Data Engineering

TP1 :Initiation

Diplôme National d'Ingénieur en Informatique

Spécialité :
Génie Logiciel

Réalisée par :
Oussama Ben Slama

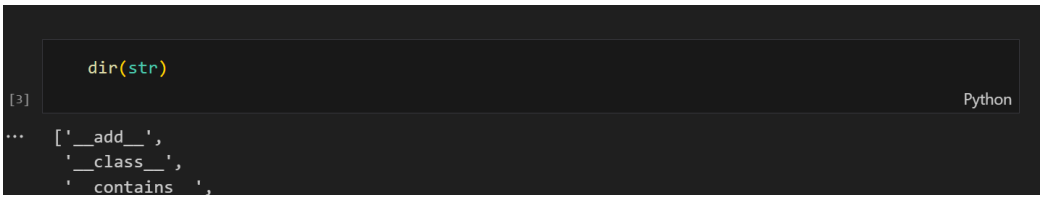
Année Universitaire 2024/2025

Chapitre 1

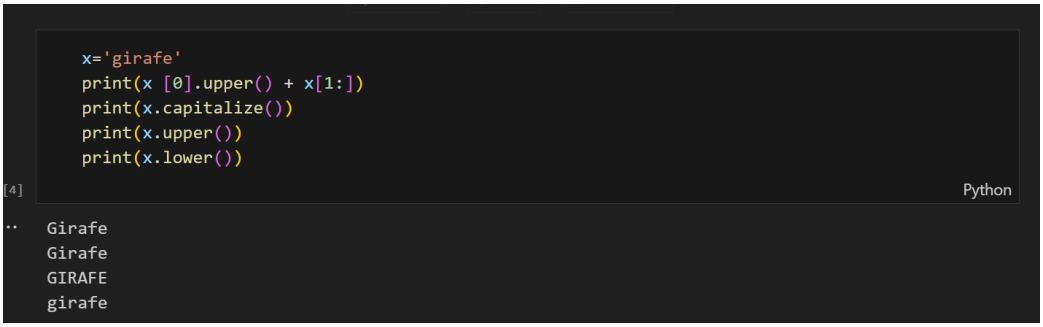
Python Functions

1.1 Basic Functions

dir(str) : It returns a list of all the attributes and methods available for the str (string) class in Python.



upper() : Converts all characters in a string to uppercase
lower() : Converts all characters in a string to lowercase.
capitalize() : Converts the first character to uppercase and the rest to lowercase.



```python
nom = "oussama"
print(nom.upper())
print(nom.lower())
print(nom.capitalize())
```

[5] Python

```
OUSSAMA
oussama
Oussama
```

split() : Splits a string into a list of substrings based on a delimiter.

```python
animaux = "girafe tigre singe souris"
animaux.split()
for animal in animaux.split():
    print(animal)
```

[6] Python

```
girafe
tigre
singe
```

```python
animaux = "girafe : tigre : singe :: souris"
animaux.split (":")
```

[8] Python

```
['girafe ', ' tigre ', ' singe ', '', ' souris']
```

```python
animaux = "girafe tigre singe souris"
animaux.split (maxsplit =2)
```

[9] Python

```
['girafe', 'tigre', 'singe souris']
```

find() : Searches for a substring in a string and returns the index of the first occurrence. If the substring is not found, it returns -1.

```python
animal = "girafe"
print(animal.find("i"))
print(animal.find("afe"))
print(animal.find("z")) # -1 for not found
print(animal.find("tig"))
```

[10] Python

```
1
3
-1
-1
```

replace() : Replaces all occurrences of a substring with another substring and returns a new string.

```
animaux = "tigre girafe tigre"
print(animaux.replace("tigre", "zarafa")) #a new string is created
print(animaux.replace("i","o"))
```

Python

```
zarafa girafe zarafa
togre gorafe togre
```

count() : Returns the number of times a substring appears in a string.

```
animaux = "girafe tigre"
print(animaux.count("i"))
print(animaux.count("z"))
print(animaux.count("tigre"))
```

Python

```
2
0
```

startswith() : Checks if a string starts with a specified prefix and returns True or False.

```
chaine = "Bonjour monsieur oussama !"
print(chaine.startswith("Bonjour"))
print(chaine.startswith("Au revoir"))
```

Python

```
True
False
```

strip() : Removes leading and trailing whitespace (or specified characters) from a string.

```
chaine = "     Comment enlever les espaces au dé but et à la fin ?     "
chaine.strip()
```

Python

```
'Comment enlever les espaces au dé but et à la fin ?'
```

```
chaine = "abboussama abb"
chaine.strip("aab")
```

Python

```
'oussama '
```

1.2 Regular expression

```python
target_string = "PYnative! dot.com; is for, Python-developer?"
result = re.split(r"[\b\W\b]+", target_string)
print(result)
```

[23] Python

```
['PYnative', 'dot', 'com', 'is', 'for', 'Python', 'developer', '']
```

1.3 Lambda Function

A lambda function is an anonymous, one-line function in Python used for short, simple operations.

Lambda function

```python
add3= lambda x:x+3
print(add3(10))
```

[24] Python

```
13
```

```python
my_list = [1, 5, 4, 6, 8, 11, 3, 12]
filtered_list = filter(lambda x: (x%2 == 0) , my_list)
new_list = list(filtered_list)

print(new_list)
print(filtered_list)
```

[25] Python

```
[4, 6, 8, 12]
<filter object at 0x000001B930D7F910>
```

1.4 Application

The goal is to divide a text based on punctuation :

```
[27]  paragraph = "Les fonctions lambda sont définies dans l'ensemble par le mot-clé lambda et elle
                                                                                          Python

[28]  re.split(r'[^\w\s]', paragraph)
                                                                                          Python

...   ['Les fonctions lambda sont définies dans l',
       'ensemble par le mot',
       'clé lambda et elles peuvent comporter n',
       'importe quel nombre d',
       'argument mais une seule expression',
```