

```
In [1]: import pandas as pd
```

```
In [2]: product = pd.read_csv("products.tsv", sep="\t")
```

C:\Users\DHRUV\AppData\Local\Temp\ipykernel_27352\1755326969.py:1: DtypeWarning: Columns (0,3,5,19,20,24,25,26,27,28,36,37,38,39,48) have mixed types. Specify dtype option on import or set low_memory=False.
 product = pd.read_csv("products.tsv", sep="\t")

```
In [3]: # top 5 entris wan to see

len(product)
```

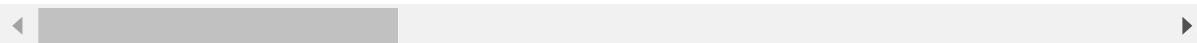
Out[3]: 356027

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

Out[4]:

	code	url	creator	created_t	created_datetime	last_m
0	3087	http://world-en.openfoodfacts.org/product/0000...	openfoodfacts-contributors	1474103866	2016-09-17T09:17:46Z	147
1	4530	http://world-en.openfoodfacts.org/product/0000...	usda-ndb-import	1489069957	2017-03-09T14:32:37Z	148
2	4559	http://world-en.openfoodfacts.org/product/0000...	usda-ndb-import	1489069957	2017-03-09T14:32:37Z	148
3	16087	http://world-en.openfoodfacts.org/product/0000...	usda-ndb-import	1489055731	2017-03-09T10:35:31Z	148
4	16094	http://world-en.openfoodfacts.org/product/0000...	usda-ndb-import	1489055653	2017-03-09T10:34:13Z	148

5 rows × 163 columns



```
In [ ]: ## what is the number of observation in the dataset?
```

```
In [5]: product.shape
```

Out[5]: (356027, 163)

```
In [7]: product.shape[0]
```

Out[7]: 356027

```
In [ ]: # print all the columns in the dataset
```

```
In [8]: product.columns
```

```
Out[8]: Index(['code', 'url', 'creator', 'created_t', 'created_datetime',  
             'last_modified_t', 'last_modified_datetime', 'product_name',  
             'generic_name', 'quantity',  
             ...  
             'fruits-vegetables-nuts_100g', 'fruits-vegetables-nuts-estimate_100g',  
             'collagen-meat-protein-ratio_100g', 'cocoa_100g', 'chlorophyll_100g',  
             'carbon-footprint_100g', 'nutrition-score-fr_100g',  
             'nutrition-score-uk_100g', 'glycemic-index_100g',  
             'water-hardness_100g'],  
            dtype='object', length=163)
```

```
In [ ]: ## what is the name of 85th column?
```

```
In [9]: product.columns[85]
```

```
Out[9]: '-eicosapentaenoic-acid_100g'
```

```
In [10]: product.columns[20]
```

```
Out[10]: 'manufacturing_places_tags'
```

```
In [13]: ## what is the dtype of 20th columns?
```

```
product.dtypes["manufacturing_places_tags"]
```

```
Out[13]: dtype('O')
```

```
In [14]: product.dtypes
```

```
Out[14]: code                object  
         url                 object  
         creator             object  
         created_t           object  
         created_datetime     object  
         ...  
         carbon-footprint_100g float64  
         nutrition-score-fr_100g float64  
         nutrition-score-uk_100g float64  
         glycemic-index_100g   float64  
         water-hardness_100g   float64  
         Length: 163, dtype: object
```

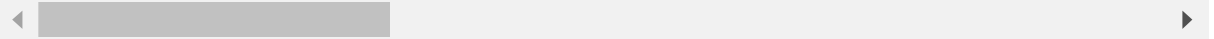
In [22]: *## what will be the last 5 records*

```
product.tail(1)
```

Out[22]:

	code	url	creator	created_t	created_datetime
356026	999990026839	http://world-en.openfoodfacts.org/product/9999...	usda-ndb-import	1489072709	2017-03-09T15:18:29Z

1 rows × 163 columns



In [23]: `product.index`

Out[23]: RangeIndex(start=0, stop=356027, step=1)

In [27]: `product.values[356026][2]`

Out[27]: 'usda-ndb-import'

In []: