Cheese is nutritious food made mostly from the milk of cows but also other mammals, including sheep, goats, buffalo, reindeer, camels, and yaks. Around 4000 thousand years ago people have started to breed animals and process their milk. That's when the cheese was born.

Explore this dataset to find out about kinds of cheese from all across the world.

248 Cheese have listed fat content. Is there a relationship between fat content and cheese type?

What about aroma, flavor, or texture?


```
In [1]: import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt
        import numpy as np
        import seaborn as sns
        import warnings
        warnings.filterwarnings('ignore')
        import plotly.express as px
        import warnings
        warnings.filterwarnings("ignore", category = DeprecationWarning)
```

```
In [2]: df=pd.read_csv('cheeses.csv')
```

```
In [3]: df.head(10)
```

Out[3]:

	cheese	url	milk	country	region	family	type	fat_content	calcium_content
0	Aarewasser	https://www.cheese.com/aarewasser/	cow	Switzerland	NaN	NaN	semi-soft	0.00%	0 gm
1	Abbaye de Belloc	https://www.cheese.com/abbaye-de-belloc/	sheep	France	Pays Basque	NaN	semi-hard, artisan	0.00%	0 gm
2	Abbaye de Belval	https://www.cheese.com/abbaye-de-belval/	cow	France	NaN	NaN	semi-hard	40-46%	0 gm
3	Abbaye de Citeaux	https://www.cheese.com/abbaye-de-citeaux/	cow	France	Burgundy	NaN	semi-soft, artisan, brined	0.00%	0 gm
4	Abbaye de Tamié	https://www.cheese.com/tamie/	cow	France	Savoie	NaN	soft, artisan	0.00%	0 gm
5	Abbaye de Timadeuc	https://www.cheese.com/abbaye-de-timadeuc/	cow	France	province of Brittany	NaN	semi-hard	0.00%	0 gm
6	Abbaye du Mont des Cats	https://www.cheese.com/abbaye-du-mont-des-cats/	cow	France	Nord-Pas-de-Calais	NaN	semi-soft, artisan, brined	50%	0 gm
7	Abbot's Gold	https://www.cheese.com/abbots-gold/	cow	England, Great Britain, United Kingdom	North Yorkshire	Cheddar	semi-hard	0.00%	0 gm
8	Abertam	https://www.cheese.com/abertam/	sheep	Czech Republic	Karlovy Vary	NaN	hard, artisan	45%	0 gm
9	Abondance	https://www.cheese.com/abondance/	cow	France	NaN	NaN	semi-hard, artisan	0.00%	0 gm

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1187 entries, 0 to 1186
Data columns (total 19 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   cheese           1187 non-null   object
 1   url              1187 non-null   object
 2   milk             1151 non-null   object
 3   country          1176 non-null   object
 4   region           854 non-null    object
 5   family           489 non-null    object
 6   type             1174 non-null   object
 7   fat_content      1187 non-null   object
 8   calcium_content  1187 non-null   object
 9   texture          1129 non-null   object
 10  rind             945 non-null    object
 11  color            1045 non-null   object
 12  flavor           1089 non-null   object
 13  aroma            929 non-null    object
 14  vegetarian       748 non-null    object
 15  vegan            748 non-null    object
 16  synonyms         294 non-null    object
 17  alt_spellings    109 non-null    object
 18  producers        787 non-null    object
dtypes: object(19)
memory usage: 176.3+ KB
```

In [5]:
```
df.columns
```

Out[5]:
```
Index(['cheese', 'url', 'milk', 'country', 'region', 'family', 'type',
       'fat_content', 'calcium_content', 'texture', 'rind', 'color', 'flavor',
       'aroma', 'vegetarian', 'vegan', 'synonyms', 'alt_spellings',
       'producers'],
      dtype='object')
```

In [6]:
```
df.describe(include='all').T
```

Out[6]:

	count	unique	top	freq
cheese	1187	1187	Aarewasser	1
url	1187	1187	https://www.cheese.com/aarewasser/	1
milk	1151	21	cow	696
country	1176	82	United States	305
region	854	347	Wisconsin	67
family	489	21	Blue	94
type	1174	84	semi-hard, artisan	133
fat_content	1187	86	0.00%	939
calcium_content	1187	25	0 gm	1162
texture	1129	309	creamy	162
rind	945	12	natural	439

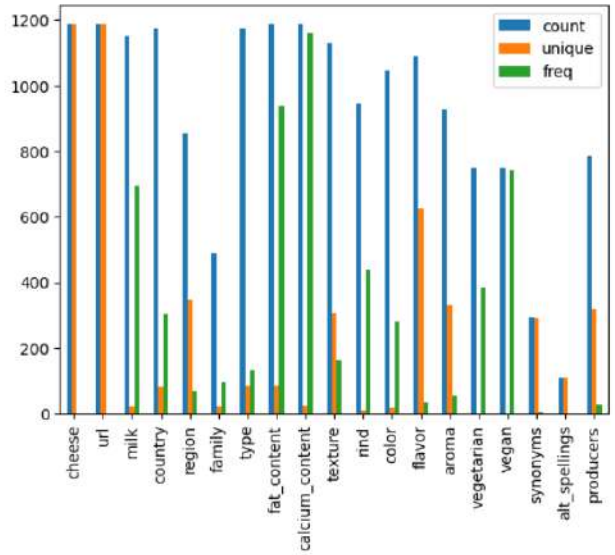
calcium_content	1187	25		0 gm	1162
click to scroll output; double click to hide				creamy	162
rind	945	12		natural	439
color	1045	17		white	281
flavor	1089	626		creamy	34
aroma	929	330		rich	56
vegetarian	748	2		False	386
vegan	748	2		False	742
synonyms	294	292		Rupp Vorarlberger Bergkäse	3
alt_spellings	109	109		Tamié, Trappiste de Tamie, Abbey of Tamie	1
producers	787	318		Sartori	27

In [7]:
```
df.describe().T
```

Out[7]:

	count	unique	top	freq
cheese	1187	1187	Aarewasser	1
url	1187	1187	https://www.cheese.com/aarewasser/	1
milk	1151	21	cow	696
country	1176	82	United States	305
region	854	347	Wisconsin	67
family	489	21	Blue	94
type	1174	84	semi-hard, artisan	133
fat_content	1187	86	0.00%	939
calcium_content	1187	25	0 gm	1162
texture	1129	309	creamy	162
rind	945	12	natural	439
color	1045	17	white	281
flavor	1089	626	creamy	34
aroma	929	330	rich	56
vegetarian	748	2	False	386
vegan	748	2	False	742
synonyms	294	292	Rupp Vorarlberger Bergkäse	3
alt_spellings	109	109	Tamié, Trappiste de Tamie, Abbey of Tamie	1
producers	787	318	Sartori	27

```
In [8]: df.describe().T.plot(kind='bar')
Out[8]: <Axes: >
```




```
In [9]: df
Out[9]:
```

	cheese	url	milk	country	region	family	type	fat_content	calcium_content	texture	rind	color	f
0	Aarewasser	https://www.cheese.com/aarewasser/	cow	Switzerland	NaN	NaN	semi-soft	0.00%	0 gm	buttery	washed	yellow	
1	Abbaye de Belloc	https://www.cheese.com/abbaye-de-belloc/	sheep	France	Pays Basque	NaN	semi-hard, artisan	0.00%	0 gm	creamy, dense, firm	natural	yellow	ca
2	Abbaye de Belval	https://www.cheese.com/abbaye-de-belval/	cow	France	NaN	NaN	semi-hard	40-46%	0 gm	elastic	washed	ivory	
3	Abbaye de Citeaux	https://www.cheese.com/abbaye-de-citeaux/	cow	France	Burgundy	NaN	semi-soft, artisan, brined	0.00%	0 gm	creamy, dense, smooth	washed	white	sr
4	Abbaye de Tamié	https://www.cheese.com/tamie/	cow	France	Savoie	NaN	soft, artisan	0.00%	0 gm	creamy, open, smooth	washed	white	
...	...	...	...	...	...	...	...	...	...	...	...	...	
1182	Sveciaost	https://www.cheese.com/sveciaost/	cow	Sweden	Low-laying regions	NaN	semi-hard, brined	45%	0 gm	creamy, supple	rindless	pale yellow	
1183	Swag	https://www.cheese.com/swag/	goat	Australia	South Australia	NaN	fresh firm, artisan	0.00%	0 gm	creamy, crumbly	ash coated	white	cr
1184	Swaledale	https://www.cheese.com/swaledale/	sheep	England	Swaledale, North Yorkshire	NaN	hard	0.00%	0 gm	semi firm	NaN	yellow	sr
1185	Sweet Style Swiss	https://www.cheese.com/sweet-style-swiss/	NaN	Switzerland	NaN	NaN	semi-hard, artisan	0.00%	0 gm	firm, supple	waxed	NaN	

1184	Swaledale	https://www.cheese.com/swaledale/	sheep	England	Swaledale, North Yorkshire	NaN	hard	0.00%	0 gm	semi firm	NaN	yellow	sr
1185	Sweet Style Swiss	https://www.cheese.com/sweet-style-swiss/	NaN	Switzerland	NaN	NaN	semi-hard, artisan	0.00%	0 gm	firm, supple	waxed	NaN	
1186	Swiss cheese	https://www.cheese.com/swiss/	cow	United States	NaN	Swiss Cheese	hard, artisan, processed	7.8 g/100g	0 gm	firm	rindless	pale yellow	

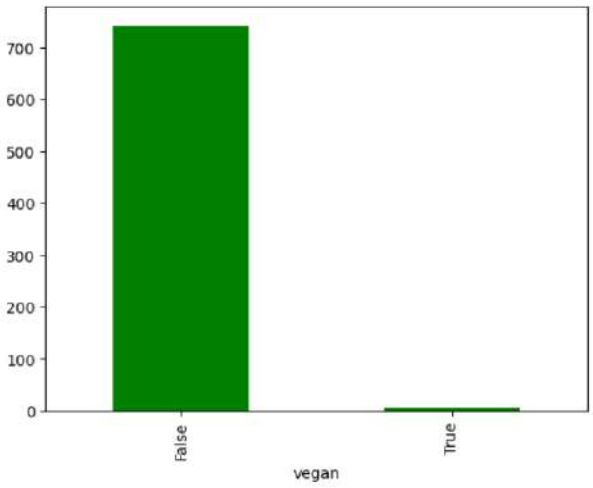
1187 rows × 19 columns

In [10]:
```
df['vegan'].value_counts()
```

Out[10]:
```
vegan
False    742
True       6
Name: count, dtype: int64
```

In [11]:
```
df['vegan'].value_counts().plot(kind='bar',color='g')
```

Out[11]:
```
<Axes: xlabel='vegan'>
```



In [12]:
```
df.columns.to_list()
```

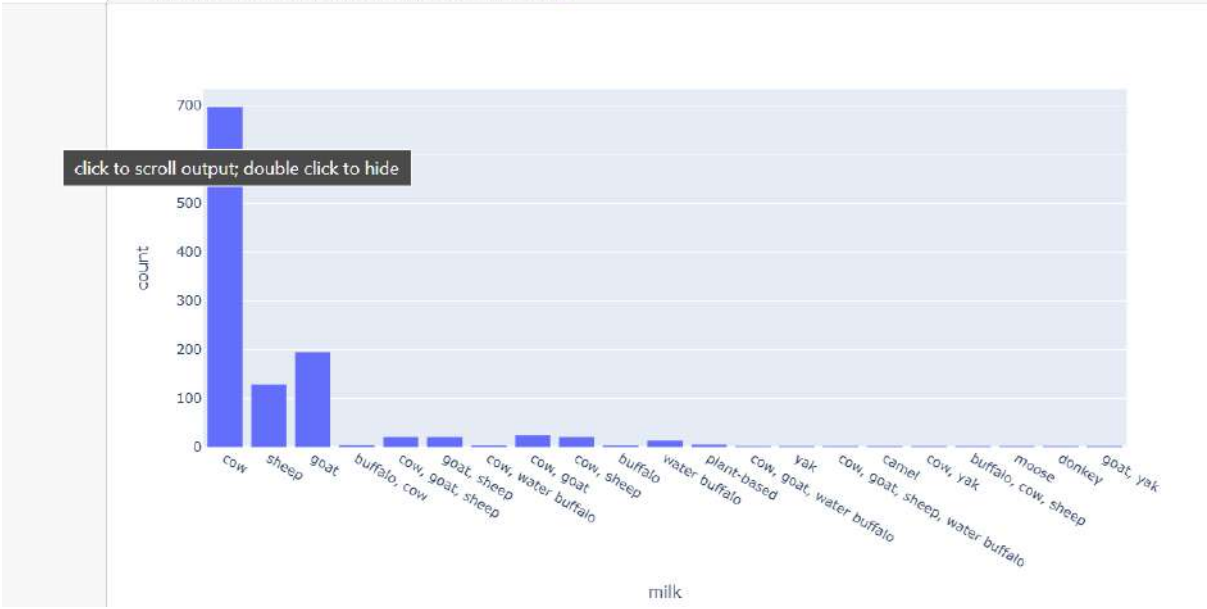
Out[12]:
```
['cheese',
 'url',
 'milk',
 'country',
 'region',
 'family',
 'type',
 'fat_content',
 'calcium_content',
 'texture',
 'rind',
 'color',
 'flavor',
 'aroma',
 'vegetarian',
 'vegan',
 'synonyms',
 'alt_spellings',
 'producers']
```

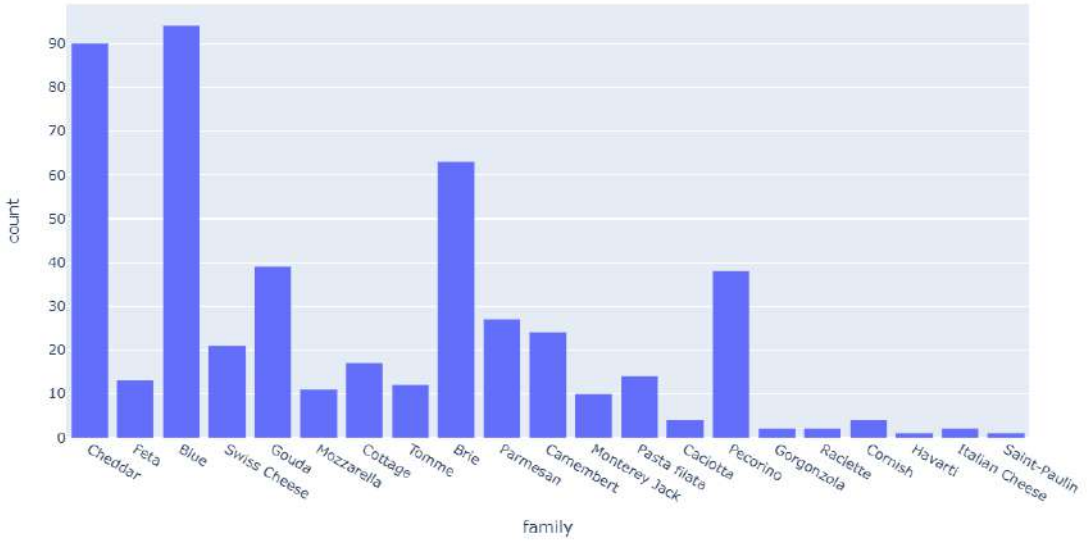
In [13]:
```
df['calcium (mg/100g)']=df['calcium_content'].apply(
    func=lambda x: x if isinstance(x, float) else int(x.split()[0]))

df.head()
```

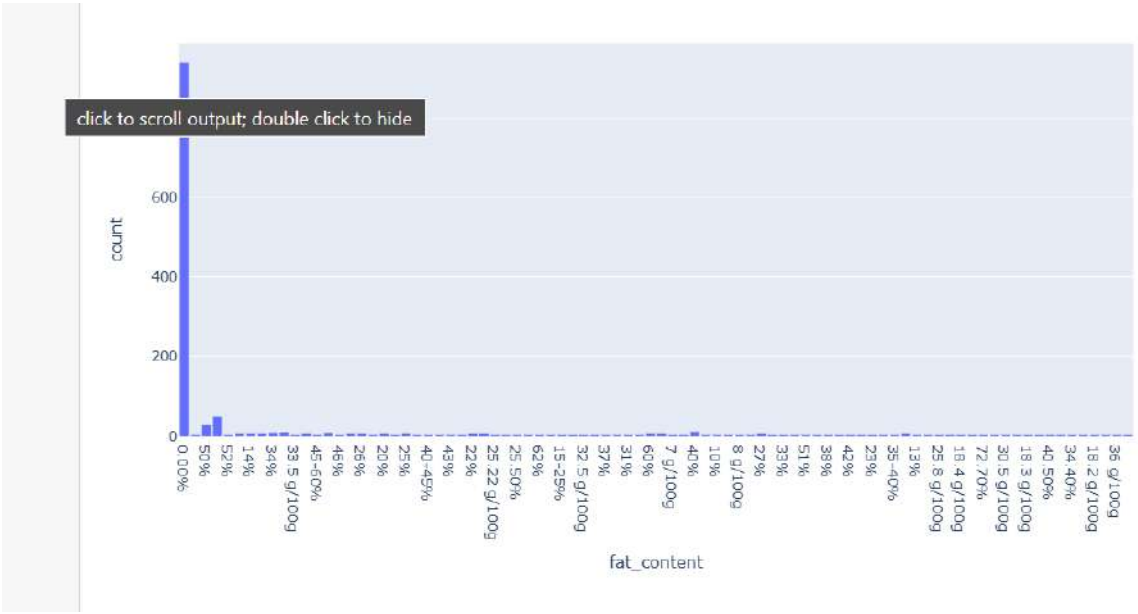
Out[13]:

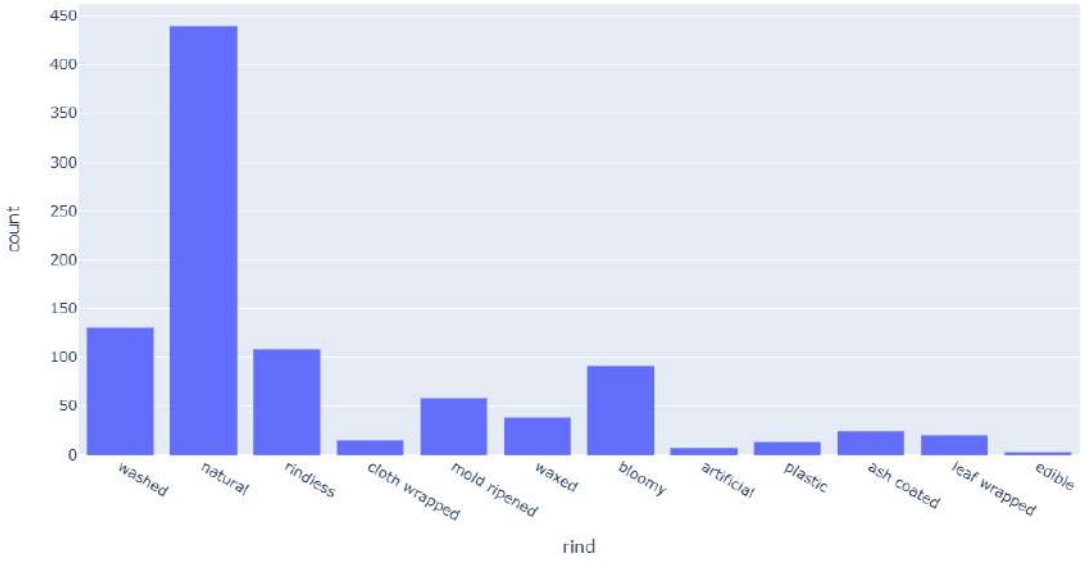
	cheese	url	milk	country	region	family	type	fat_content	calcium_content	texture	rind	color	flavor
0	Aarewasser	https://www.cheese.com/aarewasser/	cow	Switzerland	NaN	NaN	semi-soft	0.00%	0 gm	buttery	washed	yellow	sweet
1	Abbaye de Belloc	https://www.cheese.com/abbaye-de-belloc/	sheep	France	Pays Basque	NaN	semi-hard, artisan	0.00%	0 gm	creamy, dense, firm	natural	yellow	burnt caramel
2	Abbaye de Belval	https://www.cheese.com/abbaye-de-belval/	cow	France	NaN	NaN	semi-hard	40-46%	0 gm	elastic	washed	ivory	NaN
3	Abbaye de Citeaux	https://www.cheese.com/abbaye-de-citeaux/	cow	France	Burgundy	NaN	semi-soft, artisan, brined	0.00%	0 gm	creamy, dense, smooth	washed	white	acidic, milky, smooth
4	Abbaye de Tamié	https://www.cheese.com/tamie/	cow	France	Savoie	NaN	soft, artisan	0.00%	0 gm	creamy, open, smooth	washed	white	fruity, nutty

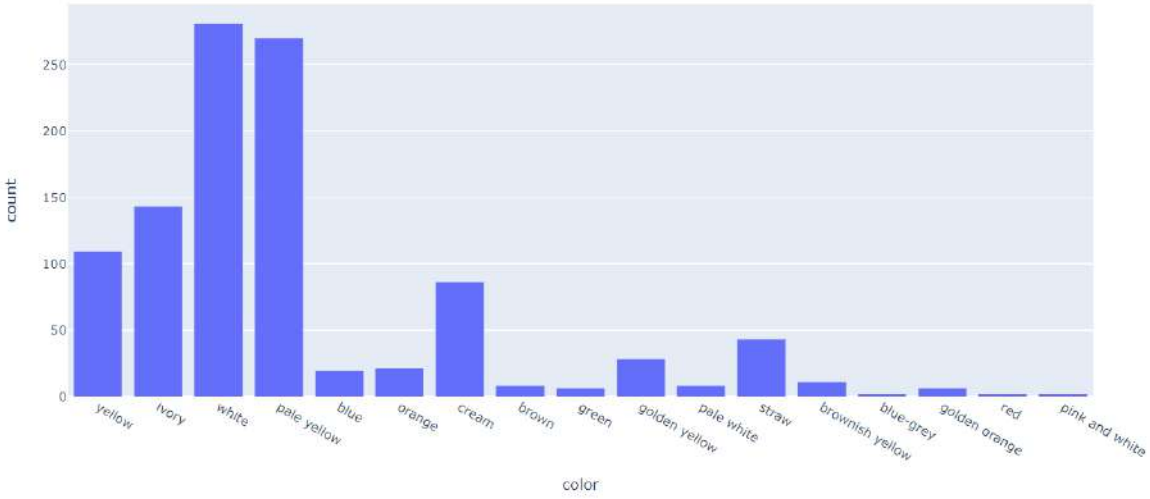
```
In [14]: from plotly import express

for column in ['milk','family','fat_content','rind','color','vegetarian','vegan']:
    express.histogram(data_frame=df, x=column).show()
```

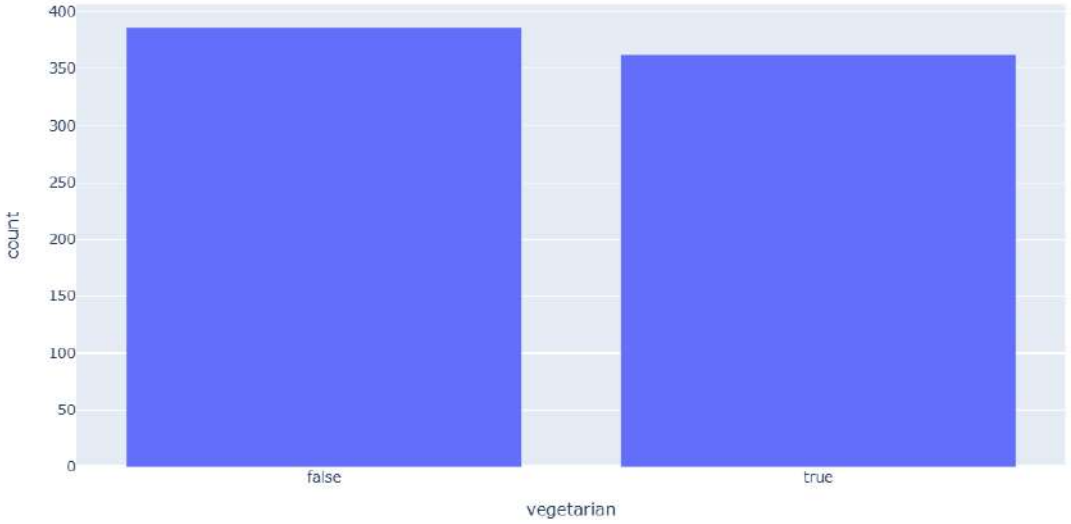


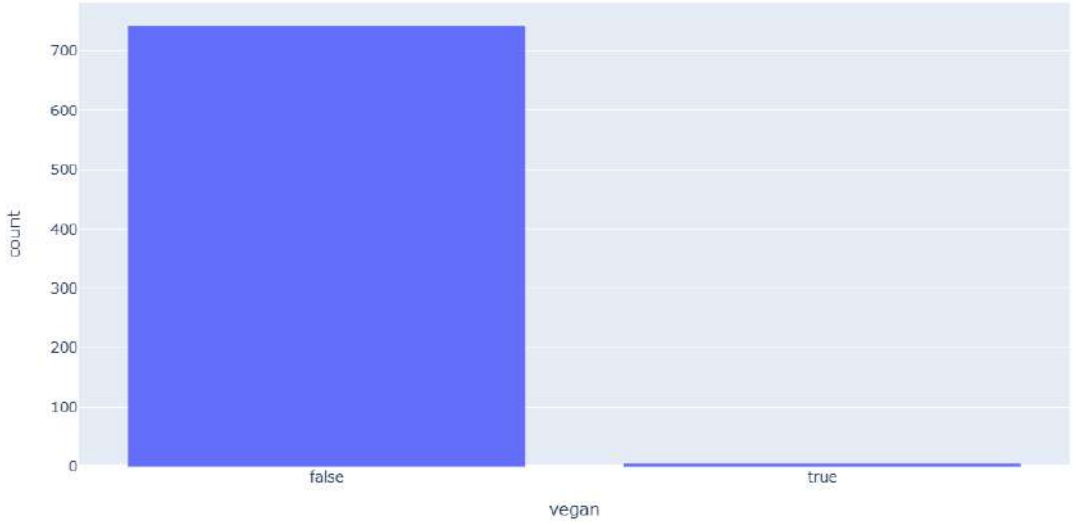



click to scroll output; double click to hide
count
600
400
200
0
0.00%
50%
52%
14%
34%
33.5 g/100g
45-60%
46%
26%
20%
25%
40-45%
43%
22%
25.22 g/100g
25.50%
62%
15-25%
32.5 g/100g
37%
31%
60%
7 g/100g
40%
10%
8 g/100g
27%
33%
51%
38%
42%
23%
35-40%
13%
25.8 g/100g
18.4 g/100g
72.70%
30.5 g/100g
18.3 g/100g
40.50%
34.40%
18.2 g/100g
36 g/100g
fat_content


count
450
400
350
300
250
200
150
100
50
0
washed
natural
rindless
cloth wrapped
mold ripened
waxed
bloomy
artificial
plastic
ash coated
leaf wrapped
edible
rind


count
250
200
150
100
50
0
yellow
ivory
white
pale yellow
blue
orange
cream
brown
green
golden yellow
pale white
straw
brownish yellow
blue-grey
golden orange
red
pink and white
color


count
400
350
300
250
200
150
100
50
0
false
true
vegetarian


count
700
600
500
400
300
200
100
0
false
true
vegan

In [15]:
```
numeric_df = df.select_dtypes(include='number')

corr_matrix = numeric_df.corr()

sns.heatmap(corr_matrix, annot=True, cmap='coolwarm')

plt.show()
```