Food_db

March 21, 2025

Food Analysis- A deep dive into food

Food and nutrition are how we obtain fuel to energize our bodies, requiring daily replenishment of essential nutrients. Mindful eating and understanding nutrition are crucial for both physical health and mental well-being. Diets rich in vegetables, fruits, and whole grains can help prevent chronic illnesses like heart disease and diabetes.

This notebook explores common foods and their nutritional profiles to highlight what fuels our bodies best.

You'll find the step-by-step guide and relevant website links at the bottom of the PDF. The original dataset is available on my GitHub: https://github.com/AlexStcR

[495]:
```
import pandas as pd
import matplotlib.pyplot as plt

#load nutrition file
foods = pd.read_csv('nutrients_csvfile.csv')
print(foods.head())
```

```
                   Food Measure  Grams Calories Protein Fat Sat.Fat Fiber  \
0             Cows' milk   1 qt.    976      660      32  40      36     0
1              Milk skim   1 qt.    984      360      36   t       t     0
2             Buttermilk   1 cup    246      127       9   5       4     0
3  Evaporated, undiluted   1 cup    252      345      16  20      18     0
4         Fortified milk  6 cups  1,419    1,373      89  42      23   1.4

  Carbs        Category
0    48  Dairy products
1    52  Dairy products
2    13  Dairy products
3    24  Dairy products
4   119  Dairy products
```

[497]:
```
fill_df = foods.fillna(0)
```

[499]:
```
#let's find out more about the t and if there are more t's on the columns
substring = 't'
foods['NameContainsSubstring'] = foods['Protein'].str.contains(substring)
filtered_df = foods[foods['NameContainsSubstring']]
```

```
print(filtered_df)
```

```
                          Food        Measure Grams Calories Protein Fat  \
29                      Butter            1T.    14      100       t  11
34                   Margarine        1/2 cup   112      806       t  91
35           Margarine, 2 pat or           1 T.    14      100       t  11
36                   Mayonnaise           1 T.    15      110       t  12
40               French dressing           1 T.    15       60       t   6
41         Thousand Island sauce           1 T.    15       75       t   8
107             Strips, from raw         1 mad.    50       20       t   t
115                    Cucumbers              8    50        6       t   0
124                      Iceberg       1/4 head   100       13       t   t
129                  Raw, green        6 small    50       22       t   t
130                      Parsley           2 T.    50        2       t   t
137               Peppers canned          1 pod    38       10       t   t
147                    Radishes        5 small    50       10       t   0
148                   Rutabagas         4 cups   100       32       t   0
158               Tomato catsup           1 T.    17       15       t   t
162          Apple juice canned          1 cup   250      125       t   0
163               Apple vinegar        1/3 cup   100       14       t   0
164                 Apples, raw          1 med   130       70       t   t
165             Stewed or canned          1 cup   240      100       t   t
177   Cranberry sauce sweetened           1 cup   277      530       t   t
189                  Lemon juice        1/2 cup   125       30       t   t
190  Lemonade concentratefrozen       6-oz. can   220      430       t   t
191  Limeade concentrate frozen       6-oz. can   218      405       t   t
203                   Pineapple  1 large slice   122       95       t   t
208             Raw, 2" diameter               1    60       30       t   t
212                  Raspberries        1/2 cup   100      100       t   t
213                     Raw, red        3/4 cup   100       57       t   t
216                          Raw          1 cup   149       54       t   t
245                  Puffed rice          1 cup    14       55       t   t
283                        Candy              5    25      104       t   3
284             Chocolate creams              2    30      130       t   4
285                        Fudge       2 pieces    90      370       t  12
286                Hard candies           1 oz.    28       90       t   0
289              Chocolate syrup           2 T.    40       80       t   t
292                        Honey           2 T.    42      120       t   0
323                         Beer         2 cups   480      228       t   0
325                        Wines        1/2 cup   120      164       t   0
326       Table (12.2% alcohol)         1/2 cup   120      100       t   0
333                       Coffee          1 cup   230        3       t   0

    Sat.Fat Fiber Carbs                         Category  \
29       10     0     t          Fats, Oils, Shortenings
34       76     0     t          Fats, Oils, Shortenings
35        9     0     t          Fats, Oils, Shortenings
36        5     0     t          Fats, Oils, Shortenings
```

40	2	0	2	Fats, Oils, Shortenings
41	3	0	1	Fats, Oils, Shortenings
107	0	0.5	5	Vegetables A-E
115	0	0.2	1	Vegetables A-E
124	0	0.5	3	Vegetables F-P
129	0	1	5	Vegetables F-P
130	0	t	t	Vegetables F-P
137	0	t	2	Vegetables R-Z
147	0	0.3	2	Vegetables R-Z
148	0	1.4	8	Vegetables R-Z
158	0	t	4	Vegetables R-Z
162	0	0	34	Fruits A-F
163	0	0	3	Fruits A-F
164	0	1	18	Fruits A-F
165	0	2	26	Fruits A-F
177	0	1.2	142	Fruits A-F
189	0	t	10	Fruits G-P
190	0	t	112	Fruits G-P
191	0	t	108	Fruits G-P
203	0	0.4	26	Fruits G-P
208	0	0.2	7	Fruits G-P
212	0	2	25	Fruits R-Z
213	0	5	14	Fruits R-Z
216	0	1.9	12	Fruits R-Z
245	0	t	12	Breads, cereals, fastfood,grains
283	3	0	19	Desserts, sweets
284	4	0	24	Desserts, sweets
285	11	0.1	80	Desserts, sweets
286	0	0	28	Desserts, sweets
289	t	0	22	Desserts, sweets
292	0	0	30	Jams, Jellies
323	0	0	8	Drinks,Alcohol, Beverages
325	0	0	9	Drinks,Alcohol, Beverages
326	0	0	5	Drinks,Alcohol, Beverages
333	0	0	1	Drinks,Alcohol, Beverages

	NameContainsSubstring
29	True
34	True
35	True
36	True
40	True
41	True
107	True
115	True
124	True
129	True
130	True

```
137                  True
147                  True
148                  True
158                  True
162                  True
163                  True
164                  True
165                  True
177                  True
189                  True
190                  True
191                  True
203                  True
208                  True
212                  True
213                  True
216                  True
245                  True
283                  True
284                  True
285                  True
286                  True
289                  True
292                  True
323                  True
325                  True
326                  True
333                  True
```

[501]:
```
substring = 't'
foods['NameContainsSubstring'] = foods['Carbs'].str.contains(substring)
filtered_df = foods[foods['NameContainsSubstring']]
print(filtered_df)
```

```
                   Food     Measure Grams Calories Protein Fat Sat.Fat Fiber  \
20               Cheddar  1-in. cube    17       70       4   6       5     0
23     Processed cheese       1 oz.    28      105       7   9       8     0
24       Roquefort type       1 oz.    28      105       6   9       8     0
25                Swiss       1 oz.    28      105       7   8       7     0
26            Eggs raw           2   100      150      12  12      10     0
28               Yolks           2    34      120       6  10       8     0
29              Butter         1T.    14      100       t  11      10     0
34           Margarine     1/2 cup   112      806       t  91      76     0
35  Margarine, 2 pat or        1 T.    14      100       t  11       9     0
36          Mayonnaise        1 T.    15      110       t  12       5     0
80             Lobster       aver.   100       92      18   1       0     0
130            Parsley        2 T.    50        2       t   t       0     t
324                Gin       1 oz.    28       70       0   0       0     0
```

```
    Carbs                  Category  NameContainsSubstring
20      t            Dairy products                   True
23      t            Dairy products                   True
24      t            Dairy products                   True
25      t            Dairy products                   True
26      t            Dairy products                   True
28      t   Fats, Oils, Shortenings                   True
29      t   Fats, Oils, Shortenings                   True
34      t   Fats, Oils, Shortenings                   True
35      t   Fats, Oils, Shortenings                   True
36      t   Fats, Oils, Shortenings                   True
80      t             Fish, Seafood                   True
130     t            Vegetables F-P                   True
324     t  Drinks,Alcohol, Beverages                   True
```

[503]:
```
#some data are showing t , I checked and it means minuscule amounts, close to␣
 ↪zero, let's replace it for zero
foods=foods.replace("t",0)
foods=foods.replace("t'",0)
```

[505]:
```
print(foods.head())
```

```
                   Food Measure  Grams Calories Protein Fat Sat.Fat Fiber  \
0            Cows' milk   1 qt.    976      660      32  40      36     0
1             Milk skim   1 qt.    984      360      36   0       0     0
2            Buttermilk   1 cup    246      127       9   5       4     0
3  Evaporated, undiluted   1 cup    252      345      16  20      18     0
4         Fortified milk  6 cups  1,419    1,373      89  42      23   1.4

  Carbs        Category  NameContainsSubstring
0    48  Dairy products                 False
1    52  Dairy products                 False
2    13  Dairy products                 False
3    24  Dairy products                 False
4   119  Dairy products                 False
```

[507]:
```
#check for non-number in calories,grams,proteins, fat,sat fat, fver,carbs
obj_cols = foods.select_dtypes(include=['object']).columns
```

[509]:
```
print(obj_cols)
```

```
Index(['Food', 'Measure', 'Grams', 'Calories', 'Protein', 'Fat', 'Sat.Fat',
       'Fiber', 'Carbs', 'Category'],
      dtype='object')
```

[511]:
```
#Let's find some nan in the df
nan_in_df = foods.isnull().sum().any()
```

```
# Print the dataframe
print(nan_in_df)
```

```
True
```

[513]:
```
#finding nan in the db
nan_in_df = foods.isnull().sum()
print(nan_in_df)
```

```
Food                    0
Measure                 0
Grams                   0
Calories                1
Protein                 0
Fat                     0
Sat.Fat                 2
Fiber                   0
Carbs                   0
Category                0
NameContainsSubstring   0
dtype: int64
```

[515]:
```
nan_rows = foods.isna().any(axis=1)
print(nan_rows)
```

```
0      False
1      False
2      False
3      False
4      False
       …
330    False
331    False
332    False
333    False
334    False
Length: 335, dtype: bool
```

[517]:
```
fill_df = foods.fillna(0)
```

[519]:
```
# dropping null value columns to avoid errors
foods.dropna(inplace = True)
```

[521]:
```
#check again, now there's no nana
nan_in_df = foods.isnull().sum()
print(nan_in_df)
```

```
Food                    0
Measure                 0
```

```
Grams                    0
Calories                 0
Protein                  0
Fat                      0
Sat.Fat                  0
Fiber                    0
Carbs                    0
Category                 0
NameContainsSubstring    0
dtype: int64
```

[ ]:

[524]:
```
is_numeric = foods["Calories"].str.isdecimal()

# Use boolean indexing to select rows that meet the condition
foods= foods[is_numeric]

# Print the result
print(foods.head())
```

```
                   Food Measure Grams Calories Protein Fat Sat.Fat Fiber  \
0             Cows' milk   1 qt.   976      660      32  40      36     0
1              Milk skim   1 qt.   984      360      36   0       0     0
2             Buttermilk   1 cup   246      127       9   5       4     0
3  Evaporated, undiluted   1 cup   252      345      16  20      18     0
5          Powdered milk   1 cup   103      515      27  28      24     0

  Carbs        Category  NameContainsSubstring
0    48  Dairy products                 False
1    52  Dairy products                 False
2    13  Dairy products                 False
3    24  Dairy products                 False
5    39  Dairy products                 False
```

[526]:
```
substring = 'NA'
foods['NameContainsSubstring'] = foods['Calories'].str.contains(substring)
filtered_df = foods[foods['NameContainsSubstring']]
print(filtered_df)
```

```
Empty DataFrame
Columns: [Food, Measure, Grams, Calories, Protein, Fat, Sat.Fat, Fiber, Carbs,
Category, NameContainsSubstring]
Index: []
```

[528]:
```
#Now we removed NA,t's and Nan from the dataset
```

[530]:
```
#Now we have to transform all the data in
foods['Grams']=pd.to_numeric(foods['Grams'])
foods['Calories']=pd.to_numeric(foods['Calories'])
foods['Protein']=pd.to_numeric(foods['Protein'])
foods['Fat']=pd.to_numeric(foods['Fat'])
foods['Sat.Fat']=pd.to_numeric(foods['Sat.Fat'])
foods['Fiber']=pd.to_numeric(foods['Fiber'])
foods['Carbs']=pd.to_numeric(foods['Carbs'])
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
File lib.pyx:2391, in pandas._libs.lib.maybe_convert_numeric()

ValueError: Unable to parse string "a"

During handling of the above exception, another exception occurred:

ValueError                                Traceback (most recent call last)
Cell In[530], line 7
      5 foods['Fat']=pd.to_numeric(foods['Fat'])
      6 foods['Sat.Fat']=pd.to_numeric(foods['Sat.Fat'])
----> 7 foods['Fiber']=pd.to_numeric(foods['Fiber'])
      8 foods['Carbs']=pd.to_numeric(foods['Carbs'])

File ~\.conda\envs\DataEnv\lib\site-packages\pandas\core\tools\numeric.py:232,␣
 ↪in to_numeric(arg, errors, downcast, dtype_backend)
    230 coerce_numeric = errors not in ("ignore", "raise")
    231 try:
--> 232     values, new_mask =␣
 ↪lib.maybe_convert_numeric(  # type: ignore[call-overload]
    233         values,
    234         set(),
    235         coerce_numeric=coerce_numeric,
    236         convert_to_masked_nullable=dtype_backend is not lib.no_default
    237         or isinstance(values_dtype, StringDtype)
    238         and not values_dtype.storage == "pyarrow_numpy",
    239     )
    240 except (ValueError, TypeError):
    241     if errors == "raise":

File lib.pyx:2433, in pandas._libs.lib.maybe_convert_numeric()

ValueError: Unable to parse string "a" at position 79
```

[531]:
```
#There's an error, an a in position 79
foods=foods.replace(",","", regex=True)
```

```
foods['Fiber']=foods['Fiber'].replace("a","", regex=True)
```

[534]:
```
#Now we have to transform all the data in numeric!
foods['Grams']=pd.to_numeric(foods['Grams'])
foods['Calories']=pd.to_numeric(foods['Calories'])
foods['Protein']=pd.to_numeric(foods['Protein'])
foods['Fat']=pd.to_numeric(foods['Fat'])
foods['Sat.Fat']=pd.to_numeric(foods['Sat.Fat'])
foods['Fiber']=pd.to_numeric(foods['Fiber'])
foods['Carbs']=pd.to_numeric(foods['Carbs'])
```

[536]:
```
#check if it's all numbers
foods.dtypes
```

[536]:
```
Food                       object
Measure                    object
Grams                       int64
Calories                    int64
Protein                     int64
Fat                         int64
Sat.Fat                     int64
Fiber                     float64
Carbs                     float64
Category                   object
NameContainsSubstring        bool
dtype: object
```

[538]:
```
print(foods.isnull().any())
print('-'*245)
print(foods.describe())
print('-'*245)
```

```
Food                      False
Measure                   False
Grams                     False
Calories                  False
Protein                   False
Fat                       False
Sat.Fat                   False
Fiber                      True
Carbs                     False
Category                  False
NameContainsSubstring     False
dtype: bool
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
-----
```

```
            Grams    Calories     Protein         Fat     Sat.Fat       Fiber  \
count  328.000000  328.000000  328.000000  328.000000  328.000000  327.000000
mean   137.878049  179.466463    8.152439    8.338415    6.399390    2.177095
std    118.362982  154.977414   17.128082   19.733038   18.625775   15.837890
min     11.000000    0.000000   -1.000000    0.000000    0.000000    0.000000
25%     59.250000   75.000000    1.000000    0.000000    0.000000    0.000000
50%    108.000000  130.000000    3.000000    1.000000    0.000000    0.200000
75%    200.000000  245.000000   12.000000   10.000000    7.000000    1.000000
max    984.000000  992.000000  232.000000  233.000000  234.000000  235.000000

            Carbs
count  328.000000
mean    23.726524
std     32.235295
min      0.000000
25%      3.000000
50%     14.000000
75%     30.000000
max    236.000000
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
-----
```

[540]:
```
foods=foods.dropna()
foods.shape
```

[540]: (327, 11)

[542]:
```
print(foods.head(20))
```

```
                     Food     Measure  Grams  Calories  Protein  Fat  \
0              Cows' milk       1 qt.    976       660       32   40
1               Milk skim       1 qt.    984       360       36    0
2              Buttermilk       1 cup    246       127        9    5
3     Evaporated undiluted       1 cup    252       345       16   20
5            Powdered milk       1 cup    103       515       27   28
6             skim instant  1 1/3 cups     85       290       30    0
7        skim non-instant     2/3 cup     85       290       30    0
8             Goats' milk       1 cup    244       165        8   10
9      (1/2 cup ice cream)      2 cups    540       690       24   24
10                   Cocoa       1 cup    252       235        8   11
11             skim. milk       1 cup    250       128       18    4
12            (cornstarch)       1 cup    248       275        9   10
13                 Custard       1 cup    248       285       13   14
14               Ice cream       1 cup    188       300        6   18
15                Ice milk       1 cup    190       275        9   10
16  Cream or half-and-half     1/2 cup    120       170        4   15
```

```
17              or whipping     1/2 cup    119       430        2   44
18                   Cheese       1 cup    225       240       30   11
19                uncreamed       1 cup    225       195       38    0
20                  Cheddar  1-in. cube     17        70        4    6

    Sat.Fat  Fiber  Carbs        Category  NameContainsSubstring
0        36    0.0   48.0  Dairy products                  False
1         0    0.0   52.0  Dairy products                  False
2         4    0.0   13.0  Dairy products                  False
3        18    0.0   24.0  Dairy products                  False
5        24    0.0   39.0  Dairy products                  False
6         0    0.0   42.0  Dairy products                  False
7         0    1.0   42.0  Dairy products                  False
8         8    0.0   11.0  Dairy products                  False
9        22    0.0   70.0  Dairy products                  False
10       10    0.0   26.0  Dairy products                  False
11        3    1.0   13.0  Dairy products                  False
12        9    0.0   40.0  Dairy products                  False
13       11    0.0   28.0  Dairy products                  False
14       16    0.0   29.0  Dairy products                  False
15        9    0.0   32.0  Dairy products                  False
16       13    0.0    5.0  Dairy products                  False
17       27    1.0    3.0  Dairy products                  False
18       10    0.0    6.0  Dairy products                  False
19        0    0.0    6.0  Dairy products                  False
20        5    0.0    0.0  Dairy products                  False
```

[544]:
```
#l;et's transform the dataset all in lowercase,it will help us to analyze it
```

[546]:
```
# Use the str.lower() method to lowercase the 'name' column
foods['Food'] = foods['Food'].str.lower()
foods['Category'] = foods['Category'].str.lower()
```

[548]:
```
print(foods.head(10))
```

```
                   Food     Measure  Grams  Calories  Protein  Fat  Sat.Fat  \
0             cows' milk       1 qt.    976       660       32   40       36
1              milk skim       1 qt.    984       360       36    0        0
2            buttermilk       1 cup    246       127        9    5        4
3   evaporated undiluted       1 cup    252       345       16   20       18
5          powdered milk       1 cup    103       515       27   28       24
6           skim instant  1 1/3 cups     85       290       30    0        0
7       skim non-instant     2/3 cup     85       290       30    0        0
8            goats' milk       1 cup    244       165        8   10        8
9    (1/2 cup ice cream)      2 cups    540       690       24   24       22
10                 cocoa       1 cup    252       235        8   11       10

    Fiber  Carbs        Category  NameContainsSubstring
```

```
0     0.0   48.0  dairy products                 False
1     0.0   52.0  dairy products                 False
2     0.0   13.0  dairy products                 False
3     0.0   24.0  dairy products                 False
5     0.0   39.0  dairy products                 False
6     0.0   42.0  dairy products                 False
7     1.0   42.0  dairy products                 False
8     0.0   11.0  dairy products                 False
9     0.0   70.0  dairy products                 False
10    0.0   26.0  dairy products                 False
```

[550]:
```
#the measures are not padronized, let's try to do it
foods['scaling_factor'] = 100 / foods['Grams']

nutritional_cols = ['Calories', 'Protein', 'Fat']
```

[552]:
```
foods[nutritional_cols] = foods[nutritional_cols].mul(foods['scaling_factor'],␣
↪axis=0)
```

[554]:
```
foods['Measure'] = '100g'

foods['Grams'] = 100
```

[556]:
```
foods = foods.drop(columns=['scaling_factor'])
```

[558]:
```
foods[nutritional_cols] = foods[nutritional_cols].round(2)
```

[560]:
```
print(foods)
```

```
                    Food Measure  Grams  Calories  Protein    Fat  Sat.Fat  \
0              cows' milk    100g    100     67.62     3.28   4.10       36
1               milk skim    100g    100     36.59     3.66   0.00        0
2              buttermilk    100g    100     51.63     3.66   2.03        4
3    evaporated undiluted    100g    100    136.90     6.35   7.94       18
5           powdered milk    100g    100    500.00    26.21  27.18       24
..                    …     …     …        …       …      …        …
330   fruit-flavored soda    100g    100     46.53     0.00   0.00        0
331            ginger ale    100g    100     30.35     0.00   0.00        0
332             root beer    100g    100     40.46     0.00   0.00        0
333                coffee    100g    100      1.30     0.00   0.00        0
334                   tea    100g    100      1.74     0.00   0.00        0

     Fiber  Carbs               Category  NameContainsSubstring
0      0.0   48.0  dairy products                 False
1      0.0   52.0  dairy products                 False
2      0.0   13.0  dairy products                 False
3      0.0   24.0  dairy products                 False
5      0.0   39.0  dairy products                 False
```

```
..     …    …                        …                        …
330    0.0   42.0  drinksalcohol beverages                  False
331    0.0   28.0  drinksalcohol beverages                  False
332    0.0   35.0  drinksalcohol beverages                  False
333    0.0    1.0  drinksalcohol beverages                  False
334    0.0    1.0  drinksalcohol beverages                  False

[327 rows x 11 columns]
```

[562]:
```
'''After some data cleaning, let's dive into plots and graphics'''
```

[562]:
```
"After some data cleaning, let's dive into plots and graphics"
```

[564]:
```
#find categories
```

[566]:
```
food_veg=foods[foods['Category'].str.contains("vegetables")]
print(food_veg.head(20))
```

```
                    Food Measure  Grams  Calories  Protein   Fat  Sat.Fat  \
92               asparagus    100g    100     18.75     1.04  0.00        0
93                   beans    100g    100     20.00     0.80  0.00        0
94                    lima    100g    100     87.50     5.00  0.00        0
95          lima dry cooked    100g    100    135.42     8.33  0.00        0
96   navy baked with pork    100g    100    125.00     5.50  3.00        6
97              red kidney    100g    100     88.46     5.77  0.38        0
98            bean sprouts    100g    100     34.00     2.00  0.00        0
99             beet greens    100g    100     27.00     2.00  0.00        0
101               broccoli    100g    100     30.00     3.33  0.00        0
102      brussels sprouts    100g    100     46.15     4.62  0.00        0
103             sauerkraut    100g    100     21.33     0.67  0.00        0
104       steamed cabbage    100g    100     23.53     1.18  0.00        0
105                carrots    100g    100     30.00     0.67  0.00        0
106             raw grated    100g    100     40.91     0.91  0.00        0
107       strips from raw    100g    100     40.00     0.00  0.00        0
108            cauliflower    100g    100     25.00     2.50  0.00        0
109                  celery    100g    100     20.00     1.00  0.00        0
110              stalk raw    100g    100     12.50     2.50  0.00        0
111          chard steamed    100g    100     20.00     1.33  0.00        0
112               collards    100g    100     34.00     3.33  0.00        0

     Fiber  Carbs        Category  NameContainsSubstring
92     0.5    3.0  vegetables a-e                  False
93     0.8    6.0  vegetables a-e                  False
94     3.0   24.0  vegetables a-e                  False
95     2.0   48.0  vegetables a-e                  False
96     2.0   37.0  vegetables a-e                  False
97     2.5   42.0  vegetables a-e                  False
98     0.3    3.0  vegetables a-e                  False
```

```
99     1.4    6.0  vegetables a-e                False
101    1.9    8.0  vegetables a-e                False
102    1.7   12.0  vegetables a-e                False
103    1.2    7.0  vegetables a-e                False
104    1.3    9.0  vegetables a-e                False
105    0.9   10.0  vegetables a-e                False
106    1.2   10.0  vegetables a-e                False
107    0.5    5.0  vegetables a-e                False
108    1.0    6.0  vegetables a-e                False
109    1.0    4.0  vegetables a-e                False
110    0.3    1.0  vegetables a-e                False
111    1.4    7.0  vegetables a-e                False
112    2.0    8.0  vegetables a-e                False
```

[567]:
```
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px

 #Let's find some foods rich in protein

#vegetables rich in protein
vegetables=['vegetables a-e',
 'vegetables f-p',
 'vegetables r-z']

prot= foods[foods['Category'].isin(vegetables)]
protein= prot.sort_values(by='Protein', ascending= False)
top_10=protein.head(10)
fig = px.bar(top_10, x='Food', y='Protein', color='Protein', title=' Top 10␣
 ↪protein rich vegetables')
fig.show()
```

[569]:
```
food_seeds=foods[foods['Category'].str.contains("seeds")]
print(food_seeds.head(50))
```

```
                    Food Measure  Grams  Calories  Protein    Fat  Sat.Fat  \
311               almonds    100g    100    607.14    18.57  54.29       28
312    roasted and salted    100g    100    627.14    18.57  57.14       31
313           brazil nuts    100g    100    652.86    14.29  67.14       31
314               cashews    100g    100    560.00    17.14  45.71       28
315     coconut sweetened    100g    100    548.00     2.00  40.00       19
316         peanut butter    100g    100    600.00    24.00  50.00       17
317  peanut butter natural    100g    100    568.00    26.00  48.00       10
318               peanuts    100g    100    580.00    26.00  50.00       16
319                pecans    100g    100    659.62     9.62  67.31       25
320          sesame seeds    100g    100    560.00    18.00  48.00       13
321       sunflower seeds    100g    100    560.00    24.00  52.00        7
```

```
322                walnuts    100g    100    650.00    14.00  64.00        7

     Fiber  Carbs        Category  NameContainsSubstring
311    1.8   13.0  seeds and nuts                  False
312    1.8   13.0  seeds and nuts                  False
313    2.0    7.0  seeds and nuts                  False
314    0.9   20.0  seeds and nuts                  False
315    2.0   26.0  seeds and nuts                  False
316    0.9    9.0  seeds and nuts                  False
317    0.9    8.0  seeds and nuts                  False
318    1.2    9.0  seeds and nuts                  False
319    1.1    7.0  seeds and nuts                  False
320    3.1   10.0  seeds and nuts                  False
321    1.9   10.0  seeds and nuts                  False
322    1.0    8.0  seeds and nuts                  False
```

[570]:
```
seeds=['seeds and nuts']

prot= foods[foods['Category'].isin(seeds)]
protein= prot.sort_values(by='Protein', ascending= False)
top_10=protein.head(10)
fig = px.bar(top_10, x='Food', y='Protein', color='Protein', title=' Top 10␣
 ↪protein rich seeds and nuts')
fig.show()
```

[571]:
```
food_seeds=foods[foods['Category'].str.contains("dairy")]
print(food_seeds.head(50))
```

```
                     Food Measure  Grams  Calories  Protein    Fat  Sat.Fat  \
0               cows' milk    100g    100     67.62     3.28   4.10       36
1                milk skim    100g    100     36.59     3.66   0.00        0
2               buttermilk    100g    100     51.63     3.66   2.03        4
3     evaporated undiluted    100g    100    136.90     6.35   7.94       18
5            powdered milk    100g    100    500.00    26.21  27.18       24
6             skim instant    100g    100    341.18    35.29   0.00        0
7         skim non-instant    100g    100    341.18    35.29   0.00        0
8              goats' milk    100g    100     67.62     3.28   4.10        8
9      (1/2 cup ice cream)    100g    100    127.78     4.44   4.44       22
10                   cocoa    100g    100     93.25     3.17   4.37       10
11              skim. milk    100g    100     51.20     7.20   1.60        3
12            (cornstarch)    100g    100    110.89     3.63   4.03        9
13                 custard    100g    100    114.92     5.24   5.65       11
14               ice cream    100g    100    159.57     3.19   9.57       16
15                ice milk    100g    100    144.74     4.74   5.26        9
16  cream or half-and-half    100g    100    141.67     3.33  12.50       13
17             or whipping    100g    100    361.34     1.68  36.97       27
18                  cheese    100g    100    106.67    13.33   4.89       10
19               uncreamed    100g    100     86.67    16.89   0.00        0
```

20	cheddar	100g	100	411.76	23.53	35.29	5
21	cheddar grated cup	100g	100	403.57	25.00	33.93	17
22	cream cheese	100g	100	375.00	7.14	39.29	10
23	processed cheese	100g	100	375.00	25.00	32.14	8
24	roquefort type	100g	100	375.00	21.43	32.14	8
25	swiss	100g	100	375.00	25.00	28.57	7
26	eggs raw	100g	100	150.00	12.00	12.00	10
27	eggs scrambled or fried	100g	100	171.88	10.16	12.50	14

	Fiber	Carbs	Category	NameContainsSubstring
0	0.0	48.0	dairy products	False
1	0.0	52.0	dairy products	False
2	0.0	13.0	dairy products	False
3	0.0	24.0	dairy products	False
5	0.0	39.0	dairy products	False
6	0.0	42.0	dairy products	False
7	1.0	42.0	dairy products	False
8	0.0	11.0	dairy products	False
9	0.0	70.0	dairy products	False
10	0.0	26.0	dairy products	False
11	1.0	13.0	dairy products	False
12	0.0	40.0	dairy products	False
13	0.0	28.0	dairy products	False
14	0.0	29.0	dairy products	False
15	0.0	32.0	dairy products	False
16	0.0	5.0	dairy products	False
17	1.0	3.0	dairy products	False
18	0.0	6.0	dairy products	False
19	0.0	6.0	dairy products	False
20	0.0	0.0	dairy products	False
21	0.0	1.0	dairy products	False
22	0.0	1.0	dairy products	False
23	0.0	0.0	dairy products	False
24	0.0	0.0	dairy products	False
25	0.0	0.0	dairy products	False
26	0.0	0.0	dairy products	False
27	0.0	1.0	dairy products	False

[572]:
```
dairy=['dairy products']

prot= foods[foods['Category'].isin(dairy)]
protein= prot.sort_values(by='Protein', ascending= False)
top_10=protein.head(10)
fig = px.bar(top_10, x='Food', y='Protein', color='Protein', title=' Top 10␣
 ↪protein rich dairy')
fig.show()
```

[573]:
```
food_seeds=foods[foods['Category'].str.contains("bread")]
print(food_seeds.head(50))
```

```
                          Food Measure  Grams  Calories  Protein    Fat  \
219                    biscuits    100g    100    342.11     7.89  10.53
220                bran flakes    100g    100    468.00    12.00   0.00
221          bread cracked wheat    100g    100    260.87     8.70   4.35
222                          rye    100g    100    239.13     8.70   4.35
225                 whole-wheat    100g    100    239.13     8.70   4.35
226       corn bread ground meal    100g    100    200.00     6.00   8.00
227                 cornflakes    100g    100    440.00     8.00   0.00
228           corn grits cooked    100g    100     49.59     3.31   0.00
229                  corn meal    100g    100    305.08     7.63   3.39
230                   crackers    100g    100    392.86     7.14   7.14
231           soda 2 1/2 square    100g    100    409.09     9.09   9.09
232                     farina    100g    100     44.12     1.26   0.00
233                      flour    100g    100    418.18    35.45  20.00
234          wheat (all purpose)    100g    100    363.64    10.91   0.91
235                wheat (whole)    100g    100    325.00    10.83   1.67
236                    macaroni    100g    100    110.71     3.57   0.71
237           baked with cheese    100g    100    215.91     8.18  11.36
238                     muffins    100g    100    281.25     8.33  10.42
239                     noodles    100g    100    125.00     4.38   1.25
240                     oatmeal    100g    100     63.56     2.12   1.27
241           pancakes 4" diam.    100g    100    231.48     6.48   8.33
242     wheat pancakes 4" diam.    100g    100    231.48     6.48   8.33
243             pizza 14" diam.    100g    100    240.00    10.67   8.00
244              popcorn salted    100g    100    542.86    10.71  25.00
245                 puffed rice    100g    100    392.86     0.00   0.00
246   puffed wheat presweetened    100g    100    375.00     3.57   0.00
247                        rice    100g    100    359.62     7.21   1.44
248                   converted    100g    100    362.03     7.49   0.00
249                       white    100g    100    362.30     7.33   0.00
250                rice flakes    100g    100    383.33     6.67   0.00
251                rice polish    100g    100    264.00    12.00  12.00
252                       rolls    100g    100    822.00     6.00  24.00
253           of refined flour    100g    100    302.63     7.89   5.26
254                 whole-wheat    100g    100    255.00    10.00   2.50
255   spaghetti with meat sauce    100g    100    114.00     5.20   4.00
256    with tomatoes and cheese    100g    100     84.00     2.40   2.00
257                spanish rice    100g    100     86.80     1.60   1.60
258      shredded wheat biscuit    100g    100    357.14    10.71   3.57
259                     waffles    100g    100    320.00    10.67  12.00
260                  wheat germ    100g    100    360.29    25.00  10.29
261   wheat-germ cereal toasted    100g    100    400.00    30.77  10.77
262  wheat meal cereal unrefined    100g    100    343.33    13.33   3.33
263                wheat cooked    100g    100    137.50     6.00   0.50
```

```
     Sat.Fat  Fiber  Carbs                       Category  \
219        3   0.00   18.0  breads cereals fastfoodgrains
220        0   0.10   32.0  breads cereals fastfoodgrains
221        1   0.10   12.0  breads cereals fastfoodgrains
222        1   0.10   12.0  breads cereals fastfoodgrains
225        0   0.31   11.0  breads cereals fastfoodgrains
226        2   0.30   15.0  breads cereals fastfoodgrains
227        0   0.10   25.0  breads cereals fastfoodgrains
228        0   0.20   27.0  breads cereals fastfoodgrains
229        2   1.60   74.0  breads cereals fastfoodgrains
230        0   0.00   10.0  breads cereals fastfoodgrains
231        0   0.00    8.0  breads cereals fastfoodgrains
232        0   8.00   22.0  breads cereals fastfoodgrains
233        0   2.90   33.0  breads cereals fastfoodgrains
234        0   0.30   84.0  breads cereals fastfoodgrains
235        0   2.80   79.0  breads cereals fastfoodgrains
236        0   0.10   32.0  breads cereals fastfoodgrains
237       24   0.00   44.0  breads cereals fastfoodgrains
238        4   0.00   19.0  breads cereals fastfoodgrains
239        2   0.10   37.0  breads cereals fastfoodgrains
240        2   4.60   26.0  breads cereals fastfoodgrains
241        0   0.10   28.0  breads cereals fastfoodgrains
242        0   0.10   28.0  breads cereals fastfoodgrains
243        5   0.00   23.0  breads cereals fastfoodgrains
244        2   0.50   20.0  breads cereals fastfoodgrains
245        0   0.00   12.0  breads cereals fastfoodgrains
246        0   0.60   26.0  breads cereals fastfoodgrains
247        0   1.20  154.0  breads cereals fastfoodgrains
248        0   0.40  142.0  breads cereals fastfoodgrains
249        0   0.30  150.0  breads cereals fastfoodgrains
250        0   0.10   26.0  breads cereals fastfoodgrains
251        0   1.20   28.0  breads cereals fastfoodgrains
252       11   0.10   23.0  breads cereals fastfoodgrains
253        2   0.00   20.0  breads cereals fastfoodgrains
254        0   0.10   20.0  breads cereals fastfoodgrains
255        6   0.50   35.0  breads cereals fastfoodgrains
256        3   0.50   36.0  breads cereals fastfoodgrains
257        0   1.20   40.0  breads cereals fastfoodgrains
258        0   0.70   23.0  breads cereals fastfoodgrains
259        1   0.10   30.0  breads cereals fastfoodgrains
260        3   2.50   34.0  breads cereals fastfoodgrains
261        3   2.50   36.0  breads cereals fastfoodgrains
262        0   0.70   25.0  breads cereals fastfoodgrains
263        0   4.40   35.0  breads cereals fastfoodgrains

     NameContainsSubstring
219                  False
220                  False
```

```
221                 False
222                 False
225                 False
226                 False
227                 False
228                 False
229                 False
230                 False
231                 False
232                 False
233                 False
234                 False
235                 False
236                 False
237                 False
238                 False
239                 False
240                 False
241                 False
242                 False
243                 False
244                 False
245                 False
246                 False
247                 False
248                 False
249                 False
250                 False
251                 False
252                 False
253                 False
254                 False
255                 False
256                 False
257                 False
258                 False
259                 False
260                 False
261                 False
262                 False
263                 False
```

[574]:
```
bread_cereals=['breads cereals fastfoodgrains']

prot= foods[foods['Category'].isin(bread_cereals)]
protein= prot.sort_values(by='Protein', ascending= False)
top_10=protein.head(10)
```

```
fig = px.bar(top_10, x='Food', y='Protein', color='Protein', title=' Top 10␣
 ↪protein rich Breads cereals fastfoodgrains')
fig.show()
```

[575]:
```
meat=foods[foods['Category'].str.contains("meat")]
print(meat.head(50))
```

```
                                  Food Measure  Grams  Calories  Protein  \
43                               bacon    100g    100    593.75    25.00
44                                beef    100g    100    288.24    27.06
45                           hamburger    100g    100    288.24    24.71
46                         ground lean    100g    100    217.65    28.24
47                           roast beef    100g    100    458.82    18.82
48                                steak    100g    100    388.24    23.53
49                  steak lean as round    100g    100    258.82    28.24
50                           corned beef    100g    100    217.65    25.88
51                corned beef hash canned    100g    100    141.18    14.12
52                 corned beef hash dried    100g    100    205.36    33.93
53                               pot-pie    100g    100    211.45     7.93
54                 corned beef hash stew    100g    100     78.72     6.38
55                               chicken    100g    100    217.65    27.06
56  fried breast or leg and thigh chicken    100g    100    288.24    29.41
57                       roasted chicken    100g    100    290.00    25.00
58                 chicken livers fried    100g    100    140.00    22.00
59                        duck domestic    100g    100    370.00    16.00
60                   lamb chop broiled    100g    100    417.39    20.87
61                         leg roasted    100g    100    365.12    23.26
62                   shoulder braised    100g    100    335.29    21.18
63                   pork chop 1 thick    100g    100    260.00    16.00
64                     ham pan-broiled    100g    100    341.18    18.82
65                              ham as    100g    100    298.25    22.81
66                   ham canned spiced    100g    100    289.47    14.04
67                          pork roast    100g    100    364.71    24.71
68                       pork sausage    100g    100    475.00    18.00
69                             turkey    100g    100    265.00    27.00
70                               veal    100g    100    217.65    27.06
71                              roast    100g    100    358.82    15.29

      Fat  Sat.Fat  Fiber  Carbs      Category  NameContainsSubstring
43  50.00       7    0.0    1.0  meat poultry                  False
44  18.82      15    0.0    0.0  meat poultry                  False
45  20.00      15    0.0    0.0  meat poultry                  False
46  11.76       9    0.0    0.0  meat poultry                  False
47  42.35      35    0.0    0.0  meat poultry                  False
48  31.76      25    0.0    0.0  meat poultry                  False
49  14.12      11    0.0    0.0  meat poultry                  False
50  11.76       9    0.0    0.0  meat poultry                  False
51   9.41       7    0.0    6.0  meat poultry                  False
```

```
52   7.14        4    0.0    0.0  meat poultry                False
53  12.33       25    0.0   32.0  meat poultry                False
54   4.26        9    0.0   15.0  meat poultry                False
55  10.59        7    0.0    0.0  meat poultry                False
56  17.65       11    0.0    0.0  meat poultry                False
57  20.00       16    0.0    0.0  meat poultry                False
58  14.00       12    0.0    2.3  meat poultry                False
59  28.00        0    0.0    0.0  meat poultry                False
60  30.43       33    0.0    0.0  meat poultry                False
61  16.28       14    0.0    0.0  meat poultry                False
62  27.06       21    0.0    0.0  meat poultry                False
63  21.00       18    0.0    0.0  meat poultry                False
64  25.88       19    0.0    0.0  meat poultry                False
65  22.81       11    0.0    0.0  meat poultry                False
66  24.56       12    0.0    1.0  meat poultry                False
67  28.24       21    0.0    0.0  meat poultry                False
68  44.00       40    0.0    0.0  meat poultry                False
69  15.00        0    0.0    0.0  meat poultry                False
70  10.59        8    0.0    0.0  meat poultry                False
71  16.47       13    0.0    0.0  meat poultry                False
```

[576]:
```
'''now let's find out more about the calories'''
```

[576]:
```
"now let's find out more about the calories"
```

[577]:
```
#plot the top 20 calorie food
```

[578]:
```
cals= foods.sort_values(by='Calories', ascending= False)
top_20_cals=cals.head(20)
fig = px.bar(top_20, x='Food', y='Calories' , color='Calories',title=' Top 20␣
 ↪calorie rich foods')
fig.show()
```

[579]:
```
#plot the top 20 fat food
```

[580]:
```
cals= foods.sort_values(by='Fat', ascending= False)
top_20_cals=cals.head(20)
fig = px.bar(top_20, x='Food', y='Calories' , color='Calories',title=' Top 20␣
 ↪calorie rich foods')
fig.show()
```

[581]:
```
#let's check the categorys
```

[584]:
```
category_dist=foods.groupby(['Category']).sum()
category_dist
```

[584]:
```
                                                                Food
\
```

```
Category
breads cereals fastfoodgrains  biscuitsbran flakesbread cracked wheatryewhole…
dairy products                 cows' milkmilk skimbuttermilkevaporated undilu…
desserts sweets                apple bettybread puddingcakeschocolate fudgecu…
drinksalcohol beverages        beerginwinestable (12.2% alcohol)carbonated dr…
fats oils shortenings          yolksbutterbutterbutterhydrogenated cooking fa…
fish seafood                   clamscodcrab meatfish sticks friedflounderhadd…
fruits a-f                     watercress stems rawapple juice cannedapple vi…
fruits g-p                     grapefruit sectionsgrapefruit fresh 5" diamete…
fruits r-z                     raisinsraspberriesraw redrhubarb sweetenedstra…
jams jellies                   honeypreservesjelliesmolassescane syrupbrown f…
meat poultry                   baconbeefhamburgerground leanroast beefsteakst…
seeds and nuts                 almondsroasted and saltedbrazil nutscashewscoc…
soups                          bean soupsbeef soupbouillonchicken soupclam ch…
vegetables a-e                 asparagusbeanslimalima dry cookednavy baked wi…
vegetables f-p                 kalekohlrabilambs quarters steamedlentilslettu…
vegetables r-z                 fresh steamed peassplit cooked peasheated peas…

                                                                        Measure
\
Category
breads cereals fastfoodgrains  100g100g100g100g100g100g100g100g100g100g100g10…
dairy products                 100g100g100g100g100g100g100g100g100g100g100g10…
desserts sweets                100g100g100g100g100g100g100g100g100g100g100g10…
drinksalcohol beverages         100g100g100g100g100g100g100g100g100g100g100g100g
fats oils shortenings          100g100g100g100g100g100g100g100g100g100g100g10…
fish seafood                   100g100g100g100g100g100g100g100g100g100g100g10…
fruits a-f                     100g100g100g100g100g100g100g100g100g100g100g10…
fruits g-p                     100g100g100g100g100g100g100g100g100g100g100g10…
fruits r-z                                      100g100g100g100g100g100g100g100g
jams jellies                                    100g100g100g100g100g100g100g100g
meat poultry                   100g100g100g100g100g100g100g100g100g100g100g10…
seeds and nuts                  100g100g100g100g100g100g100g100g100g100g100g100g
soups                                  100g100g100g100g100g100g100g100g100g100g
vegetables a-e                 100g100g100g100g100g100g100g100g100g100g100g10…
vegetables f-p                 100g100g100g100g100g100g100g100g100g100g100g10…
vegetables r-z                 100g100g100g100g100g100g100g100g100g100g100g10…

                               Grams  Calories  Protein      Fat  Sat.Fat  \
Category
breads cereals fastfoodgrains   4300  12620.88   380.37   241.71       77
dairy products                  2700   5678.63   333.12   360.49      299
desserts sweets                 2900   8177.22    87.10   196.02      150
drinksalcohol beverages         1200    677.48     0.00     0.00        0
fats oils shortenings           1400   8581.67   221.23  1146.49      536
fish seafood                    1800   2814.46   484.47   213.35      252
fruits a-f                      2200   2183.48    25.66    17.75       12
```

```
fruits g-p                      2800   2110.54    15.68    36.30       21
fruits r-z                       800    751.86     4.18     0.11        0
jams jellies                     800   2181.16     0.00     0.00        0
meat poultry                    2900   8641.89   632.35   606.27      427
seeds and nuts                  1200   7172.76   212.19   643.59      232
soups                           1000    475.99    23.64    16.35       43
vegetables a-e                  2600   1147.20    63.62     4.94        6
vegetables f-p                  1400    511.30    26.95     1.56        0
vegetables r-z                  2700   2729.76    74.18    91.20       44

                                Fiber   Carbs  NameContainsSubstring
Category
breads cereals fastfoodgrains   39.41  1614.0                      0
dairy products                   3.00   532.0                      0
desserts sweets                 20.50  1184.0                      0
drinksalcohol beverages          0.00   167.0                      0
fats oils shortenings          234.00   239.0                      0
fish seafood                   235.00   263.0                      0
fruits a-f                      33.50   812.0                      0
fruits g-p                      21.10  1009.0                      0
fruits r-z                      17.40   330.0                      0
jams jellies                     8.00   345.0                      0
meat poultry                     0.00    57.3                      0
seeds and nuts                  18.60   140.0                      0
soups                            4.00   155.0                      0
vegetables a-e                  34.30   346.0                      0
vegetables f-p                  16.90   142.0                      0
vegetables r-z                  26.20   447.0                      0
```

[585]:
```
#let's check the amount in each category

category_dist=foods.groupby(['Category']).count()
category_dist
```

[585]:
```
                                Food  Measure  Grams  Calories  Protein  Fat  \
Category
breads cereals fastfoodgrains     43       43     43        43       43   43
dairy products                    27       27     27        27       27   27
desserts sweets                   29       29     29        29       29   29
drinksalcohol beverages           12       12     12        12       12   12
fats oils shortenings             14       14     14        14       14   14
fish seafood                      18       18     18        18       18   18
fruits a-f                        22       22     22        22       22   22
fruits g-p                        28       28     28        28       28   28
fruits r-z                         8        8      8         8        8    8
jams jellies                       8        8      8         8        8    8
meat poultry                      29       29     29        29       29   29
```

```
seeds and nuts                  12       12     12       12       12   12
soups                           10       10     10       10       10   10
vegetables a-e                  26       26     26       26       26   26
vegetables f-p                  14       14     14       14       14   14
vegetables r-z                  27       27     27       27       27   27

                              Sat.Fat  Fiber  Carbs  NameContainsSubstring
Category
breads cereals fastfoodgrains      43     43     43                    43
dairy products                     27     27     27                    27
desserts sweets                    29     29     29                    29
drinksalcohol beverages            12     12     12                    12
fats oils shortenings              14     14     14                    14
fish seafood                       18     18     18                    18
fruits a-f                         22     22     22                    22
fruits g-p                         28     28     28                    28
fruits r-z                          8      8      8                     8
jams jellies                        8      8      8                     8
meat poultry                       29     29     29                    29
seeds and nuts                     12     12     12                    12
soups                              10     10     10                    10
vegetables a-e                     26     26     26                    26
vegetables f-p                     14     14     14                    14
vegetables r-z                     27     27     27                    27
```

[586]:
```
#I will use some pie graphic to show the amount of different foods containing␣
 ↪each nutrient
fig = px.pie(values=foods["Calories"],title='Calories', names=foods["Category"])
fig.show()
```

[587]:
```
fig = px.pie(values=foods["Protein"],title='Protein' ,names=foods["Category"])
fig.show()
```

[588]:
```
fig = px.pie(values=foods["Carbs"] ,title='Carbs',names=foods["Category"])
fig.show()
```

[589]:
```
fig = px.pie(values=foods["Fat"],title='Fat', names=foods["Category"])
fig.show()
```

[590]:
```
fig = px.pie(values=foods["Sat.Fat"],title='Sat. Fat', names=foods["Category"])
fig.show()
```

[591]:
```
'''Next, I will use 3D graphics to better illustrate the amounts of the three␣
 ↪macronutrients we consume daily: carbohydrates, protein, and fat.
I will focus on the most popular food groups-meat, bread and cereals, and dairy␣
 ↪products-without categorizing them alphabetically.
```

```
The 3D visuals will provide a clear representation of these foods and their␣
 ↪respective macronutrient content'''
```

[591]:
```
'Next, I will use 3D graphics to better illustrate the amounts of the three
macronutrients we consume daily: carbohydrates, protein, and fat. \nI will focus
on the most popular food groups-meat, bread and cereals, and dairy
products-without categorizing them alphabetically. \nThe 3D visuals will provide
a clear representation of these foods and their respective macronutrient
content'
```

[592]:
```
# Filter meat category
meat = ['meat poultry']
dessert_df = foods[foods['Category'].isin(meat)]

#3d graphic taking in consideration carb, fat and calories of the meat category
fig = px.scatter_3d(dessert_df,x='Carbs',y='Fat', z='Calories',␣
 ↪hover_name='Food',color='Calories', title='Meats: 3D Nutritional Analysis')


fig.update_layout(
    scene=dict(
        xaxis_title='Carbohydrates (g)',
        yaxis_title='Fat (g)',
        zaxis_title='Calories'
    ),
    margin=dict(l=0, r=0, b=0, t=30)
)

fig.show()
```

[593]:
```
        #

# Filter meat category
breads= ['breads cereals fastfoodgrains']
dessert_df = foods[foods['Category'].isin(breads)]

#3d graphic taking in consideration carb, fat and calories of the breads␣
 ↪cereals fastfoodgrains category
fig = px.scatter_3d(dessert_df,x='Carbs',y='Fat', z='Calories',␣
 ↪hover_name='Food',color='Calories', title='Breads cereals fastfoodgrains: 3D␣
 ↪Nutritional Analysis')


fig.update_layout(
    scene=dict(
        xaxis_title='Carbohydrates (g)',
        yaxis_title='Fat (g)',
```

```
        zaxis_title='Calories'
    ),
    margin=dict(l=0, r=0, b=0, t=30)
)

fig.show()
```

[594]:
```
# Filter meat category
dairy = ['dairy products']
dessert_df = foods[foods['Category'].isin(dairy)]

#3d graphic taking in consideration carb, fat and calories of  dairy category
fig = px.scatter_3d(
    dessert_df,x='Carbs',y='Fat',␣
 ↪z='Calories',hover_name='Food',color='Calories', title='dairy products: 3D␣
 ↪Nutritional Analysis')


fig.update_layout(
    scene=dict(
        xaxis_title='Carbohydrates (g)',
        yaxis_title='Fat (g)',
        zaxis_title='Calories'
    ),
    margin=dict(l=0, r=0, b=0, t=30)
)

fig.show()
```

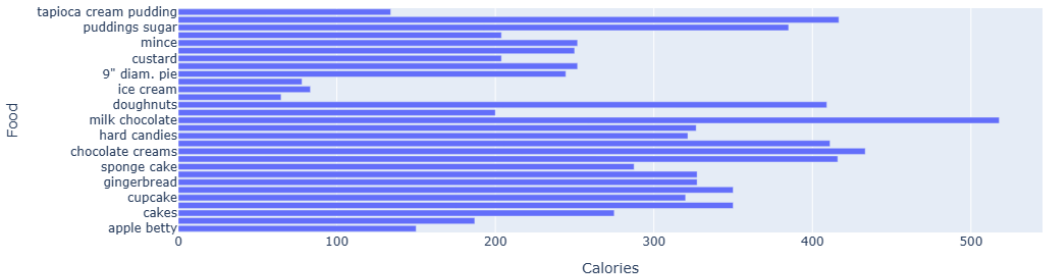
[595]:
```
'''We can reach some conclusions based on the graphics above :
   - Fish and seafood and meats are the winner in protein ammount , but the␣
 ↪bread and cereal have a good amount too
   -breads and cereals are rich in carbs, follow by the sweets
   -Seeds and nuts have are full of fat content
   -Dairy products have more saturated fat content than seafood '''
```

[595]:
```
'We can reach some conclusions based on the graphics above :\n   - Fish and
seafood and meats are the winner in protein ammount , but the bread and cereal
have a good amount too\n   -breads and cereals are rich in carbs, follow by the
sweets\n   -Seeds and nuts have are full of fat content\n   -Dairy products have
more saturated fat content than seafood '
```

[596]:
```
#Now let's analyse some desserts to see what we can eat after a good meal
desserts = ['desserts sweets']
dessert_and_sweets = foods[foods['Category'].isin(desserts)]

# Create a funnel plot of desserts vs calories
```

```
fig = px.bar(dessert_and_sweets, x='Calories', y='Food')
fig.show()
```



[597]:
```
# Filter desserts category
desserts = ['desserts sweets']
dessert_df = foods[foods['Category'].isin(desserts)]

#3d graphic taking in consideration carb, fat and calories of the dessert and␣
 ↪sweet category
fig = px.scatter_3d(dessert_df,x='Carbs',  y='Fat',z='Calories',␣
 ↪hover_name='Food',color='Calories', title='Desserts & Sweets: 3D Nutritional␣
 ↪Analysis')


fig.update_layout(
    scene=dict(
        xaxis_title='Carbohydrates (g)',
        yaxis_title='Fat (g)',
        zaxis_title='Calories'
    ),
    margin=dict(l=0, r=0, b=0, t=30)
)

fig.show()
```

Conclusion

Key Findings:

Protein: Meat/seafood > dairy > breads/cereals.

Carbs: Breads/cereals > sweets > vegetables.

Fat: Seeds/nuts > dairy > meats.

Saturated Fat: Dairy > processed foods > meats.

Dietary Observations:

Highlighted nutrient-dense vs. calorie-dense foods.

Compared saturated fat content across categories.

Find the web links below.

[653]:
```
#https://www.geeksforgeeks.org/python-plotly-tutorial/
#https://www.geeksforgeeks.org/pandas-tutorial/
#https://www.w3schools.com/python/pandas/pandas_cleaning.asp
#https://plotly.com/graphing-libraries/
#https://www.w3schools.com/python/matplotlib_pyplot.asp
```

Find the step-by-step process below.

[ ]:
```
'''
1. Data Loading & Initial Inspection
File Import: Loaded nutrients_csvenile.csv using pd.read_csv().

Head Preview: Displayed the first 5 rows to understand the structure


2. Data Cleaning
Missing Values:

Identified NaN values with isnull().sum().

Filled missing values using fillna(0).

Placeholder Handling:

Replaced "t" (trace amounts) and "a" (erroneous entry) with 0.

Removed commas from numeric fields (e.g., 1,419 → 1419).

Type Conversion:

Converted columns (Grams, Calories, etc.) to numeric using pd.to_numeric().

Dropped rows with residual NaN values after cleaning.


3. Data Standardization
Normalization:

Created a scaling_factor = 100 / Grams to standardize all nutritional values to␣
↪per 100g.
```

```
Scaled Calories, Protein, Fat, and Carbs using this factor.

Updated Measure to "100g" and Grams to 100 for consistency.

Text Formatting:

Lowercased Food and Category columns for uniformity.

4. Exploratory Data Analysis (EDA)
Category Filtering: Extracted specific food groups (e.g., vegetables, seeds,␣
 ↪dairy)
Nutrient Ranking:

Visualization:

Created 3D scatter plots (Carbs vs. Fat vs. Calories) for categories like meat,␣
 ↪breads, and dairy.

Generated bar charts for top protein/fat sources.

Built pie charts to show macronutrient distribution across categories
Sorted foods by protein/fat/carb content to identify top sources

5. Insights & Conclusions
Key Findings:

Protein: Meat/seafood > dairy > breads/cereals.

Carbs: Breads/cereals > sweets > vegetables.

Fat: Seeds/nuts > dairy > meats.

Saturated Fat: Dairy > processed foods > meats.

Dietary Observations:

Highlighted nutrient-dense vs. calorie-dense foods.

Compared saturated fat content across categories.

Key Tools Used
Pandas: Data manipulation (filtering, scaling, type conversion).

Plotly: Interactive 3D plots (px.scatter_3d), bar/pie charts.

Seaborn/Matplotlib: Supplementary visualization.
```

```
This process transformed raw, messy data into actionable insights about␣
↪macronutrient distribution across food groups.'''
```