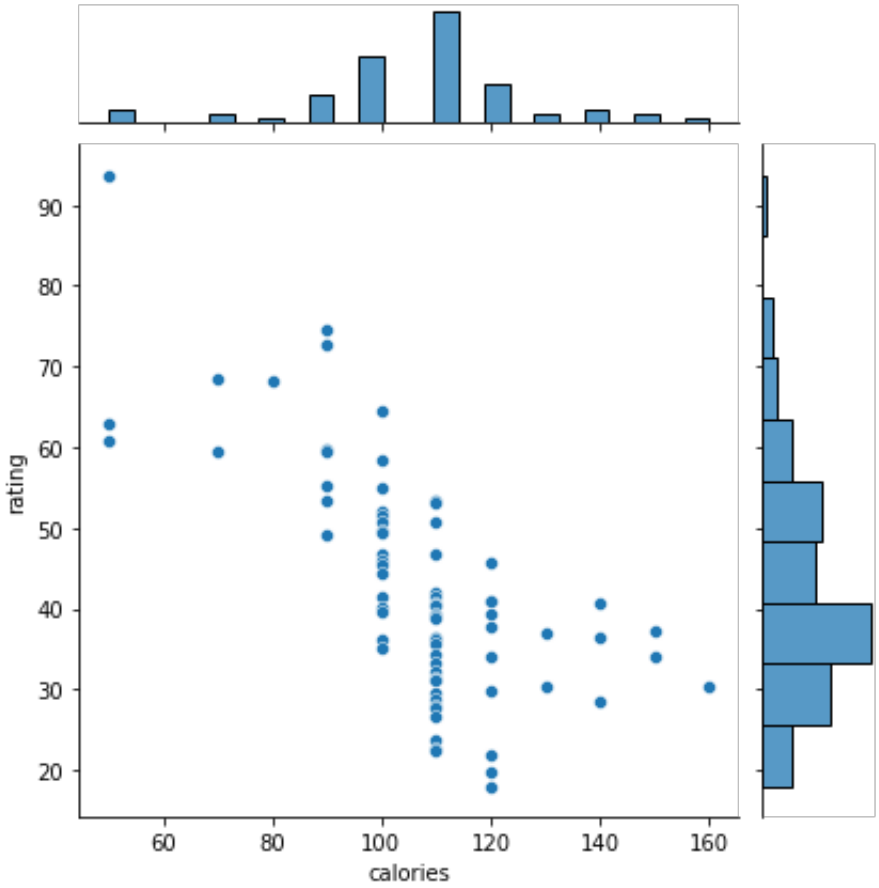
In [3]:
```
# (b)
df.sort_values(by='rating', ascending=False, inplace = True)
tops_bottoms = df.head().append(df.tail())
sns.barplot(x="rating", y="name", data=tops_bottoms)
```

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



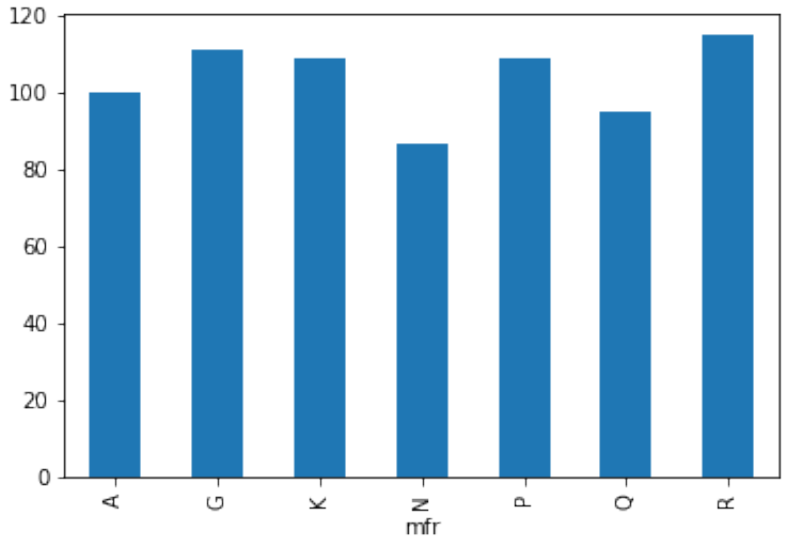
In [4]:
```
# (c)
sns.jointplot(data=df, x='calories', y='rating')
```

Out[4]:
```
<seaborn.axisgrid.JointGrid at 0x7fa628524fa0>
```



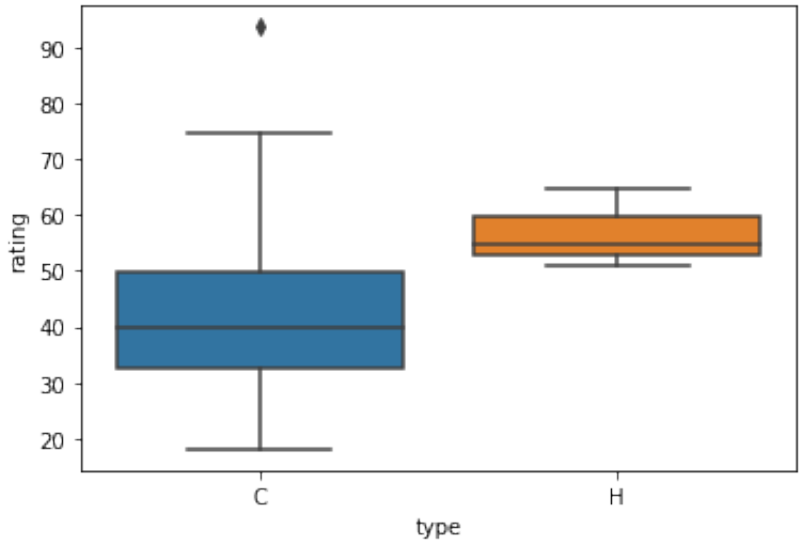
In [5]:
```
# (d)
mfr_cal = df.groupby('mfr')['calories'].mean().reset_index()
mfr_cal.plot.bar(x='mfr', y='calories', legend=False)
```

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



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

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



In [ ]: