**PYTHON BASICS FOR DATA SCIENCE**

Did you know? IBM Watson Studio lets you build and deploy an AI solution, using the best of open source and IBM software and giving your team a single environment to work in. Learn more here.

[Watson Studio Tutorial](Watson%20Studio%20Tutorial.docx)

Introducción Jupyter LAB (atajos teclado) How to Use JupyterLab

Tipos de variables Python: int, float, bool, str (cadena de caracteres).

Para convertir un numero “int(1.1)” -> 1 (da 1 como resultado, un int). También se pueden convertir str en int, int(“1”) -> 1.

Operadores de python, a parte de los típicos: “//” división de enteros, redondeando a la baja y da como resultado un entero (25//6 = 4). y \*\* es la potencia y % es el resto.

En python a parte de los típicos operadores lógicos, tenemos “and” (en C++ es &&), “or” (||) y “not” (!).

Para saber el tipo de variable type().

Para saber la versión de python:

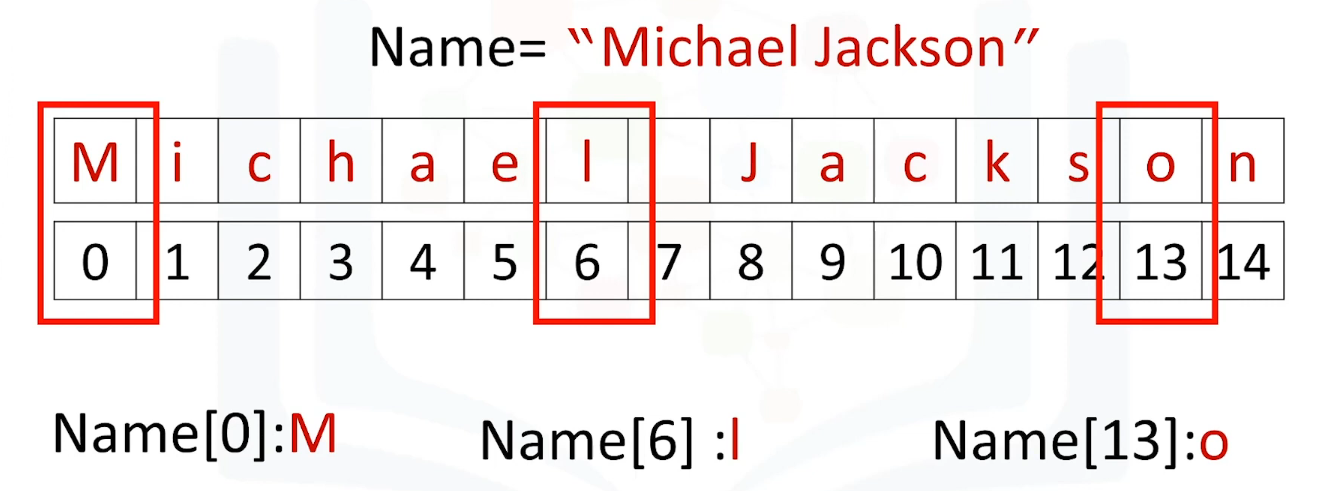
# Check the Python Version

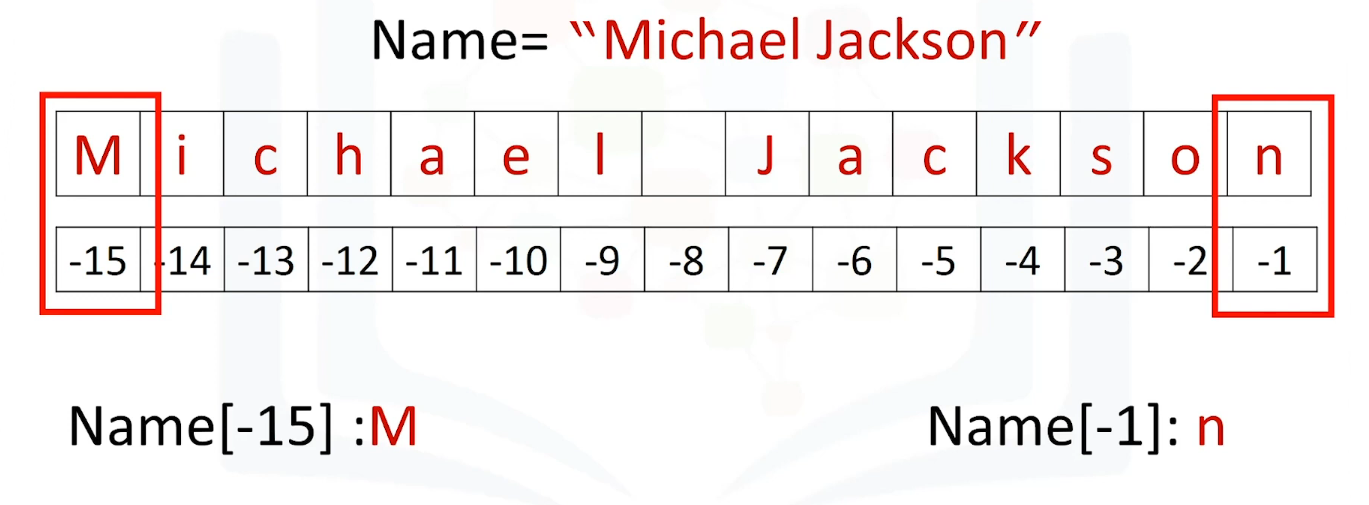
import sys

print(sys.version)

-> 3.9 etc...

Se puede acceder a un string de dos formas, con indices de 0 a n-1, o de -n a -1.





“len” para obtener la longitud del array (n)

Name[posicion inicial:posición final(no incluida):salto]

Para ordenar una lista se puede usar sorted(). Para usar sorted tienes que asignar el output a una variable nueva (no deja sobreescribir), pero si quieres sobreescribir puedes usar sort(), y sobreescribe en el antiguo list por el nuevo y ordenado list.

list() convierte un tuple en lista, y tuple() convierte una lista en tuple

set() es otro tipo de colección/lista, pero los elementos son únicos, y no están ordenados.

Los diccionarios son listas con keys, en vez de indices, y values en vez de items. Los keys pueden ser cadenas de texto o lo que quieras.

En Python se puede hacer un for i,s in enumerate(squares): s i, y te da el indice (i) del enumerate y el valor (s) de cada posición del squares.

Functions

Se puede poner una descripción en las funciones con tres comillas dobles y luego se puede reclamar la ayuda de esa funcion escribiendo help(function) y te devolverá la descripción.

“””

Yo sumo

“””

help(functon)

Yo sumo

Al usar “pass” en python, no hace nada, y si lo pones dentro de una funcion, esta devuelve “None”

Para usar una funcion con independencia de la longitud de la lista en python, se usa el asterisco:

def function(\*variable)

Y así se le pueden pasar listas de diferentes longitudes.

Para poder usar una variable fuera de una función y que se reconozca y se guarde incluso fuera de la funcion, se puede usar dentro de la funcion “global” para definirla como una variable global, de forma que también se reconocerá aunque no se esté usando la función.

En las funciones que están dentro de una clase, al definirlas siempre se pone self como atributo pero cuando llamas a la función eso no es necesario, ya que lo hace solo python, solo le pasas el resto de atributos si es que tiene, sino pues vacia funcion().

Para definir parametros por default en las funciones dentro de objetos se le puede asignar un valor a la variable al definir la funcion tal que asi def funcion(saludo= “hola”) y si no ponemos ningun parametro cogera el default que en este caso es “hola”.

Eventhough the function is defined like car(self,make,model,color) you can input in idfferent order the atributes like this *Car(model="Accord",make="Honda",color="blue")*.

matplotlib es una libreria para dibujar cosas

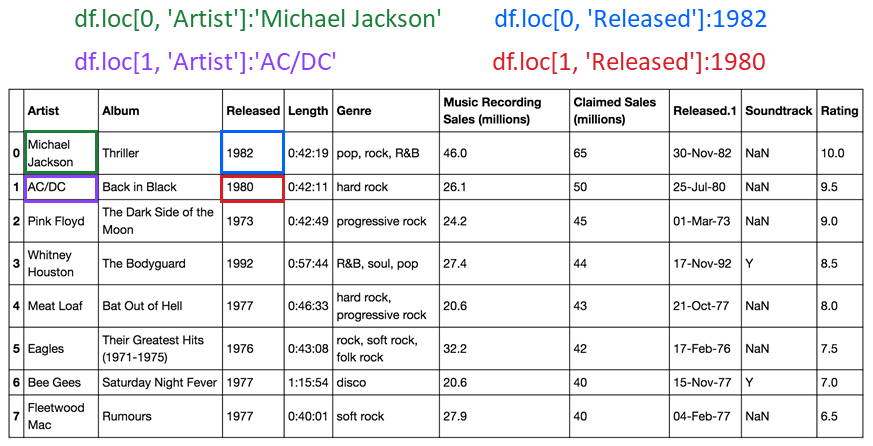
**Using loc, iloc and ix**

There are three ways to select data from a data frame in Pandas: *loc*, *iloc*, and *ix*.

**loc**

*loc*is primarily label based; when two arguments are used, you use column headers and row indexes to select the data you want. *loc*can also take an integer as a row or column number.

Examples of *loc*usage:

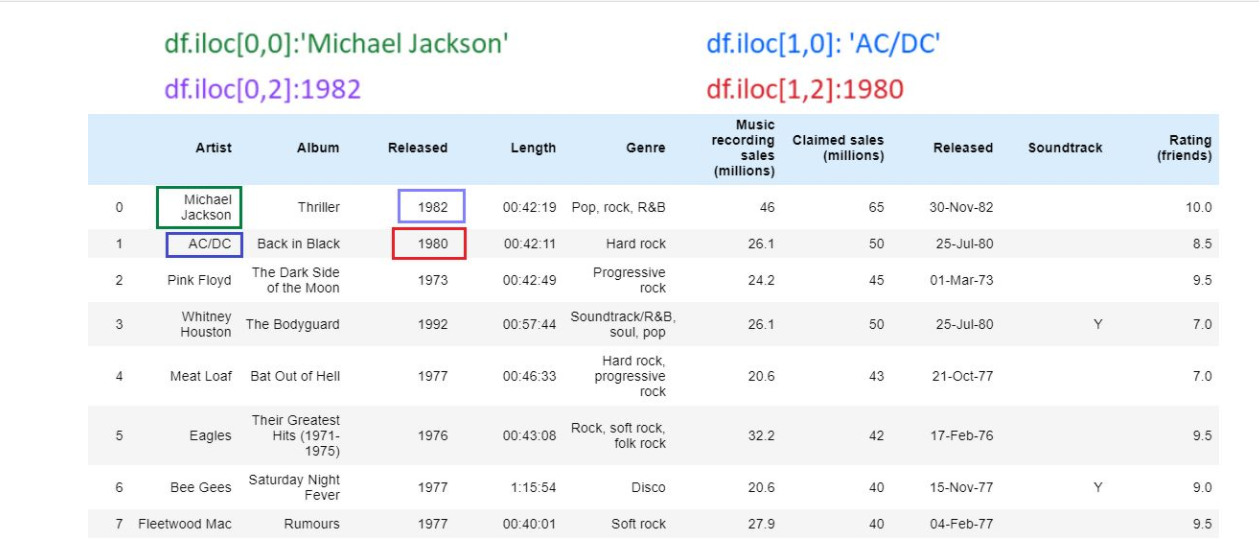


*loc*will return a *KeyError*if the requested items are not found.

**iloc**

*iloc*is integer-based. You use column numbers and row numbers to get rows or columns at particular positions in the data frame.

Examples of *iloc*usage:



*iloc*will return an *IndexError*if the requested indexer is out-of-bounds.

**ix**

By default, *ix*looks for a label. If ix doesn't find a label, it will use an integer. This means you can select data by using either column numbers and row numbers or column headers and row names using *ix*.

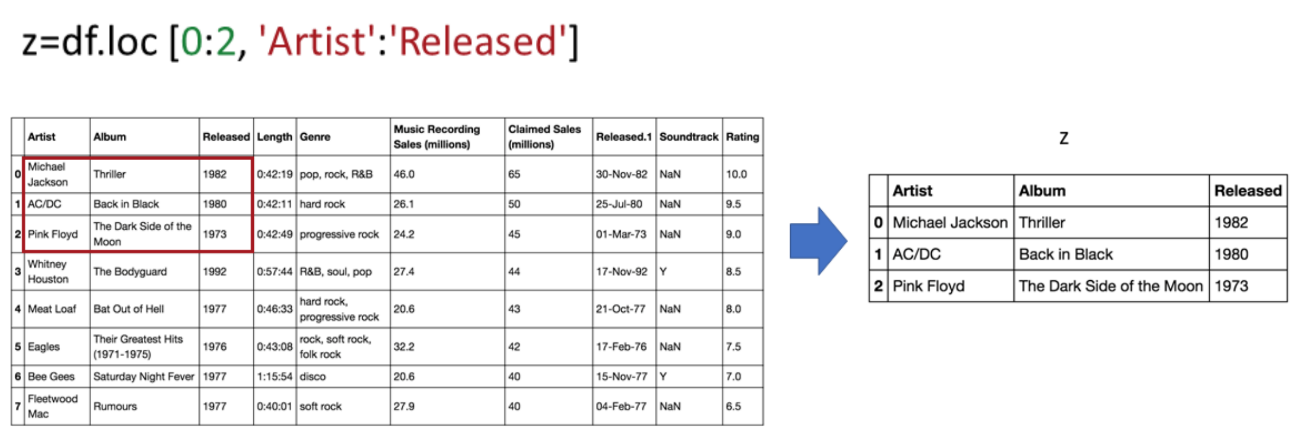
In Pandas version 0.20.0 and later, *ix*is deprecated.

**Using loc and iloc for slicing**

You can also use *loc*and *iloc*to slice data frames and assign the values to a new data frame.

**Creating a new dataframe with loc slicing**

You can also slice data frames and assign the values to a new data frame using the column names. The code assigns the first three rows and all columns in between to the columns named Artist and Released. The result is a new data frame Z with the corresponding values.



**Creating a new dataframe with iloc slicing**

In this example, we assign the first two rows and the first three columns to the variable Z. The result is a data frame comprised of the selected rows and columns.

