STA130 Week 01 Homework

September 12, 2024

```python
[6]: import pandas as pd

     # Load your dataset
     url = "https://raw.githubusercontent.com/datasets/commodity-prices/master/data/
      ↪commodity-prices.csv"
     com = pd.read_csv(url)
```

```python
[5]: com
```

```
[5]:          Date  All Commodity Price Index  Non-Fuel Price Index  \
     0    1980-01-01                        NaN                   NaN
     1    1980-02-01                        NaN                   NaN
     2    1980-03-01                        NaN                   NaN
     3    1980-04-01                        NaN                   NaN
     4    1980-05-01                        NaN                   NaN
     ..          ...                        ...                   ...
     445  2017-02-01                 118.615999            145.232866
     446  2017-03-01                 113.314945            142.940071
     447  2017-04-01                 113.226876            138.528012
     448  2017-05-01                 110.589462            138.865166
     449  2017-06-01                 106.840977            138.622118

          Food and Beverage Price Index  Food Price Index  Beverage Price Index  \
     0                              NaN               NaN            189.310021
     1                              NaN               NaN            190.387928
     2                              NaN               NaN            194.060355
     3                              NaN               NaN            186.137888
     4                              NaN               NaN            185.070241
     ..                             ...               ...                   ...
     445                     152.198420        151.815075            155.737047
     446                     148.145483        147.637196            152.837443
     447                     145.566652        145.227898            148.693674
     448                     150.514652        150.880372            147.138716
     449                     152.565099        153.117377            147.467065

          Industrial Inputs Price Index  Agricultural Raw Materials Index  \
     0                        81.889650                         78.900146
```

```
1                        83.048367                          75.715154
2                        75.228895                          69.002475
3                        72.471253                          67.877108
4                        69.588103                          65.879671
..                          …                                  …
445                     138.197090                         120.680238
446                     137.682182                         121.436062
447                     131.418412                         119.724719
448                     127.098235                         117.028866
449                     124.538572                         113.649909

     Metals Price Index  Fuel Energy Index  …  Sugar Free Market  \
0            84.049004                NaN  …          17.299999
1            88.345232                NaN  …          22.750000
2            79.726312                NaN  …          19.630001
3            75.789658                NaN  …          21.250000
4            72.266748                NaN  …          30.939999
..               …                 …   …              …
445         150.849718         103.044209  …          20.348421
446         149.416945          95.983218  …          18.059565
447         139.864904          98.424834  …          16.404737
448         134.371459          94.047194  …          15.689091
449         132.403581          88.247907  …          13.748182

     Sugar U.S. import price  Sunflower oil         Tea           Tin  \
0                  19.660000     566.926957  225.179901  16973.587904
1                  24.689999     573.958609  233.094498  17090.213164
2                  21.180000     535.284522  226.833298  17460.590400
3                  22.669998     486.062957  221.806793  17041.712600
4                  31.889999     502.763130  229.611206  17180.604736
..                     …             …           …           …
445                30.401579     966.971042  382.874800  19446.475000
446                29.829130     936.780525  356.102841  19875.195652
447                28.719474     939.952292  362.096533  19910.319444
448                28.429545     949.833062  371.482435  20200.333333
449                27.827273     936.240625  381.217333  19658.840909

       Uranium       Wheat  Wool coarse   Wool fine        Zinc
0    40.000000  175.634750   553.120940  684.277366   773.821533
1    38.000000  172.695236   568.154805  722.567093   868.620361
2    35.000000  163.509323   552.745109  695.956860   740.752441
3    32.000000  156.528030   510.650280  688.130355   707.683105
4    32.000000  161.304703   524.932443  720.760976   701.069092
..       …           …            …           …            …
445  25.058877  147.313095  1029.581497  1368.141078  2845.550000
446  24.552024  146.387567  1059.601984  1454.829350  2776.880435
447  23.165761  138.412425   991.122950  1404.982250  2614.916667
```

```
448  21.562894  146.486143  1019.954554  1433.465267  2590.214286
449  19.678442  157.196087  1065.808020  1403.827722  2573.397727

[450 rows x 64 columns]
```

[11]: `com.shape`

[11]: (450, 64)

[3]: `com.describe()`

[3]:
```
        All Commodity Price Index  Non-Fuel Price Index  \
count                 306.000000            318.000000
mean                  100.796302            115.066289
std                    50.113245             35.294688
min                    41.977549             70.367583
25%                    58.535131             84.577086
50%                    83.431606            100.435879
75%                   134.400893            142.700852
max                   220.030251            206.483885

        Food and Beverage Price Index  Food Price Index  Beverage Price Index  \
count                      318.000000        318.000000            450.000000
mean                       119.879623        119.596359            125.658430
std                         33.364661         33.166343             37.433449
min                         75.831061         76.044302             60.645581
25%                         94.558640         95.624823             90.239453
50%                        106.323867        106.677349            126.875813
75%                        145.553982        144.008989            156.104855
max                        196.003967        193.769425            221.989467

        Industrial Inputs Price Index  Agricultural Raw Materials Index  \
count                      450.000000                        450.000000
mean                        98.198455                        100.918462
std                         38.484350                         22.816541
min                         50.918535                         56.872307
25%                         72.048698                         89.668998
50%                         86.136822                         99.956156
75%                        119.460374                        115.762298
max                        217.069457                        171.362510

        Metals Price Index  Fuel Energy Index  Crude Oil petroleum  …  \
count           450.000000         306.000000           450.000000  …
mean             96.233763          91.801289            77.003297  …
std              53.126663          59.861121            55.687369  …
min              44.165892          22.088529            18.445490  …
25%              57.493306          38.270759            34.608027  …
```

```
50%                70.359960            72.341040              54.648132  …
75%               129.755124           134.324103             101.974867  …
max               256.236049           249.607241             249.662602  …

       Sugar Free Market  Sugar U.S. import price  Sunflower oil          Tea  \
count          450.000000               450.000000     450.000000   450.000000
mean            12.219680                23.070477     767.572572   233.166576
std              6.117983                 4.121527     354.825494    61.524454
min              2.780000                15.490002     332.545455   134.050003
25%              7.941287                21.079318     554.931949   187.388680
50%             10.817750                21.880001     637.594565   212.841148
75%             14.954955                23.171478     981.476563   261.518372
max             40.549999                41.700001    2300.190476   428.754486

                Tin      Uranium        Wheat   Wool coarse     Wool fine  \
count    450.000000   450.000000   450.000000    450.000000    450.000000
mean   10988.283345    24.917038   165.459005    582.860121    838.285276
std     6340.344540    20.085725    54.560825    267.707111    279.162646
min     3698.368421     7.100000    88.552237    247.089050    417.466580
25%     5676.724937    10.253125   128.185215    394.507217    639.457158
50%     8102.787500    17.000000   154.739189    503.195973    733.646305
75%    15148.243000    34.998541   179.309097    717.054265   1013.089075
max    32347.694444   136.222222   403.812897   1391.473685   1865.439360

              Zinc
count    450.000000
mean    1398.266704
std      693.671853
min      597.451904
25%      881.296814
50%     1107.272826
75%     1846.990693
max     4381.447368

[8 rows x 63 columns]
```

[4]: `com.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 450 entries, 0 to 449
Data columns (total 64 columns):
 #   Column                                                         Non-Null
Count  Dtype
---  ------
-------------  -----
 0   Date                                                           450 non-
null    object
```

```
 1   All Commodity Price Index                    306 non-
null    float64
 2   Non-Fuel Price Index                          318 non-
null    float64
 3   Food and Beverage Price Index                 318 non-
null    float64
 4   Food Price Index                              318 non-
null    float64
 5   Beverage Price Index                          450 non-
null    float64
 6   Industrial Inputs Price Index                 450 non-
null    float64
 7   Agricultural Raw Materials Index              450 non-
null    float64
 8   Metals Price Index                            450 non-
null    float64
 9   Fuel Energy Index                             306 non-
null    float64
 10  Crude Oil petroleum                           450 non-
null    float64
 11  Aluminum                                      450 non-
null    float64
 12  Bananas                                       450 non-
null    float64
 13  Barley                                        450 non-
null    float64
 14  Beef                                          450 non-
null    float64
 15  Coal                                          450 non-
null    float64
 16  Cocoa beans                                   450 non-
null    float64
 17  Coffee Other Mild Arabicas                    450 non-
null    float64
 18  Coffee Robusta                                450 non-
null    float64
 19  Rapeseed oil                                  450 non-
null    float64
 20  Copper                                        450 non-
null    float64
 21  Cotton                                        450 non-
null    float64
 22  Fishmeal                                      450 non-
null    float64
 23  Groundnuts peanuts                            450 non-
null    float64
 24  Hides                                         450 non-
null    float64
```

```
 25  China import Iron Ore Fines 62% FE spot                          450 non-
null    float64
 26  Lamb                                                             450 non-
null    float64
 27  Lead                                                             450 non-
null    float64
 28  Soft Logs                                                        450 non-
null    float64
 29  Hard Logs                                                        450 non-
null    float64
 30  Maize corn                                                       450 non-
null    float64
 31  Natural Gas - Russian Natural Gas border price in Germany        390 non-
null    float64
 32  Natural Gas - Indonesian Liquefied Natural Gas in Japan          306 non-
null    float64
 33  Natural Gas - Spot price at the Henry Hub terminal in Louisiana  318 non-
null    float64
 34  Nickel                                                           450 non-
null    float64
 35  Crude Oil - petroleum-simple average of three spot prices        450 non-
null    float64
 36  Crude Oil - petroleum - Dated Brent light blend                  450 non-
null    float64
 37  Oil Dubai                                                        450 non-
null    float64
 38  Crude Oil petroleum - West Texas Intermediate 40 API             450 non-
null    float64
 39  Olive Oil                                                        450 non-
null    float64
 40  Oranges                                                          450 non-
null    float64
 41  Palm oil                                                         450 non-
null    float64
 42  Swine - pork                                                     450 non-
null    float64
 43  Poultry chicken                                                  450 non-
null    float64
 44  Rice                                                             450 non-
null    float64
 45  Rubber                                                           450 non-
null    float64
 46  Fish salmon                                                      450 non-
null    float64
 47  Hard Sawnwood                                                    450 non-
null    float64
 48  Soft Sawnwood                                                    450 non-
null    float64
```

```
 49  Shrimp                                    450 non-
null    float64
 50  Soybean Meal                              450 non-
null    float64
 51  Soybean Oil                               450 non-
null    float64
 52  Soybeans                                  450 non-
null    float64
 53  Sugar European import price               318 non-
null    float64
 54  Sugar Free Market                         450 non-
null    float64
 55  Sugar U.S. import price                   450 non-
null    float64
 56  Sunflower oil                             450 non-
null    float64
 57  Tea                                       450 non-
null    float64
 58  Tin                                       450 non-
null    float64
 59  Uranium                                   450 non-
null    float64
 60  Wheat                                     450 non-
null    float64
 61  Wool coarse                               450 non-
null    float64
 62  Wool fine                                 450 non-
null    float64
 63  Zinc                                      450 non-
null    float64
dtypes: float64(63), object(1)
memory usage: 225.1+ KB
```

[5]: `com['Date'].value_counts()`

```
[5]: Date
     1980-01-01    1
     2008-02-01    1
     2005-08-01    1
     2005-07-01    1
     2005-06-01    1
                  ..
     1992-03-01    1
     1992-02-01    1
     1992-01-01    1
     1991-12-01    1
     2017-06-01    1
```

```
Name: count, Length: 450, dtype: int64
```

[6]:
```python
com['All Commodity Price Index'].describe()
```

[6]:
```
count    306.000000
mean     100.796302
std       50.113245
min       41.977549
25%       58.535131
50%       83.431606
75%      134.400893
max      220.030251
Name: All Commodity Price Index, dtype: float64
```

[7]:
```python
com.isnull().sum()
```

[7]:
```
Date                             0
All Commodity Price Index      144
Non-Fuel Price Index           132
Food and Beverage Price Index  132
Food Price Index               132
                               ...
Uranium                          0
Wheat                            0
Wool coarse                      0
Wool fine                        0
Zinc                             0
Length: 64, dtype: int64
```

[9]:
```python
com = com.dropna()
```

[7]:
```python
del.com['col']
```

```
  Cell In[7], line 1
    del.com['col']
       ^
SyntaxError: invalid syntax
```

[8]:
```python
com.describe()
```

[8]:
```
       All Commodity Price Index  Non-Fuel Price Index  \
count                306.000000            318.000000
mean                 100.796302            115.066289
std                   50.113245             35.294688
min                   41.977549             70.367583
```

25%	58.535131	84.577086
50%	83.431606	100.435879
75%	134.400893	142.700852
max	220.030251	206.483885

	Food and Beverage Price Index	Food Price Index	Beverage Price Index \
count	318.000000	318.000000	450.000000
mean	119.879623	119.596359	125.658430
std	33.364661	33.166343	37.433449
min	75.831061	76.044302	60.645581
25%	94.558640	95.624823	90.239453
50%	106.323867	106.677349	126.875813
75%	145.553982	144.008989	156.104855
max	196.003967	193.769425	221.989467

	Industrial Inputs Price Index	Agricultural Raw Materials Index \
count	450.000000	450.000000
mean	98.198455	100.918462
std	38.484350	22.816541
min	50.918535	56.872307
25%	72.048698	89.668998
50%	86.136822	99.956156
75%	119.460374	115.762298
max	217.069457	171.362510

	Metals Price Index	Fuel Energy Index	Crude Oil petroleum	… \
count	450.000000	306.000000	450.000000	…
mean	96.233763	91.801289	77.003297	…
std	53.126663	59.861121	55.687369	…
min	44.165892	22.088529	18.445490	…
25%	57.493306	38.270759	34.608027	…
50%	70.359960	72.341040	54.648132	…
75%	129.755124	134.324103	101.974867	…
max	256.236049	249.607241	249.662602	…

	Sugar Free Market	Sugar U.S. import price	Sunflower oil	Tea \
count	450.000000	450.000000	450.000000	450.000000
mean	12.219680	23.070477	767.572572	233.166576
std	6.117983	4.121527	354.825494	61.524454
min	2.780000	15.490002	332.545455	134.050003
25%	7.941287	21.079318	554.931949	187.388680
50%	10.817750	21.880001	637.594565	212.841148
75%	14.954955	23.171478	981.476563	261.518372
max	40.549999	41.700001	2300.190476	428.754486

	Tin	Uranium	Wheat	Wool coarse	Wool fine \
count	450.000000	450.000000	450.000000	450.000000	450.000000

```
mean    10988.283345    24.917038   165.459005    582.860121    838.285276
std      6340.344540    20.085725    54.560825    267.707111    279.162646
min      3698.368421     7.100000    88.552237    247.089050    417.466580
25%      5676.724937    10.253125   128.185215    394.507217    639.457158
50%      8102.787500    17.000000   154.739189    503.195973    733.646305
75%     15148.243000    34.998541   179.309097    717.054265   1013.089075
max     32347.694444   136.222222   403.812897   1391.473685   1865.439360

              Zinc
count   450.000000
mean   1398.266704
std     693.671853
min     597.451904
25%     881.296814
50%    1107.272826
75%    1846.990693
max    4381.447368

[8 rows x 63 columns]
```

[10]: `com.describe()`

[10]:
```
        All Commodity Price Index   Non-Fuel Price Index  \
count                306.000000             306.000000
mean                 100.796302             116.171481
std                   50.113245              35.521939
min                   41.977549              70.367583
25%                   58.535131              85.202530
50%                   83.431606             101.042369
75%                  134.400893             143.278330
max                  220.030251             206.483885

        Food and Beverage Price Index  Food Price Index  Beverage Price Index  \
count                      306.000000        306.000000            306.000000
mean                       120.745196        120.377035            124.143664
std                         33.708971         33.559628             40.205564
min                         75.831061         76.044302             60.645581
25%                         94.567236         95.479311             87.017161
50%                        107.995015        108.061076            118.487880
75%                        146.239117        145.110330            159.839051
max                        196.003967        193.769425            221.989467

        Industrial Inputs Price Index  Agricultural Raw Materials Index  \
count                      306.000000                        306.000000
mean                       111.551655                        111.483056
std                         39.044612                         17.203357
min                         64.463893                         81.445057
```

```
25%                                   76.705439                    98.707152
50%                                   96.450915                   109.770960
75%                                  148.351758                   120.454762
max                                  217.069457                   171.362510


        Metals Price Index  Fuel Energy Index  Crude Oil petroleum   …  \
count           306.000000         306.000000           306.000000   …
mean            111.601206          91.801289            91.286259   …
std              57.500463          59.861121            61.772177   …
min              48.018317          22.088529            19.536436   …
25%              60.715941          38.270759            36.433344   …
50%              84.367944          72.341040            70.214583   …
75%             168.322455         134.324103           136.244451   …
max             256.236049         249.607241           249.662602   …


        Sugar Free Market  Sugar U.S. import price  Sunflower oil          Tea  \
count          306.000000               306.000000     306.000000   306.000000
mean            12.911900                23.482727     868.499172   243.850481
std              5.511843                 4.321716     387.002967    62.340190
min              5.111200                17.084000     332.545455   143.370102
25%              8.723328                21.120854     583.051136   192.591274
50%             11.698964                21.940046     699.987813   230.472175
75%             16.457923                24.395422    1093.929545   279.725000
max             29.736842                40.367368    2300.190476   403.032348


                 Tin      Uranium        Wheat  Wool coarse     Wool fine  \
count     306.000000   306.000000   306.000000   306.000000    306.000000
mean    11427.789744    28.514516   174.300314   644.394825    854.520921
std      7211.691487    23.075775    62.288807   301.125862    291.283457
min      3698.368421     7.100000    88.552237   247.089050    417.466580
25%      5478.321415    10.124845   129.746083   400.358493    646.737966
50%      7415.171443    18.550453   158.628471   541.588789    747.554633
75%     17673.906579    41.888021   206.429419   883.484252   1030.052806
max     32347.694444   136.222222   403.812897  1391.473685   1865.439360


             Zinc
count   306.000000
mean   1589.249010
std     737.032802
min     748.809524
25%    1019.070350
50%    1274.896066
75%    2045.714380
max    4381.447368


[8 rows x 63 columns]
```

```
[2]: import neon as pdf

     url = "https://raw.githubusercontent.com/mwaskom/seaborn-data/master/titanic.
      ↪csv"

     dp = pd.read_csv(url)
```

```
    ---------------------------------------------------------------------------
    ModuleNotFoundError                       Traceback (most recent call last)
    Cell In[2], line 1
    ----> 1 import neon as pdf
          3 url = "https://raw.githubusercontent.com/mwaskom/seaborn-data/master/
      ↪titanic.csv"
          5 dp = pd.read_csv(url)

    ModuleNotFoundError: No module named 'neon'
```

```
[3]: import pandas as pd

     url = "https://raw.githubusercontent.com/mwaskom/seaborn-data/master/titanic.
      ↪csv"

     df = pd.read_csv(url)
```

```
[4]: com
```

```
    ---------------------------------------------------------------------------
    NameError                                 Traceback (most recent call last)
    Cell In[4], line 1
    ----> 1 com

    NameError: name 'com' is not defined
```

```
[5]: df
```

```
[5]:      survived  pclass     sex   age  sibsp  parch     fare embarked    class  \
     0           0       3    male  22.0      1      0   7.2500        S    Third
     1           1       1  female  38.0      1      0  71.2833        C    First
     2           1       3  female  26.0      0      0   7.9250        S    Third
     3           1       1  female  35.0      1      0  53.1000        S    First
     4           0       3    male  35.0      0      0   8.0500        S    Third
     ..        ...     ...     ...   ...    ...    ...      ...      ...      ...
     886         0       2    male  27.0      0      0  13.0000        S   Second
     887         1       1  female  19.0      0      0  30.0000        S    First
     888         0       3  female   NaN      1      2  23.4500        S    Third
```

12

```
889         1     1   male  26.0    0       0  30.0000      C  First
890         0     3   male  32.0    0       0   7.7500      Q  Third

        who  adult_male deck  embark_town alive  alone
0       man        True  NaN  Southampton    no  False
1     woman       False    C    Cherbourg   yes  False
2     woman       False  NaN  Southampton   yes   True
3     woman       False    C  Southampton   yes  False
4       man        True  NaN  Southampton    no   True
..      ...         ...  ...          ...   ...    ...
886     man        True  NaN  Southampton    no   True
887   woman       False    B  Southampton   yes   True
888   woman       False  NaN  Southampton    no  False
889     man        True    C    Cherbourg   yes   True
890     man        True  NaN   Queenstown    no   True

[891 rows x 15 columns]
```

[7]: `com`

[7]:
```
           Date  All Commodity Price Index  Non-Fuel Price Index  \
0    1980-01-01                        NaN                   NaN
1    1980-02-01                        NaN                   NaN
2    1980-03-01                        NaN                   NaN
3    1980-04-01                        NaN                   NaN
4    1980-05-01                        NaN                   NaN
..          ...                        ...                   ...
445  2017-02-01                 118.615999            145.232866
446  2017-03-01                 113.314945            142.940071
447  2017-04-01                 113.226876            138.528012
448  2017-05-01                 110.589462            138.865166
449  2017-06-01                 106.840977            138.622118

     Food and Beverage Price Index  Food Price Index  Beverage Price Index  \
0                              NaN               NaN            189.310021
1                              NaN               NaN            190.387928
2                              NaN               NaN            194.060355
3                              NaN               NaN            186.137888
4                              NaN               NaN            185.070241
..                             ...               ...                   ...
445                     152.198420        151.815075            155.737047
446                     148.145483        147.637196            152.837443
447                     145.566652        145.227898            148.693674
448                     150.514652        150.880372            147.138716
449                     152.565099        153.117377            147.467065

     Industrial Inputs Price Index  Agricultural Raw Materials Index  \
```

```
0                          81.889650                          78.900146
1                          83.048367                          75.715154
2                          75.228895                          69.002475
3                          72.471253                          67.877108
4                          69.588103                          65.879671
..                             …                                  …
445                       138.197090                         120.680238
446                       137.682182                         121.436062
447                       131.418412                         119.724719
448                       127.098235                         117.028866
449                       124.538572                         113.649909

        Metals Price Index  Fuel Energy Index  …  Sugar Free Market  \
0                84.049004                NaN  …          17.299999
1                88.345232                NaN  …          22.750000
2                79.726312                NaN  …          19.630001
3                75.789658                NaN  …          21.250000
4                72.266748                NaN  …          30.939999
..                     …                  …  …               …
445             150.849718         103.044209  …          20.348421
446             149.416945          95.983218  …          18.059565
447             139.864904          98.424834  …          16.404737
448             134.371459          94.047194  …          15.689091
449             132.403581          88.247907  …          13.748182

        Sugar U.S. import price  Sunflower oil          Tea           Tin  \
0                     19.660000     566.926957   225.179901   16973.587904
1                     24.689999     573.958609   233.094498   17090.213164
2                     21.180000     535.284522   226.833298   17460.590400
3                     22.669998     486.062957   221.806793   17041.712600
4                     31.889999     502.763130   229.611206   17180.604736
..                          …            …            …             …
445                   30.401579     966.971042   382.874800   19446.475000
446                   29.829130     936.780525   356.102841   19875.195652
447                   28.719474     939.952292   362.096533   19910.319444
448                   28.429545     949.833062   371.482435   20200.333333
449                   27.827273     936.240625   381.217333   19658.840909

         Uranium      Wheat   Wool coarse    Wool fine        Zinc
0      40.000000  175.634750    553.120940   684.277366   773.821533
1      38.000000  172.695236    568.154805   722.567093   868.620361
2      35.000000  163.509323    552.745109   695.956860   740.752441
3      32.000000  156.528030    510.650280   688.130355   707.683105
4      32.000000  161.304703    524.932443   720.760976   701.069092
..          …          …            …            …            …
445    25.058877  147.313095   1029.581497  1368.141078  2845.550000
446    24.552024  146.387567   1059.601984  1454.829350  2776.880435
```

```
447  23.165761  138.412425   991.122950  1404.982250  2614.916667
448  21.562894  146.486143  1019.954554  1433.465267  2590.214286
449  19.678442  157.196087  1065.808020  1403.827722  2573.397727

[450 rows x 64 columns]
```

[8]: `df`

```
[8]:      survived  pclass     sex   age  sibsp  parch     fare embarked   class  \
     0           0       3    male  22.0      1      0   7.2500        S   Third
     1           1       1  female  38.0      1      0  71.2833        C   First
     2           1       3  female  26.0      0      0   7.9250        S   Third
     3           1       1  female  35.0      1      0  53.1000        S   First
     4           0       3    male  35.0      0      0   8.0500        S   Third
     ..        ...     ...     ...   ...    ...    ...      ...      ...     ...
     886         0       2    male  27.0      0      0  13.0000        S  Second
     887         1       1  female  19.0      0      0  30.0000        S   First
     888         0       3  female   NaN      1      2  23.4500        S   Third
     889         1       1    male  26.0      0      0  30.0000        C   First
     890         0       3    male  32.0      0      0   7.7500        Q   Third

            who  adult_male deck  embark_town alive  alone
     0      man        True  NaN  Southampton    no  False
     1    woman       False    C    Cherbourg   yes  False
     2    woman       False  NaN  Southampton   yes   True
     3    woman       False    C  Southampton   yes  False
     4      man        True  NaN  Southampton    no   True
     ..     ...         ...  ...          ...   ...    ...
     886    man        True  NaN  Southampton    no   True
     887  woman       False    B  Southampton   yes   True
     888  woman       False  NaN  Southampton    no  False
     889    man        True    C    Cherbourg   yes   True
     890    man        True  NaN   Queenstown    no   True

     [891 rows x 15 columns]
```

[9]: `df.columns`

```
[9]: Index(['survived', 'pclass', 'sex', 'age', 'sibsp', 'parch', 'fare',
            'embarked', 'class', 'who', 'adult_male', 'deck', 'embark_town',
            'alive', 'alone'],
           dtype='object')
```

[12]: `df.groupby("age")["fare"].describe()`

```
[12]:        count    mean     std     min     25%     50%     75%  \
       age
```

```
        0.42     1.0      8.5167          NaN      8.5167      8.5167      8.5167      8.5167
        0.67     1.0     14.5000          NaN     14.5000     14.5000     14.5000     14.5000
        0.75     2.0     19.2583     0.000000     19.2583     19.2583     19.2583     19.2583
        0.83     2.0     23.8750     7.247845     18.7500     21.3125     23.8750     26.4375
        0.92     1.0    151.5500          NaN    151.5500    151.5500    151.5500    151.5500
        ...      ...        ...          ...         ...         ...         ...         ...
       70.00     2.0     40.7500    42.779960     10.5000     25.6250     40.7500     55.8750
       70.50     1.0      7.7500          NaN      7.7500      7.7500      7.7500      7.7500
       71.00     2.0     42.0792    10.500536     34.6542     38.3667     42.0792     45.7917
       74.00     1.0      7.7750          NaN      7.7750      7.7750      7.7750      7.7750
       80.00     1.0     30.0000          NaN     30.0000     30.0000     30.0000     30.0000

                 max
        age
        0.42     8.5167
        0.67    14.5000
        0.75    19.2583
        0.83    29.0000
        0.92   151.5500
        ...        ...
       70.00    71.0000
       70.50     7.7500
       71.00    49.5042
       74.00     7.7750
       80.00    30.0000

        [88 rows x 8 columns]
```

[15]: `df.groupby("sibsp")["age"].describe()`

[15]:

sibsp	count	mean	std	min	25%	50%	75%	max
0	471.0	31.397558	13.647767	0.42	22.00	29.0	39.00	80.0
1	183.0	30.089727	14.645033	0.67	20.00	30.0	39.00	70.0
2	25.0	22.620000	14.679230	0.75	16.00	23.0	28.00	53.0
3	12.0	13.916667	11.317391	2.00	3.75	9.5	23.25	33.0
4	18.0	7.055556	4.880601	1.00	3.25	6.5	9.00	17.0
5	5.0	10.200000	5.805170	1.00	9.00	11.0	14.00	16.0
8	0.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN

[16]: `df.describe()`

[16]:

	survived	pclass	age	sibsp	parch	fare
count	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000

25%	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

```python
[17]: df.groupby("Sex")["age"].describe()
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
Cell In[17], line 1
----> 1 df.groupby("Sex")["age"].describe()

File /opt/conda/lib/python3.11/site-packages/pandas/core/frame.py:8869, in
 DataFrame.groupby(self, by, axis, level, as_index, sort, group_keys, observed,
 dropna)
   8866 if level is None and by is None:
   8867     raise TypeError("You have to supply one of 'by' and 'level'")
-> 8869 return DataFrameGroupBy(
   8870     obj=self,
   8871     keys=by,
   8872     axis=axis,
   8873     level=level,
   8874     as_index=as_index,
   8875     sort=sort,
   8876     group_keys=group_keys,
   8877     observed=observed,
   8878     dropna=dropna,
   8879 )

File /opt/conda/lib/python3.11/site-packages/pandas/core/groupby/groupby.py:
 1278, in GroupBy.__init__(self, obj, keys, axis, level, grouper, exclusions,
 selection, as_index, sort, group_keys, observed, dropna)
   1275 self.dropna = dropna
   1277 if grouper is None:
-> 1278     grouper, exclusions, obj = get_grouper(
   1279         obj,
   1280         keys,
   1281         axis=axis,
   1282         level=level,
   1283         sort=sort,
   1284         observed=False if observed is lib.no_default else observed,
   1285         dropna=self.dropna,
   1286     )
   1288 if observed is lib.no_default:
   1289     if any(ping._passed_categorical for ping in grouper.groupings):

File /opt/conda/lib/python3.11/site-packages/pandas/core/groupby/grouper.py:
 1009, in get_grouper(obj, key, axis, level, sort, observed, validate, dropna)
```

```
1007            in_axis, level, gpr = False, gpr, None
1008        else:
-> 1009            raise KeyError(gpr)
1010 elif isinstance(gpr, Grouper) and gpr.key is not None:
1011        # Add key to exclusions
1012        exclusions.add(gpr.key)

KeyError: 'Sex'
```

[18]: `df.groupby("sex")["age"].describe()`

[18]:
	count	mean	std	min	25%	50%	75%	max
sex								
female	261.0	27.915709	14.110146	0.75	18.0	27.0	37.0	63.0
male	453.0	30.726645	14.678201	0.42	21.0	29.0	39.0	80.0

[ ]: