In [2]:
```
!pip install pandas
```

```
Defaulting to user installation because normal site-packages is not writeable
Collecting pandas
  Downloading pandas-2.0.3-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (18 kB)
Requirement already satisfied: python-dateutil>=2.8.2 in /home/kaanthanvishnu/.local/lib/python3.8/site-packages (fr
om pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /home/kaanthanvishnu/.local/lib/python3.8/site-packages (from pandas)
(2025.1)
Collecting tzdata>=2022.1 (from pandas)
  Downloading tzdata-2025.1-py2.py3-none-any.whl.metadata (1.4 kB)
Collecting numpy>=1.20.3 (from pandas)
  Downloading numpy-1.24.4-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (5.6 kB)
Requirement already satisfied: six>=1.5 in /usr/lib/python3/dist-packages (from python-dateutil>=2.8.2->pandas) (1.1
4.0)
Downloading pandas-2.0.3-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (12.4 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 12.4/12.4 MB 52.8 MB/s eta 0:00:00:00:01
Downloading numpy-1.24.4-cp38-cp38-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.3 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 17.3/17.3 MB 63.4 MB/s eta 0:00:00:00:01
Downloading tzdata-2025.1-py2.py3-none-any.whl (346 kB)
Installing collected packages: tzdata, numpy, pandas
Successfully installed numpy-1.24.4 pandas-2.0.3 tzdata-2025.1
```

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

In [16]:
```
data = pd.read_csv("~/dataset/climate-change-earth-surface-temperature-data/GlobalTemperatures.csv")
print(data.info())
print(data.describe())
print(data.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3192 entries, 0 to 3191
Data columns (total 9 columns):
 #   Column                                     Non-Null Count  Dtype  
---  ------                                     --------------  -----  
 0   dt                                         3192 non-null   object 
 1   LandAverageTemperature                     3180 non-null   float64
 2   LandAverageTemperatureUncertainty          3180 non-null   float64
 3   LandMaxTemperature                         1992 non-null   float64
 4   LandMaxTemperatureUncertainty              1992 non-null   float64
 5   LandMinTemperature                         1992 non-null   float64
 6   LandMinTemperatureUncertainty              1992 non-null   float64
 7   LandAndOceanAverageTemperature             1992 non-null   float64
 8   LandAndOceanAverageTemperatureUncertainty  1992 non-null   float64
dtypes: float64(8), object(1)
memory usage: 224.6+ KB
None
       LandAverageTemperature  LandAverageTemperatureUncertainty  \
count             3180.000000                        3180.000000   
mean                 8.374731                           0.938468   
std                  4.381310                           1.096440   
min                 -2.080000                           0.034000   
25%                  4.312000                           0.186750   
50%                  8.610500                           0.392000   
75%                 12.548250                           1.419250   
max                 19.021000                           7.880000   

       LandMaxTemperature  LandMaxTemperatureUncertainty  LandMinTemperature  \
count         1992.000000                    1992.000000         1992.000000   
mean            14.350601                       0.479782            2.743595   
std              4.309579                       0.583203            4.155835   
min              5.900000                       0.044000           -5.407000   
25%             10.212000                       0.142000           -1.334500   
50%             14.760000                       0.252000            2.949500   
75%             18.451500                       0.539000            6.778750   
max             21.320000                       4.373000            9.715000   

       LandMinTemperatureUncertainty  LandAndOceanAverageTemperature  \
count                    1992.000000                     1992.000000   
mean                        0.431849                       15.212566   
std                         0.445838                        1.274093   
```

```
min                          0.045000                      12.475000
25%                          0.155000                      14.047000
50%                          0.279000                      15.251000
75%                          0.458250                      16.396250
max                          3.498000                      17.611000

       LandAndOceanAverageTemperatureUncertainty
count                                1992.000000
mean                                    0.128532
std                                     0.073587
min                                     0.042000
25%                                     0.063000
50%                                     0.122000
75%                                     0.151000
max                                     0.457000
           dt  LandAverageTemperature  LandAverageTemperatureUncertainty  \
0  1750-01-01                   3.034                              3.574
1  1750-02-01                   3.083                              3.702
2  1750-03-01                   5.626                              3.076
3  1750-04-01                   8.490                              2.451
4  1750-05-01                  11.573                              2.072

   LandMaxTemperature  LandMaxTemperatureUncertainty  LandMinTemperature  \
0                 NaN                            NaN                 NaN
1                 NaN                            NaN                 NaN
2                 NaN                            NaN                 NaN
3                 NaN                            NaN                 NaN
4                 NaN                            NaN                 NaN

   LandMinTemperatureUncertainty  LandAndOceanAverageTemperature  \
0                            NaN                             NaN
1                            NaN                             NaN
2                            NaN                             NaN
3                            NaN                             NaN
4                            NaN                             NaN

   LandAndOceanAverageTemperatureUncertainty
0                                        NaN
1                                        NaN
2                                        NaN
```

```
3                                       NaN
4                                       NaN
```

In [17]:
```
print(data.isnull().sum())
```

```
dt                                          0
LandAverageTemperature                     12
LandAverageTemperatureUncertainty          12
LandMaxTemperature                       1200
LandMaxTemperatureUncertainty            1200
LandMinTemperature                       1200
LandMinTemperatureUncertainty            1200
LandAndOceanAverageTemperature           1200
LandAndOceanAverageTemperatureUncertainty    1200
dtype: int64
```

In [8]:
```
country_data = pd.read_csv("~/dataset/climate-change-earth-surface-temperature-data/GlobalLandTemperaturesByCountry
print(country_data.info())
print(country_data.describe())
print(country_data.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 577462 entries, 0 to 577461
Data columns (total 4 columns):
 #   Column                         Non-Null Count   Dtype
---  ------                         --------------   -----
 0   dt                             577462 non-null  object
 1   AverageTemperature             544811 non-null  float64
 2   AverageTemperatureUncertainty  545550 non-null  float64
 3   Country                        577462 non-null  object
dtypes: float64(2), object(2)
memory usage: 17.6+ MB
None
       AverageTemperature  AverageTemperatureUncertainty
count       544811.000000                  545550.000000
mean            17.193354                       1.019057
std             10.953966                       1.201930
min            -37.658000                       0.052000
25%             10.025000                       0.323000
50%             20.901000                       0.571000
75%             25.814000                       1.206000
max             38.842000                      15.003000
           dt  AverageTemperature  AverageTemperatureUncertainty Country
0  1743-11-01                4.384                          2.294   Åland
1  1743-12-01                  NaN                            NaN   Åland
2  1744-01-01                  NaN                            NaN   Åland
3  1744-02-01                  NaN                            NaN   Åland
4  1744-03-01                  NaN                            NaN   Åland
```

In [9]:
```
print(country_data.isnull().sum())
```

```
dt                                 0
AverageTemperature             32651
AverageTemperatureUncertainty  31912
Country                            0
dtype: int64
```

In [14]:
```
state_data = pd.read_csv("~/dataset/climate-change-earth-surface-temperature-data/GlobalLandTemperaturesByState.csv
print(state_data.info())
print(state_data.describe())
print(state_data.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 645675 entries, 0 to 645674
Data columns (total 5 columns):
 #   Column                         Non-Null Count   Dtype  
---  ------                         --------------   -----  
 0   dt                             645675 non-null  object 
 1   AverageTemperature             620027 non-null  float64
 2   AverageTemperatureUncertainty  620027 non-null  float64
 3   State                          645675 non-null  object 
 4   Country                        645675 non-null  object 
dtypes: float64(2), object(3)
memory usage: 24.6+ MB
None
       AverageTemperature  AverageTemperatureUncertainty
count       620027.000000                  620027.000000
mean             8.993111                       1.287647
std             13.772150                       1.360392
min            -45.389000                       0.036000
25%             -0.693000                       0.316000
50%             11.199000                       0.656000
75%             19.899000                       1.850000
max             36.339000                      12.646000
           dt  AverageTemperature  AverageTemperatureUncertainty State Country
0  1855-05-01              25.544                          1.171  Acre  Brazil
1  1855-06-01              24.228                          1.103  Acre  Brazil
2  1855-07-01              24.371                          1.044  Acre  Brazil
3  1855-08-01              25.427                          1.073  Acre  Brazil
4  1855-09-01              25.675                          1.014  Acre  Brazil
```

In [15]:
```
print(state_data.isnull().sum())
```

```
dt                                 0
AverageTemperature             25648
AverageTemperatureUncertainty  25648
State                              0
Country                            0
dtype: int64
```

In [10]:
```
city_data = pd.read_csv("~/dataset/climate-change-earth-surface-temperature-data/GlobalLandTemperaturesByCity.csv")
print(city_data.info())
```

```
print(city_data.describe())
print(city_data.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8599212 entries, 0 to 8599211
Data columns (total 7 columns):
 #   Column                         Dtype
---  ------                         -----
 0   dt                             object
 1   AverageTemperature             float64
 2   AverageTemperatureUncertainty  float64
 3   City                           object
 4   Country                        object
 5   Latitude                       object
 6   Longitude                      object
dtypes: float64(2), object(5)
memory usage: 459.2+ MB
None
       AverageTemperature  AverageTemperatureUncertainty
count        8.235082e+06                   8.235082e+06
mean         1.672743e+01                   1.028575e+00
std          1.035344e+01                   1.129733e+00
min         -4.270400e+01                   3.400000e-02
25%          1.029900e+01                   3.370000e-01
50%          1.883100e+01                   5.910000e-01
75%          2.521000e+01                   1.349000e+00
max          3.965100e+01                   1.539600e+01
           dt  AverageTemperature  AverageTemperatureUncertainty   City  \
0  1743-11-01                6.068                          1.737  Århus
1  1743-12-01                  NaN                            NaN  Århus
2  1744-01-01                  NaN                            NaN  Århus
3  1744-02-01                  NaN                            NaN  Århus
4  1744-03-01                  NaN                            NaN  Århus

   Country Latitude Longitude
0  Denmark   57.05N    10.33E
1  Denmark   57.05N    10.33E
2  Denmark   57.05N    10.33E
3  Denmark   57.05N    10.33E
4  Denmark   57.05N    10.33E
```

In [11]:
```
print(city_data.isnull().sum())
```

```
dt                                 0
AverageTemperature            364130
AverageTemperatureUncertainty 364130
City                               0
Country                            0
Latitude                           0
Longitude                          0
dtype: int64
```

In [12]:
```
major_city_data = pd.read_csv("~/dataset/climate-change-earth-surface-temperature-data/GlobalLandTemperaturesByMajo
print(major_city_data.info())
print(major_city_data.describe())
print(major_city_data.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 239177 entries, 0 to 239176
Data columns (total 7 columns):
 #   Column                         Non-Null Count   Dtype
---  ------                         --------------   -----
 0   dt                             239177 non-null  object
 1   AverageTemperature             228175 non-null  float64
 2   AverageTemperatureUncertainty  228175 non-null  float64
 3   City                           239177 non-null  object
 4   Country                        239177 non-null  object
 5   Latitude                       239177 non-null  object
 6   Longitude                      239177 non-null  object
dtypes: float64(2), object(5)
memory usage: 12.8+ MB
None
       AverageTemperature  AverageTemperatureUncertainty
count       228175.000000                  228175.000000
mean            18.125969                       0.969343
std             10.024800                       0.979644
min            -26.772000                       0.040000
25%             12.710000                       0.340000
50%             20.428000                       0.592000
75%             25.918000                       1.320000
max             38.283000                      14.037000
           dt  AverageTemperature  AverageTemperatureUncertainty     City  \
0  1849-01-01              26.704                          1.435  Abidjan
1  1849-02-01              27.434                          1.362  Abidjan
2  1849-03-01              28.101                          1.612  Abidjan
3  1849-04-01              26.140                          1.387  Abidjan
4  1849-05-01              25.427                          1.200  Abidjan

         Country Latitude Longitude
0  Côte D'Ivoire    5.63N     3.23W
1  Côte D'Ivoire    5.63N     3.23W
2  Côte D'Ivoire    5.63N     3.23W
3  Côte D'Ivoire    5.63N     3.23W
4  Côte D'Ivoire    5.63N     3.23W
```

In [13]:
```
print(major_city_data.isnull().sum())
```

```
dt                                0
AverageTemperature            11002
AverageTemperatureUncertainty 11002
City                              0
Country                           0
Latitude                          0
Longitude                         0
dtype: int64
```