

```
In [1]: import matplotlib.pyplot as plt
```

```
In [2]: from mlxtend.frequent_patterns import apriori, association_rules
```

```
In [3]: import pandas as pd
```

```
In [4]: from wordcloud import WordCloud
```

```
In [5]: import seaborn as sns
```

```
In [6]: import squarify
```

```
In [7]: import networkx as nx
import numpy as np
```

C:\Users\samir\Anaconda3\lib\site-packages\networkx\classes\reportviews.py:95: DeprecationWarning: Using or importing the ABCs from 'collections' instead of from 'collections.abc' is deprecated, and in 3.8 it will stop working  
from collections import Mapping, Set, Iterable

```
In [8]: from mlxtend.preprocessing import TransactionEncoder
```

```
In [9]: dataset = pd.read_csv("C:/Abhay/Market_Basket.csv - Market_Basket.csv", header = None)
```

```
In [10]: print(dataset.head())
```

	0	1	2	3	4	\
0	shrimp	almonds	avocado	vegetables mix	green grapes	
1	burgers	meatballs	eggs	NaN	NaN	
2	chutney	NaN	NaN	NaN	NaN	
3	turkey	avocado	NaN	NaN	NaN	
4	mineral water	milk	energy bar	whole wheat rice	green tea	

	5	6	7	8	9	\
0	whole weat flour	yams	cottage cheese	energy drink	tomato juice	
1	NaN	NaN	NaN	NaN	NaN	
2	NaN	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	

	10	11	12	13	14	15	\
0	low fat yogurt	green tea	honey	salad	mineral water	salmon	
1	NaN	NaN	NaN	NaN	NaN	NaN	
2	NaN	NaN	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	NaN	

	16	17	18	19
0	antioxydant juice	frozen smoothie	spinach	olive oil
1	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN

```
In [11]: print(dataset.shape)
```

(7501, 20)

```
In [12]: print(dataset.info)
```

```

<bound method DataFrame.info of                                     0                1                2
3  \
0      shrimp      almonds      avocado      vegetables mix
1      burgers      meatballs      eggs      NaN
2      chutney      NaN      NaN      NaN
3      turkey      avocado      NaN      NaN
4  mineral water      milk      energy bar      whole wheat rice
...      ...      ...      ...      ...
7496      butter      light mayo      fresh bread      NaN
7497      burgers      frozen vegetables      eggs      french fries
7498      chicken      NaN      NaN      NaN
7499      escalope      green tea      NaN      NaN
7500      eggs      frozen smoothie      yogurt cake      low fat yogurt

              4              5              6              7              8  \
0  green grapes  whole weat flour  yams  cottage cheese  energy drink
1      NaN      NaN      NaN      NaN      NaN
2      NaN      NaN      NaN      NaN      NaN
3      NaN      NaN      NaN      NaN      NaN
4      green tea      NaN      NaN      NaN      NaN
...      ...      ...      ...      ...      ...
7496      NaN      NaN      NaN      NaN      NaN
7497      magazines      green tea      NaN      NaN      NaN
7498      NaN      NaN      NaN      NaN      NaN
7499      NaN      NaN      NaN      NaN      NaN
7500      NaN      NaN      NaN      NaN      NaN

              9              10              11              12              13              14  \
0  tomato juice  low fat yogurt  green tea  honey  salad  mineral water
1      NaN      NaN      NaN      NaN      NaN      NaN
2      NaN      NaN      NaN      NaN      NaN      NaN
3      NaN      NaN      NaN      NaN      NaN      NaN
4      NaN      NaN      NaN      NaN      NaN      NaN
...      ...      ...      ...      ...      ...      ...
7496      NaN      NaN      NaN      NaN      NaN      NaN
7497      NaN      NaN      NaN      NaN      NaN      NaN
7498      NaN      NaN      NaN      NaN      NaN      NaN
7499      NaN      NaN      NaN      NaN      NaN      NaN
7500      NaN      NaN      NaN      NaN      NaN      NaN

              15              16              17              18              19
0  salmon  antioxydant juice  frozen smoothie  spinach  olive oil
1      NaN      NaN      NaN      NaN      NaN
2      NaN      NaN      NaN      NaN      NaN
3      NaN      NaN      NaN      NaN      NaN
4      NaN      NaN      NaN      NaN      NaN
...      ...      ...      ...      ...      ...
7496      NaN      NaN      NaN      NaN      NaN
7497      NaN      NaN      NaN      NaN      NaN
7498      NaN      NaN      NaN      NaN      NaN
7499      NaN      NaN      NaN      NaN      NaN
7500      NaN      NaN      NaN      NaN      NaN

```

```
[7501 rows x 20 columns]>
```

```
In [13]: print(dataset.describe())
```

	0	1	2	3	4	\
count	7501	5747	4389	3345	2529	
unique	115	117	115	114	110	
top	mineral water	mineral water	mineral water	mineral water	green tea	
freq	577	484	375	201	153	

	5	6	7	8	9	\
count	1864	1369	981	654	395	
unique	106	102	97	88	80	
top	french fries	green tea	green tea	green tea	green tea	
freq	107	96	67	57	31	

	10	11	12	13	14	\
count	256	154	87	47	25	
unique	66	50	43	28	19	
top	low fat yogurt	green tea	green tea	green tea	magazines	
freq	22	15	8	4	3	

	15	16	17	18	19
count	8	4	4	3	1
unique	8	3	3	3	1
top	frozen smoothie	frozen smoothie	protein bar	spinach	olive oil
freq	1	2	2	1	1

```
In [14]: plt.rcParams['figure.figsize'] = (14, 14)
```

```
In [15]: wordcloud = WordCloud(background_color = 'white', width = 1200, height = 1200, max>
```

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]



```
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)  
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: DeprecationWarning: fontsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use  
bboxsize or getbbox() instead.  
    box_size = draw.textsize(word, font=transposed_font)
```

[illegible]

[illegible]

[illegible]



[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

[illegible]



[illegible]

[illegible]

[illegible]

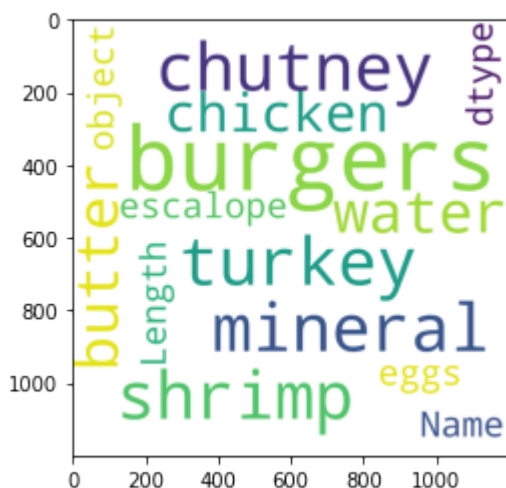
```

C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: Deprecation
Warning: textsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use
textbbox or textlength instead.
    box_size = draw.textsize(word, font=transposed_font)
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: Deprecation
Warning: textsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use
textbbox or textlength instead.
    box_size = draw.textsize(word, font=transposed_font)
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: Deprecation
Warning: textsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use
textbbox or textlength instead.
    box_size = draw.textsize(word, font=transposed_font)
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: Deprecation
Warning: textsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use
textbbox or textlength instead.
    box_size = draw.textsize(word, font=transposed_font)
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:519: Deprecation
Warning: ROTATE_90 is deprecated and will be removed in Pillow 10 (2023-07-01). Us
e Transpose.ROTATE_90 instead.
    orientation = (Image.ROTATE_90 if orientation is None else
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: Deprecation
Warning: textsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use
textbbox or textlength instead.
    box_size = draw.textsize(word, font=transposed_font)
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:499: Deprecation
Warning: ROTATE_90 is deprecated and will be removed in Pillow 10 (2023-07-01). Us
e Transpose.ROTATE_90 instead.
    orientation = Image.ROTATE_90
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: Deprecation
Warning: textsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use
textbbox or textlength instead.
    box_size = draw.textsize(word, font=transposed_font)
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:499: Deprecation
Warning: ROTATE_90 is deprecated and will be removed in Pillow 10 (2023-07-01). Us
e Transpose.ROTATE_90 instead.
    orientation = Image.ROTATE_90
C:\Users\samir\Anaconda3\lib\site-packages\wordcloud\wordcloud.py:508: Deprecation
Warning: textsize is deprecated and will be removed in Pillow 10 (2023-07-01). Use
textbbox or textlength instead.
    box_size = draw.textsize(word, font=transposed_font)

```

```
In [16]: plt.imshow(wordcloud)
```

```
Out[16]: <matplotlib.image.AxesImage at 0x21aa56882b0>
```



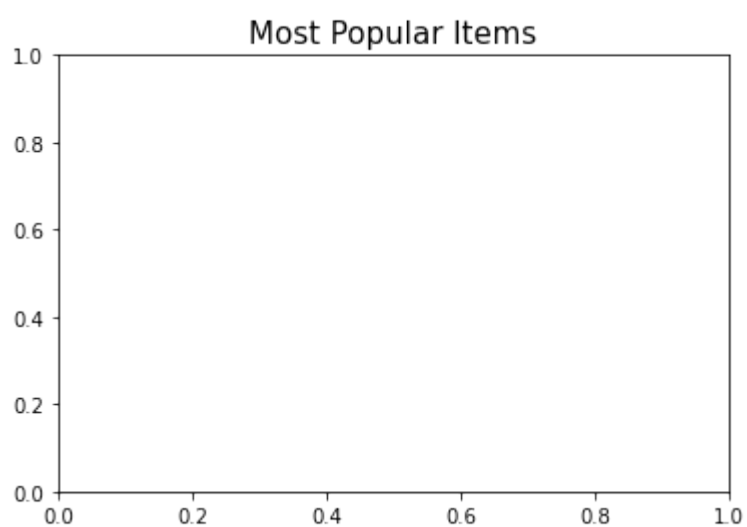
```
In [17]: plt.axis('off')
```



```
Out[17]: (0.0, 1.0, 0.0, 1.0)
```

```
In [18]: plt.title('Most Popular Items',fontsize = 15)
```

```
Out[18]: Text(0.5, 1.0, 'Most Popular Items')
```



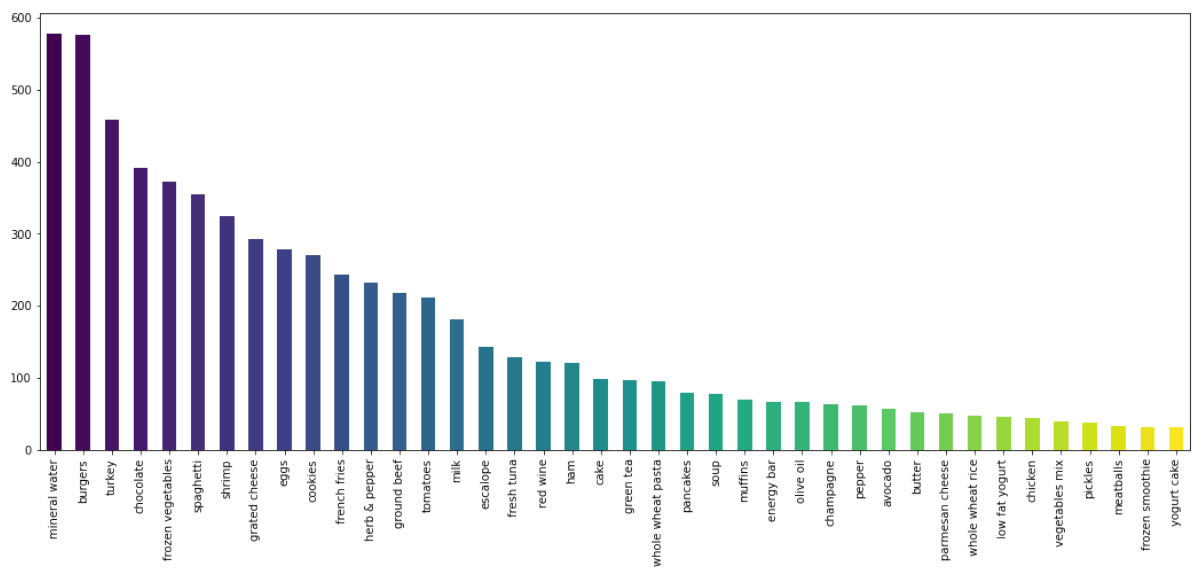
```
In [19]: plt.show()
```

```
In [20]: plt.rcParams['figure.figsize'] = (18, 7)
```

```
In [21]: color = plt.cm.viridis(np.linspace(0, 1, 40))
```

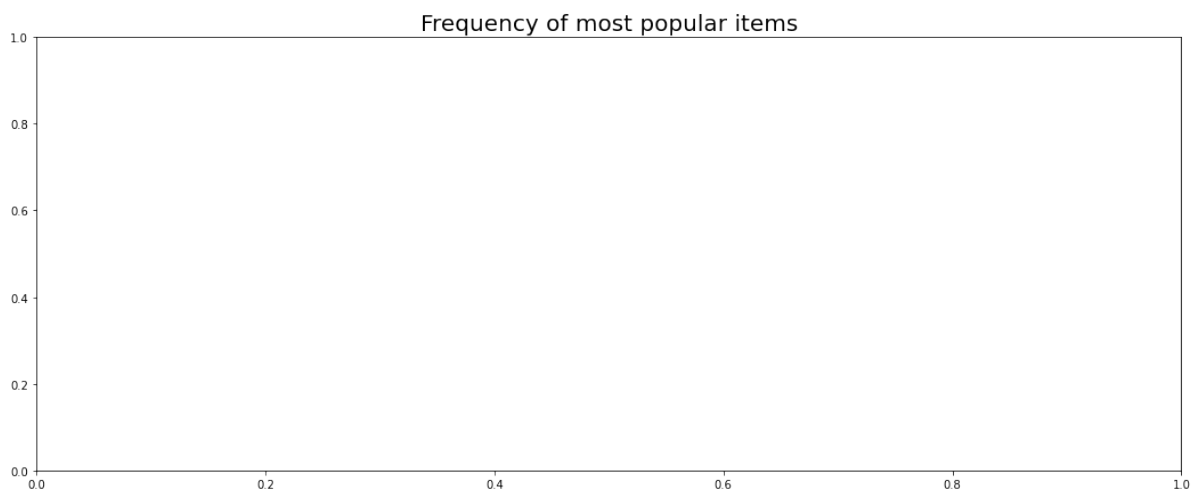
```
In [22]: dataset[0].value_counts().head(40).plot.bar(color = color)
```

```
Out[22]: <AxesSubplot:>
```



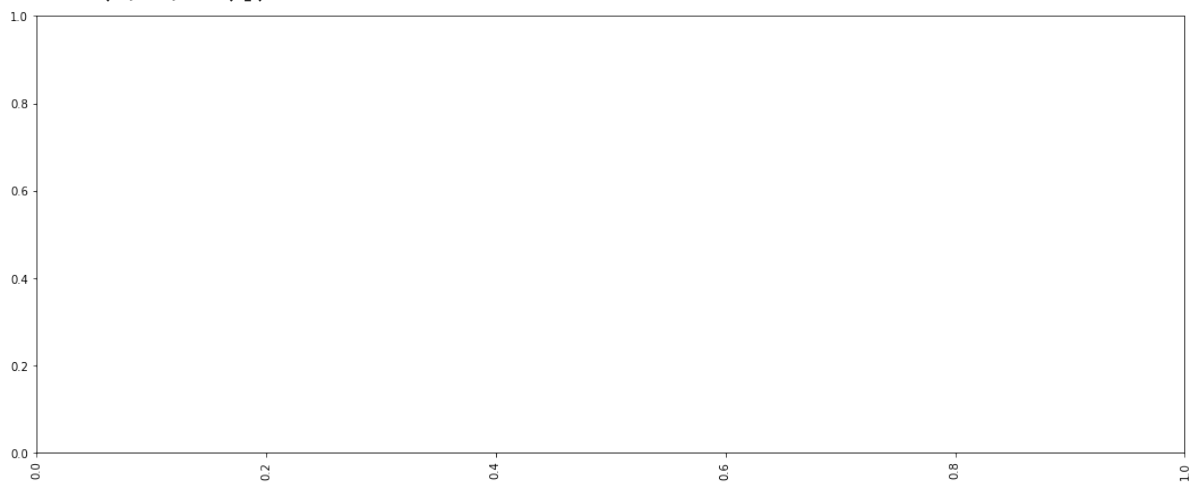
```
In [23]: plt.title('Frequency of most popular items', fontsize = 20)
```

```
Out[23]: Text(0.5, 1.0, 'Frequency of most popular items')
```

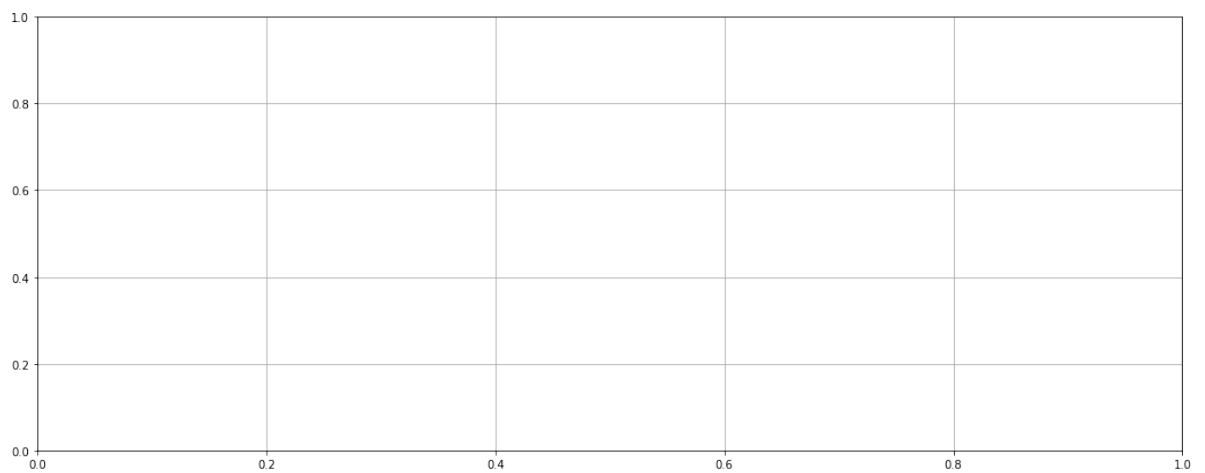


```
In [24]: plt.xticks(rotation = 90 )
```

```
Out[24]: (array([0. , 0.2, 0.4, 0.6, 0.8, 1. ]),
 [Text(0, 0, ''),
  Text(0, 0, ''),
  Text(0, 0, ''),
  Text(0, 0, ''),
  Text(0, 0, ''),
  Text(0, 0, '')])
```



```
In [25]: plt.grid()
```



```
In [26]: plt.show()
```

```
In [27]: transactions = []
for i in range(0, 7501):
    transactions.append([str(dataset.values[i,j]) for j in range(0, 20)])

transac = TransactionEncoder()
dataset = transac.fit_transform(transactions)
print(dataset)

[[ True  True False ...  True False False]
 [False False False ... False False False]
 [False False False ... False False False]
 ...
 [False False False ... False False False]
 [False False False ... False False False]
 [False False False ... False  True False]]
```

```
In [28]: data = pd.DataFrame(dataset, columns= transac.columns_)
print(data.head())
```

	almonds	antioxydant juice	asparagus	avocado	babies food	bacon	\
0	True	True	False	True	False	False	
1	False	False	False	False	False	False	
2	False	False	False	False	False	False	
3	False	False	False	True	False	False	
4	False	False	False	False	False	False	

	barbecue sauce	black tea	blueberries	body spray	...	turkey	\
0	False	False	False	False	...	False	
1	False	False	False	False	...	False	
2	False	False	False	False	...	False	
3	False	False	False	False	...	True	
4	False	False	False	False	...	False	

	vegetables mix	water spray	white wine	whole weat flour	\
0	True	False	False	True	
1	False	False	False	False	
2	False	False	False	False	
3	False	False	False	False	
4	False	False	False	False	

	whole wheat pasta	whole wheat rice	yams	yogurt cake	zucchini
0	False	False	True	False	False
1	False	False	False	False	False
2	False	False	False	False	False
3	False	False	False	False	False
4	False	True	False	False	False

[5 rows x 120 columns]

```
In [29]: frequent_itemsets = apriori(data, min_support=0.003, use_colnames=True)
frequent_itemsets['length'] = frequent_itemsets['itemsets'].apply(lambda x : len(x))
print(frequent_itemsets.head())
```

	support	itemsets	length
0	0.020397	(almonds)	1
1	0.008932	(antioxydant juice)	1
2	0.004799	(asparagus)	1
3	0.033329	(avocado)	1
4	0.004533	(babies food)	1

```
In [30]: print(frequent_itemsets[frequent_itemsets['length'] >= 3].head(10))
```

	support	itemsets	length
1017	0.005199	(almonds, nan, burgers)	3
1018	0.003066	(cake, almonds, nan)	3
1019	0.005999	(chocolate, almonds, nan)	3
1020	0.006532	(eggs, almonds, nan)	3
1021	0.004399	(french fries, almonds, nan)	3
1022	0.003066	(almonds, frozen vegetables, nan)	3
1023	0.004933	(green tea, almonds, nan)	3
1024	0.003866	(ground beef, almonds, nan)	3
1025	0.005199	(milk, almonds, nan)	3
1026	0.007466	(mineral water, almonds, nan)	3

```
In [31]: rules = association_rules(frequent_itemsets, metric='lift', min_threshold=1)
```

```
In [32]: print(rules.head(20))
```



	antecedents	consequents	antecedent support \
0	(almonds)	(burgers)	0.020397
1	(burgers)	(almonds)	0.087188
2	(cake)	(almonds)	0.081056
3	(almonds)	(cake)	0.020397
4	(chocolate)	(almonds)	0.163845
5	(almonds)	(chocolate)	0.020397
6	(eggs)	(almonds)	0.179709
7	(almonds)	(eggs)	0.020397
8	(french fries)	(almonds)	0.170911
9	(almonds)	(french fries)	0.020397
10	(almonds)	(frozen vegetables)	0.020397
11	(frozen vegetables)	(almonds)	0.095321
12	(green tea)	(almonds)	0.132116
13	(almonds)	(green tea)	0.020397
14	(ground beef)	(almonds)	0.098254
15	(almonds)	(ground beef)	0.020397
16	(almonds)	(milk)	0.020397
17	(milk)	(almonds)	0.129583
18	(mineral water)	(almonds)	0.238368
19	(almonds)	(mineral water)	0.020397

	consequent support	support	confidence	lift	leverage	conviction
0	0.087188	0.005199	0.254902	2.923577	0.003421	1.225089
1	0.020397	0.005199	0.059633	2.923577	0.003421	1.041724
2	0.020397	0.003066	0.037829	1.854607	0.001413	1.018117
3	0.081056	0.003066	0.150327	1.854607	0.001413	1.081527
4	0.020397	0.005999	0.036615	1.795099	0.002657	1.016834
5	0.163845	0.005999	0.294118	1.795099	0.002657	1.184553
6	0.020397	0.006532	0.036350	1.782108	0.002867	1.016555
7	0.179709	0.006532	0.320261	1.782108	0.002867	1.206774
8	0.020397	0.004399	0.025741	1.261983	0.000913	1.005485
9	0.170911	0.004399	0.215686	1.261983	0.000913	1.057089
10	0.095321	0.003066	0.150327	1.577065	0.001122	1.064738
11	0.020397	0.003066	0.032168	1.577065	0.001122	1.012162
12	0.020397	0.005066	0.038345	1.879913	0.002371	1.018663
13	0.132116	0.005066	0.248366	1.879913	0.002371	1.154663
14	0.020397	0.003866	0.039349	1.929116	0.001862	1.019728
15	0.098254	0.003866	0.189542	1.929116	0.001862	1.112639
16	0.129583	0.005199	0.254902	1.967098	0.002556	1.168192
17	0.020397	0.005199	0.040123	1.967098	0.002556	1.020551
18	0.020397	0.007599	0.031879	1.562914	0.002737	1.011860
19	0.238368	0.007599	0.372549	1.562914	0.002737	1.213851

In [33]: `print(rules.shape)`

(17372, 9)

In [34]: `print(rules[(rules['lift'] >= 6) & (rules['confidence'] >= 0.4)])`

	antecedents	consequents \
8190	(whole wheat pasta, mineral water)	(olive oil)
15897	(whole wheat pasta, mineral water, nan)	(olive oil)
15900	(whole wheat pasta, mineral water)	(olive oil, nan)

	antecedent support	consequent support	support	confidence	lift \
8190	0.009599	0.065858	0.003866	0.402778	6.115863
15897	0.009599	0.065858	0.003866	0.402778	6.115863
15900	0.009599	0.065725	0.003866	0.402778	6.128268

	leverage	conviction
8190	0.003234	1.564145
15897	0.003234	1.564145
15900	0.003235	1.564368

