

```
In [2]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as se
```

```
In [4]: weather = pd.read_csv("/home/student/Downloads/weatherAUS.csv")
```

```
In [5]: weather.head(6)
```

```
Out[5]:
```

| | Date | Location | MinTemp | MaxTemp | Rainfall | Evaporation | Sunshine | WindGustDir | WindGustSpeed | WindDir |
|---|------------|----------|---------|---------|----------|-------------|----------|-------------|---------------|---------|
| 0 | 2008-12-01 | Albury | 13.4 | 22.9 | 0.6 | NaN | NaN | W | 44.0 | |
| 1 | 2008-12-02 | Albury | 7.4 | 25.1 | 0.0 | NaN | NaN | WNW | 44.0 | |
| 2 | 2008-12-03 | Albury | 12.9 | 25.7 | 0.0 | NaN | NaN | WSW | 46.0 | |
| 3 | 2008-12-04 | Albury | 9.2 | 28.0 | 0.0 | NaN | NaN | NE | 24.0 | |
| 4 | 2008-12-05 | Albury | 17.5 | 32.3 | 1.0 | NaN | NaN | W | 41.0 | |
| 5 | 2008-12-06 | Albury | 14.6 | 29.7 | 0.2 | NaN | NaN | WNW | 56.0 | |

6 rows × 24 columns

```
In [6]: weather.tail(4)
```

```
Out[6]:
```

| | Date | Location | MinTemp | MaxTemp | Rainfall | Evaporation | Sunshine | WindGustDir | WindGustSpeed | WindDir |
|--------|------------|----------|---------|---------|----------|-------------|----------|-------------|---------------|---------|
| 142189 | 2017-06-21 | Uluru | 2.8 | 23.4 | 0.0 | NaN | NaN | E | 31.0 | |
| 142190 | 2017-06-22 | Uluru | 3.6 | 25.3 | 0.0 | NaN | NaN | NNW | 22.0 | |
| 142191 | 2017-06-23 | Uluru | 5.4 | 26.9 | 0.0 | NaN | NaN | N | 37.0 | |
| 142192 | 2017-06-24 | Uluru | 7.8 | 27.0 | 0.0 | NaN | NaN | SE | 28.0 | |

4 rows × 24 columns

```
In [7]: weather.index
```

```
Out[7]: RangeIndex(start=0, stop=142193, step=1)
```

```
In [8]: weather.columns
```

```
Out[8]: 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 [9]: weather.shape
```

Out[9]: (142193, 24)

In [10]: weather.dtypes

Out[10]: 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 [11]: weather.columns.values

Out[11]: 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 [12]: weather.describe(include='all')

| | Date | Location | MinTemp | MaxTemp | Rainfall | Evaporation | Sunshine | WindGust |
|--------|------------|----------|---------------|---------------|---------------|--------------|--------------|----------|
| count | 142193 | 142193 | 141556.000000 | 141871.000000 | 140787.000000 | 81350.000000 | 74377.000000 | 1328 |
| unique | 3436 | 49 | NaN | NaN | NaN | NaN | NaN | |
| top | 2013-12-01 | Canberra | NaN | NaN | NaN | NaN | NaN | |
| freq | 49 | 3418 | NaN | NaN | NaN | NaN | NaN | 97 |
| mean | NaN | NaN | 12.186400 | 23.226784 | 2.349974 | 5.469824 | 7.624853 | N |
| std | NaN | NaN | 6.403283 | 7.117618 | 8.465173 | 4.188537 | 3.781525 | N |
| min | NaN | NaN | -8.500000 | -4.800000 | 0.000000 | 0.000000 | 0.000000 | N |
| 25% | NaN | NaN | 7.600000 | 17.900000 | 0.000000 | 2.600000 | 4.900000 | N |
| 50% | NaN | NaN | 12.000000 | 22.600000 | 0.000000 | 4.800000 | 8.500000 | N |
| 75% | NaN | NaN | 16.800000 | 28.200000 | 0.800000 | 7.400000 | 10.600000 | N |
| max | NaN | NaN | 33.900000 | 48.100000 | 371.000000 | 145.000000 | 14.500000 | N |

11 rows × 24 columns

```
In [15]: weather['Date']

Out[15]: 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 [16]: weather.iloc[5]

Out[16]: 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 [17]: weather[0:3]

Out[17]:
```

| | Date | Location | MinTemp | MaxTemp | Rainfall | Evaporation | Sunshine | WindGustDir | WindGustSpeed | WindDi |
|---|------------|----------|---------|---------|----------|-------------|----------|-------------|---------------|--------|
| 0 | 2008-12-01 | Albury | 13.4 | 22.9 | 0.6 | NaN | NaN | W | 44.0 | |
| 1 | 2008-12-02 | Albury | 7.4 | 25.1 | 0.0 | NaN | NaN | WNW | 44.0 | |
| 2 | 2008-12-03 | Albury | 12.9 | 25.7 | 0.0 | NaN | NaN | WSW | 46.0 | |

3 rows × 24 columns

```
In [18]: weather.loc[:,["Date", "Location"]]
```

Out[18]:

| | 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 [19]: weather.iloc[:2,]

| | Date | Location | MinTemp | MaxTemp | Rainfall | Evaporation | Sunshine | WindGustDir | WindGustSpeed | WindDi |
|---|------------|----------|---------|---------|----------|-------------|----------|-------------|---------------|--------|
| 0 | 2008-12-01 | Albury | 13.4 | 22.9 | 0.6 | NaN | NaN | W | 44.0 | |
| 1 | 2008-12-02 | Albury | 7.4 | 25.1 | 0.0 | NaN | NaN | WNW | 44.0 | |

2 rows × 24 columns

In [20]: weather.iloc[:2,5]

Out[20]: 0 NaN
1 NaN
Name: Evaporation, dtype: float64

In [21]: weather.iloc[4:,]

Out[21]:

| | | Date | Location | MinTemp | MaxTemp | Rainfall | Evaporation | Sunshine | WindGustDir | WindGustSpeed | W |
|--|--------|------------|----------|---------|---------|----------|-------------|----------|-------------|---------------|---|
| | 4 | 2008-12-05 | Albury | 17.5 | 32.3 | 1.0 | NaN | NaN | W | 41.0 | |
| | 5 | 2008-12-06 | Albury | 14.6 | 29.7 | 0.2 | NaN | NaN | WNW | 56.0 | |
| | 6 | 2008-12-07 | Albury | 14.3 | 25.0 | 0.0 | NaN | NaN | W | 50.0 | |
| | 7 | 2008-12-08 | Albury | 7.7 | 26.7 | 0.0 | NaN | NaN | W | 35.0 | |
| | 8 | 2008-12-09 | Albury | 9.7 | 31.9 | 0.0 | NaN | NaN | NNW | 80.0 | |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| | 142188 | 2017-06-20 | Uluru | 3.5 | 21.8 | 0.0 | NaN | NaN | E | 31.0 | |
| | 142189 | 2017-06-21 | Uluru | 2.8 | 23.4 | 0.0 | NaN | NaN | E | 31.0 | |
| | 142190 | 2017-06-22 | Uluru | 3.6 | 25.3 | 0.0 | NaN | NaN | NNW | 22.0 | |
| | 142191 | 2017-06-23 | Uluru | 5.4 | 26.9 | 0.0 | NaN | NaN | N | 37.0 | |
| | 142192 | 2017-06-24 | Uluru | 7.8 | 27.0 | 0.0 | NaN | NaN | SE | 28.0 | |

142189 rows × 24 columns

In [22]:

weather.iloc[:2,:2]

Out[22]:

| | | Date | Location |
|---|------------|--------|----------|
| 0 | 2008-12-01 | Albury | |
| 1 | 2008-12-02 | Albury | |

In [23]:

weather.iloc[3:5,0:2]

Out[23]:

| | | Date | Location |
|---|------------|--------|----------|
| 3 | 2008-12-04 | Albury | |
| 4 | 2008-12-05 | Albury | |

In []: