

# pandas weather

In [1]:

```
import numpy as np
import pandas as pd

def header(msg):
    print('-' * 50)
    print('[' + msg + '']')
```

In [3]:

```
df_list = {'Month':['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec'],
           'avg_high':[45, 54, 76, 76, 23, 86, 66, 77, 87, 54, 55, 76],
           'avg_low':[23, 44, 23, 23, 23, 43, 44, 23, 56, 46, 54, 66],
           'record_high':[77, 65, 76, 54, 23, 77, 65, 56, 87, 77, 55, 89],
           'record_low':[22, 12, 15, 45, 20, 26, 34, 23, 22, 11, 21, 32],
           'avg_precipitation':[4.5, 5.04, 3.76, 2.76, 1.23, 3.86,1.66, 1.7, 3.87, 2.54, 5.5, 4.76]}

df = pd.DataFrame(df_list)
header('Weather Prescription')
df
```

-----  
[Weather Prescription]

Out[3]:

	Month	avg_high	avg_low	record_high	record_low	avg_precipitation
0	Jan	45	23	77	22	4.50
1	Feb	54	44	65	12	5.04
2	Mar	76	23	76	15	3.76
3	Apr	76	23	54	45	2.76
4	May	23	23	23	20	1.23
5	Jun	86	43	77	26	3.86
6	Jul	66	44	65	34	1.66
7	Aug	77	23	56	23	1.70
8	Sep	87	56	87	22	3.87
9	Oct	54	46	77	11	2.54
10	Nov	55	54	55	21	5.50
11	Dec	76	66	89	32	4.76

In [9]:

```
header('2 - Weather Prescription - Data from Text File')
filename = 'weather.txt'
df = pd.read_csv(filename)
df
```

-----  
[2 - Weather Prescription - Data from Text File]

Out[9]:

	Month	avg_high	avg_low	record_high	record_low	avg_precipitation
0	Jan	45	23	77	22	4.50
1	Feb	54	44	65	12	5.04
2	Mar	76	23	76	15	3.76
3	Apr	76	23	54	45	2.76
4	May	23	23	23	20	1.23
5	Jun	86	43	77	26	3.86
6	Jul	66	44	65	34	1.66
7	Aug	77	23	56	23	1.70
8	Sep	87	56	87	22	3.87
9	Oct	54	46	77	11	2.54
10	Nov	55	54	55	21	5.50
11	Dec	76	66	89	32	4.76

In [10]:

```
df.head(5) # prints first 5
```

Out[10]:

	Month	avg_high	avg_low	record_high	record_low	avg_precipitation
0	Jan	45	23	77	22	4.50
1	Feb	54	44	65	12	5.04
2	Mar	76	23	76	15	3.76
3	Apr	76	23	54	45	2.76
4	May	23	23	23	20	1.23

In [11]:

```
df.tail(3) # prints last 3
```

Out[11]:

	Month	avg_high	avg_low	record_high	record_low	avg_precipitation
9	Oct	54	46	77	11	2.54
10	Nov	55	54	55	21	5.50
11	Dec	76	66	89	32	4.76

In [12]:

```
df.dtypes
```

Out[12]:

```
Month                object
avg_high             int64
avg_low              int64
record_high          int64
record_low           int64
avg_precipitation    float64
dtype: object
```

In [13]:

```
df.index
```

Out[13]:

```
RangeIndex(start=0, stop=12, step=1)
```

In [14]:

```
df.columns
```

Out[14]:

```
Index(['Month', ' avg_high', ' avg_low', ' record_high', ' record_low',
      ' avg_precipitation'],
      dtype='object')
```

In [15]:

```
df.values
```

Out[15]:

```
array([[ 'Jan', 45, 23, 77, 22, 4.5],
       [ 'Feb', 54, 44, 65, 12, 5.04],
       [ 'Mar', 76, 23, 76, 15, 3.76],
       [ 'Apr', 76, 23, 54, 45, 2.76],
       [ 'May', 23, 23, 23, 20, 1.23],
       [ 'Jun', 86, 43, 77, 26, 3.86],
       [ 'Jul', 66, 44, 65, 34, 1.66],
       [ 'Aug', 77, 23, 56, 23, 1.7],
       [ 'Sep', 87, 56, 87, 22, 3.87],
       [ 'Oct', 54, 46, 77, 11, 2.54],
       [ 'Nov', 55, 54, 55, 21, 5.5],
       [ 'Dec', 76, 66, 89, 32, 4.76]], dtype=object)
```

In [16]:

```
df.describe()
```

Out[16]:

	avg_high	avg_low	record_high	record_low	avg_precipitation
<b>count</b>	12.000000	12.000000	12.000000	12.000000	12.000000
<b>mean</b>	64.583333	39.000000	66.750000	23.583333	3.431667
<b>std</b>	18.918285	15.486064	18.196528	9.699656	1.429716
<b>min</b>	23.000000	23.000000	23.000000	11.000000	1.230000
<b>25%</b>	54.000000	23.000000	55.750000	18.750000	2.330000
<b>50%</b>	71.000000	43.500000	70.500000	22.000000	3.810000
<b>75%</b>	76.250000	48.000000	77.000000	27.500000	4.565000
<b>max</b>	87.000000	66.000000	89.000000	45.000000	5.500000

## Slicing

In [23]:

```
print(df[2:4]) #rows 2 to 3
```

	Month	avg_high	avg_low	record_high	record_low	avg_precipitation
2	Mar	76	23	76	15	3.76
3	Apr	76	23	54	45	2.76

In [49]:

```
df.iloc[3:5, [0,3]]
```

Out[49]:

	Month	record_high
3	Apr	54
4	May	23

In [53]:

```
df[df['Month'].isin(['Jun','Jul'])] #filters months in the dataframe
```

Out[53]:

	Month	avg_high	avg_low	record_high	record_low	avg_precipitation
5	Jun	86	43	77	26	3.86
6	Jul	66	44	65	34	1.66

## Rename column

In [69]:

```
df.rename(columns = {'avg_low':'ave_low'}, inplace=True)  
df.head()
```

Out[69]:

	Month	avg_high	avg_low	record_high	record_low	avg_precipitation
0	Jan	45	23	77	22	4.50
1	Feb	54	44	65	12	5.04
2	Mar	76	23	76	15	3.76
3	Apr	76	23	54	45	2.76
4	May	23	23	23	20	1.23

In [70]:

```
df.rename(columns = {'avg_low':'ave_low'}, inplace=True)
```

In [73]:

```
df
```

Out[73]:

	Month	avg_high	avg_low	record_high	record_low	avg_precipitation
0	Jan	45	23	77	22	4.50
1	Feb	54	44	65	12	5.04
2	Mar	76	23	76	15	3.76
3	Apr	76	23	54	45	2.76
4	May	23	23	23	20	1.23
5	Jun	86	43	77	26	3.86
6	Jul	66	44	65	34	1.66
7	Aug	77	23	56	23	1.70
8	Sep	87	56	87	22	3.87
9	Oct	54	46	77	11	2.54
10	Nov	55	54	55	21	5.50
11	Dec	76	66	89	32	4.76

In [ ]:

In [ ]: