learn_series2

August 9, 2023

```
[2]: import pandas as pd
     import numpy as np
```

```
[36]: pokemon = pd.read_csv('../github_source/pokemon.csv')
      pokemon
```

```
[36]:           Pokemon          Type
      0        Bulbasaur  Grass / Poison
      1          Ivysaur  Grass / Poison
      2         Venusaur  Grass / Poison
      3       Charmander            Fire
      4       Charmeleon            Fire
      ..             ...             ...
      804      Stakataka    Rock / Steel
      805    Blacephalon    Fire / Ghost
      806        Zeraora        Electric
      807         Meltan           Steel
      808       Melmetal           Steel

      [809 rows x 2 columns]
```

```
[64]: #
      pokemon = pd.read_csv('pokemon.csv', index_col='Pokemon')
      pokemon
```

```
[64]:                      Type
      Pokemon
      Bulbasaur    Grass / Poison
      Ivysaur      Grass / Poison
      Venusaur     Grass / Poison
      Charmander             Fire
      Charmeleon            Fire
      ...                     ...
      Stakataka      Rock / Steel
      Blacephalon    Fire / Ghost
      Zeraora            Electric
      Meltan                Steel
```

```
Melmetal                Steel

[809 rows x 1 columns]
```

[54]:
```python
type(pokemon) #
```

[54]:
```
pandas.core.frame.DataFrame
```

[65]:
```python
#
pokemon = pokemon.squeeze('columns')
```

[66]:
```python
pokemon
```

[66]:
```
Pokemon
Bulbasaur       Grass / Poison
Ivysaur         Grass / Poison
Venusaur        Grass / Poison
Charmander               Fire
Charmeleon               Fire
                      …
Stakataka         Rock / Steel
Blacephalon       Fire / Ghost
Zeraora               Electric
Meltan                   Steel
Melmetal                 Steel
Name: Type, Length: 809, dtype: object
```

[67]:
```python
type(pokemon)
```

[67]:
```
pandas.core.series.Series
```

[40]:
```python
google = pd.read_csv('../github_source/google_stocks.csv')
type(google['Date'][0])
```

[40]:
```
str
```

[41]:
```python
 #
google = pd.read_csv('../github_source/google_stocks.csv', parse_dates=['Date'])
type(google['Date'][0])
```

[41]:
```
pandas._libs.tslibs.timestamps.Timestamp
```

[42]:
```python
#
google = pd.read_csv('../github_source/google_stocks.csv',␣
  ↪parse_dates=['Date'], index_col='Date').squeeze('columns')
google
```

```
[42]: Date
      2004-08-19      49.98
      2004-08-20      53.95
      2004-08-23      54.50
      2004-08-24      52.24
      2004-08-25      52.80
                       …
      2019-10-21    1246.15
      2019-10-22    1242.80
      2019-10-23    1259.13
      2019-10-24    1260.99
      2019-10-25    1265.13
      Name: Close, Length: 3824, dtype: float64
```

```
[45]: battles = pd.read_csv('../github_source/revolutionary_war.csv',␣
      ↪parse_dates=['Start Date'], index_col='Start Date')
      battles
```

```
[45]:                                          Battle          State
      Start Date
      1774-09-01                         Powder Alarm  Massachusetts
      1774-12-14  Storming of Fort William and Mary  New Hampshire
      1775-04-19   Battles of Lexington and Concord  Massachusetts
      1775-04-19                     Siege of Boston  Massachusetts
      1775-04-20                  Gunpowder Incident        Virginia
      …                                         …              …
      1782-09-11                Siege of Fort Henry        Virginia
      1782-09-13          Grand Assault on Gibraltar            NaN
      1782-10-18           Action of 18 October 1782            NaN
      1782-12-06            Action of 6 December 1782            NaN
      1783-01-22           Action of 22 January 1783        Virginia

      [232 rows x 2 columns]
```

```
[47]: #
      battles = pd.read_csv('../github_source/revolutionary_war.csv',␣
      ↪parse_dates=['Start Date'], index_col='Start Date', usecols=['State', 'Start␣
      ↪Date']).squeeze()
      battles
```

```
[47]: Start Date
      1774-09-01    Massachusetts
      1774-12-14    New Hampshire
      1775-04-19    Massachusetts
      1775-04-19    Massachusetts
      1775-04-20         Virginia
                          …
```

```
1782-09-11         Virginia
1782-09-13              NaN
1782-10-18              NaN
1782-12-06              NaN
1783-01-22         Virginia
Name: State, Length: 232, dtype: object
```

[68]: `type(battles), type(google), type(pokemon)`

[68]: (pandas.core.series.Series,
       pandas.core.series.Series,
       pandas.core.series.Series)

[72]:
```
    #
battles.sort_values(ascending=False, na_position='first')
```

[72]:
```
Start Date
1775-09-17              NaN
1775-12-31              NaN
1776-03-03              NaN
1776-03-25              NaN
1776-05-18              NaN
                        …
1777-05-17          Florida
1777-09-03         Delaware
1781-09-06      Connecticut
1779-07-05      Connecticut
1777-04-27      Connecticut
Name: State, Length: 232, dtype: object
```

[73]:
```
#    NAN
battles.dropna().sort_values(ascending=False)
```

[73]:
```
Start Date
1783-01-22         Virginia
1781-09-28         Virginia
1779-05-10         Virginia
1781-09-05         Virginia
1781-07-06         Virginia
                        …
1777-05-17          Florida
1777-09-03         Delaware
1781-09-06      Connecticut
1779-07-05      Connecticut
1777-04-27      Connecticut
Name: State, Length: 162, dtype: object
```

```
[76]: #
      battles.sort_index()
```

```
[76]: Start Date
      1774-09-01     Massachusetts
      1774-12-14     New Hampshire
      1775-04-19     Massachusetts
      1775-04-19     Massachusetts
      1775-04-20          Virginia
                          …
      1783-01-22          Virginia
      NaT              New Jersey
      NaT                Virginia
      NaT                     NaN
      NaT                     NaN
      Name: State, Length: 232, dtype: object
```

```
[80]: #
      google.nlargest(3)
```

```
[80]: Date
      2019-04-29    1287.58
      2019-04-26    1272.18
      2018-07-26    1268.33
      Name: Close, dtype: float64
```

```
[81]: google.nsmallest() #defualt 5
```

```
[81]: Date
      2004-09-03    49.82
      2004-09-01    49.94
      Name: Close, dtype: float64
```

```
[82]: google.max(), google.min()
```

```
[82]: (1287.58, 49.82)
```

```
[90]: #
      battles.value_counts()
```

```
[90]: State
      South Carolina    31
      New York          28
      New Jersey        24
      Virginia          21
      Massachusetts     11
      Pennsylvania      10
```

```
North Carolina     9
Florida            8
Georgia            6
Rhode Island       3
Connecticut        3
Vermont            3
New Hampshire      1
Delaware           1
Indiana            1
Louisiana          1
Ohio               1
Name: count, dtype: int64
```

[95]:
```python
#
#battles.value_counts(normalize=True)
(battles.value_counts(normalize=True)*100).round(2)   #
```

[95]:
```
State
South Carolina    19.14
New York          17.28
New Jersey        14.81
Virginia          12.96
Massachusetts      6.79
Pennsylvania       6.17
North Carolina     5.56
Florida            4.94
Georgia            3.70
Rhode Island       1.85
Connecticut        1.85
Vermont            1.85
New Hampshire      0.62
Delaware           0.62
Indiana            0.62
Louisiana          0.62
Ohio               0.62
Name: proportion, dtype: float64
```

[91]:
```python
#
battles.nunique()
```

[91]: 17

[97]:
```python
#                    :
bins = [0, 200, 400, 600, 800, 1300]
google.value_counts(bins=bins).sort_index()
```

```
[97]: Close
      (-0.001, 200.0]      595
      (200.0, 400.0]      1568
      (400.0, 600.0]       575
      (600.0, 800.0]       380
      (800.0, 1300.0]      706
      Name: count, dtype: int64
```

```python
[98]: #                                           :
      google.value_counts(bins=4).sort_index()
```

```
[98]: Close
      (48.581, 359.26]      2086
      (359.26, 668.7]        729
      (668.7, 978.14]        495
      (978.14, 1287.58]      514
      Name: count, dtype: int64
```

```python
[99]: #                                           :
      battles.index.value_counts()
```

```
[99]: Start Date
      1781-04-25    2
      1781-05-22    2
      1780-08-18    2
      1781-09-13    2
      1782-03-16    2
                   ..
      1778-06-30    1
      1778-07-03    1
      1778-07-27    1
      1778-08-21    1
      1783-01-22    1
      Name: count, Length: 217, dtype: int64
```

```python
[101]: #
       google.apply(round) #
```

```
[101]: Date
       2004-08-19      50
       2004-08-20      54
       2004-08-23      54
       2004-08-24      52
       2004-08-25      53
                      …
       2019-10-21    1246
       2019-10-22    1243
```

```
2019-10-23    1259
2019-10-24    1261
2019-10-25    1265
Name: Close, Length: 3824, dtype: int64
```

[ ]: