In [3]:
```
#import python libs
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv("E:/Data Analytics Project/New folder/cereal.csv")
print(df)
```

```
                         name mfr type  calories  protein  fat  sodium  fiber  \
0                   100% Bran   N    C        70        4    1     130   10.0
1           100% Natural Bran   Q    C       120        3    5      15    2.0
2                    All-Bran   K    C        70        4    1     260    9.0
3   All-Bran with Extra Fiber   K    C        50        4    0     140   14.0
4              Almond Delight   R    C       110        2    2     200    1.0
..                        ...  ..  ...       ...      ...  ...     ...    ...
72                    Triples   G    C       110        2    1     250    0.0
73                       Trix   G    C       110        1    1     140    0.0
74                 Wheat Chex   R    C       100        3    1     230    3.0
75                   Wheaties   G    C       100        3    1     200    3.0
76        Wheaties Honey Gold   G    C       110        2    1     200    1.0

    carbo  sugars  potass  vitamins  shelf  weight  cups     rating
0     5.0       6     280        25      3     1.0  0.33  68.402973
1     8.0       8     135         0      3     1.0  1.00  33.983679
2     7.0       5     320        25      3     1.0  0.33  59.425505
3     8.0       0     330        25      3     1.0  0.50  93.704912
4    14.0       8      -1        25      3     1.0  0.75  34.384843
..    ...     ...     ...       ...    ...     ...   ...        ...
72   21.0       3      60        25      3     1.0  0.75  39.106174
73   13.0      12      25        25      2     1.0  1.00  27.753301
74   17.0       3     115        25      1     1.0  0.67  49.787445
75   17.0       3     110        25      1     1.0  1.00  51.592193
76   16.0       8      60        25      1     1.0  0.75  36.187559

[77 rows x 16 columns]
```

In [4]:
```
df.head()
```

Out[4]:

	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 [6]:
```
print(df.isnull().sum())
```

```
name        0
mfr         0
type        0
calories    0
protein     0
fat         0
sodium      0
fiber       0
carbo       0
sugars      0
potass      0
vitamins    0
shelf       0
weight      0
cups        0
rating      0
dtype: int64
```

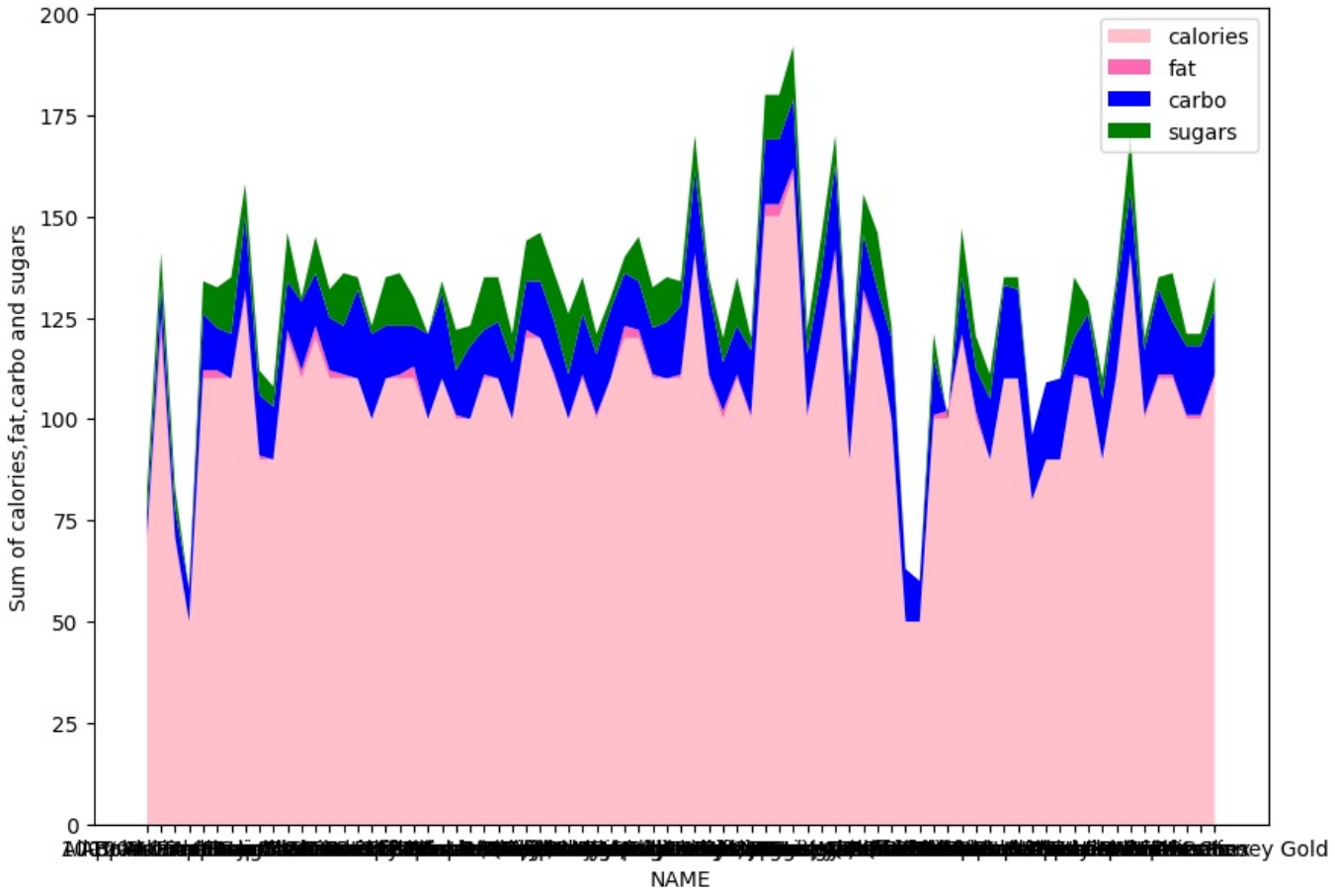
In [39]:
```
gp = df.groupby("name").agg({"calories":"sum","fat":"sum","protein":"sum","sodium":"sum","fiber":"sum","carbo":
#print(gp)
print(gp)
```

```
                           calories  fat  protein  sodium  fiber  carbo  \
name
100% Bran                        70    1        4     130   10.0    5.0
100% Natural Bran               120    5        3      15    2.0    8.0
All-Bran                         70    1        4     260    9.0    7.0
All-Bran with Extra Fiber        50    0        4     140   14.0    8.0
Almond Delight                  110    2        2     200    1.0   14.0
...                             ...  ...      ...     ...    ...    ...
Triples                         110    1        2     250    0.0   21.0
Trix                            110    1        1     140    0.0   13.0
Wheat Chex                      100    1        3     230    3.0   17.0
Wheaties                        100    1        3     200    3.0   17.0
Wheaties Honey Gold             110    1        2     200    1.0   16.0

                           sugars  potass  vitamins  shelf
name
100% Bran                       6     280        25      3
100% Natural Bran               8     135         0      3
All-Bran                        5     320        25      3
All-Bran with Extra Fiber       0     330        25      3
Almond Delight                  8      -1        25      3
...                           ...     ...       ...    ...
Triples                         3      60        25      3
Trix                           12      25        25      2
Wheat Chex                      3     115        25      1
Wheaties                        3     110        25      1
Wheaties Honey Gold             8      60        25      1

[77 rows x 10 columns]
```

In [111…
```
gp = df.groupby("name").agg({"calories":"mean","fat":"mean","protein":"mean","sodium":"mean","fiber":"mean","ca
#print(gp)
print((gp.values > 0).any())
```

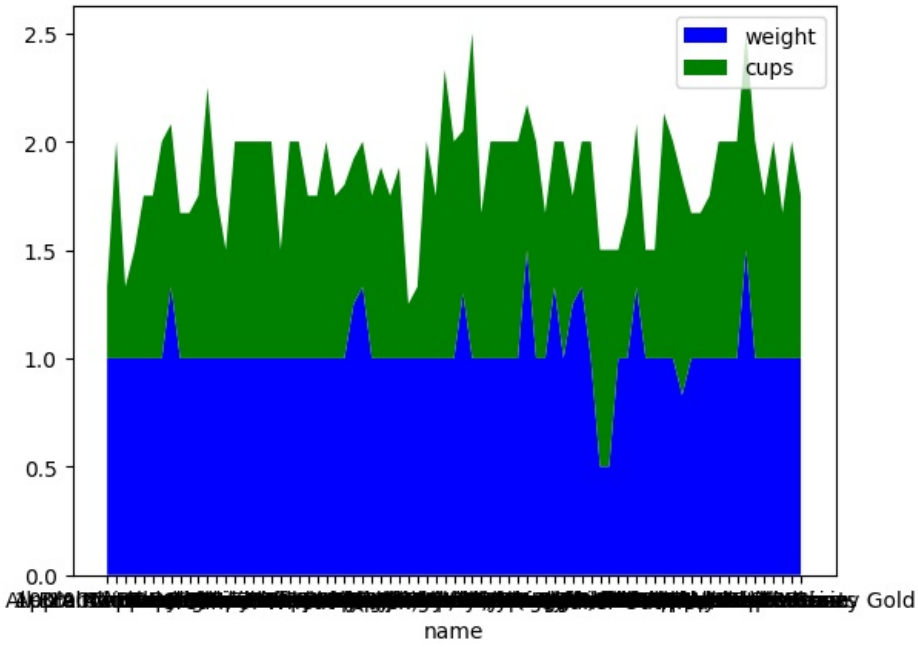
```
True
```

In [203…
```
df = pd.read_csv("E:/Data Analytics Project/New folder/cereal.csv")
#print(df)
gp = df.groupby("name").agg({"calories":"mean","fat":"mean","carbo":"mean","sugars":"mean"})
plt.figure(figsize=(10, 7))
plt.stackplot(df["name"],df["calories"], df["fat"], df["carbo"],df["sugars"],colors = ["pink","hotpink","blue",
plt.xlabel("NAME")
plt.ylabel("Sum of calories,fat,carbo and sugars")
plt.legend(loc=1)
plt.show()
```

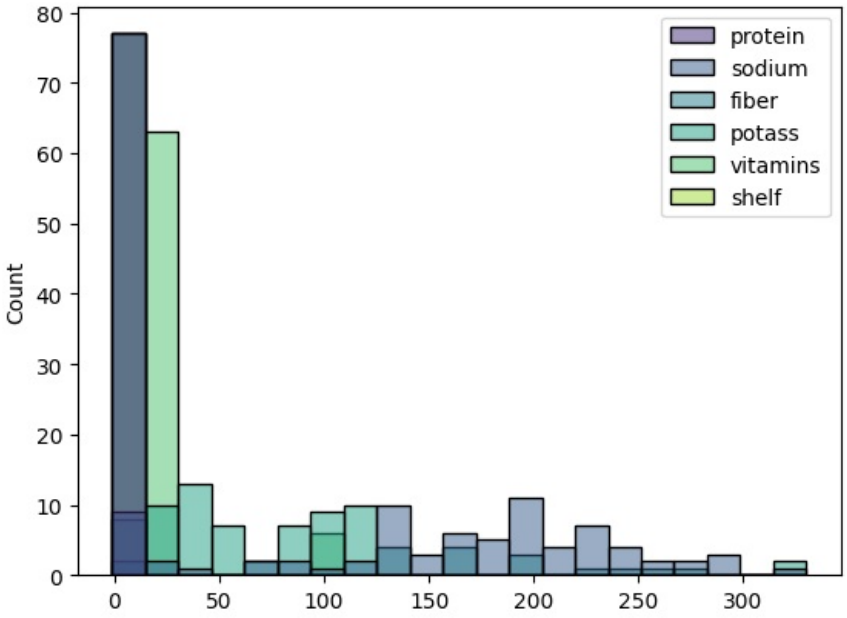


In [173…
```
plt.stackplot(df["name"], df["weight"], df["cups"],colors = ["blue","green"], labels = ["weight","cups"])
plt.xlabel("name")
```

```
plt.legend(loc=1)
plt.show()
```

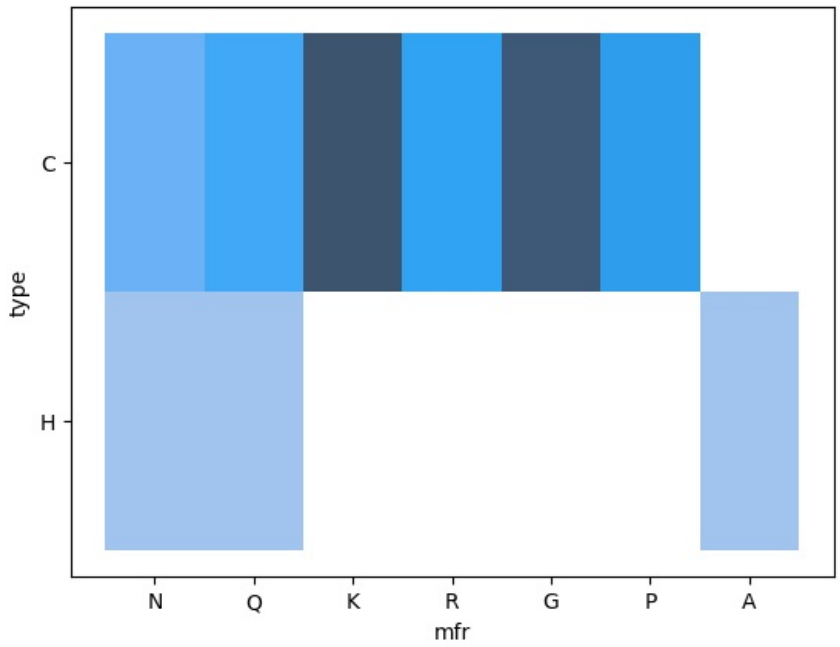


In [110…
```
gb = df.groupby("name").agg({"protein":"sum","sodium":"sum","fiber":"sum","potass":"sum","vitamins":"sum","shel
#print(gb)
#color = ["blue","gray","green","yellow"]
sns.histplot(gb, palette = "viridis")
plt.show()
```



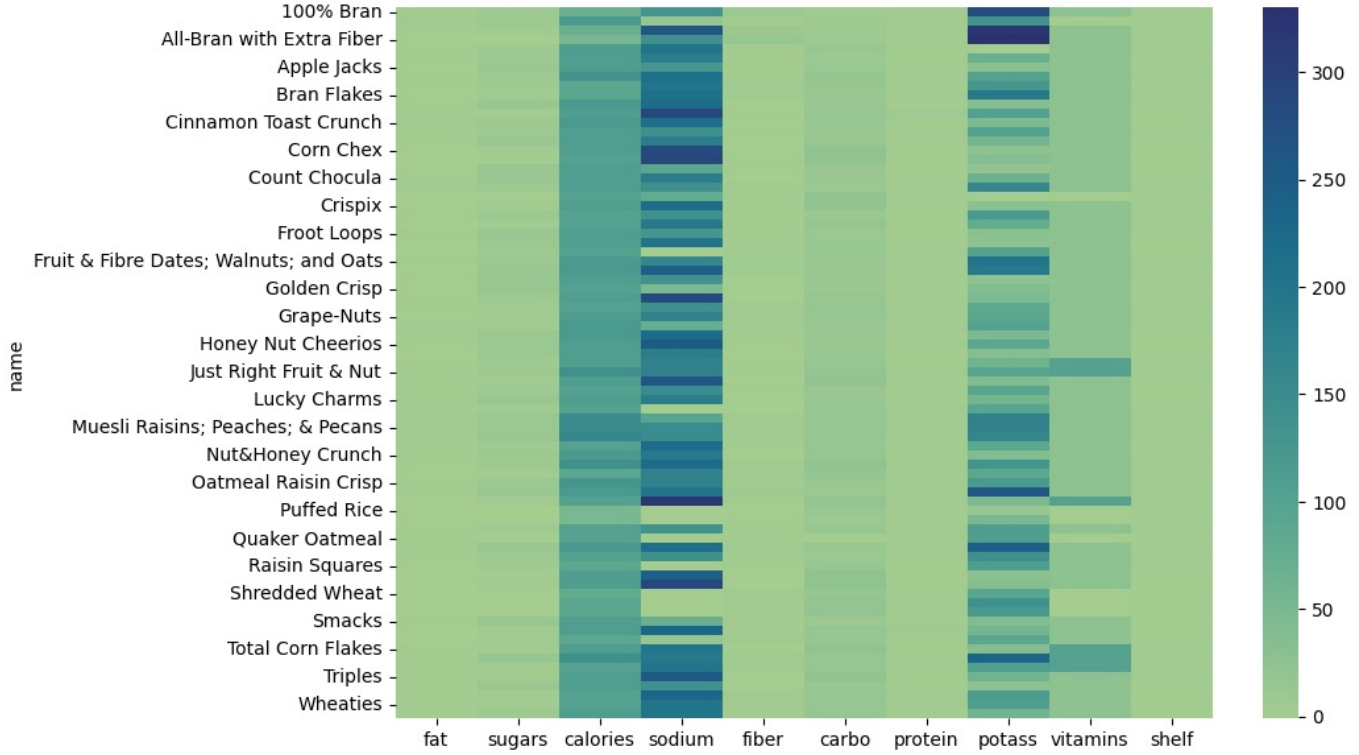
In [142…
```
import matplotlib.pyplot as plt
import seaborn as sns
gb = df.groupby(["mfr","type"]).agg({"name":"count"})
print(gb)
plt.stem(gb)
#sns.histplot(data = data , x = "mfr", y = "type")
plt.show()
```

```
          name
mfr type
A   H        1
G   C       22
K   C       23
N   C        5
    H        1
P   C        9
Q   C        7
    H        1
R   C        8
```



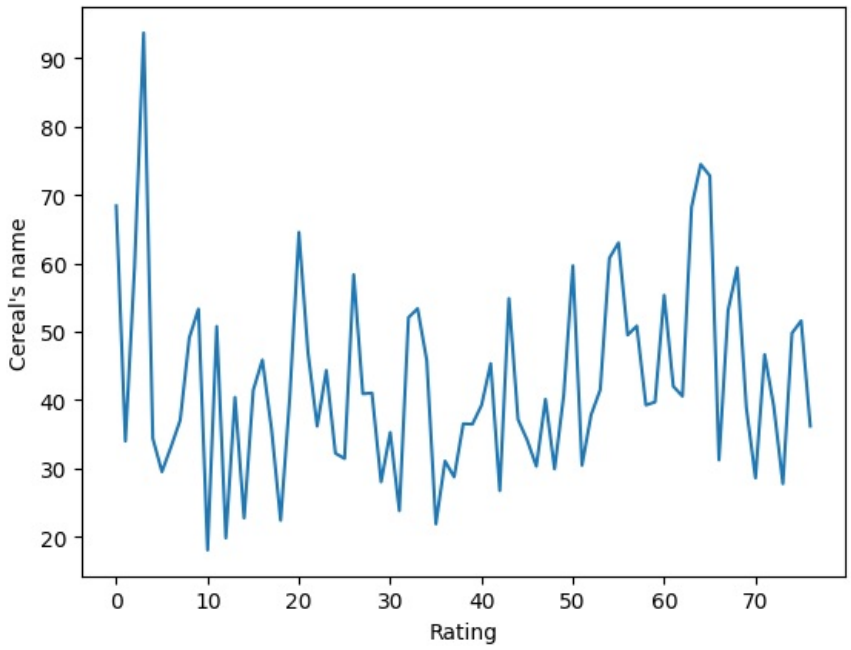
In [196…

```
import seaborn as sns
import pandas as pd

gp = data.groupby("name").agg({"fat":"sum","sugars":"sum","calories":"sum","sodium":"sum","fiber":"sum","carbo"

plt.figure(figsize=(10, 7))
sns.heatmap(gp, cmap="crest")
plt.show()
```



In [214…

```
plt.plot(df["rating"])
plt.xlabel("Rating")
plt.ylabel("Cereal's name")
plt.show()
```



In [ ]: