```
import pandas as pd
from google.colab import files
from io import StringIO

# Carregar o arquivo CSV
uploaded = files.upload()

# Ler o conteúdo do arquivo CSV
filename = next(iter(uploaded))
df = pd.read_csv(filename, sep=';', quoting=1)

# Verificar se os dados foram importados adequadamente
print(df.info())
print(df.head())
print(df.tail())

# Criar uma cópia do conjunto de dados
df_copy = df.copy()

# Substituir valores nulos na coluna ‘Calories’ por 0
df_copy['Calories'].fillna(0, inplace=True)
print("\nApós substituir valores nulos em 'Calories':")
print(df_copy)

# Substituir valores nulos na coluna ‘Date’ por ‘1900/01/01’
df_copy['Date'].fillna('1900/01/01', inplace=True)
print("\nApós substituir valores nulos em 'Date':")
print(df_copy)

# Transformar a coluna ‘Date’ em datetime
df_copy['Date'] = pd.to_datetime(df_copy['Date'], errors='coerce')

# Substituir ‘1900/01/01’ por NaN
df_copy['Date'].replace(pd.Timestamp('1900-01-01'), pd.NaT, inplace=True)
print("\nApós substituir '1900/01/01' por NaT:")
print(df_copy)

# Corrigir valor ‘20201226’
df_copy['Date'] = df_copy['Date'].astype(str).replace('20201226', '2020/12/26')
df_copy['Date'] = pd.to_datetime(df_copy['Date'], errors='coerce')
print("\nApós corrigir '20201226':")
print(df_copy)

# Remover registros com valores nulos
df_copy.dropna(inplace=True)
print("\nApós remover registros com valores nulos:")
print(df_copy)
```

Escolher arquivos arq_dados.csv

- **arq_dados.csv**(text/csv) - 1130 bytes, last modified: 22/10/2024 - 100% done

```
Saving arq_dados.csv to arq_dados.csv
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 32 entries, 0 to 31
Data columns (total 6 columns):
 #   Column    Non-Null Count  Dtype 
---  ------    --------------  ----- 
 0   ID        32 non-null     int64 
 1   Duration  32 non-null     int64 
 2   Date      31 non-null     object
 3   Pulse     32 non-null     int64 
 4   Maxpulse  32 non-null     int64 
 5   Calories  30 non-null     object
dtypes: int64(4), object(2)
memory usage: 1.6+ KB
None
   ID  Duration          Date  Pulse  Maxpulse Calories
0   0        60  '2020/12/01'    110       130     4091
1   1        60  '2020/12/02'    117       145     4790
2   2        60  '2020/12/03'    103       135     3400
3   3        45  '2020/12/04'    109       175     2824
4   4        45  '2020/12/05'    117       148     4060
    ID  Duration          Date  Pulse  Maxpulse Calories
27  27        60  '2020/12/27'     92       118     2410
28  28        60  '2020/12/28'    103       132      NaN
29  29        60  '2020/12/29'    100       132     2800
30  30        60  '2020/12/30'    102       129     3803
31  31        60  '2020/12/31'     92       115     2430

Após substituir valores nulos em 'Calories':
    ID  Duration          Date  Pulse  Maxpulse Calories
0    0        60  '2020/12/01'    110       130     4091
1    1        60  '2020/12/02'    117       145     4790
2    2        60  '2020/12/03'    103       135     3400
3    3        45  '2020/12/04'    109       175     2824
4    4        45  '2020/12/05'    117       148     4060
5    5        60  '2020/12/06'    102       127     3000
6    6        60  '2020/12/07'    110       136     3740
7    7       450  '2020/12/08'    104       134     2533
8    8        30  '2020/12/09'    109       133     1951
9    9        60  '2020/12/10'     98       124     2690
10  10        60  '2020/12/11'    103       147     3293
11  11        60  '2020/12/12'    100       120     2507
12  12        60  '2020/12/12'    100       120     2507
13  13        60  '2020/12/13'    106       128     3453
14  14        60  '2020/12/14'    104       132     3793
15  15        60  '2020/12/15'     98       123     2750
16  16        60  '2020/12/16'     98       120     2152
17  17        60  '2020/12/17'    100       120     3000
18  18        45  '2020/12/18'     90       112        0
19  19        60  '2020/12/19'    103       123     3230
20  20        45  '2020/12/20'     97       125   2430 2
21   1        60  '2020/12/21'    108       131     3642
22  22        45           NaN    100       119     2820
23  23        60  '2020/12/23'    130       101     3000
24  24        45  '2020/12/24'    105       132     2460
25  25        60  '2020/12/25'    102       126     3345
26  26        60      20201226    100       120     2500
27  27        60  '2020/12/27'     92       118     2410
28  28        60  '2020/12/28'    103       132        0
29  29        60  '2020/12/29'    100       132     2800
30  30        60  '2020/12/30'    102       129     3803
31  31        60  '2020/12/31'     92       115     2430

Após substituir valores nulos em 'Date':
    ID  Duration          Date  Pulse  Maxpulse Calories
0    0        60  '2020/12/01'    110       130     4091
1    1        60  '2020/12/02'    117       145     4790
2    2        60  '2020/12/03'    103       135     3400
3    3        45  '2020/12/04'    109       175     2824
4    4        45  '2020/12/05'    117       148     4060
5    5        60  '2020/12/06'    102       127     3000
6    6        60  '2020/12/07'    110       136     3740
7    7       450  '2020/12/08'    104       134     2533
8    8        30  '2020/12/09'    109       133     1951
9    9        60  '2020/12/10'     98       124     2690
10  10        60  '2020/12/11'    103       147     3293
11  11        60  '2020/12/12'    100       120     2507
12  12        60  '2020/12/12'    100       120     2507
13  13        60  '2020/12/13'    106       128     3453
14  14        60  '2020/12/14'    104       132     3793
15  15        60  '2020/12/15'     98       123     2750
16  16        60  '2020/12/16'     98       120     2152
17  17        60  '2020/12/17'    100       120     3000
18  18        45  '2020/12/18'     90       112        0
19  19        60  '2020/12/19'    103       123     3230
20  20        45  '2020/12/20'     97       125   2430 2
21   1        60  '2020/12/21'    108       131     3642
22  22        45    1900/01/01    100       119     2820
```

```
23  23        60  '2020/12/23'    130       101     3000
24  24        45  '2020/12/24'    105       132     2460
25  25        60  '2020/12/25'    102       126     3345
26  26        60      20201226    100       120     2500
27  27        60  '2020/12/27'     92       118     2410
28  28        60  '2020/12/28'    103       132        0
29  29        60  '2020/12/29'    100       132     2800
30  30        60  '2020/12/30'    102       129     3803
31  31        60  '2020/12/31'     92       115     2430

Após substituir '1900/01/01' por NaT:
    ID  Duration       Date  Pulse  Maxpulse Calories
0    0        60 2020-12-01    110       130     4091
1    1        60 2020-12-02    117       145     4790
2    2        60 2020-12-03    103       135     3400
3    3        45 2020-12-04    109       175     2824
4    4        45 2020-12-05    117       148     4060
5    5        60 2020-12-06    102       127     3000
6    6        60 2020-12-07    110       136     3740
7    7       450 2020-12-08    104       134     2533
8    8        30 2020-12-09    109       133     1951
9    9        60 2020-12-10     98       124     2690
10  10        60 2020-12-11    103       147     3293
11  11        60 2020-12-12    100       120     2507
12  12        60 2020-12-12    100       120     2507
13  13        60 2020-12-13    106       128     3453
14  14        60 2020-12-14    104       132     3793
15  15        60 2020-12-15     98       123     2750
16  16        60 2020-12-16     98       120     2152
17  17        60 2020-12-17    100       120     3000
18  18        45 2020-12-18     90       112        0
19  19        60 2020-12-19    103       123     3230
20  20        45 2020-12-20     97       125   2430 2
21   1        60 2020-12-21    108       131     3642
22  22        45        NaT    100       119     2820
23  23        60 2020-12-23    130       101     3000
24  24        45 2020-12-24    105       132     2460
25  25        60 2020-12-25    102       126     3345
26  26        60        NaT    100       120     2500
27  27        60 2020-12-27     92       118     2410
28  28        60 2020-12-28    103       132        0
29  29        60 2020-12-29    100       132     2800
30  30        60 2020-12-30    102       129     3803
31  31        60 2020-12-31     92       115     2430

Após corrigir '20201226':
    ID  Duration       Date  Pulse  Maxpulse Calories
0    0        60 2020-12-01    110       130     4091
1    1        60 2020-12-02    117       145     4790
2    2        60 2020-12-03    103       135     3400
3    3        45 2020-12-04    109       175     2824
4    4        45 2020-12-05    117       148     4060
5    5        60 2020-12-06    102       127     3000
6    6        60 2020-12-07    110       136     3740
7    7       450 2020-12-08    104       134     2533
8    8        30 2020-12-09    109       133     1951
9    9        60 2020-12-10     98       124     2690
10  10        60 2020-12-11    103       147     3293
11  11        60 2020-12-12    100       120     2507
12  12        60 2020-12-12    100       120     2507
13  13        60 2020-12-13    106       128     3453
14  14        60 2020-12-14    104       132     3793
15  15        60 2020-12-15     98       123     2750
16  16        60 2020-12-16     98       120     2152
17  17        60 2020-12-17    100       120     3000
18  18        45 2020-12-18     90       112        0
19  19        60 2020-12-19    103       123     3230
20  20        45 2020-12-20     97       125   2430 2
21   1        60 2020-12-21    108       131     3642
22  22        45        NaT    100       119     2820
23  23        60 2020-12-23    130       101     3000
24  24        45 2020-12-24    105       132     2460
25  25        60 2020-12-25    102       126     3345
26  26        60        NaT    100       120     2500
27  27        60 2020-12-27     92       118     2410
28  28        60 2020-12-28    103       132        0
29  29        60 2020-12-29    100       132     2800
30  30        60 2020-12-30    102       129     3803
31  31        60 2020-12-31     92       115     2430

Após remover registros com valores nulos:
    ID  Duration       Date  Pulse  Maxpulse Calories
0    0        60 2020-12-01    110       130     4091
1    1        60 2020-12-02    117       145     4790
2    2        60 2020-12-03    103       135     3400
3    3        45 2020-12-04    109       175     2824
4    4        45 2020-12-05    117       148     4060
5    5        60 2020-12-06    102       127     3000
6    6        60 2020-12-07    110       136     3740
7    7       450 2020-12-08    104       134     2533
```

```
8    8        30 2020-12-09    109       133     1951
9    9        60 2020-12-10     98       124     2690
10  10        60 2020-12-11    103       147     3293
11  11        60 2020-12-12    100       120     2507
12  12        60 2020-12-12    100       120     2507
13  13        60 2020-12-13    106       128     3453
14  14        60 2020-12-14    104       132     3793
15  15        60 2020-12-15     98       123     2750
16  16        60 2020-12-16     98       120     2152
17  17        60 2020-12-17    100       120     3000
18  18        45 2020-12-18     90       112        0
19  19        60 2020-12-19    103       123     3230
20  20        45 2020-12-20     97       125   2430 2
21   1        60 2020-12-21    108       131     3642
23  23        60 2020-12-23    130       101     3000
24  24        45 2020-12-24    105       132     2460
25  25        60 2020-12-25    102       126     3345
27  27        60 2020-12-27     92       118     2410
28  28        60 2020-12-28    103       132        0
29  29        60 2020-12-29    100       132     2800
30  30        60 2020-12-30    102       129     3803
31  31        60 2020-12-31     92       115     2430
<ipython-input-18-a363b87482ca>:22: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained as
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col


  df_copy['Calories'].fillna(0, inplace=True)
<ipython-input-18-a363b87482ca>:27: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained as
The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col
```