

Obtención de estadísticas descriptivas

Cantidad de datos que tengo:

Columnas: 23

Filas: 145,460

Variables de cada vector de datos:

Date (DD/MM/YY)

String

Boolean

Int

Float

Analiza las variables para saber qué representan y en qué rangos se encuentran.

Cada una representa distintos parámetros que definen el clima de los días en Australia para determinar con facilidad los días en los que lloverá, a dónde se dirige el aire, cuáles son las temperaturas, los lapsos de fechas en los que ocurre, la velocidad del viento, la presión atmosférica, el estado de las nubes, etc, a varias horas del día. Además, se incluyen datos booleanos de pronósticos que determinan si lloverá hoy o mañana.

Datos: (Min), (Max)

MinTemp: -8.5, 33.9

MaxTemp: -4.8, 48.1

Rainfall: 0, 371

Evaporation: 0, 145

Sunshine: 0, 14.5

WindGustDir: String

WindGustSpeed: 6, 135

WindDir9am: String

WinDir3pm: String
WindSpeed9am: 0, 130
WindSpeed3pm: 0, 87
Humidity9am: 0, 100
Humidity3pm: 0, 100
Pressure9am: 980.5, 1041
Pressure3pm: 977.1, 1039.6
Cloud9am: 0, 9
Cloud3pm: 0, 9
Temp9am: -7.2, 40.2
Temp3pm: -5.4, 46.7
RainToday: Boolean
RainTomorrow: Boolean

Basándote en la media, mediana y desviación estándar de cada variable, ¿qué conclusiones puedes entregar de los datos?

Los datos representan las distintas condiciones climáticas de Australia que dependiendo de la fecha o estación junto a los factores como nubes, temperatura y velocidad del viento determinan si habrá lluvia y que al mismo tiempo se confirma de cierta forma si hubo para poder hacer predicciones y tomar decisiones en un futuro.

Screenshot - Carga de datos con Pandas

```
main.py x
1 import pandas as pd # Se importa la librería pandas
2
3 inf=pd.read_csv("weatherAUS.csv") # Se declara una variable inf que va dirigida a pandas como lectura de archivos
4 print(inf.head()) # Se imprime lo que hay al principio del documento
5 print("-----") #Espacio
6 print(inf.describe()) # Se imprimen los datos estadísticos importantes de varias columnas
7
8
```

```
Shell x
>>> %Run main.py

      Date Location  MinTemp  ...  Temp3pm  RainToday  RainTomorrow
0  2008-12-01  Albury      13.4  ...      21.8         No           No
1  2008-12-02  Albury       7.4  ...      24.3         No           No
2  2008-12-03  Albury      12.9  ...      23.2         No           No
3  2008-12-04  Albury       9.2  ...      26.5         No           No
4  2008-12-05  Albury      17.5  ...      29.7         No           No

[5 rows x 23 columns]
-----
      MinTemp      MaxTemp  ...      Temp9am      Temp3pm
count  143975.000000  144199.000000  ...  143693.000000  141851.000000
mean     12.194034     23.221348  ...     16.990631     21.68339
std       6.398495       7.119049  ...       6.488753       6.93665
min      -8.500000     -4.800000  ...      -7.200000     -5.40000
25%       7.600000     17.900000  ...     12.300000     16.60000
50%      12.000000     22.600000  ...     16.700000     21.10000
75%      16.900000     28.200000  ...     21.600000     26.40000
max      33.900000     48.100000  ...     40.200000     46.70000

[8 rows x 16 columns]

>>>
```