

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as se
```

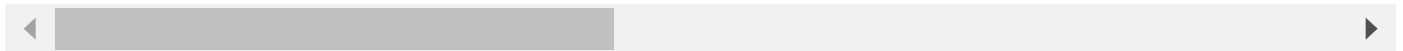
```
In [2]: weather=pd.read_csv("/home/student/Downloads/weatherAUS.csv")
```

```
In [3]: weather.head(6)
```

```
Out[3]:
```

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpe
0	2008-12-01	Albury	13.4	22.9	0.6	NaN	NaN	W	4.
1	2008-12-02	Albury	7.4	25.1	0.0	NaN	NaN	WNW	4.
2	2008-12-03	Albury	12.9	25.7	0.0	NaN	NaN	WSW	4.
3	2008-12-04	Albury	9.2	28.0	0.0	NaN	NaN	NE	2.
4	2008-12-05	Albury	17.5	32.3	1.0	NaN	NaN	W	4.
5	2008-12-06	Albury	14.6	29.7	0.2	NaN	NaN	WNW	5.

6 rows × 24 columns



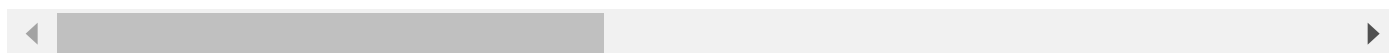
```
In [4]: weather.tail(10)
```

Out[4]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpeed
--	------	----------	---------	---------	----------	-------------	----------	-------------	---------------

<b>142183</b>	2017-06-15	Uluru	2.6	22.5	0.0	NaN	NaN	S	1.0
<b>142184</b>	2017-06-16	Uluru	5.2	24.3	0.0	NaN	NaN	E	1.0
<b>142185</b>	2017-06-17	Uluru	6.4	23.4	0.0	NaN	NaN	ESE	1.0
<b>142186</b>	2017-06-18	Uluru	8.0	20.7	0.0	NaN	NaN	ESE	1.0
<b>142187</b>	2017-06-19	Uluru	7.4	20.6	0.0	NaN	NaN	E	1.0
<b>142188</b>	2017-06-20	Uluru	3.5	21.8	0.0	NaN	NaN	E	1.0
<b>142189</b>	2017-06-21	Uluru	2.8	23.4	0.0	NaN	NaN	E	1.0
<b>142190</b>	2017-06-22	Uluru	3.6	25.3	0.0	NaN	NaN	NNW	1.0
<b>142191</b>	2017-06-23	Uluru	5.4	26.9	0.0	NaN	NaN	N	1.0
<b>142192</b>	2017-06-24	Uluru	7.8	27.0	0.0	NaN	NaN	SE	1.0

10 rows × 24 columns



In [5]: `weather.index`

Out[5]: `RangeIndex(start=0, stop=142193, step=1)`

In [6]: `weather.columns`

Out[6]: `Index(['Date', 'Location', 'MinTemp', 'MaxTemp', 'Rainfall', 'Evaporation', 'Sunshine', 'WindGustDir', 'WindGustSpeed', 'WindDir9am', 'WindDir3pm', 'WindSpeed9am', 'WindSpeed3pm', 'Humidity9am', 'Humidity3pm', 'Pressure9am', 'Pressure3pm', 'Cloud9am', 'Cloud3pm', 'Temp9am', 'Temp3pm', 'RainToday', 'RISK_MM', 'RainTomorrow'], dtype='object')`

In [7]: `weather.shape`

Out[7]: `(142193, 24)`

In [8]: `weather.dtypes`

```
Out[8]: Date          object
Location         object
MinTemp          float64
MaxTemp          float64
Rainfall         float64
Evaporation      float64
Sunshine         float64
WindGustDir      object
WindGustSpeed    float64
WindDir9am      object
WindDir3pm      object
WindSpeed9am     float64
WindSpeed3pm     float64
Humidity9am      float64
Humidity3pm      float64
Pressure9am      float64
Pressure3pm      float64
Cloud9am         float64
Cloud3pm         float64
Temp9am          float64
Temp3pm          float64
RainToday        object
RISK_MM          float64
RainTomorrow     object
dtype: object
```

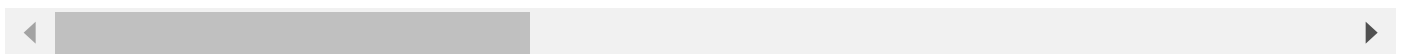
```
In [9]: weather.columns.values
```

```
Out[9]: array(['Date', 'Location', 'MinTemp', 'MaxTemp', 'Rainfall',
              'Evaporation', 'Sunshine', 'WindGustDir', 'WindGustSpeed',
              'WindDir9am', 'WindDir3pm', 'WindSpeed9am', 'WindSpeed3pm',
              'Humidity9am', 'Humidity3pm', 'Pressure9am', 'Pressure3pm',
              'Cloud9am', 'Cloud3pm', 'Temp9am', 'Temp3pm', 'RainToday',
              'RISK_MM', 'RainTomorrow'], dtype=object)
```

```
In [10]: weather.describe(include='all')
```

Out[10]:	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	1
<b>count</b>	142193	142193	141556.000000	141871.000000	140787.000000	81350.000000	74377.000000	
<b>unique</b>	3436	49	NaN	NaN	NaN	NaN	NaN	
<b>top</b>	2013-12-01	Canberra	NaN	NaN	NaN	NaN	NaN	
<b>freq</b>	49	3418	NaN	NaN	NaN	NaN	NaN	
<b>mean</b>	NaN	NaN	12.186400	23.226784	2.349974	5.469824	7.624853	
<b>std</b>	NaN	NaN	6.403283	7.117618	8.465173	4.188537	3.781525	
<b>min</b>	NaN	NaN	-8.500000	-4.800000	0.000000	0.000000	0.000000	
<b>25%</b>	NaN	NaN	7.600000	17.900000	0.000000	2.600000	4.900000	
<b>50%</b>	NaN	NaN	12.000000	22.600000	0.000000	4.800000	8.500000	
<b>75%</b>	NaN	NaN	16.800000	28.200000	0.800000	7.400000	10.600000	
<b>max</b>	NaN	NaN	33.900000	48.100000	371.000000	145.000000	14.500000	

11 rows × 24 columns



In [11]: `weather['Date']`

Out[11]:

```

0      2008-12-01
1      2008-12-02
2      2008-12-03
3      2008-12-04
4      2008-12-05
...
142188 2017-06-20
142189 2017-06-21
142190 2017-06-22
142191 2017-06-23
142192 2017-06-24
Name: Date, Length: 142193, dtype: object

```

In [12]: `weather.sort_index(axis=1,ascending=False)`

Out[12]:

	WindSpeed9am	WindSpeed3pm	WindGustSpeed	WindGustDir	WindDir9am	WindDir3pm
0	20.0	24.0	44.0	W	W	WNW
1	4.0	22.0	44.0	WNW	NNW	WSW
2	19.0	26.0	46.0	WSW	W	WSW
3	11.0	9.0	24.0	NE	SE	E
4	7.0	20.0	41.0	W	ENE	NW
...	...	...	...	...	...	...
142188	15.0	13.0	31.0	E	ESE	E
142189	13.0	11.0	31.0	E	SE	ENE
142190	13.0	9.0	22.0	NNW	SE	N
142191	9.0	9.0	37.0	N	SE	WNW
142192	13.0	7.0	28.0	SE	SSE	N

142193 rows × 24 columns

◀

▶

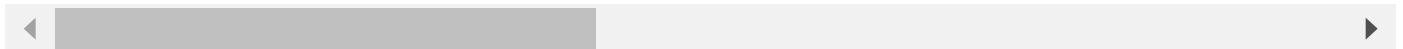
In [13]:

weather.sort\_values(by="Date")

Out[13]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	Win
<b>44351</b>	2007-11-01	Canberra	8.0	24.3	0.0	3.4	6.3	NW	
<b>44352</b>	2007-11-02	Canberra	14.0	26.9	3.6	4.4	9.7	ENE	
<b>44353</b>	2007-11-03	Canberra	13.7	23.4	3.6	5.8	3.3	NW	
<b>44354</b>	2007-11-04	Canberra	13.3	15.5	39.8	7.2	9.1	NW	
<b>44355</b>	2007-11-05	Canberra	7.6	16.1	2.8	5.6	10.6	SSE	
...	...	...	...	...	...	...	...	...	...
<b>38814</b>	2017-06-25	WaggaWagga	-0.1	12.7	0.0	NaN	NaN	SSW	
<b>32833</b>	2017-06-25	Sydney	7.6	19.3	0.0	3.4	9.4	W	
<b>29496</b>	2017-06-25	Richmond	2.7	20.0	0.2	NaN	NaN	W	
<b>26545</b>	2017-06-25	Penrith	3.1	20.0	0.0	NaN	NaN	SW	
<b>35838</b>	2017-06-25	SydneyAirport	6.8	19.5	0.0	3.4	9.4	WSW	

142193 rows × 24 columns



In [14]:

```
weather.iloc[5]
```

```
Out[14]: Date                2008-12-06
Location                Albury
MinTemp                 14.6
MaxTemp                 29.7
Rainfall                0.2
Evaporation             NaN
Sunshine                NaN
WindGustDir             WNW
WindGustSpeed           56.0
WindDir9am              W
WindDir3pm              W
WindSpeed9am            19.0
WindSpeed3pm            24.0
Humidity9am             55.0
Humidity3pm             23.0
Pressure9am             1009.2
Pressure3pm             1005.4
Cloud9am                NaN
Cloud3pm                NaN
Temp9am                 20.6
Temp3pm                 28.9
RainToday               No
RISK_MM                 0.0
RainTomorrow            No
Name: 5, dtype: object
```

```
In [15]: weather[0:3]
```

```
Out[15]:
```

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpe
0	2008-12-01	Albury	13.4	22.9	0.6	NaN	NaN	W	4
1	2008-12-02	Albury	7.4	25.1	0.0	NaN	NaN	WNW	4
2	2008-12-03	Albury	12.9	25.7	0.0	NaN	NaN	WSW	4

3 rows × 24 columns

```
In [17]: weather.loc[:, ["Date", "Location"]]
```

Out[17]:

	Date	Location
0	2008-12-01	Albury
1	2008-12-02	Albury
2	2008-12-03	Albury
3	2008-12-04	Albury
4	2008-12-05	Albury
...	...	...
142188	2017-06-20	Uluru
142189	2017-06-21	Uluru
142190	2017-06-22	Uluru
142191	2017-06-23	Uluru
142192	2017-06-24	Uluru

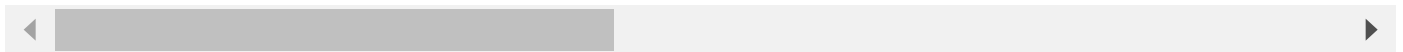
142193 rows × 2 columns

In [18]: `weather.iloc[:,2:]`

Out[18]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGustSpe
0	2008-12-01	Albury	13.4	22.9	0.6	NaN	NaN	W	4
1	2008-12-02	Albury	7.4	25.1	0.0	NaN	NaN	WNW	4

2 rows × 24 columns



In [19]: `weather.iloc[:,2]`



Out[19]:

	Date	Location
0	2008-12-01	Albury
1	2008-12-02	Albury
2	2008-12-03	Albury
3	2008-12-04	Albury
4	2008-12-05	Albury
...	...	...
142188	2017-06-20	Uluru
142189	2017-06-21	Uluru
142190	2017-06-22	Uluru
142191	2017-06-23	Uluru
142192	2017-06-24	Uluru

142193 rows × 2 columns

In [20]:

```
weather.iloc[:,3]
```

Out[20]:

	Date	Location	MinTemp
0	2008-12-01	Albury	13.4
1	2008-12-02	Albury	7.4

In [22]:

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

Out[22]:

	Date	Location	MinTemp
3	2008-12-04	Albury	9.2
4	2008-12-05	Albury	17.5

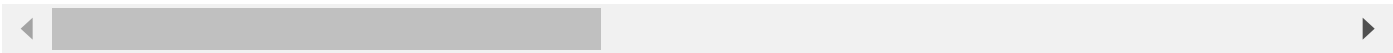
In [25]:

```
weather.isnull()
```

Out[25]:

	Date	Location	MinTemp	MaxTemp	Rainfall	Evaporation	Sunshine	WindGustDir	WindGu
0	False	False	False	False	False	True	True	False	
1	False	False	False	False	False	True	True	False	
2	False	False	False	False	False	True	True	False	
3	False	False	False	False	False	True	True	False	
4	False	False	False	False	False	True	True	False	
...	...	...	...	...	...	...	...	...	
142188	False	False	False	False	False	True	True	False	
142189	False	False	False	False	False	True	True	False	
142190	False	False	False	False	False	True	True	False	
142191	False	False	False	False	False	True	True	False	
142192	False	False	False	False	False	True	True	False	

142193 rows × 24 columns



In [ ]: