NHANES_FPED_Component_Stats

April 16, 2021

Objectives

Provide statistical distribution plots of FPED components, for comparison on how they are distributed among the seafood vs non-seafood meals.

Applied Data Filters

The dataframe included in this analysis contains the following modifications of the original data set:

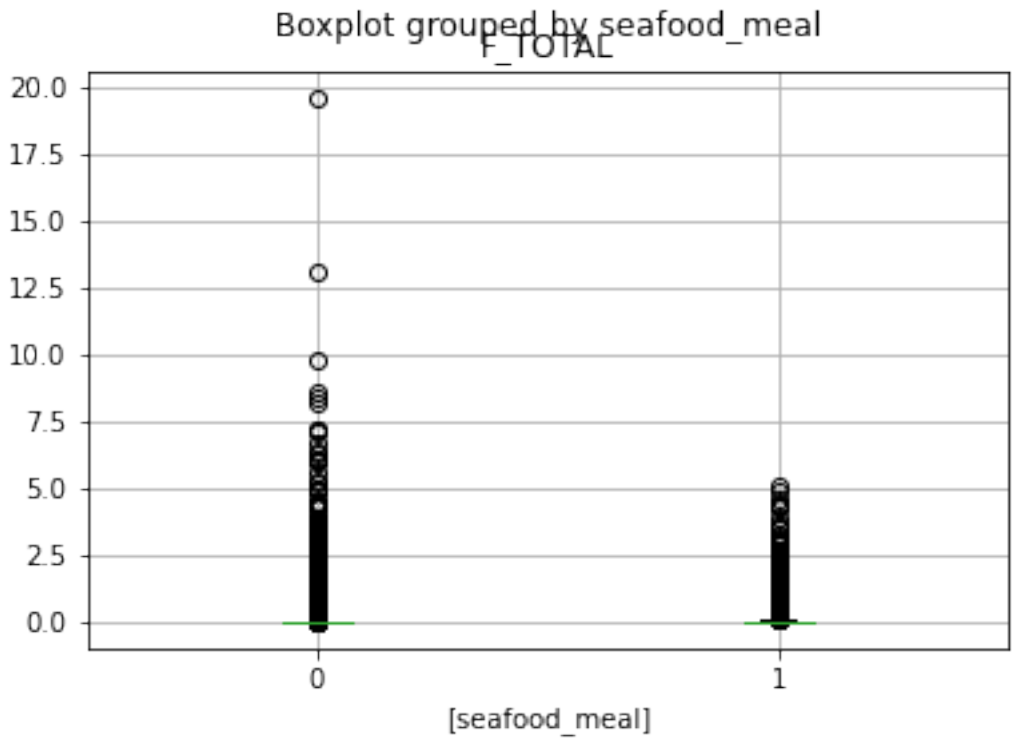
1. Meal level aggregation
2. Meals that are only lunch or dinner
3. Meals that have both seafood and meat, where there is a grey area in the ratio between the two, are dropped
4. Meals that are more than 0 KCAL
5. Meals of participants older than 18 years of age
6. Meals that are consumed at home
7. Meals that are non-vegeterian

Section 1: Fruits

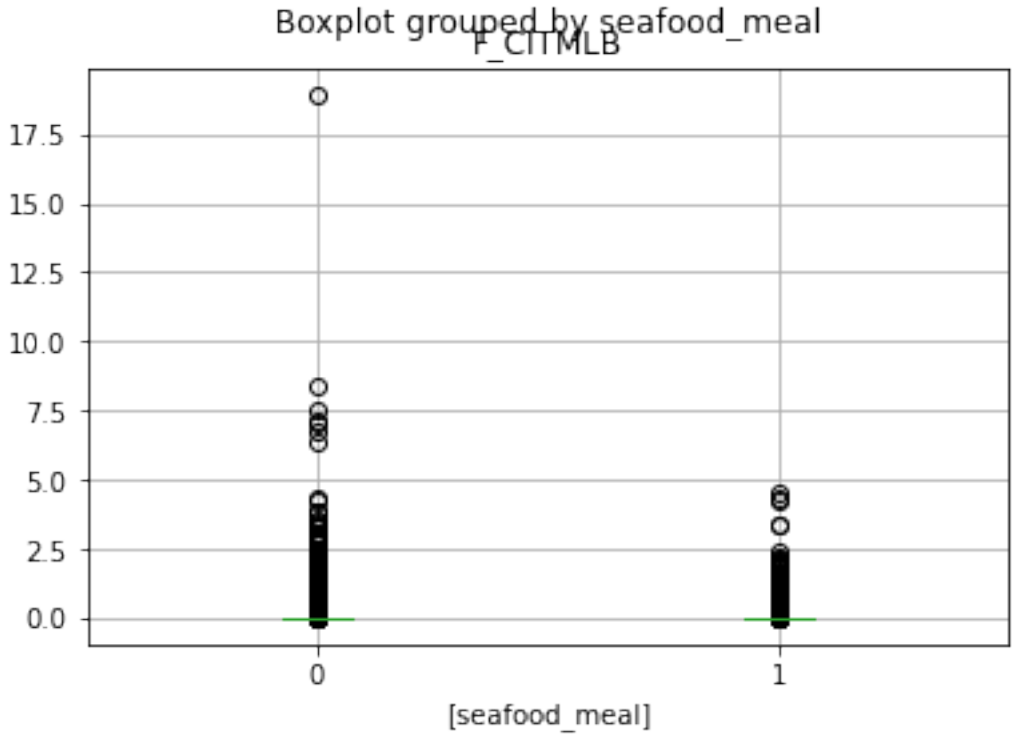
This section provides boxplots and density plots of the Fruit FPED components in the seafood meal and non seafood meal groups. The code for seafood meal is 1 if meal contains seafood, and 0 if meal does not contain seafood.

[1]:
```
import pandas as pd

#Read data frame and add plant pf total variable
df = pd.read_csv('../../Data/nhanes_full_pre_proc.csv')
```

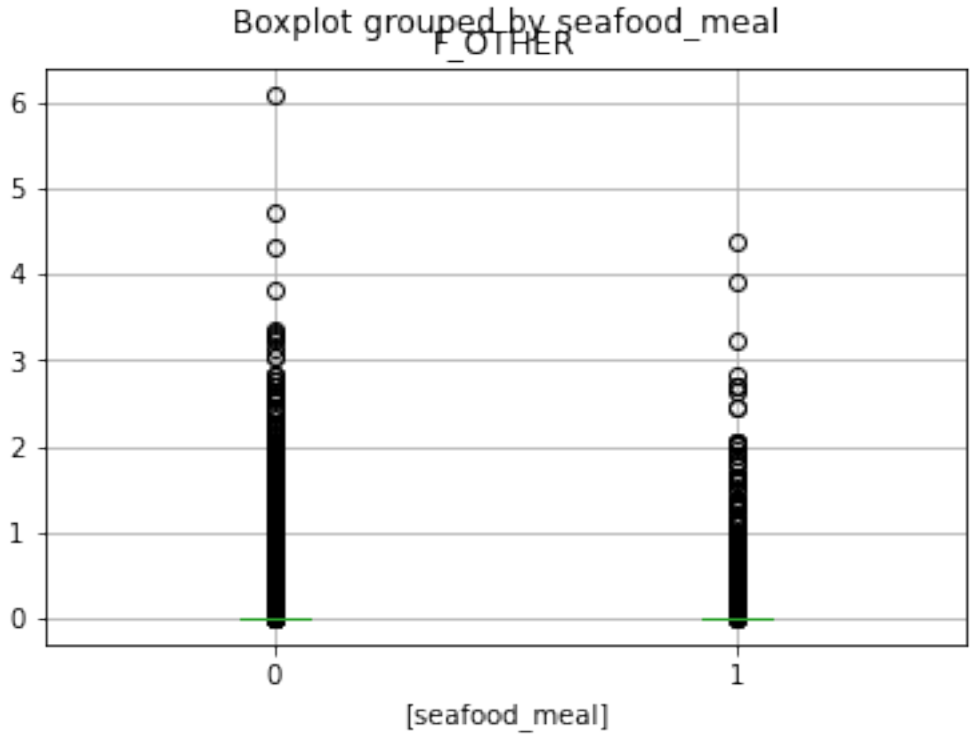
[2]:
```
import matplotlib.pyplot as plt

fruits = ['F_TOTAL', 'F_CITMLB', 'F_OTHER', 'F_JUICE']

for var in fruits:
    z = df.boxplot(column=var,by=['seafood_meal'])
    plt.show(z)
    plt.clf()
```



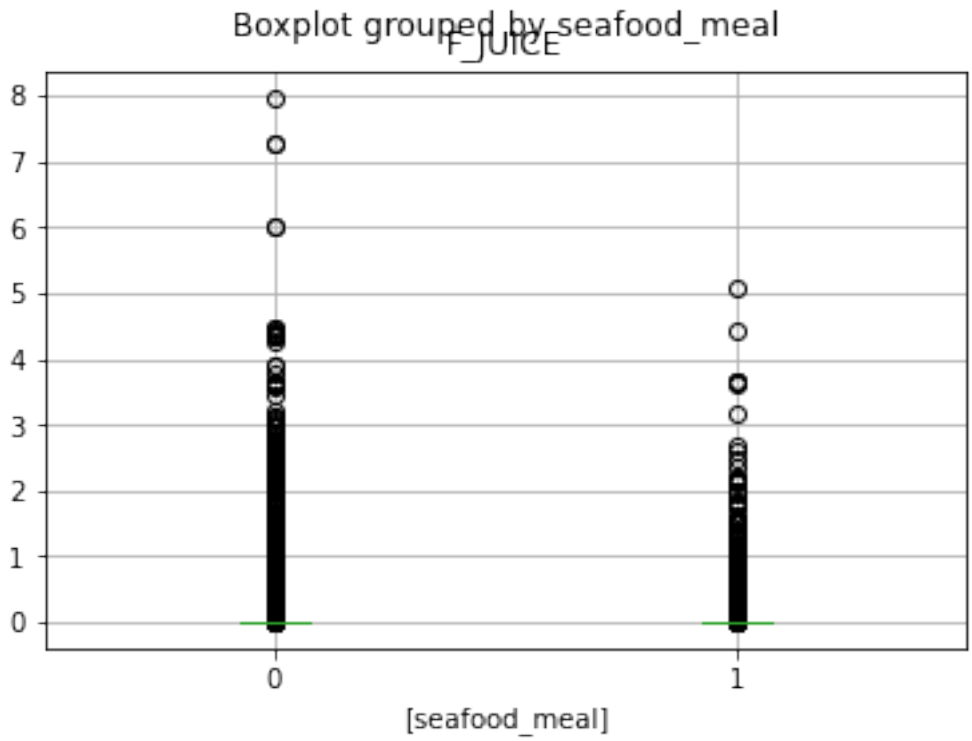

```
<Figure size 432x288 with 0 Axes>
```



```
<Figure size 432x288 with 0 Axes>
```

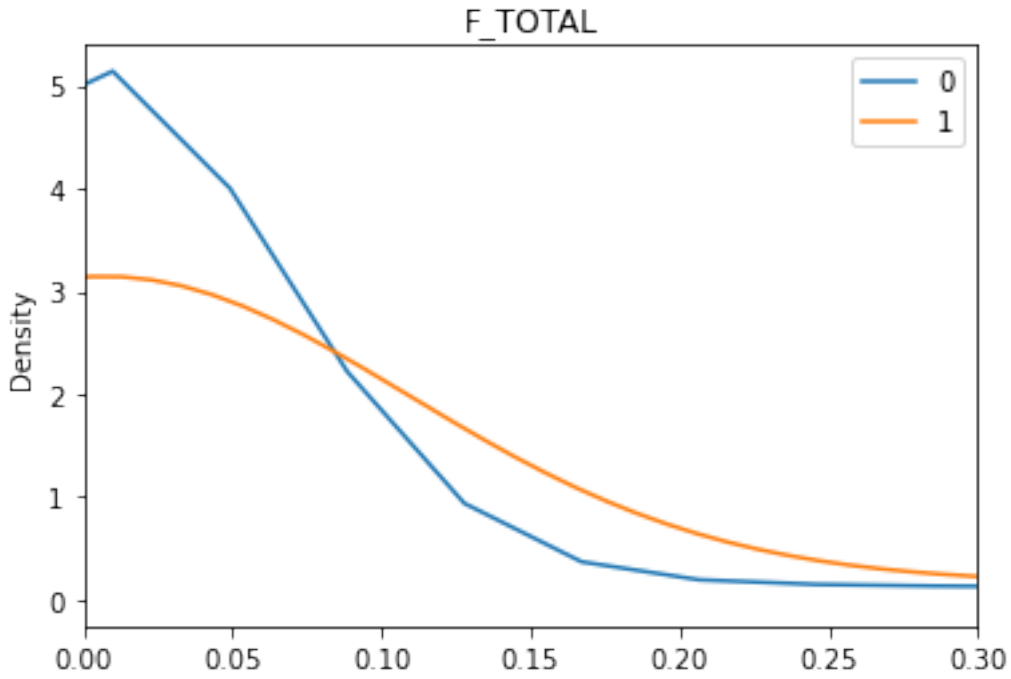


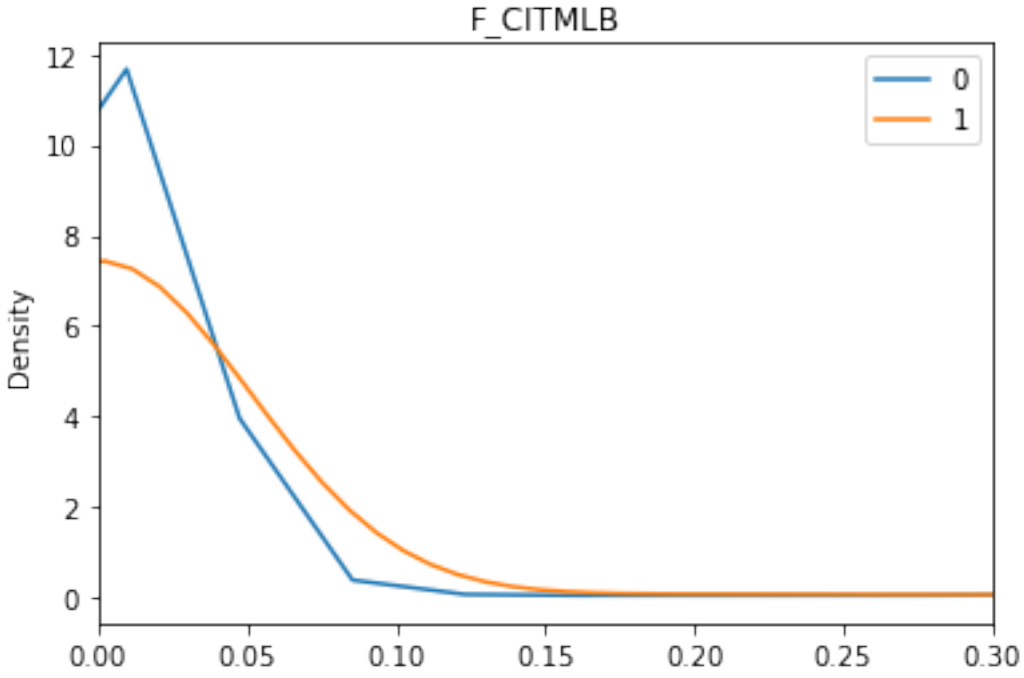

```
<Figure size 432x288 with 0 Axes>
```

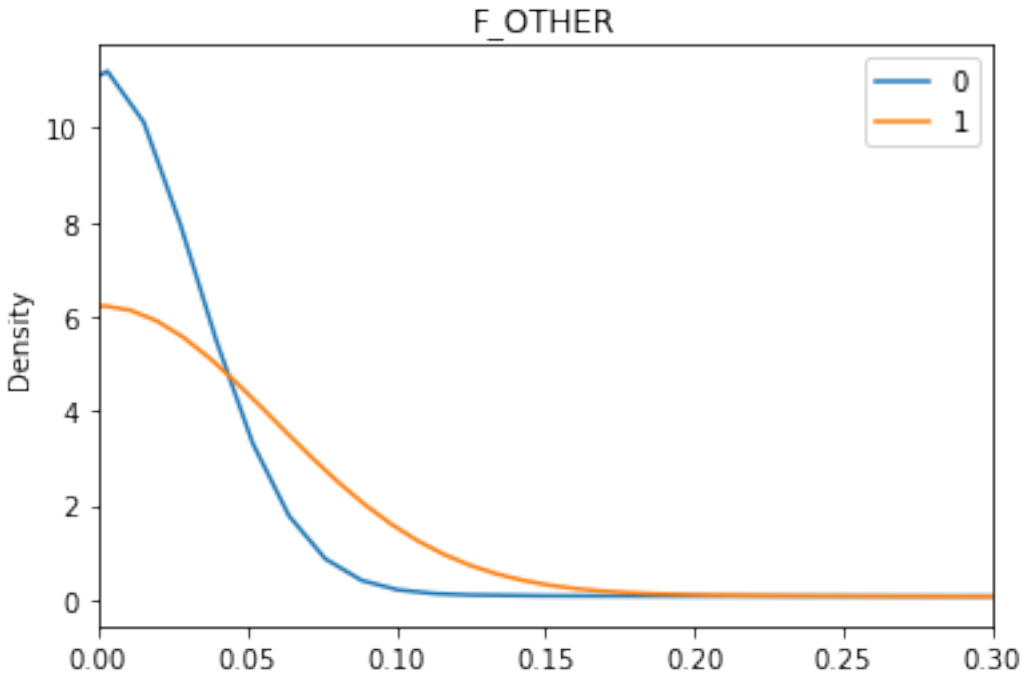


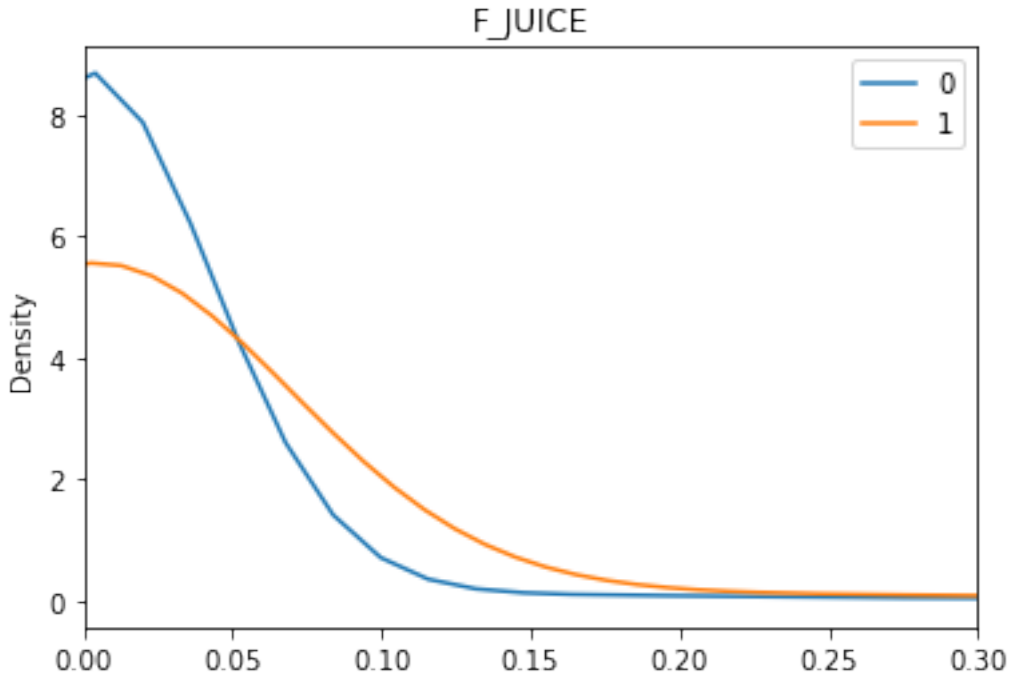

```
<Figure size 432x288 with 0 Axes>
```

[3]:
```
for var in fruits:
    z = df.groupby('seafood_meal')[var].plot.kde(title = var, legend='x')
    plt.show(z[0].set_xlim(0, 0.3))
    plt.clf()
```


F_TOTAL
Density
0
1
0
1
2
3
4
5
0.00
0.05
0.10
0.15
0.20
0.25
0.30


F_CITMLB
Density
0
1
0
2
4
6
8
10
12
0.00
0.05
0.10
0.15
0.20
0.25
0.30






```
<Figure size 432x288 with 0 Axes>
```

[4]:
```
for var in fruits:
    z = df.groupby('seafood_meal')[var].describe()
    print("Statistics for "+var+'\n')
    print(z)
    print('\n')
```

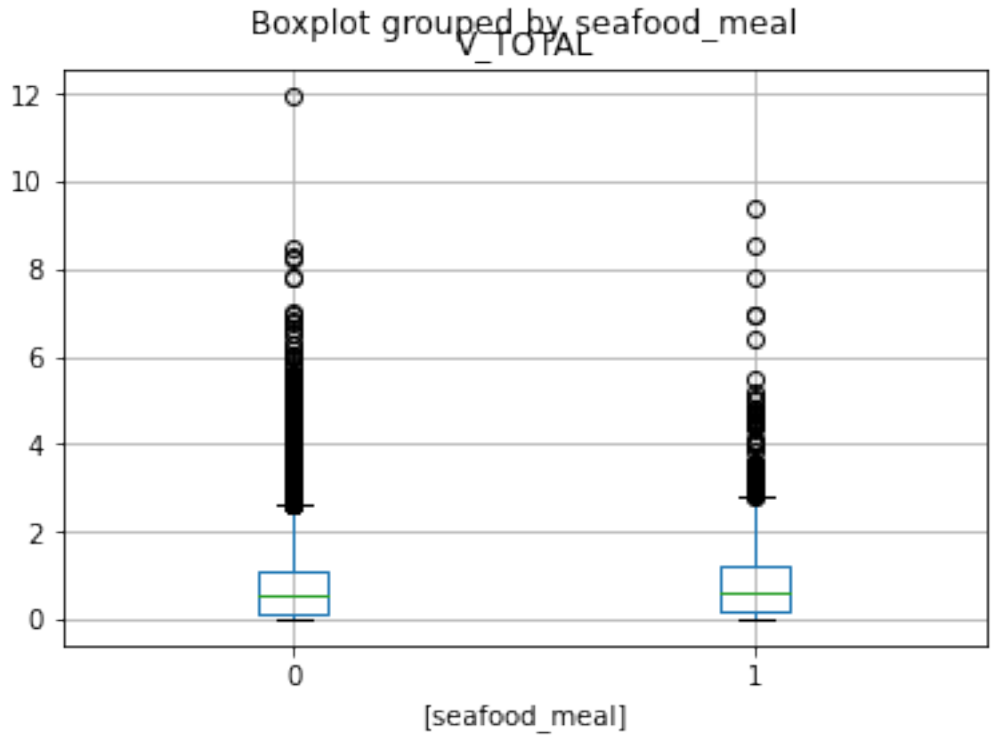
```
Statistics for F_TOTAL

                count      mean       std  min  25%  50%   75%    max
seafood_meal
0             26011.0  0.148838  0.491428  0.0  0.0  0.0  0.00  19.64
1              3232.0  0.187642  0.513770  0.0  0.0  0.0  0.03   5.08


Statistics for F_CITMLB

                count      mean       std  min  25%  50%  75%    max
seafood_meal
0             26011.0  0.029057  0.239037  0.0  0.0  0.0  0.0  18.94
1              3232.0  0.043815  0.255454  0.0  0.0  0.0  0.0   4.55


Statistics for F_OTHER

                count      mean       std  min  25%  50%  75%   max
seafood_meal
0             26011.0  0.054642  0.249359  0.0  0.0  0.0  0.0  6.09
1              3232.0  0.069378  0.293768  0.0  0.0  0.0  0.0  4.37


Statistics for F_JUICE

                count      mean       std  min  25%  50%  75%   max
seafood_meal
0             26011.0  0.065132  0.322374  0.0  0.0  0.0  0.0  7.97
1              3232.0  0.074459  0.324533  0.0  0.0  0.0  0.0  5.08
```

Section 2: Vegetables

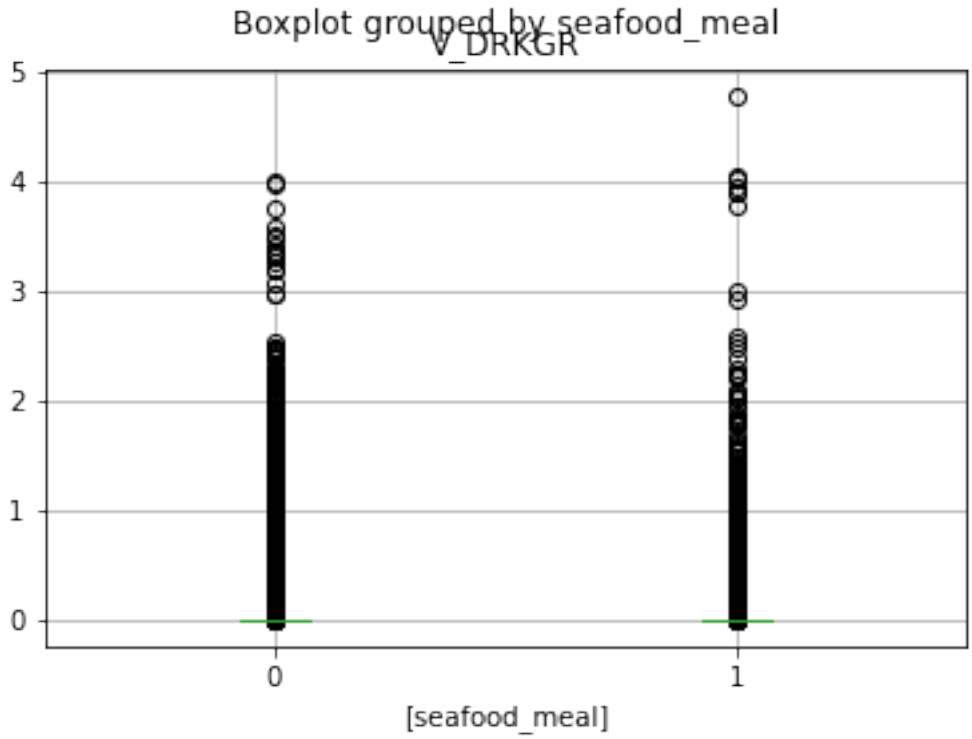
This section provides boxplots and density plots of the Vegetable FPED components in the seafood meal and non seafood meal groups. The code for seafood meal is 1 if meal contains seafood, and 0 if meal does not contain seafood.

[5]:
```
veggie = ['V_TOTAL', 'V_DRKGR', 'V_REDOR_TOMATO', 'V_REDOR_OTHER',␣
 ↪'V_STARCHY_POTATO',
                   'V_STARCHY_OTHER', 'V_OTHER', 'V_LEGUMES']
```

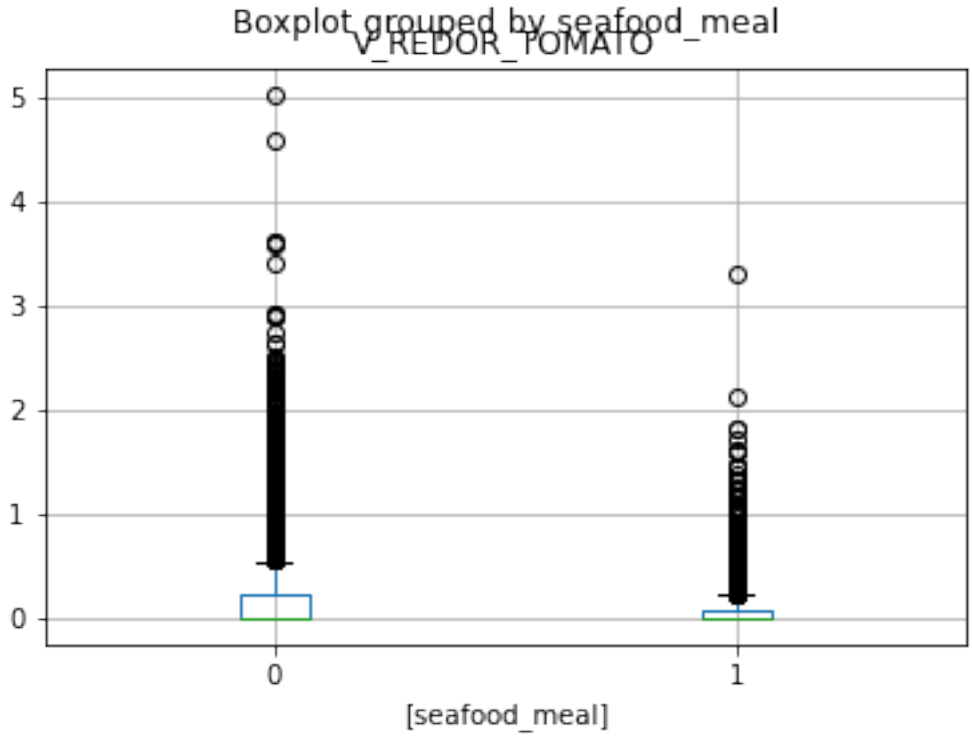
```
for var in veggie:
    z = df.boxplot(column=var,by=['seafood_meal'])
    plt.show(z)
    plt.clf()
```



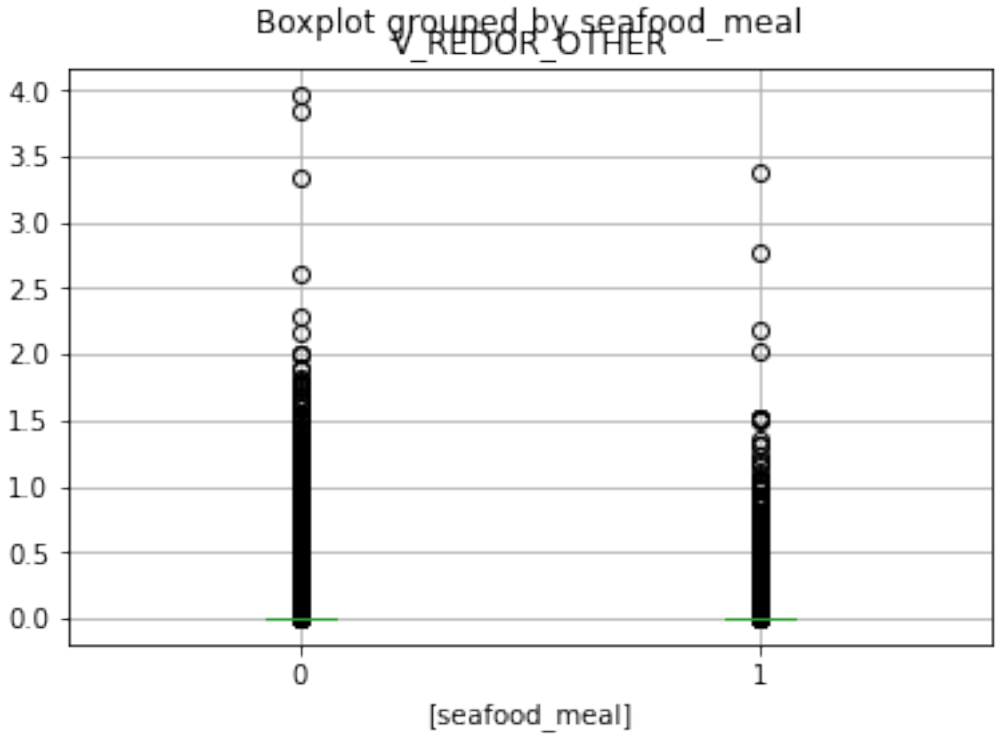

```
<Figure size 432x288 with 0 Axes>
```



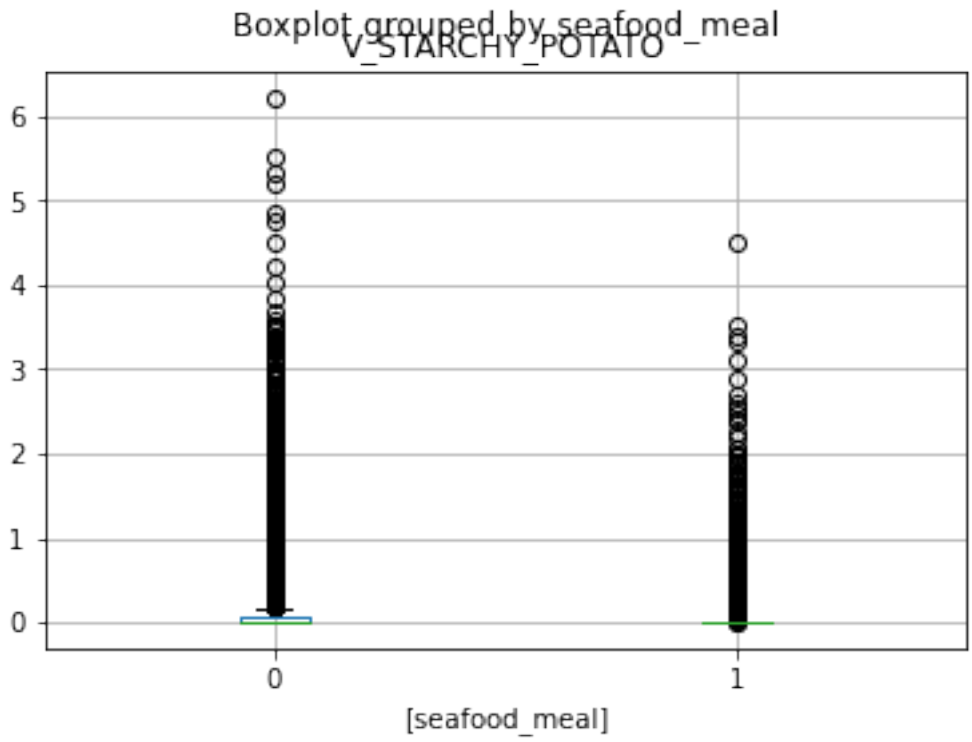
<Figure size 432x288 with 0 Axes>



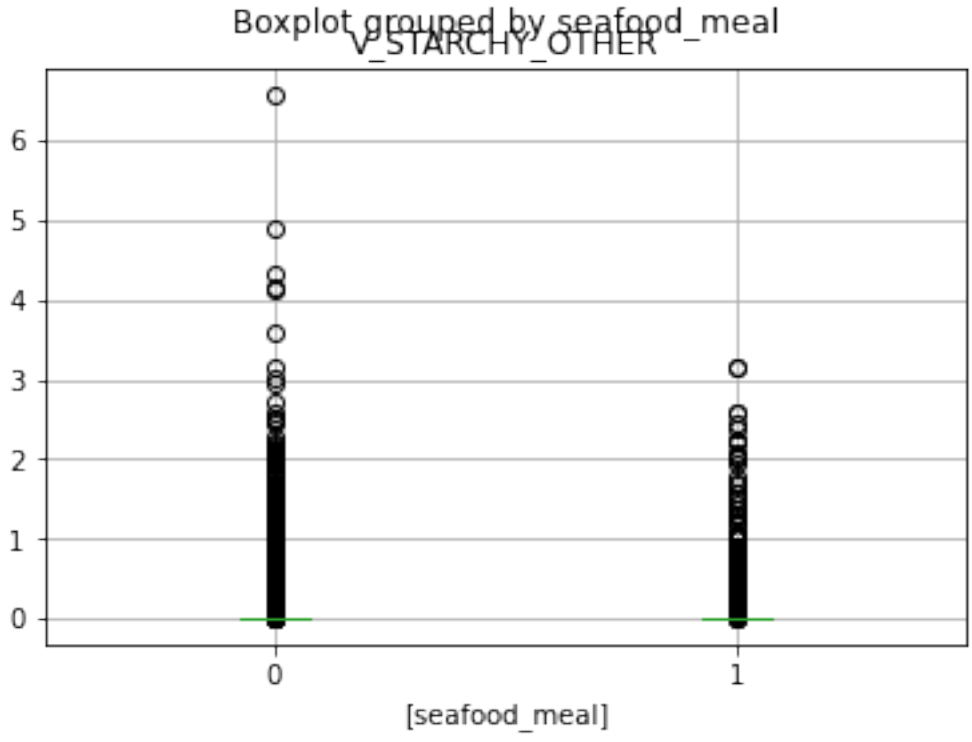
```
<Figure size 432x288 with 0 Axes>
```



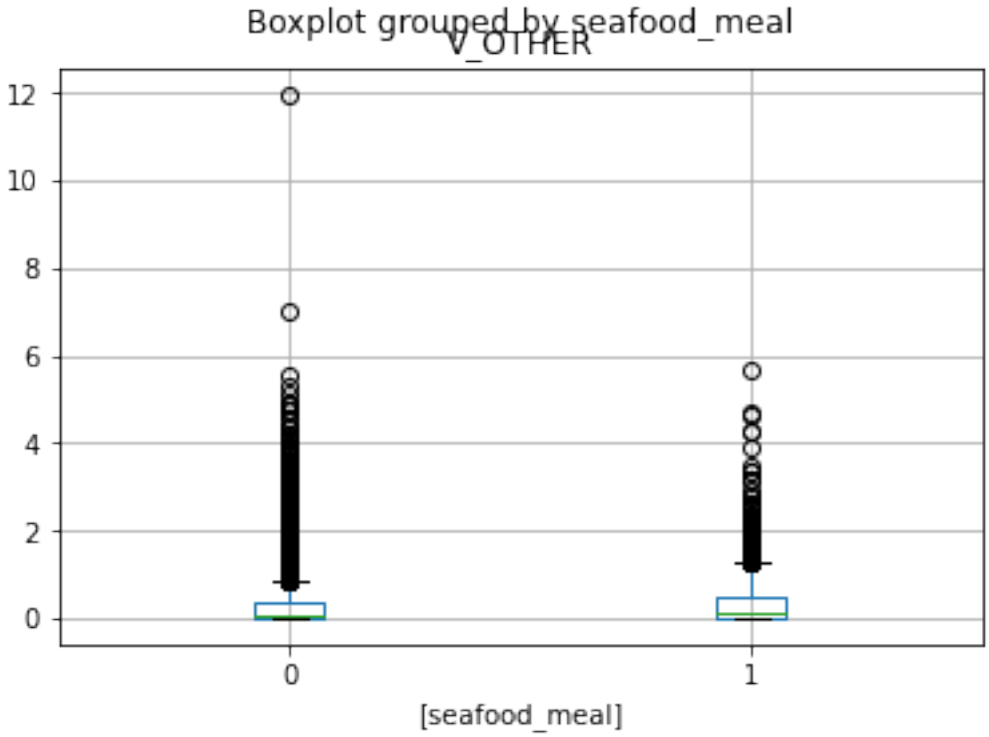

```
<Figure size 432x288 with 0 Axes>
```



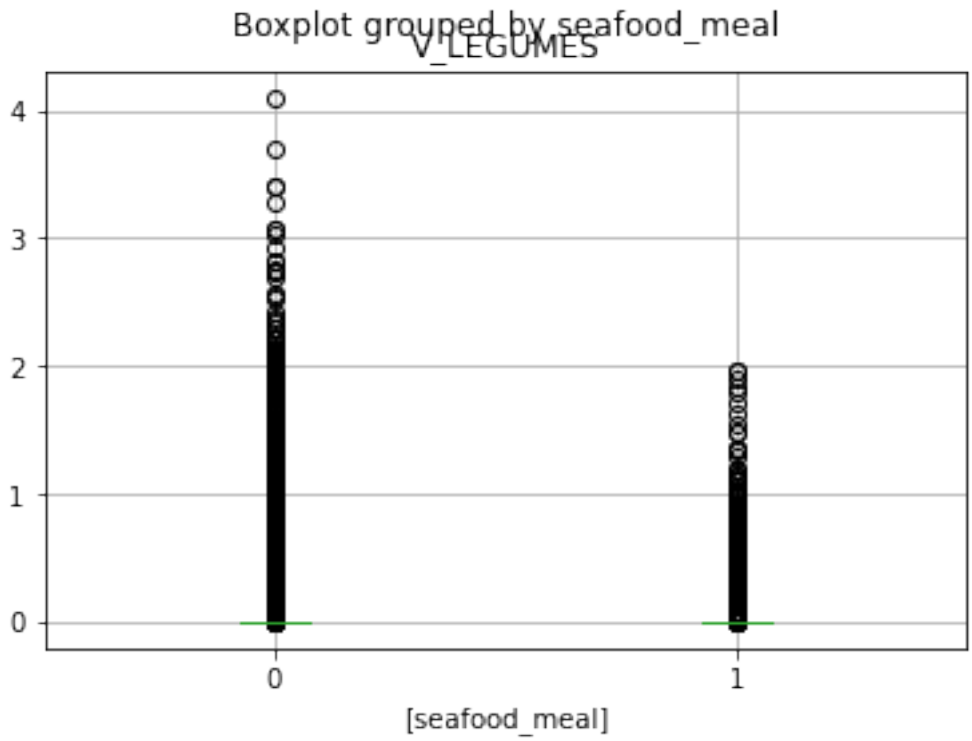

```
<Figure size 432x288 with 0 Axes>
```



```
<Figure size 432x288 with 0 Axes>
```

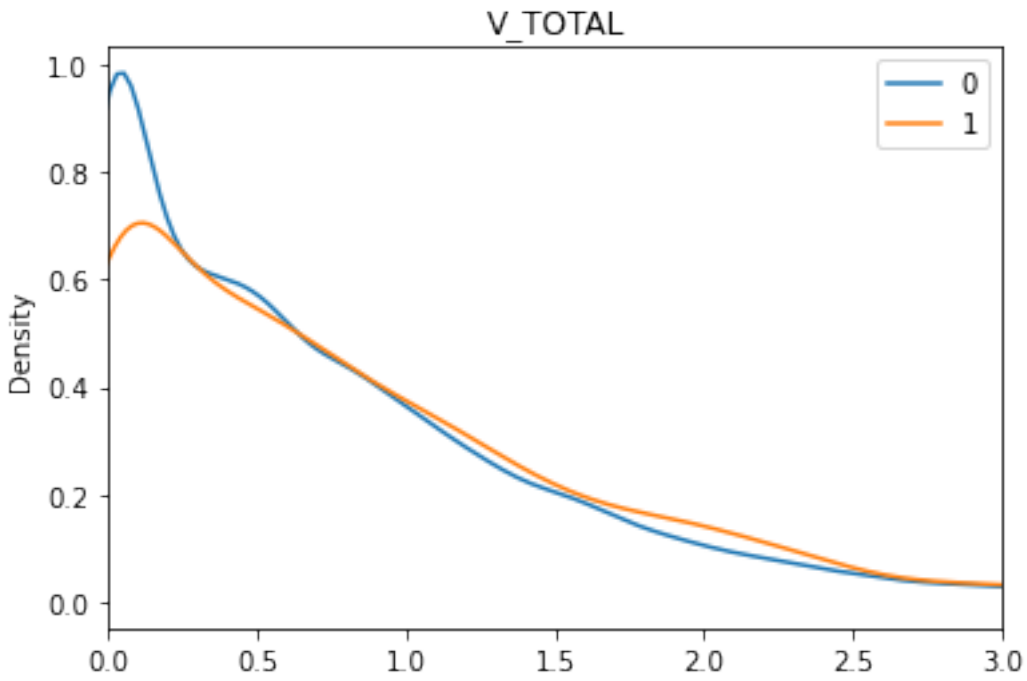



```
<Figure size 432x288 with 0 Axes>
```



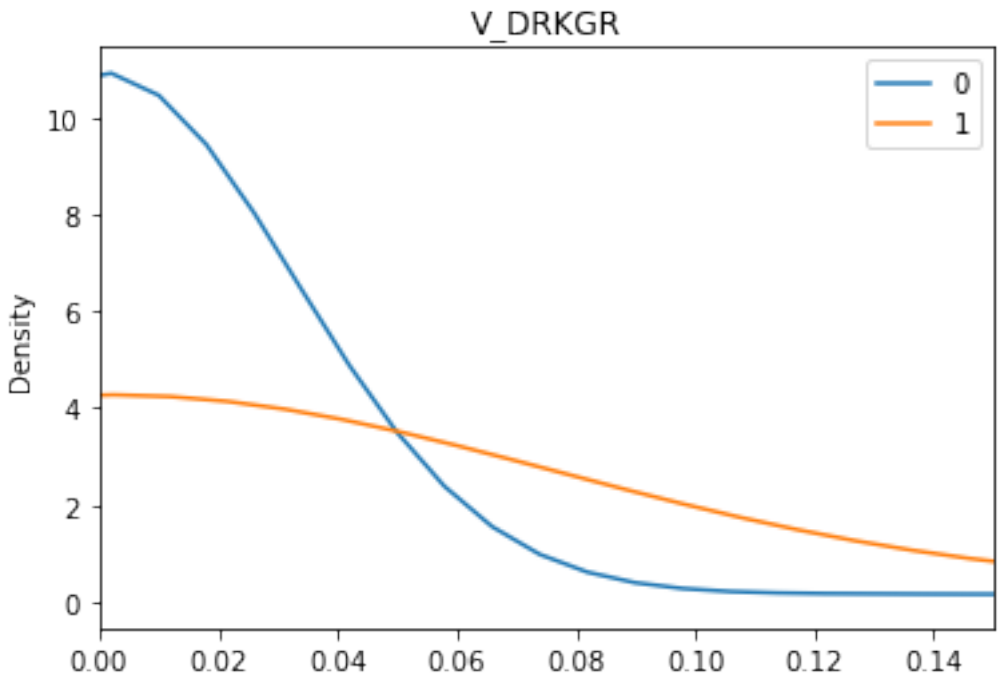

```
<Figure size 432x288 with 0 Axes>
```

[6]:
```
z = df.groupby('seafood_meal')[veggie[0]].plot.kde(title = veggie[0],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 3))
plt.clf()
```



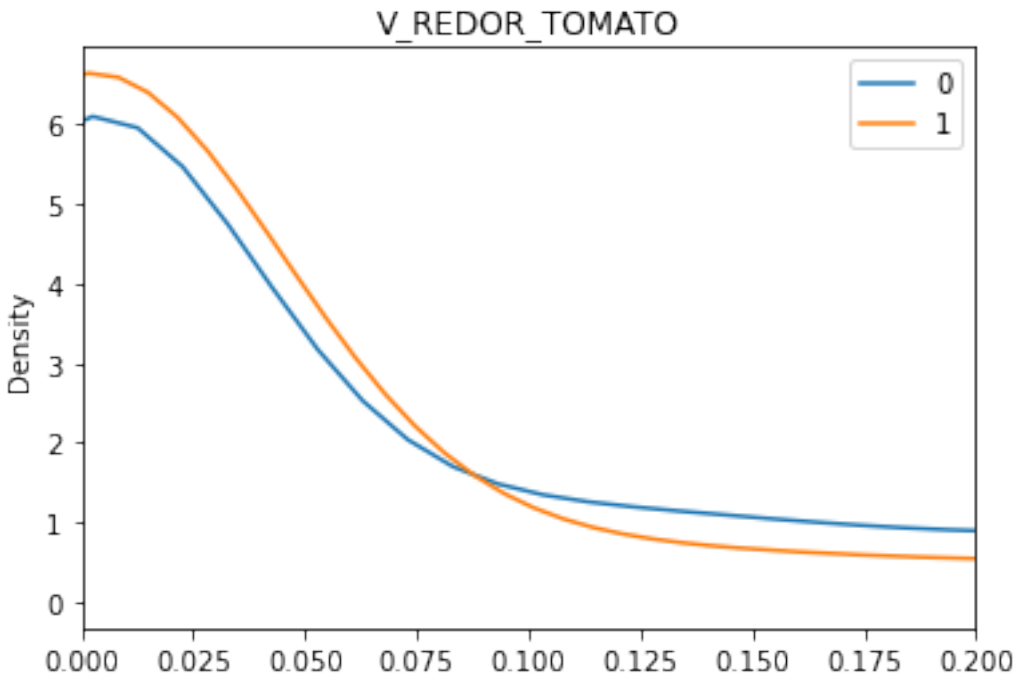

```
<Figure size 432x288 with 0 Axes>
```

[7]:
```
z = df.groupby('seafood_meal')[veggie[1]].plot.kde(title = veggie[1],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 0.15))
plt.clf()
```



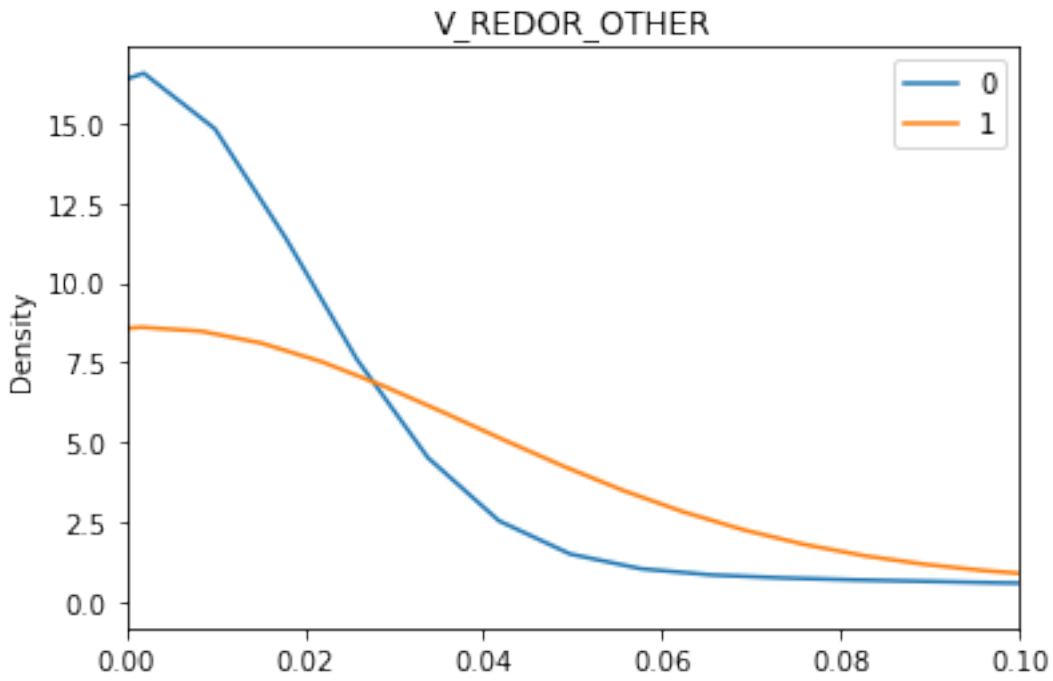

```
<Figure size 432x288 with 0 Axes>
```

[8]:
```
z = df.groupby('seafood_meal')[veggie[2]].plot.kde(title = veggie[2],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 0.2))
plt.clf()
```



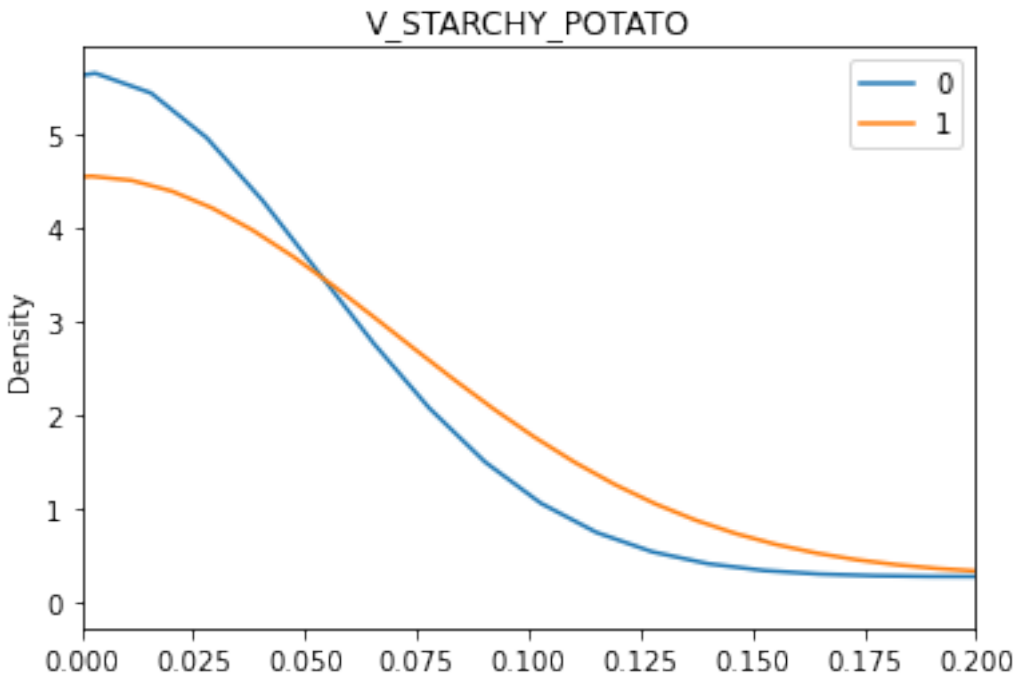

```
<Figure size 432x288 with 0 Axes>
```

[9]:
```
z = df.groupby('seafood_meal')[veggie[3]].plot.kde(title = veggie[3],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 0.1))
plt.clf()
```



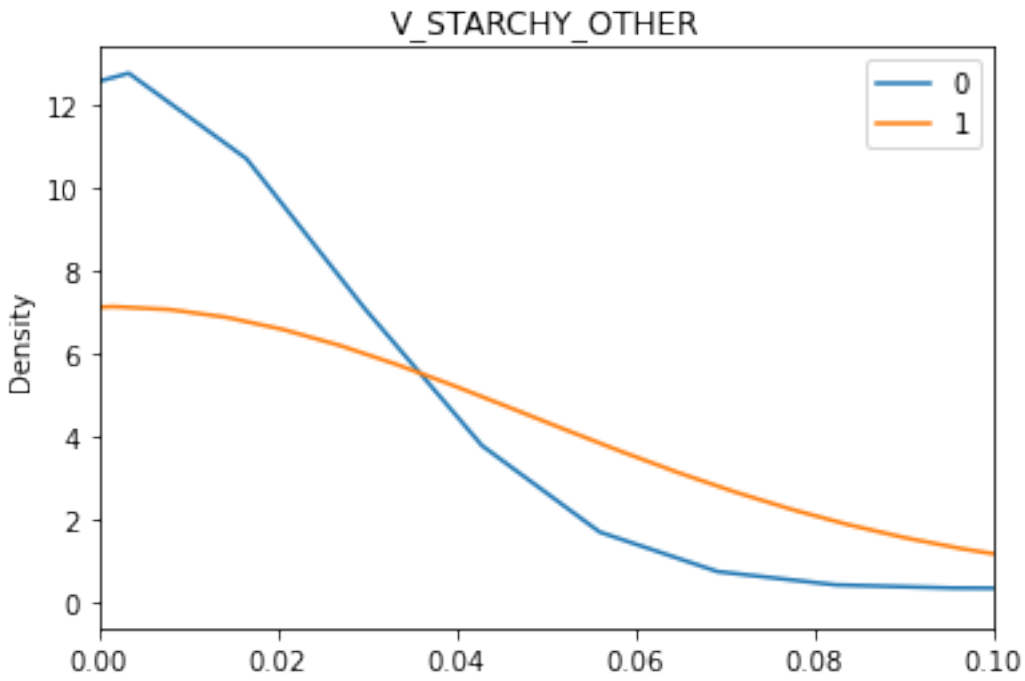

```
<Figure size 432x288 with 0 Axes>
```

[10]:
```
z = df.groupby('seafood_meal')[veggie[4]].plot.kde(title = veggie[4],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 0.2))
plt.clf()
```



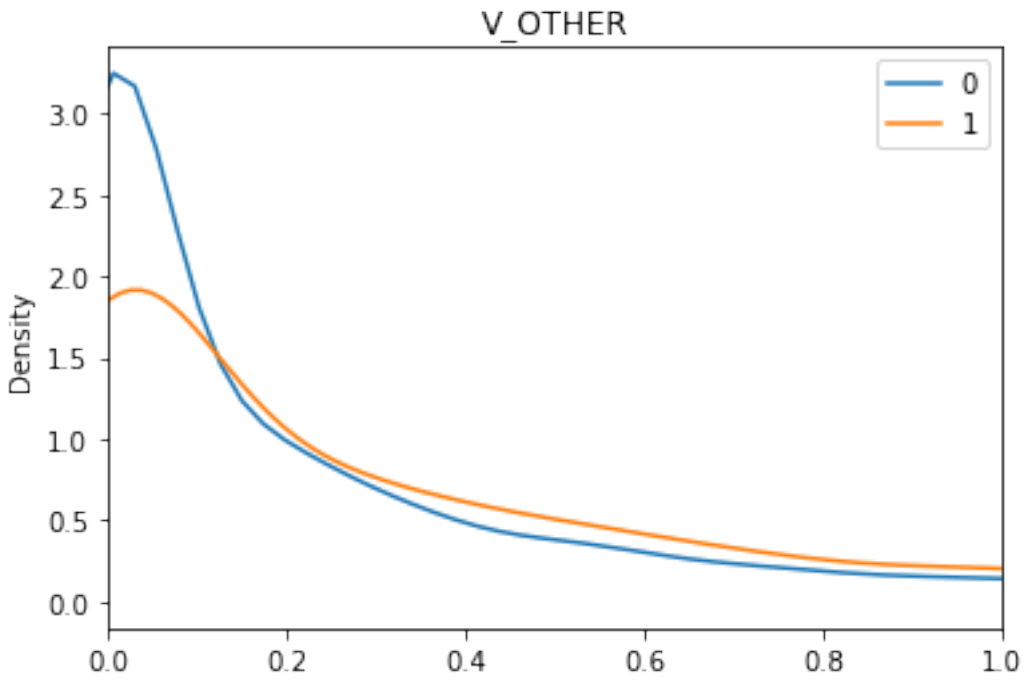

```
<Figure size 432x288 with 0 Axes>
```

[11]:
```
z = df.groupby('seafood_meal')[veggie[5]].plot.kde(title = veggie[5],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 0.1))
plt.clf()
```



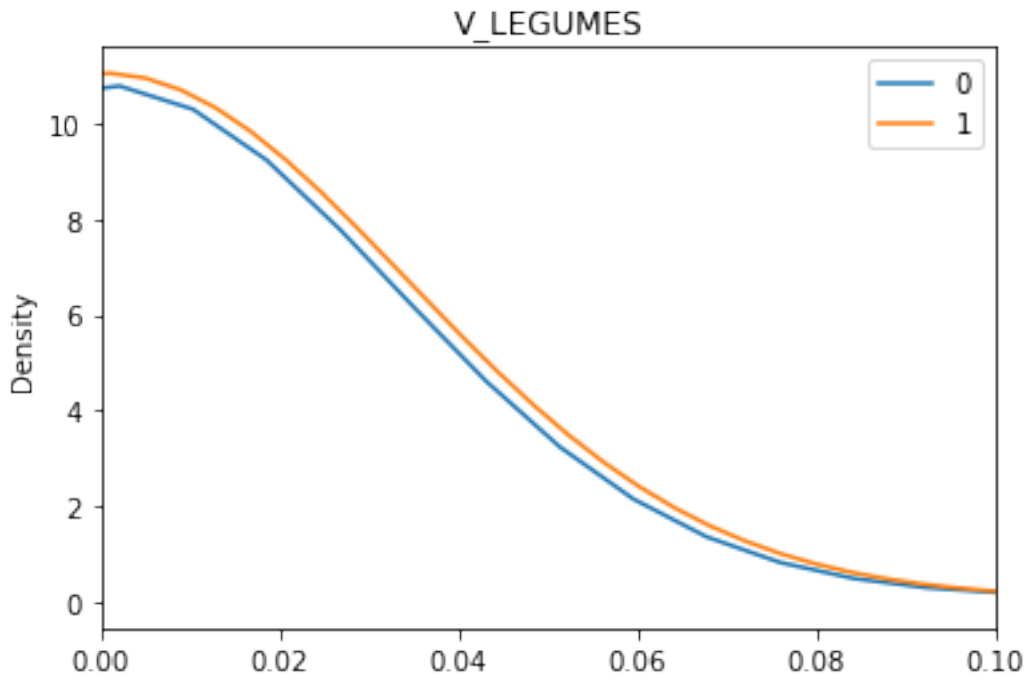

```
<Figure size 432x288 with 0 Axes>
```

[12]:
```
z = df.groupby('seafood_meal')[veggie[6]].plot.kde(title = veggie[6],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 1))
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

[13]:
```
z = df.groupby('seafood_meal')[veggie[7]].plot.kde(title = veggie[7],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 0.1))
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

[14]:
```
for var in veggie:
    z = df.groupby('seafood_meal')[var].describe()
    print("Statistics for "+var+'\n')
    print(z)
    print('\n')
```

```
Statistics for V_TOTAL

                count      mean       std  min   25%   50%   75%    max
seafood_meal
0             26011.0  0.766811  0.830492  0.0  0.13  0.54  1.12  11.97
1              3232.0  0.839502  0.884815  0.0  0.17  0.61  1.23   9.40


Statistics for V_DRKGR

                count      mean       std  min  25%  50%  75%   max
seafood_meal
0             26011.0  0.064261  0.248228  0.0  0.0  0.0  0.0  3.99
1              3232.0  0.133026  0.378411  0.0  0.0  0.0  0.0  4.78
```

Statistics for V_REDOR_TOMATO

seafood_meal	count	mean	std	min	25%	50%	75%	max
0	26011.0	0.155829	0.280726	0.0	0.0	0.01	0.22	5.03
1	3232.0	0.096572	0.217002	0.0	0.0	0.00	0.09	3.31

Statistics for V_REDOR_OTHER

seafood_meal	count	mean	std	min	25%	50%	75%	max
0	26011.0	0.044952	0.150933	0.0	0.0	0.0	0.0	3.97
1	3232.0	0.061105	0.187205	0.0	0.0	0.0	0.0	3.37

Statistics for V_STARCHY_POTATO

seafood_meal	count	mean	std	min	25%	50%	75%	max
0	26011.0	0.172294	0.404003	0.0	0.0	0.0	0.07	6.22
1	3232.0	0.131043	0.351846	0.0	0.0	0.0	0.00	4.50

Statistics for V_STARCHY_OTHER

seafood_meal	count	mean	std	min	25%	50%	75%	max
0	26011.0	0.055785	0.204900	0.0	0.0	0.0	0.0	6.58
1	3232.0	0.066473	0.240464	0.0	0.0	0.0	0.0	3.16

Statistics for V_OTHER

seafood_meal	count	mean	std	min	25%	50%	75%	max
0	26011.0	0.273681	0.465530	0.0	0.0	0.08	0.35	11.97
1	3232.0	0.351272	0.506081	0.0	0.0	0.15	0.51	5.65

Statistics for V_LEGUMES

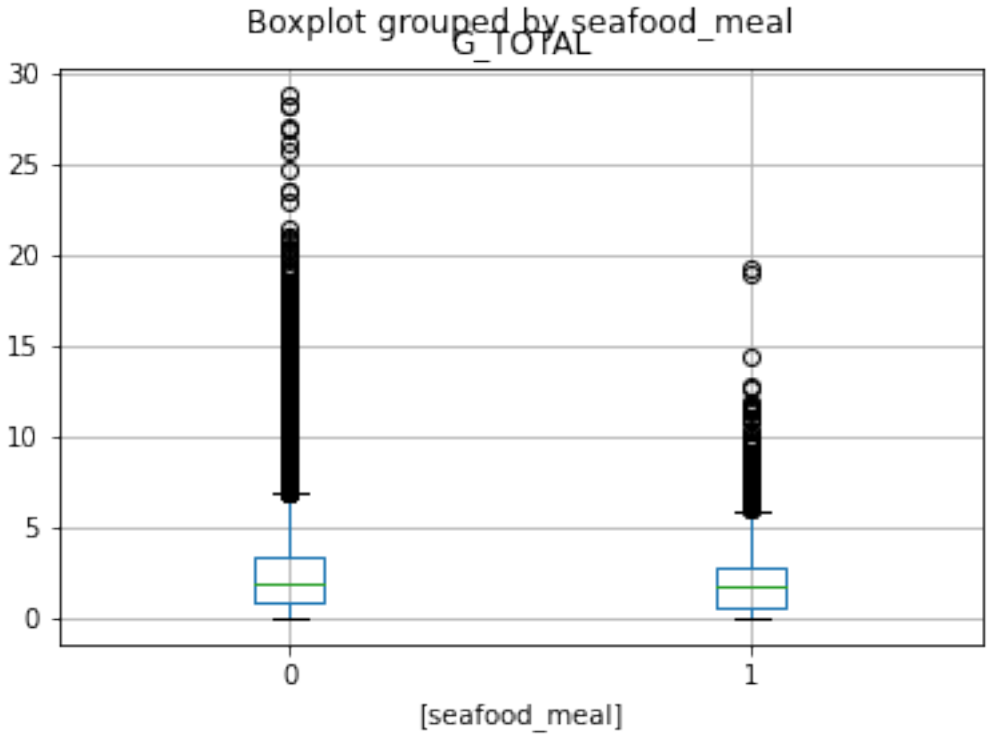
seafood_meal	count	mean	std	min	25%	50%	75%	max
0	26011.0	0.070522	0.249932	0.0	0.0	0.0	0.0	4.10
1	3232.0	0.034558	0.170488	0.0	0.0	0.0	0.0	1.97

Section 3: Grains

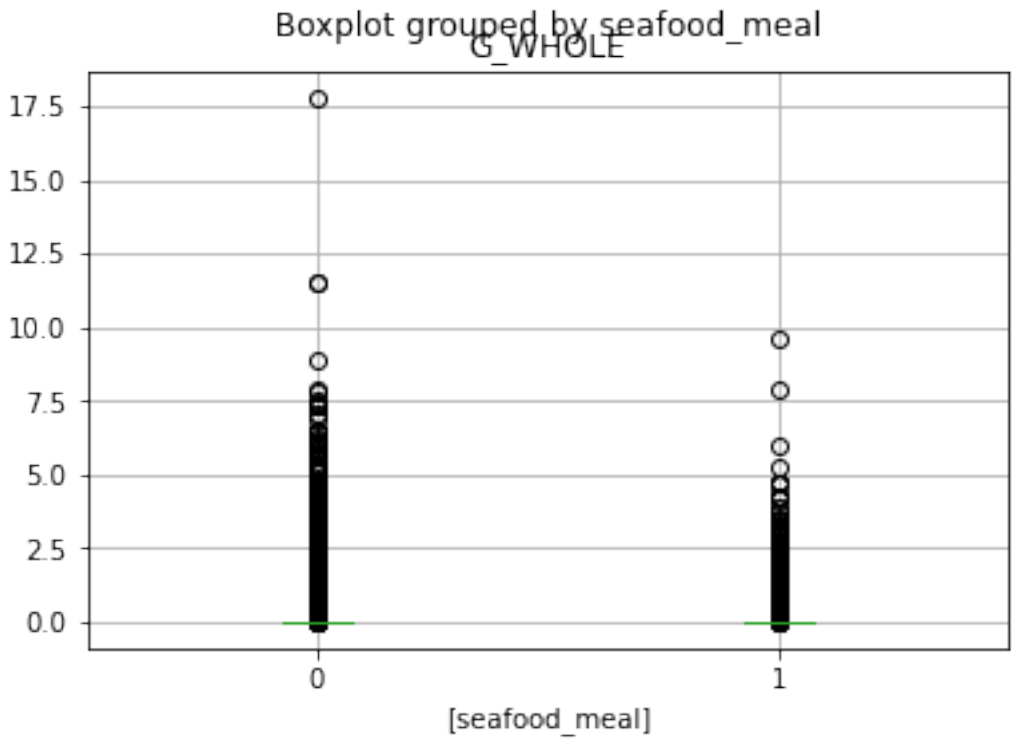
This section provides boxplots and density plots of the Grains FPED components in the seafood meal and non seafood meal groups. The code for seafood meal is 1 if meal contains seafood, and 0 if meal does not contain seafood.

Plot: Meal calories distribution with meals that are 0 KCAL removed, split by 'eathome' groups.

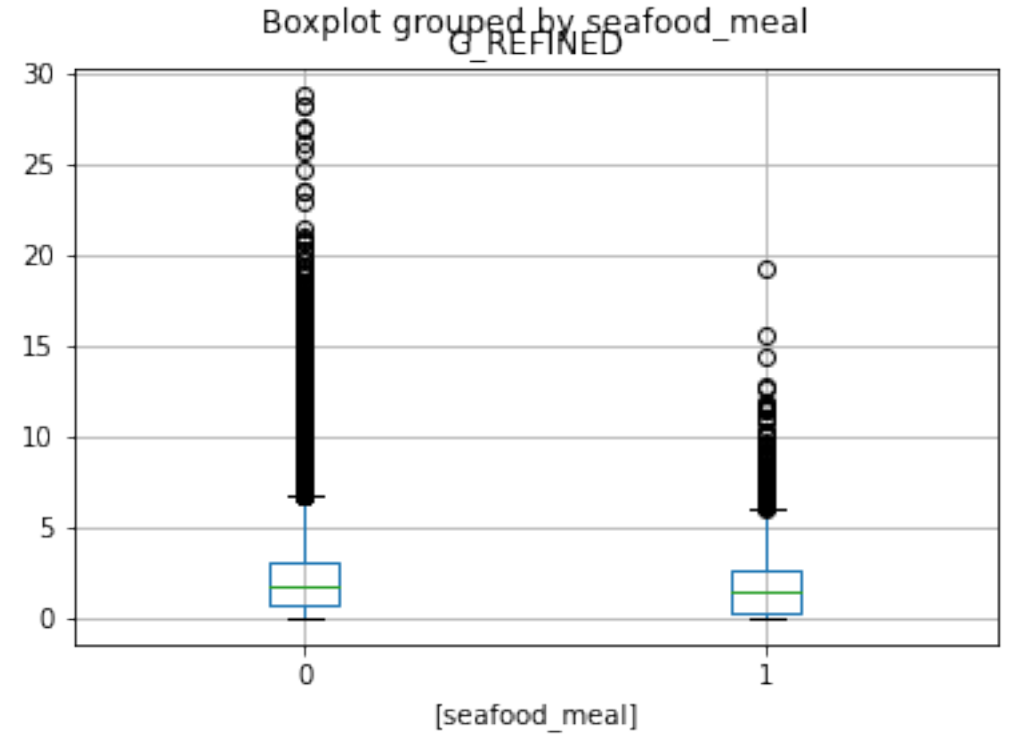
[15]:
```
grains = ['G_TOTAL', 'G_WHOLE','G_REFINED']

for var in grains:
    z = df.boxplot(column=var,by=['seafood_meal'])
    plt.show(z)
    plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

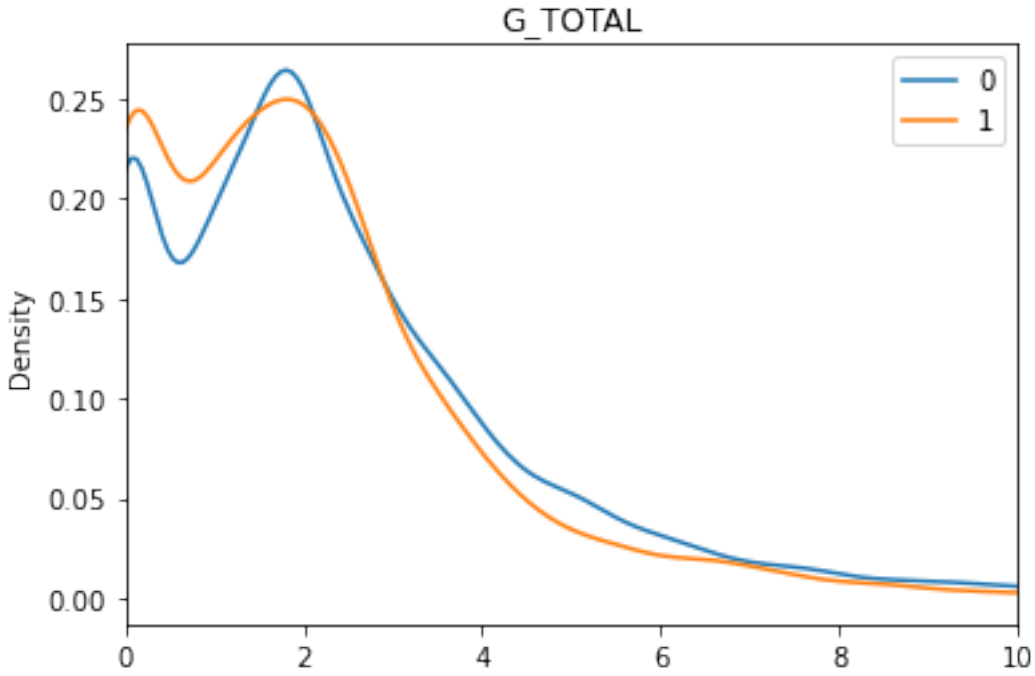



```
<Figure size 432x288 with 0 Axes>
```



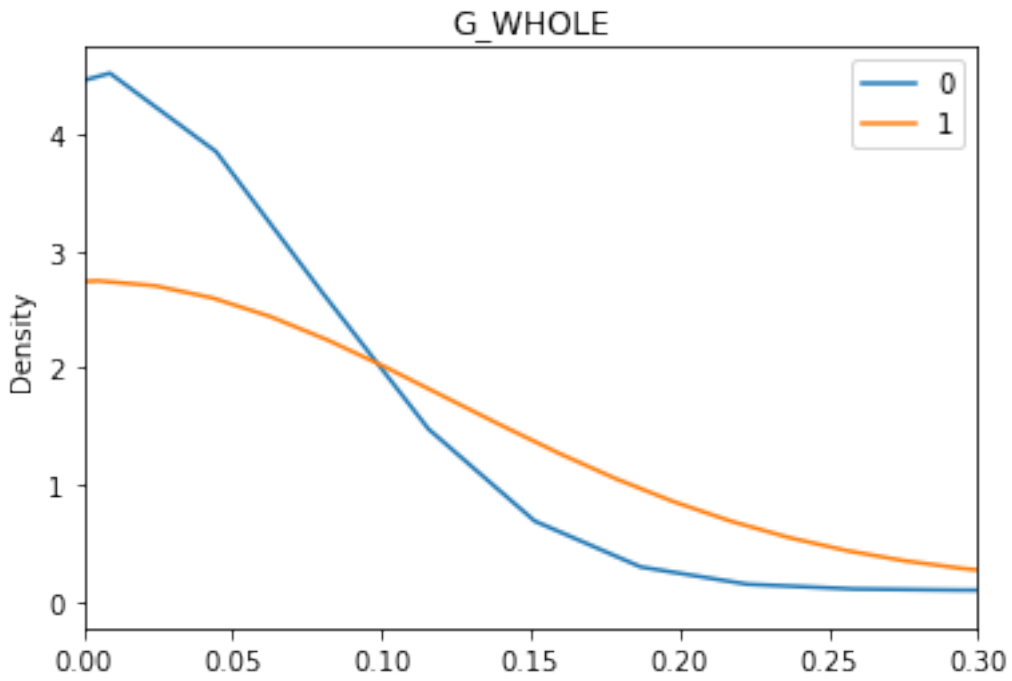
```
<Figure size 432x288 with 0 Axes>
```

[16]:
```
z = df.groupby('seafood_meal')[grains[0]].plot.kde(title = grains[0],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 10))
plt.clf()
```



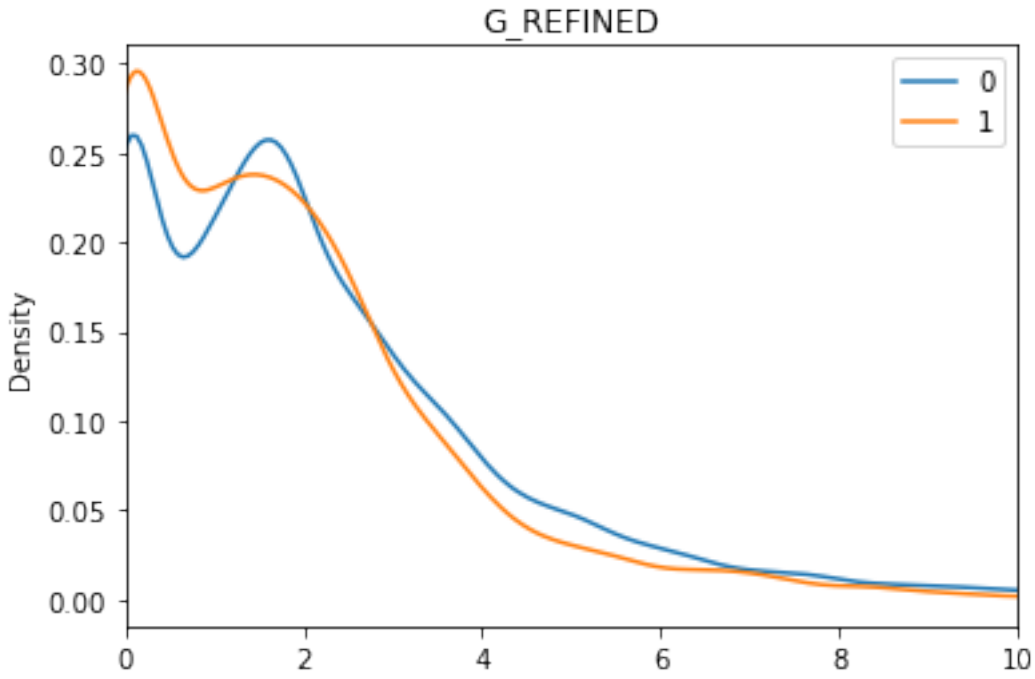

```
<Figure size 432x288 with 0 Axes>
```

[17]:
```
z = df.groupby('seafood_meal')[grains[1]].plot.kde(title = grains[1],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 0.3))
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

[18]:
```
z = df.groupby('seafood_meal')[grains[2]].plot.kde(title = grains[2],␣
 ↪legend='x')
plt.show(z[0].set_xlim(0, 10))
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

[19]:
```
for var in grains:
    z = df.groupby('seafood_meal')[var].describe()
    print("Statistics for "+var+'\n')
    print(z)
    print('\n')
```

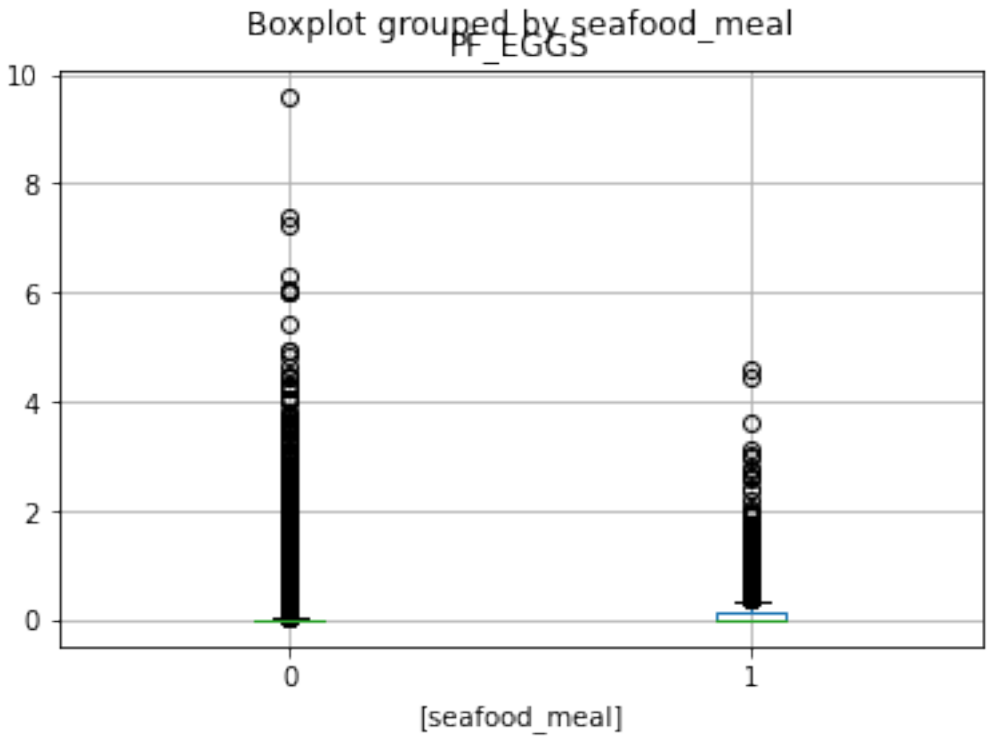
```
Statistics for G_TOTAL

                count      mean       std  min   25%   50%     75%    max
seafood_meal
0             26011.0  2.480072  2.398177  0.0  0.94  1.95  3.3200  28.83
1              3232.0  2.053747  1.958012  0.0  0.67  1.75  2.7825  19.26


Statistics for G_WHOLE

                count      mean       std  min  25%  50%  75%    max
seafood_meal
0             26011.0  0.158145  0.580842  0.0  0.0  0.0  0.0  17.79
1              3232.0  0.184836  0.620747  0.0  0.0  0.0  0.0   9.65
```

```
Statistics for G_REFINED

                count      mean       std  min   25%   50%     75%    max
seafood_meal
0             26011.0  2.321905  2.39243  0.0  0.74  1.76  3.1600  28.83
1              3232.0  1.868917  1.93630  0.0  0.36  1.46  2.6025  19.26
```

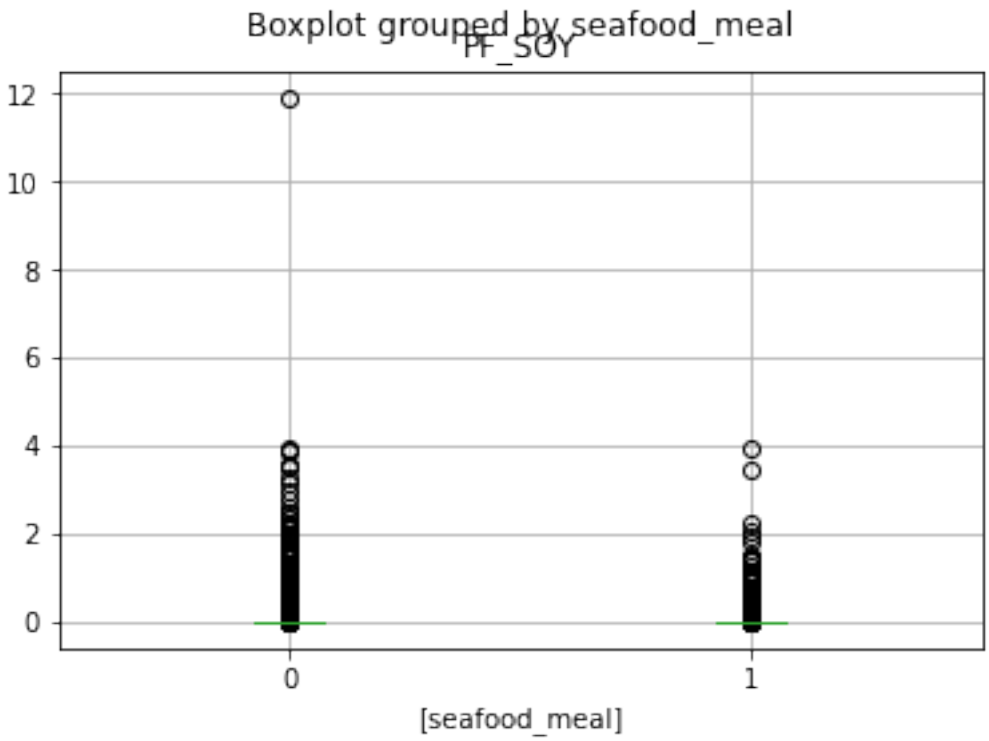
Section 4: Non-Meat Proteins

This section provides boxplots and density plots of the non-meat protein FPED components in the seafood meal and non seafood meal groups. The code for seafood meal is 1 if meal contains seafood, and 0 if meal does not contain seafood.

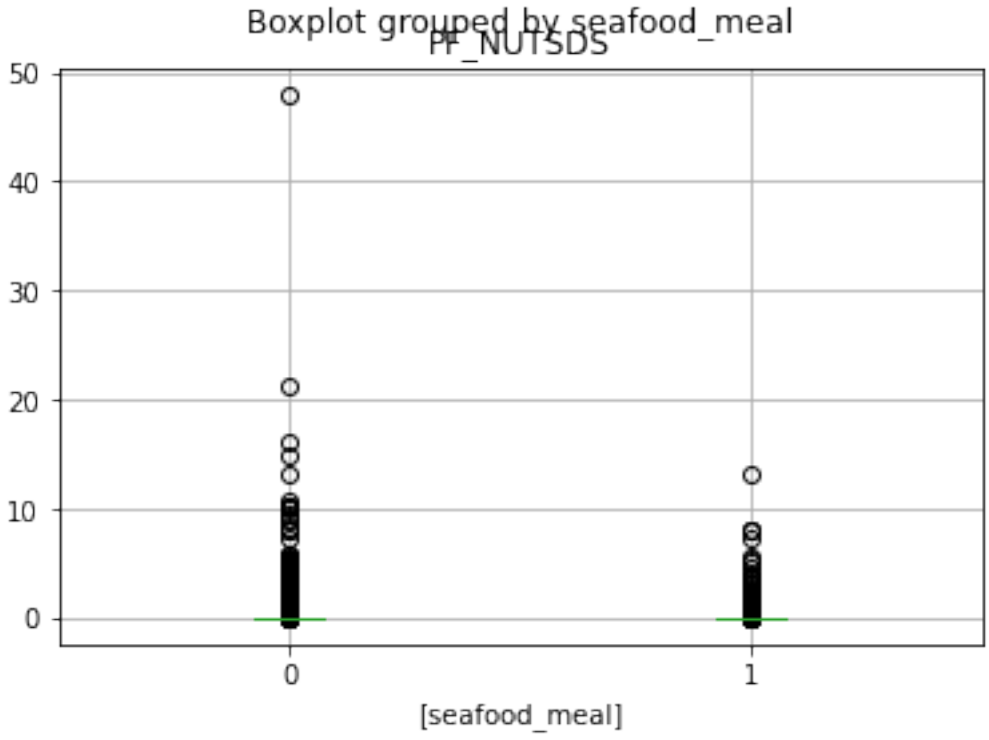
[20]:
```
non_meat_protein = ['PF_EGGS', 'PF_SOY', 'PF_NUTSDS', 'PF_LEGUMES']

for var in non_meat_protein:
    z = df.boxplot(column=var,by=['seafood_meal'])
    plt.show(z)
    plt.clf()
```



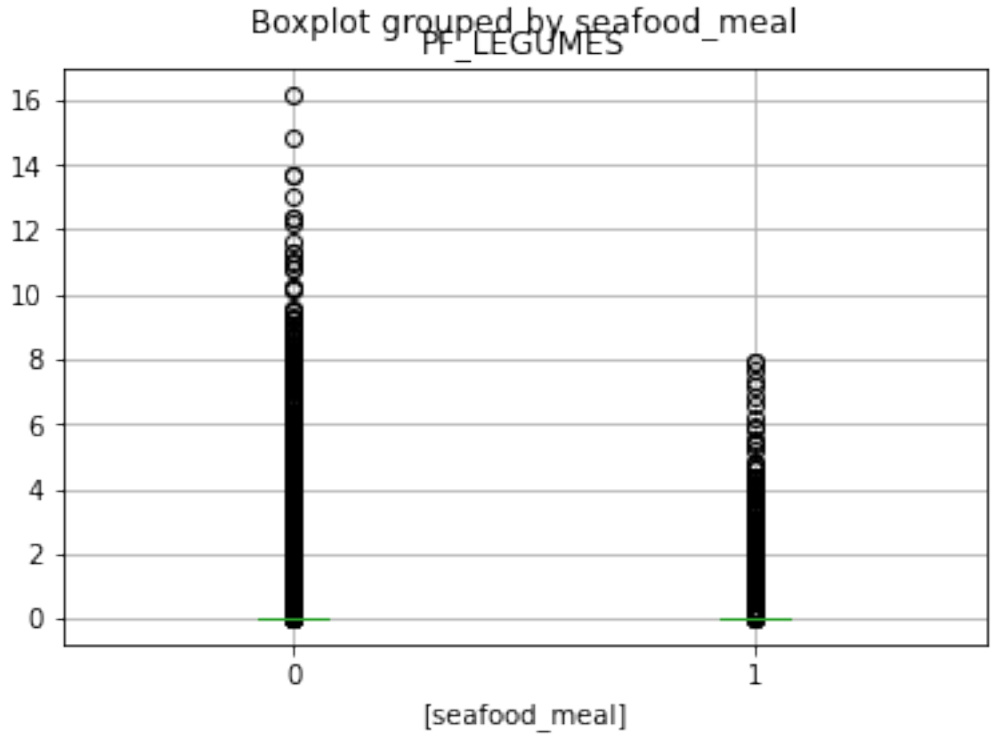

```
<Figure size 432x288 with 0 Axes>
```




```
<Figure size 432x288 with 0 Axes>
```

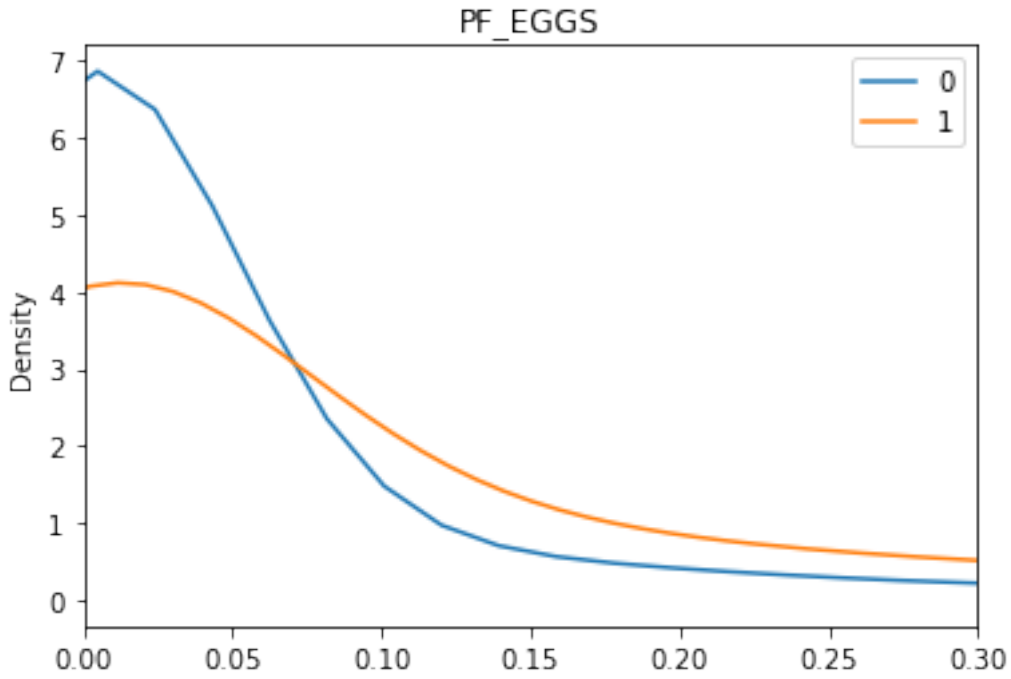


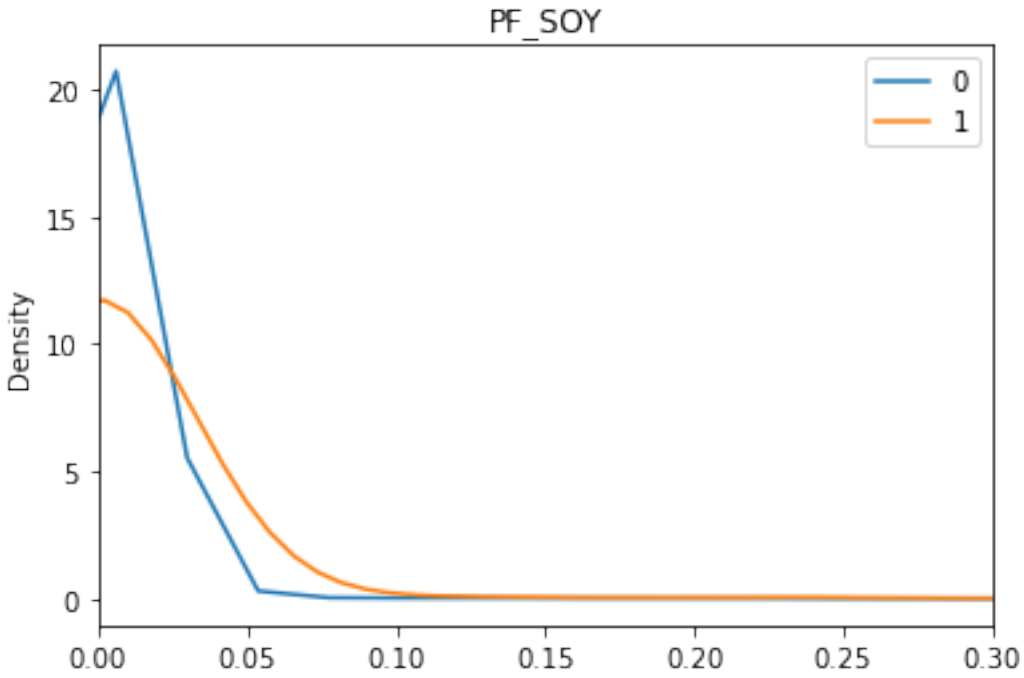
```
<Figure size 432x288 with 0 Axes>
```

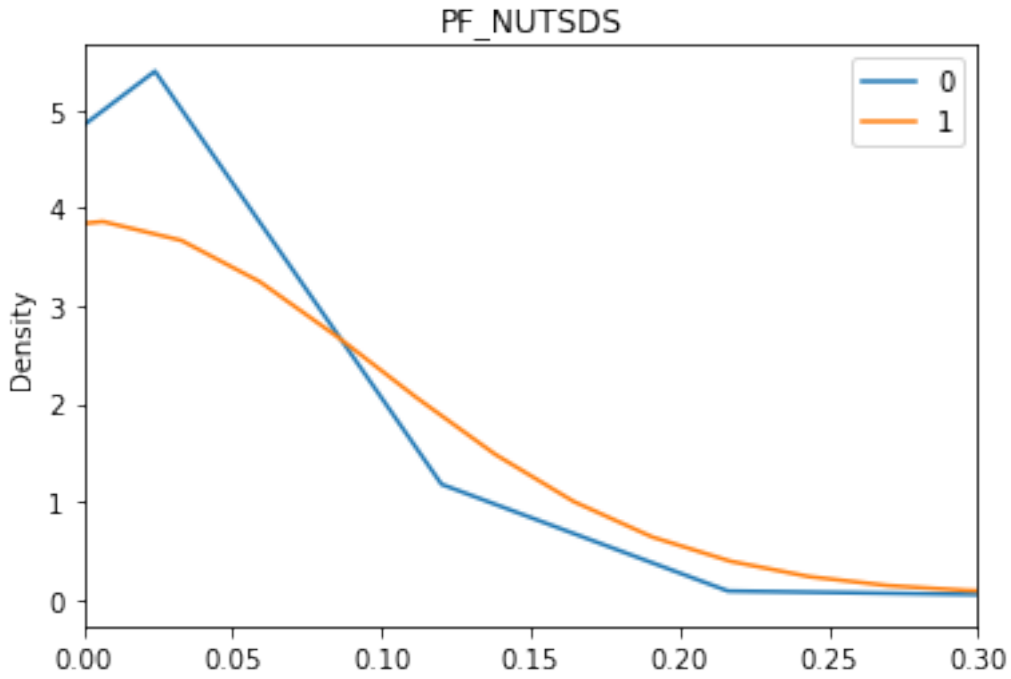


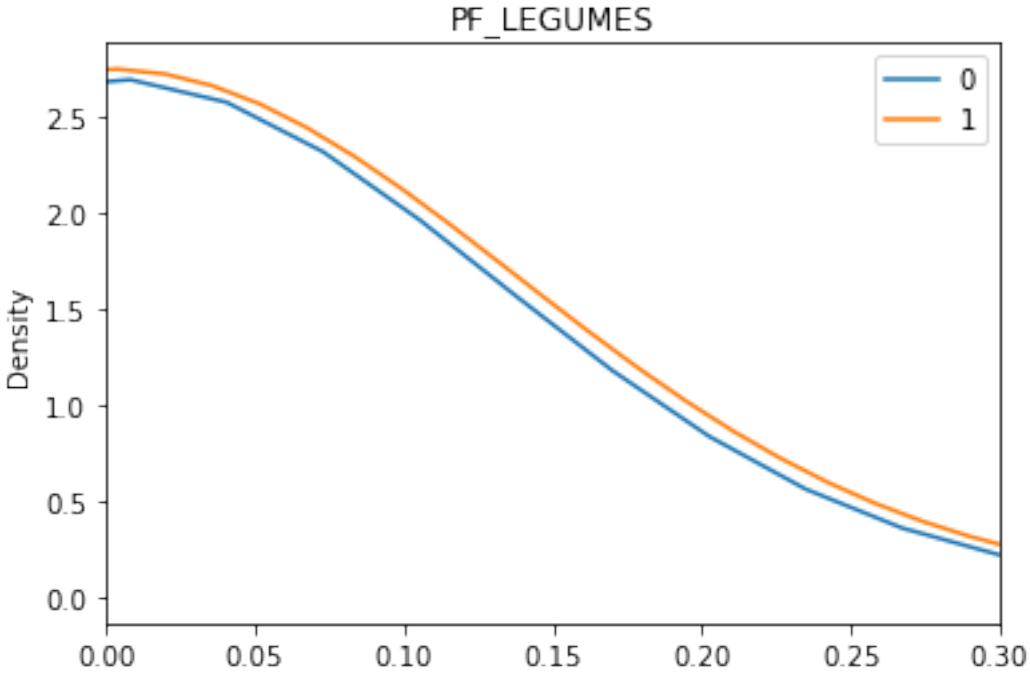

```
<Figure size 432x288 with 0 Axes>
```

[21]:
```
for var in non_meat_protein:
    z = df.groupby('seafood_meal')[var].plot.kde(title = var, legend='x')
    plt.show(z[0].set_xlim(0, 0.3))
    plt.clf()
```


PF_EGGS
Density
0
1
0
1
2
3
4
5
6
7
0.00
0.05
0.10
0.15
0.20
0.25
0.30


PF_SOY
Density
0
1
0
5
10
15
20
0.00
0.05
0.10
0.15
0.20
0.25
0.30






```
<Figure size 432x288 with 0 Axes>
```

[22]:
```
for var in non_meat_protein:
    z = df.groupby('seafood_meal')[var].describe()
    print("Statistics for "+var+'\n')
    print(z)
    print('\n')
```

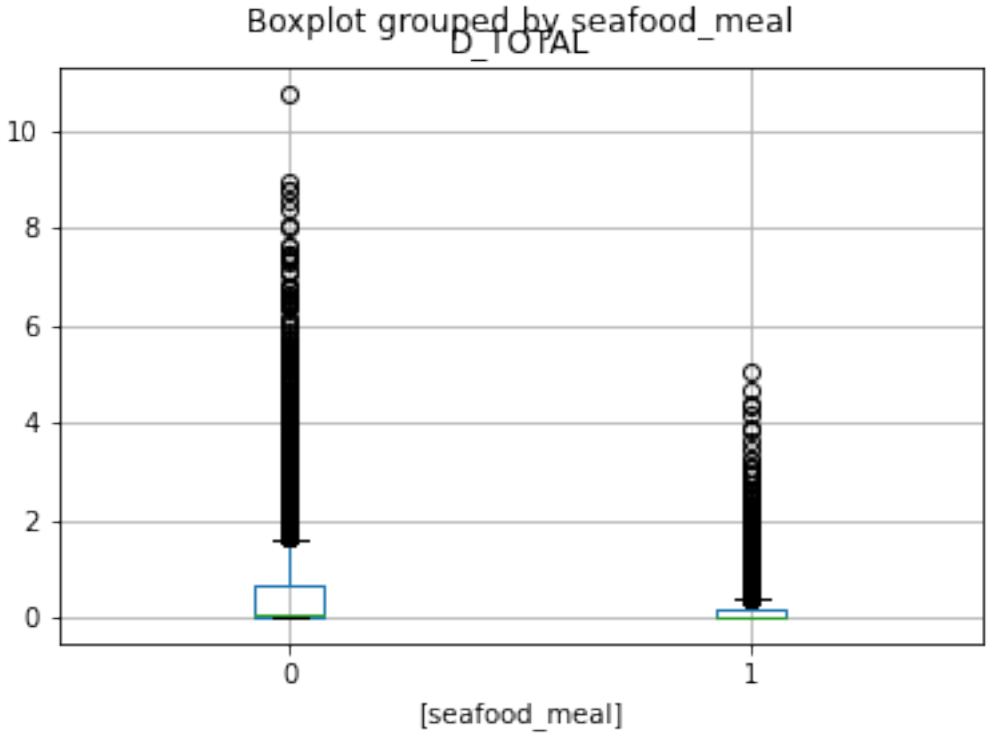
```
Statistics for PF_EGGS

                count      mean       std  min  25%   50%   75%   max
seafood_meal
0             26011.0  0.092293  0.356726  0.0  0.0  0.00  0.02  9.60
1              3232.0  0.143728  0.324184  0.0  0.0  0.01  0.15  4.62


Statistics for PF_SOY

                count      mean       std  min  25%  50%  75%    max
seafood_meal
0             26011.0  0.009890  0.136564  0.0  0.0  0.0  0.0  11.91
1              3232.0  0.021417  0.164274  0.0  0.0  0.0  0.0   3.97


Statistics for PF_NUTSDS

                count      mean       std  min  25%  50%  75%    max
seafood_meal
0             26011.0  0.046389  0.508116  0.0  0.0  0.0  0.0  48.00
1              3232.0  0.066460  0.493075  0.0  0.0  0.0  0.0  13.14


Statistics for PF_LEGUMES

                count      mean       std  min  25%  50%  75%    max
seafood_meal
0             26011.0  0.282282  1.000122  0.0  0.0  0.0  0.0  16.18
1              3232.0  0.139044  0.685143  0.0  0.0  0.0  0.0   7.93
```

Section 5: Dairy

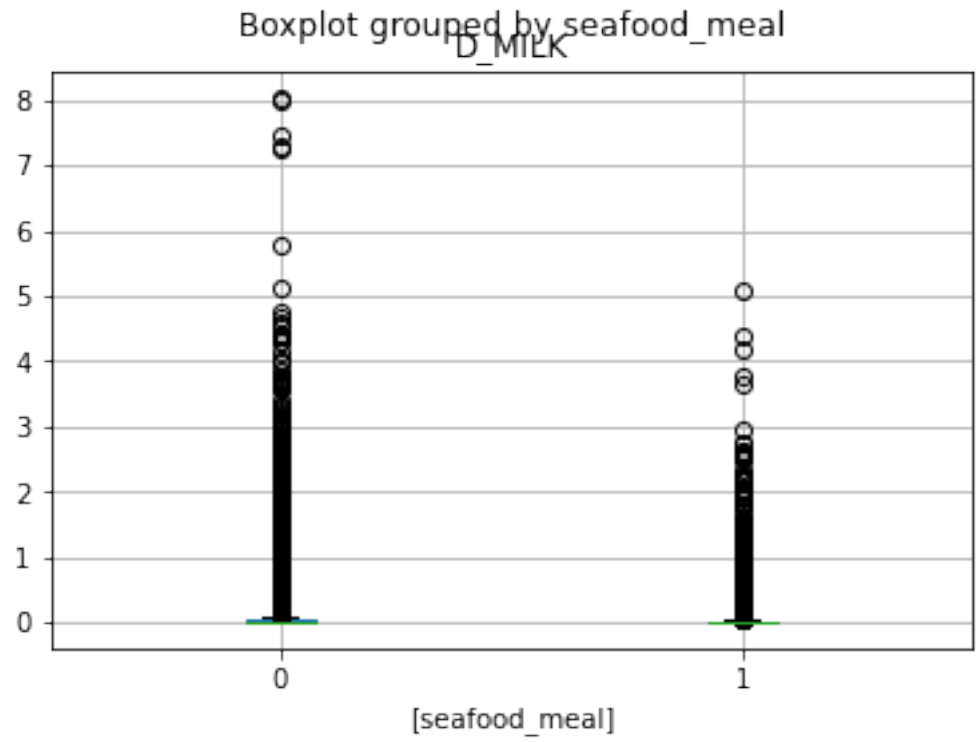
This section provides boxplots and density plots of the dairy FPED components in the seafood meal and non seafood meal groups. The code for seafood meal is 1 if meal contains seafood, and 0 if meal does not contain seafood.

[23]:
```
dairy = ['D_TOTAL', 'D_MILK', 'D_YOGURT', 'D_CHEESE']

for var in dairy:
```

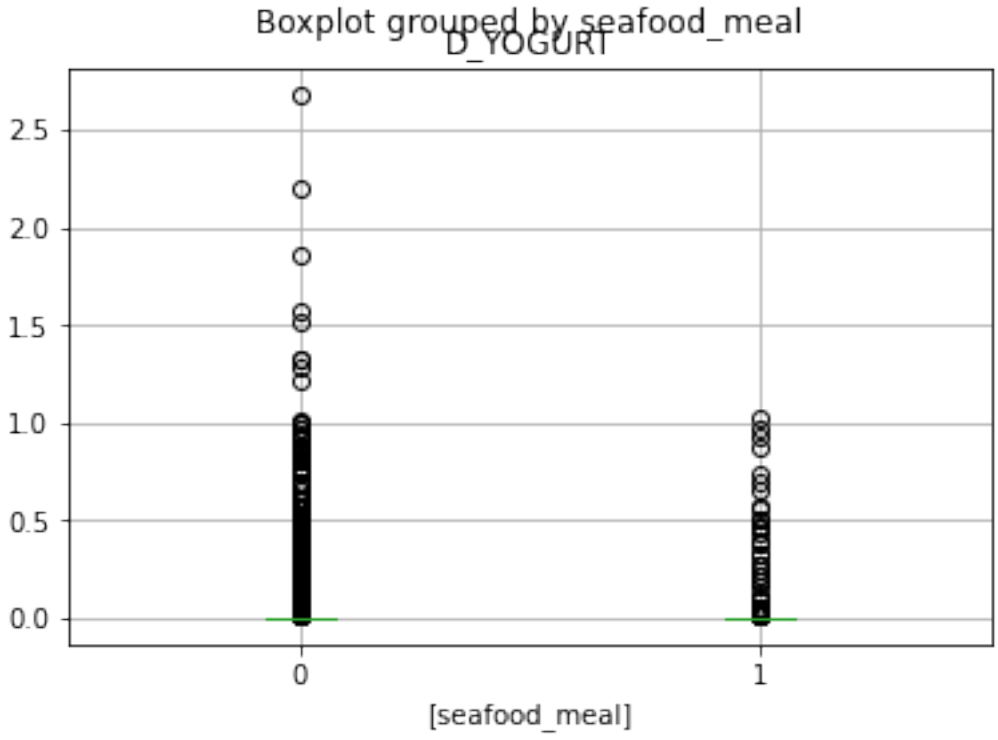
```
    z = df.boxplot(column=var,by=['seafood_meal'])
    plt.show(z)
    plt.clf()
```



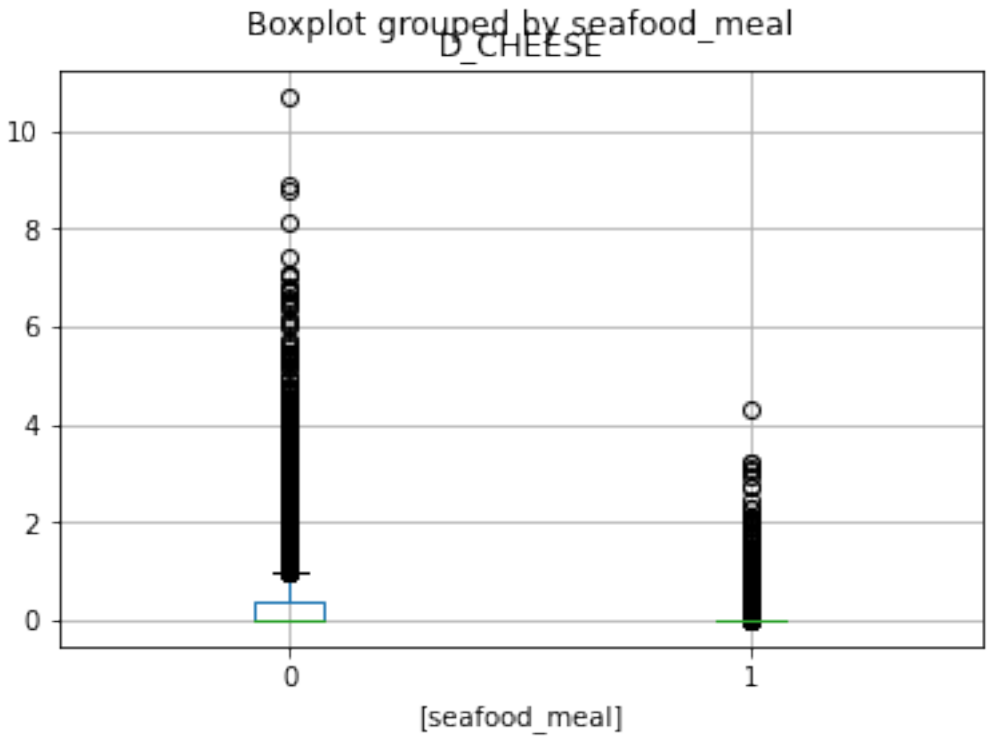

```
<Figure size 432x288 with 0 Axes>
```




```
<Figure size 432x288 with 0 Axes>
```

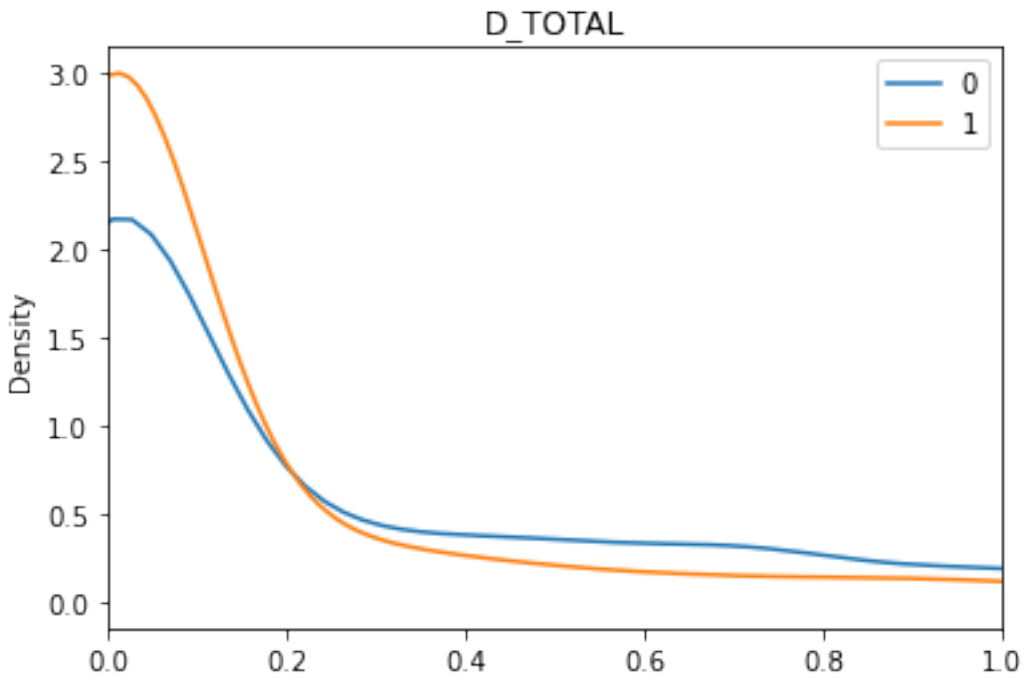


```
<Figure size 432x288 with 0 Axes>
```



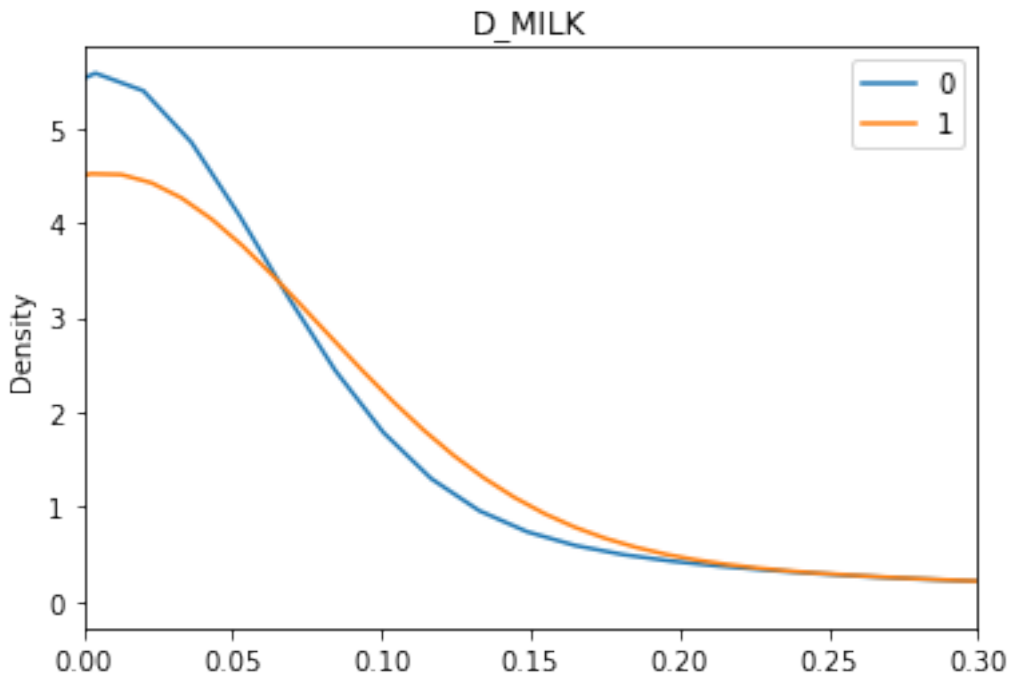

```
<Figure size 432x288 with 0 Axes>
```

[24]:
```
z = df.groupby('seafood_meal')[dairy[0]].plot.kde(title = dairy[0], legend='x')
plt.show(z[0].set_xlim(0, 1))
plt.clf()
```



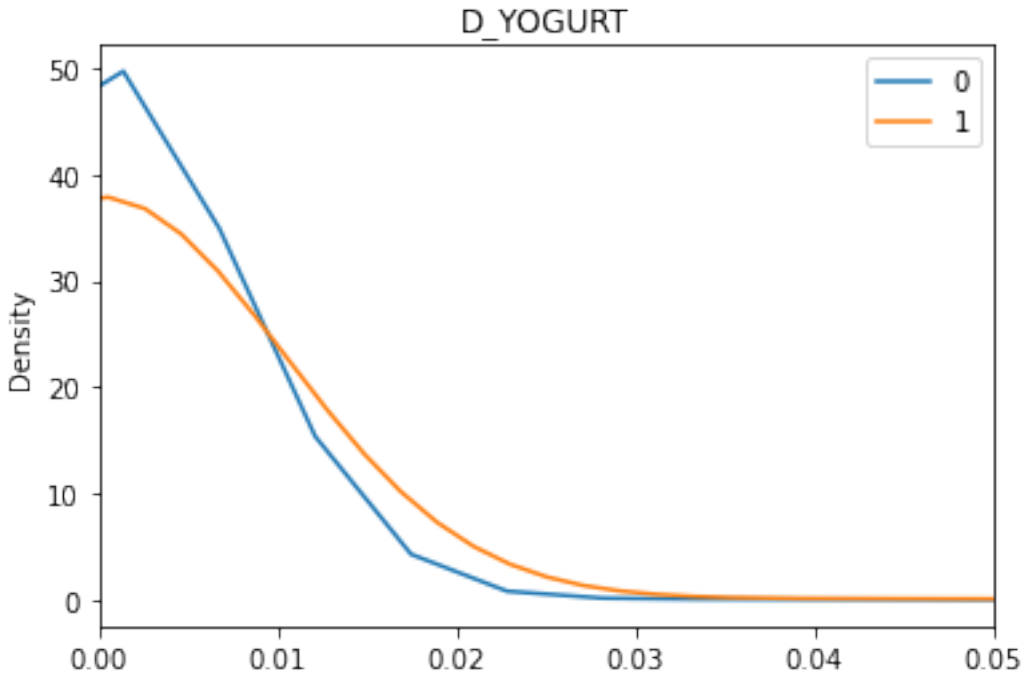

```
<Figure size 432x288 with 0 Axes>
```

[25]:
```
z = df.groupby('seafood_meal')[dairy[1]].plot.kde(title = dairy[1], legend='x')
plt.show(z[0].set_xlim(0, 0.3))
plt.clf()
```



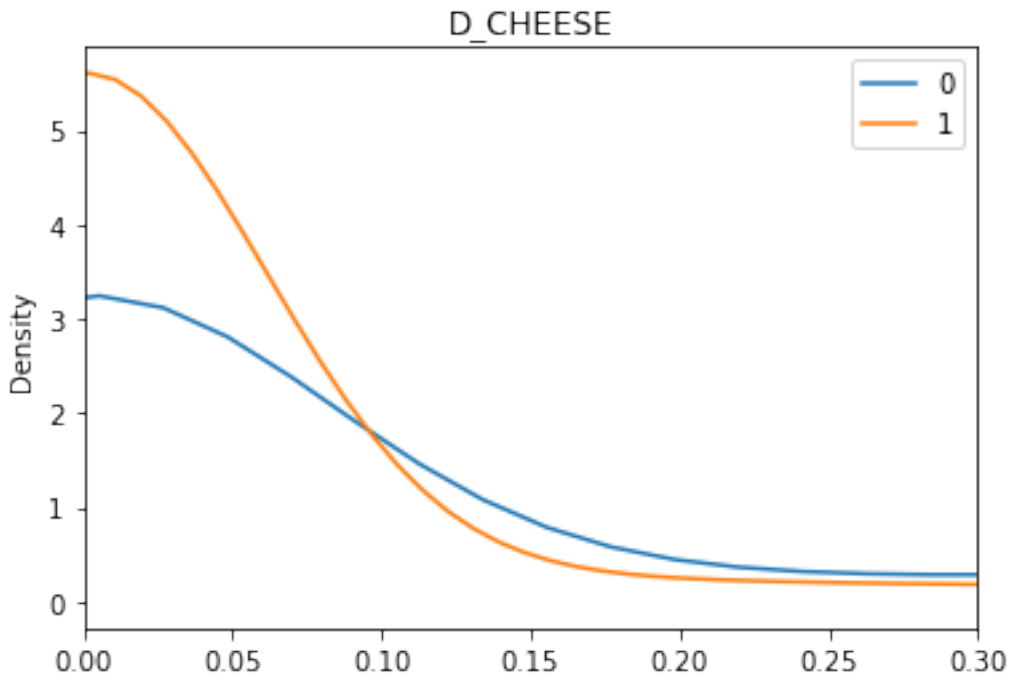

```
<Figure size 432x288 with 0 Axes>
```

[26]:
```
z = df.groupby('seafood_meal')[dairy[2]].plot.kde(title = dairy[2], legend='x')
plt.show(z[0].set_xlim(0, 0.05))
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

[27]:
```
z = df.groupby('seafood_meal')[dairy[3]].plot.kde(title = dairy[3], legend='x')
plt.show(z[0].set_xlim(0, 0.3))
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

[28]:
```
for var in dairy:
    z = df.groupby('seafood_meal')[var].describe()
    print("Statistics for "+var+'\n')
    print(z)
    print('\n')
```

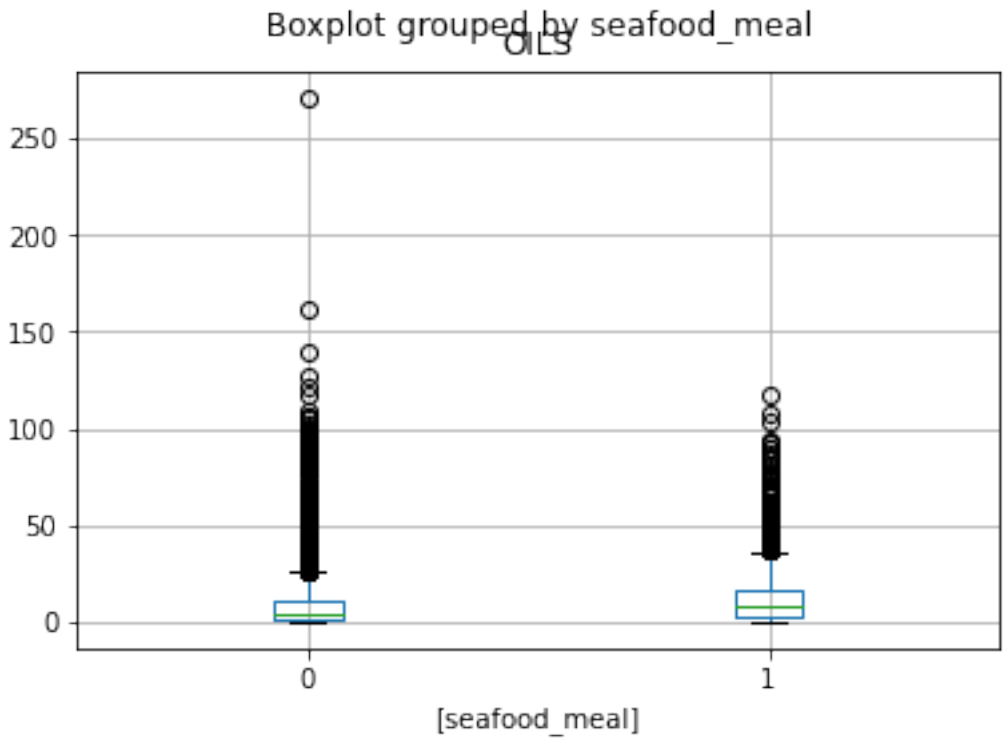
```
Statistics for D_TOTAL

                count      mean       std  min  25%   50%   75%    max
seafood_meal
0             26011.0  0.443461  0.764443  0.0  0.0  0.06  0.65  10.76
1              3232.0  0.209084  0.490321  0.0  0.0  0.00  0.15   5.07


Statistics for D_MILK

                count      mean       std  min  25%  50%   75%   max
seafood_meal
0             26011.0  0.136574  0.428739  0.0  0.0  0.0  0.04  8.04
1              3232.0  0.111470  0.365981  0.0  0.0  0.0  0.02  5.07
```

```
Statistics for D_YOGURT

                count      mean       std  min  25%  50%  75%   max
seafood_meal
0             26011.0  0.004826  0.059659  0.0  0.0  0.0  0.0  2.68
1              3232.0  0.004947  0.052113  0.0  0.0  0.0  0.0  1.02


Statistics for D_CHEESE

                count      mean       std  min  25%  50%  75%    max
seafood_meal
0             26011.0  0.299695  0.623149  0.0  0.0  0.0  0.4  10.71
1              3232.0  0.089947  0.306277  0.0  0.0  0.0  0.0   4.29
```

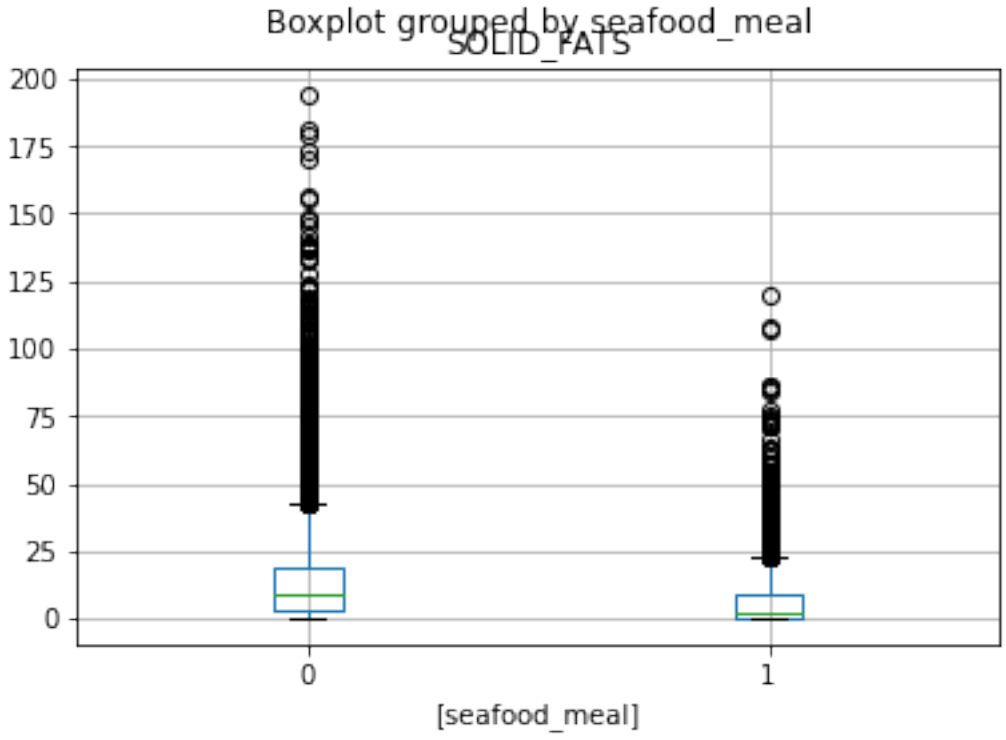
Section 6: Others

This section provides boxplots and density plots of the other FPED components in the seafood meal and non seafood meal groups. The code for seafood meal is 1 if meal contains seafood, and 0 if meal does not contain seafood.

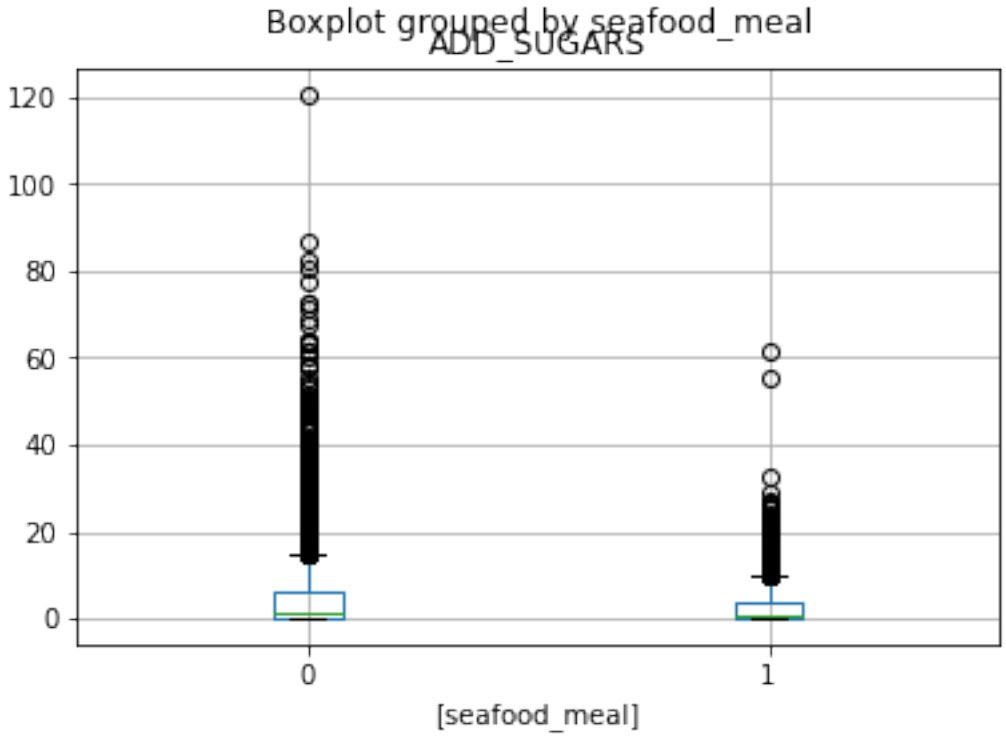
[29]:
```
other = ['OILS', 'SOLID_FATS', 'ADD_SUGARS', 'A_DRINKS']

for var in other:
    z = df.boxplot(column=var,by=['seafood_meal'])
    plt.show(z)
    plt.clf()
```



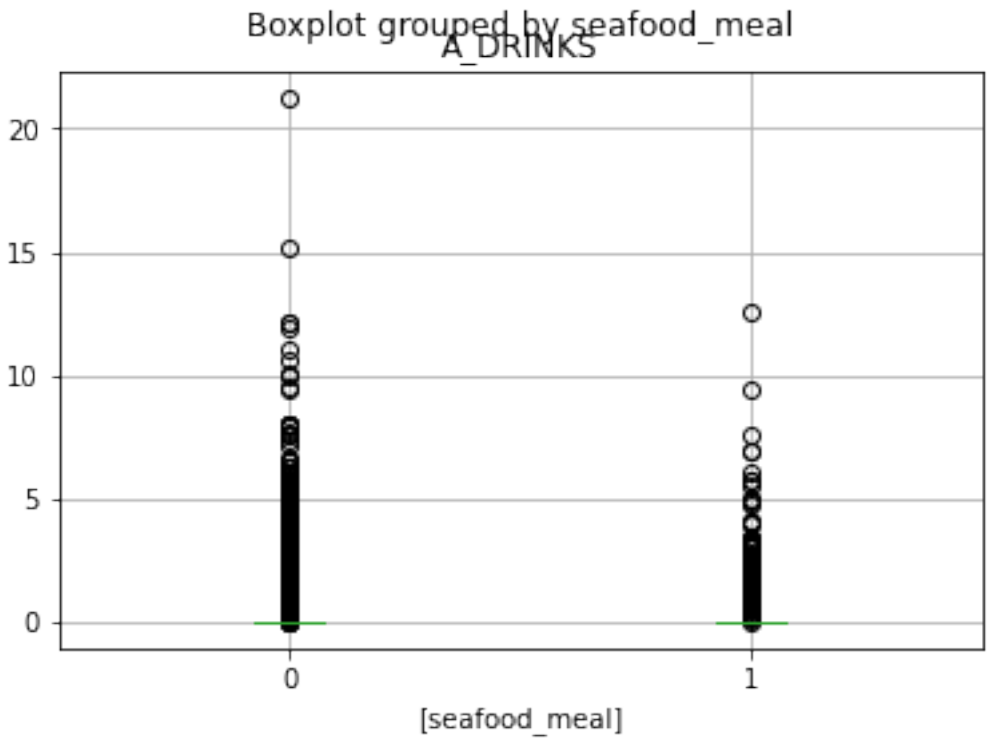

```
<Figure size 432x288 with 0 Axes>
```



```
<Figure size 432x288 with 0 Axes>
```

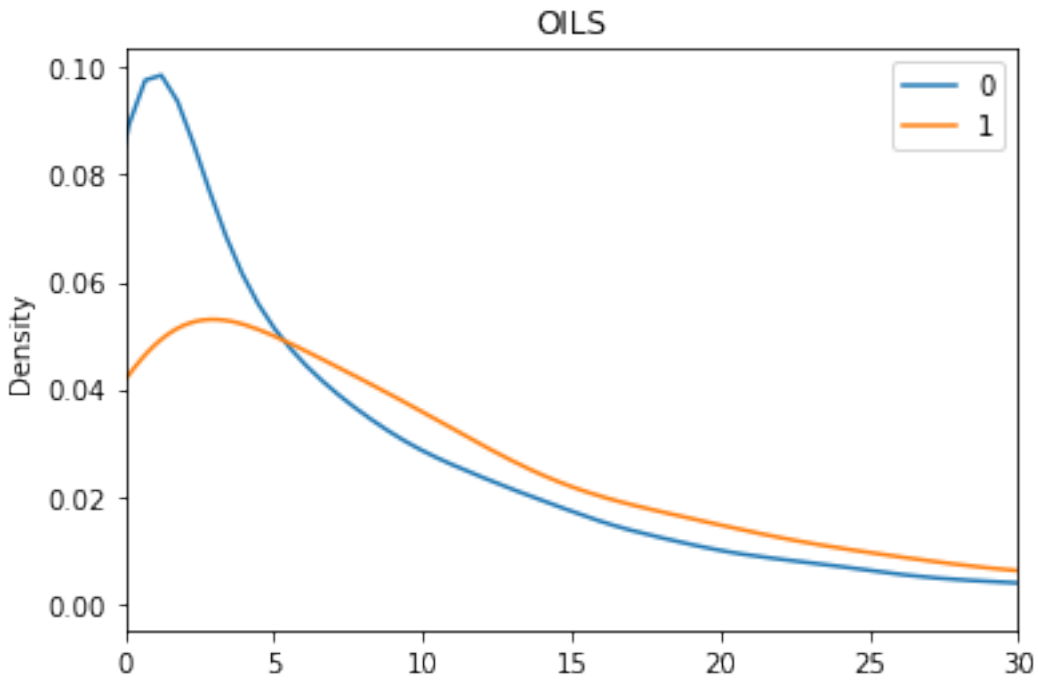



```
<Figure size 432x288 with 0 Axes>
```



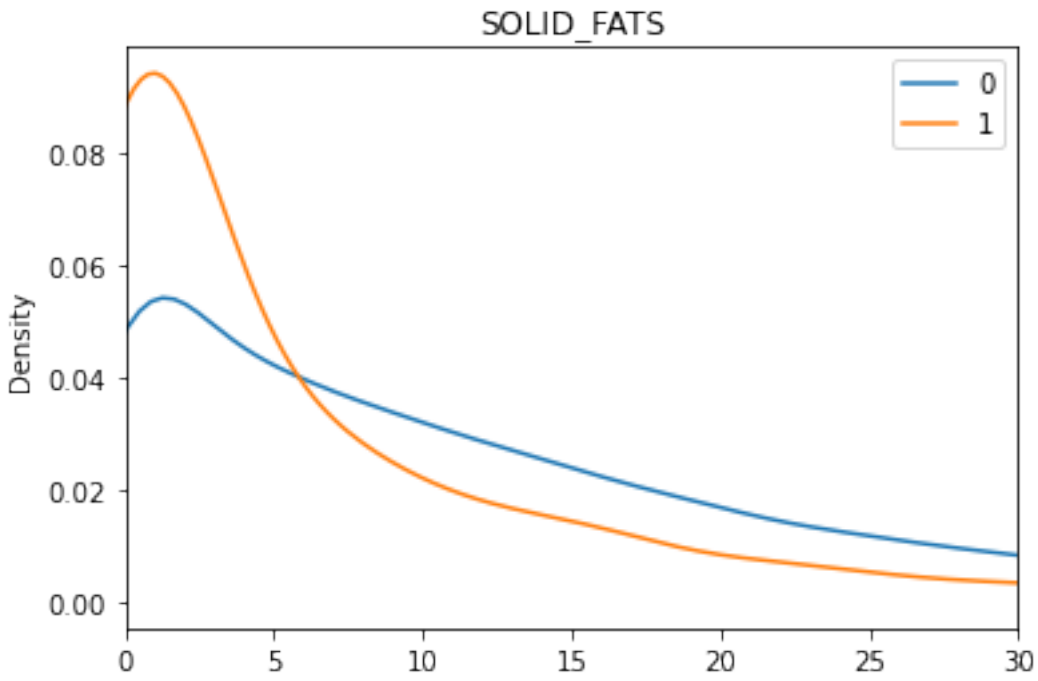

```
<Figure size 432x288 with 0 Axes>
```

[30]:
```
z = df.groupby('seafood_meal')[other[0]].plot.kde(title = other[0], legend='x')
plt.show(z[0].set_xlim(0, 30))
plt.clf()
```



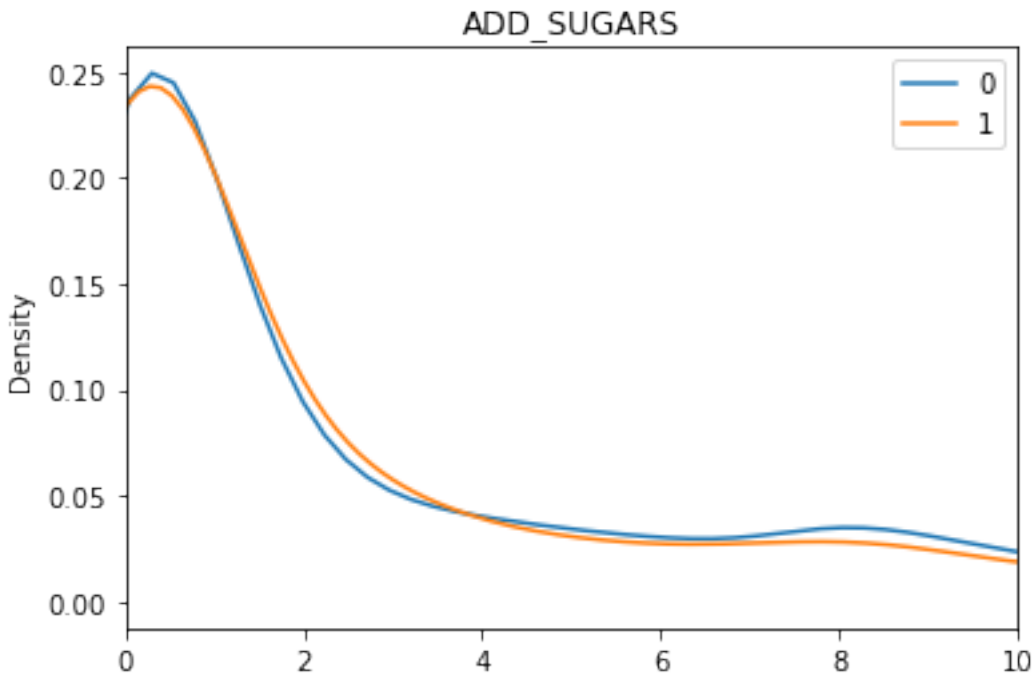
<Figure size 432x288 with 0 Axes>

[31]:
```
z = df.groupby('seafood_meal')[other[1]].plot.kde(title = other[1], legend='x')
plt.show(z[0].set_xlim(0, 30))
plt.clf()
```



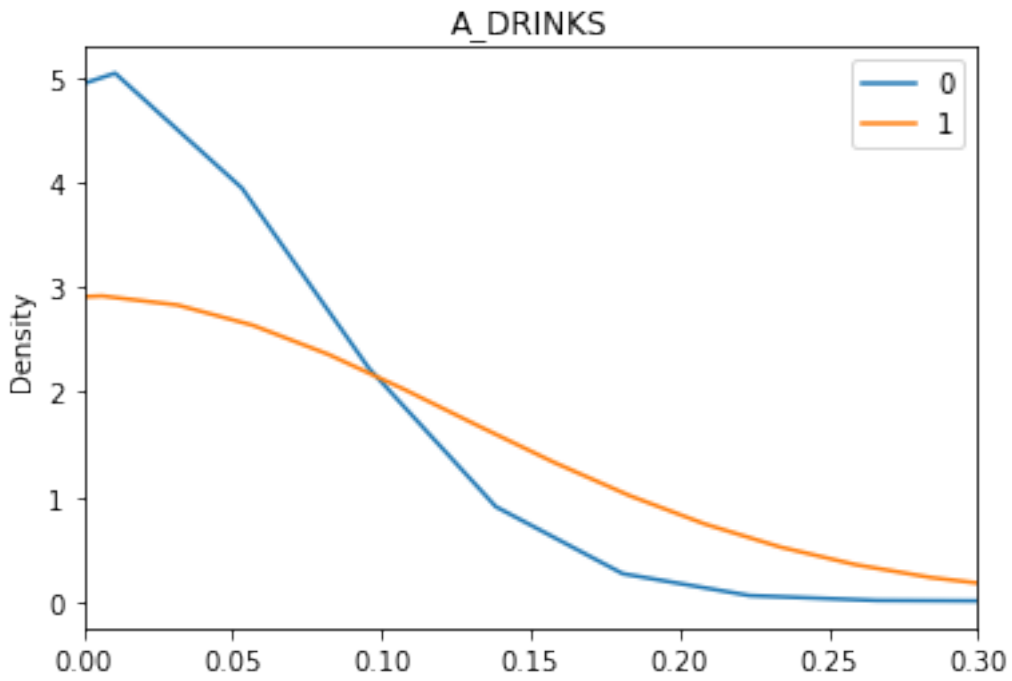

```
<Figure size 432x288 with 0 Axes>
```

[32]:
```
z = df.groupby('seafood_meal')[other[2]].plot.kde(title = other[2], legend='x')
plt.show(z[0].set_xlim(0, 10))
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

[33]:
```
z = df.groupby('seafood_meal')[other[3]].plot.kde(title = other[3], legend='x')
plt.show(z[0].set_xlim(0, 0.3))
plt.clf()
```



<Figure size 432x288 with 0 Axes>

[34]:
```
for var in other:
    z = df.groupby('seafood_meal')[var].describe()
    print("Statistics for "+var+'\n')
    print(z)
    print('\n')
```

Statistics for OILS

seafood_meal	count	mean	std	min	25%	50%	75%	max
0	26011.0	8.186339	10.844504	0.0	1.23	4.51	11.3700	271.03
1	3232.0	11.997413	13.827625	0.0	2.74	7.84	16.3275	117.17

Statistics for SOLID_FATS

seafood_meal	count	mean	std	min	25%	50%	75%	max
0	26011.0	13.294004	14.985754	0.0	2.77	9.000	18.83	194.29
1	3232.0	7.082955	11.323490	0.0	0.02	2.535	9.24	119.91

```
Statistics for ADD_SUGARS

                count      mean       std  min  25%   50%   75%     max
seafood_meal
0             26011.0  3.867824  5.930227  0.0  0.1  1.12  5.92  120.61
1              3232.0  2.961179  4.698025  0.0  0.0  0.75  4.01   61.64


Statistics for A_DRINKS

                count      mean       std  min  25%  50%  75%    max
seafood_meal
0             26011.0  0.098922  0.568399  0.0  0.0  0.0  0.0  21.26
1              3232.0  0.143902  0.632477  0.0  0.0  0.0  0.0  12.62
```

[ ]:

[ ]: