Outline essential Python for data

- review data structure
- OOP : Object Oriented Programming
- request API
- read SQLite
- Library : numpy

In [13]:
```
##list, tuple, dictionary, set

friend = ["toy","john","Mary",42,50,75,[1,2,3],{"orange","banana"}]
```

In [15]:
```
len(friend)
```

Out[15]: 8

In [17]:
```
##dictinary ; key values

euro = {
    "name" : "Euro",
    "age" : 25,
    "company" : "KPMG",
    "position" : "Data scientist"
}
```

In [18]:
```
euro["company"]
```

Out[18]: 'KPMG'

In [24]:
```
##loop in dictionary

fruits = ["orange","orange","orange","banana","banana"]

result = {} # empty dict

for fruit in fruits :
    if fruit in result :
        result[fruit] += 1
    else :
        result[fruit] = 1

print(result)
```

```
{'orange': 3, 'banana': 2}
```

In [25]:
```
##OOP object oriented programming

class Moodeng():
    pass
```

In [32]:
```
md = Moodeng()
mdang = Moodeng()
```

In [34]:
```
print(type(md)),
print(type(mdang))
```

```
<class '__main__.Moodeng'>
<class '__main__.Moodeng'>
```

In [35]:
```
class Moodeng():
    def __init__(self, name, age, species):
        self.name = name
        self.age = age
        self.species = species
```

In [39]:
```
md = Moodeng("moodeng",1,"hippo")
euro = Moodeng("euro", 25, "human")
```

In [40]:
```
md
```

Out[40]:
```
<__main__.Moodeng at 0x18e84c94070>
```

In [41]:
```
print(md.age, md.name, md.species)
print(euro.age, euro.name, euro.species)
```

```
1 moodeng hippo
25 euro human
```

In [51]:
```
class Moodeng():
    def __init__(self, name, age, species):
        self.name = name
        self.age = age
        self.species = species

    def hello(self) :
        print("I am Moodeng!")

    def sleep(self) :
        print("I am going to sleep now.")
```

In [52]:
```
md = Moodeng("moodeng",1,"hippo")
```

In [54]:
```
md.hello()
md.sleep()
```

```
I am Moodeng!
I am going to sleep now.
```

In [71]:
```
class User():
    def __init__(self,name,age,gender,city):
        self.name = name
        self.age = age
        self.gender = gender
        self.city = city

    def upload_image(self):
        ##take impage from user
        print("Upload image successfully!")

    def add_age(self):
        self.age += 1

    def substract_age(self):
        self.age -= 1

    ##string representation
    def __str__(self):
        text = f"{self.name} is a {self.gender}, {self.age} years old lives in {self.ci
        return(text)
```

In [72]:
```
user1 =User("euro",25,"male","Bangkok")
```

In [74]:
```
print(user1.city)
user1.upload_image()

user1.add_age()
print(user1.age)

print(user1)
```

```
Bangkok
Upload image successfully!
27
euro is a male, 27 years old lives in Bangkok.
```

Read CSV file

In [78]:
```
!cd
```

```
C:\Users\ADMIN
```

In [81]:
```
!dir
```

```
 Volume in drive C has no label.
 Volume Serial Number is 4ACB-CF27

 Directory of C:\Users\ADMIN

09/16/2024  10:12 AM    <DIR>          .
09/16/2024  10:12 AM    <DIR>          ..
09/16/2024  09:24 AM    <DIR>          .ipynb_checkpoints
07/01/2021  03:13 PM    <DIR>          .ipython
09/05/2024  08:44 PM    <DIR>          .jupyter
```

```
09/13/2024  08:53 PM    <DIR>          .matplotlib
02/06/2021  08:02 PM    <DIR>          3D Objects
02/06/2021  08:02 PM    <DIR>          Contacts
09/05/2024  10:57 PM    <DIR>          Desktop
06/24/2024  05:18 PM    <DIR>          Documents
09/13/2024  08:19 PM    <DIR>          Downloads
09/16/2024  09:24 AM            66,670 DSB_10_practice_01.ipynb
09/05/2024  09:16 PM            12,071 DSB_10_Python_01.ipynb
09/05/2024  10:14 PM            27,462 DSB_10_Python_02.ipynb
09/07/2024  09:28 PM             4,461 DSB_10_Python_HW_game.ipynb
09/07/2024  12:25 PM            38,586 DSB_10_Python_live_01.ipynb
09/16/2024  10:12 AM             8,154 DSB_10_Python_live_02.ipynb
02/06/2021  08:02 PM    <DIR>          Favorites
07/09/2021  03:51 PM                88 file8
04/18/2023  11:58 PM    <DIR>          iCloudDrive
02/06/2021  08:02 PM    <DIR>          Links
11/22/2021  09:35 PM    <DIR>          Music
09/05/2024  10:07 PM               125 mydata.csv
01/12/2023  12:00 AM    <DIR>          OneDrive
05/25/2023  12:03 AM    <DIR>          Pictures
07/09/2021  05:04 PM             7,538 Practice10.ipynb
07/09/2021  05:22 PM             5,321 Practice11.ipynb
07/03/2021  03:02 PM             7,111 Practice4.ipynb
07/03/2021  03:03 PM             4,437 Practice5.ipynb
07/03/2021  04:42 PM             7,275 Practice6.ipynb
07/09/2021  03:32 PM             4,042 Practice7.ipynb
07/09/2021  03:54 PM             1,909 Practice8.ipynb
07/09/2021  04:43 PM             3,458 Practice9.ipynb
02/25/2023  01:16 PM    <DIR>          Saved Games
09/13/2024  08:53 PM    <DIR>          seaborn-data
02/06/2021  08:03 PM    <DIR>          Searches
04/22/2024  02:23 PM             3,000 Test01.ipynb
02/06/2021  08:31 PM    <DIR>          Videos
07/09/2021  05:21 PM                40 worldcapital
              18 File(s)        201,748 bytes
              21 Dir(s)  245,945,823,232 bytes free
```

In [82]:
```
!mkdir newFolder
```

In [83]:
```
!rmdir newFolder
```

In [91]:
```
import csv

data = []

try: ##to check error

    ##with : context manager -> open and close file automatically
    with file = open("mydata.csv","r") as file:
        reader = csv.reader(file)
        for row in reader:
            data.append(row)
    print(data)

except:
    print("Recheck code")
```

```
[['', 'nickname', 'age', 'sex', 'city'], ['0', 'Euro', '22', 'M', 'London'], ['1', 'Fluk', '23', 'M', 'Bangkok'], ['2', 'Anna', '25', 'F', 'London'], ['3', 'Toy', '27', 'M', 'Bangkok'], ['4', 'Mary', '28', 'F', 'London']]
```

In [90]:

```
import csv

data = []

try:
    with file = open("mydata.csv","r") as file:
        reader = csv.reader(file)
        for row in readererr:
            data.append(row)
    print(data)

except:
    print("Recheck code")
```

```
Recheck code
```

In [92]:

```
import pandas as pd
```

In [96]:

```
pd.read_csv("mydata.csv")
```

Out[96]:

	Unnamed: 0	nickname	age	sex	city
0	0	Euro	22	M	London
1	1	Fluk	23	M	Bangkok
2	2	Anna	25	F	London
3	3	Toy	27	M	Bangkok
4	4	Mary	28	F	London

In [97]:

```
!pip install gazpacho
```

```
Collecting gazpacho
  Downloading gazpacho-1.1.tar.gz (7.9 kB)
  Installing build dependencies: started
  Installing build dependencies: finished with status 'done'
  Getting requirements to build wheel: started
  Getting requirements to build wheel: finished with status 'done'
    Preparing wheel metadata: started
    Preparing wheel metadata: finished with status 'done'
Building wheels for collected packages: gazpacho
  Building wheel for gazpacho (PEP 517): started
  Building wheel for gazpacho (PEP 517): finished with status 'done'
  Created wheel for gazpacho: filename=gazpacho-1.1-py3-none-any.whl size=7487 sha256=2fb2222047e8881311026cf2c08bf2360c9088b3874c822515aa112470a14786
  Stored in directory: c:\users\admin\appdata\local\pip\cache\wheels\ec\45\e0\490eb5e25601b4f9425fcde4a0034601c492a29e82268be4d3
Successfully built gazpacho
Installing collected packages: gazpacho
Successfully installed gazpacho-1.1
```

In [106…
```
##how to wirte csv

import csv

header = ["Name","Company","Position"]

data = [
    ["Euro","KPMG","Data scientist"],
    ["Fluke","EY","Accountant"],
    ["Bird","PWC","Lawyer"]
]

with open("example_data.csv", "w") as file :
    writer = csv.writer(file)
    writer.writerow(header)
    writer.writerows(data)
```

In [107…
```
pd.read_csv("example_data.csv")
```

Out[107…

	Name	Company	Position
0	Euro	KPMG	Data scientist
1	Fluke	EY	Accountant
2	Bird	PWC	Lawyer

In [108…
```
df = pd.read_csv("example_data.csv")
```

In [113…
```
df["Country"] = ["Austria", "USA", "Japan"]
df["Age"] = [25,26,26]
```

In [114…
```
df
```

Out[114…

	Name	Company	Position	Country	Age
0	Euro	KPMG	Data scientist	Austria	25
1	Fluke	EY	Accountant	USA	26
2	Bird	PWC	Lawyer	Japan	26

In [115…
```
df["Age"] += 1
```

In [116…
```
df
```

Out[116…

	Name	Company	Position	Country	Age
0	Euro	KPMG	Data scientist	Austria	26

	Name	Company	Position	Country	Age
1	Fluke	EY	Accountant	USA	27
2	Bird	PWC	Lawyer	Japan	27

In [117…
```
df.to_csv("updated_data.csv")
```

API

In [118…
```
## import requests
from requests import get
```

In [121…
```
response = get("https://swapi.dev/api/people/1")
```

In [122…
```
response.status_code
```

Out[122… 200

In [123…
```
response.json()
```

Out[123…
```
{'name': 'Luke Skywalker',
 'height': '172',
 'mass': '77',
 'hair_color': 'blond',
 'skin_color': 'fair',
 'eye_color': 'blue',
 'birth_year': '19BBY',
 'gender': 'male',
 'homeworld': 'https://swapi.dev/api/planets/1/',
 'films': ['https://swapi.dev/api/films/1/',
  'https://swapi.dev/api/films/2/',
  'https://swapi.dev/api/films/3/',
  'https://swapi.dev/api/films/6/'],
 'species': [],
 'vehicles': ['https://swapi.dev/api/vehicles/14/',
  'https://swapi.dev/api/vehicles/30/'],
 'starships': ['https://swapi.dev/api/starships/12/',
  'https://swapi.dev/api/starships/22/'],
 'created': '2014-12-09T13:50:51.644000Z',
 'edited': '2014-12-20T21:17:56.891000Z',
 'url': 'https://swapi.dev/api/people/1/'}
```

In [129…
```
##get data from id 1-5
from requests import get
from time import sleep

based_url = "https://swapi.dev/api/people/"

for i in range(1,6):
    api_url = based_url + str(i)
    response = get(api_url)
```

```
    print(response.json()["name"])
    sleep(5) ##similar as break
```

```
Luke Skywalker
C-3PO
R2-D2
Darth Vader
Leia Organa
```

In [132…

```
characters = []

for i in range(1,6):
    api_url = based_url + str(i)
    response = get(api_url)
    response_js = response.json()
    name = response_js["name"]
    height = response_js["height"]
    mass = response_js["mass"]
    result = [name, height, mass]
    characters.append(result)
    sleep(2)
```

In [133…

```
print(characters)
```

```
[['Luke Skywalker', '172', '77'], ['C-3PO', '167', '75'], ['R2-D2', '96', '32'], ['Darth Vader', '202', '136'], ['Leia Organa', '150', '49']]
```

In [134…

```
header = ["name", "height", "mass"]
with open("starwars.csv", "w") as file:
    writer = csv.writer(file)
    writer.writerow(header)
    writer.writerows(characters)
```

In [137…

```
pd.read_csv("starwars.csv")
```

Out[137…

	name	height	mass
0	Luke Skywalker	172	77
1	C-3PO	167	75
2	R2-D2	96	32
3	Darth Vader	202	136
4	Leia Organa	150	49

Gazpacho

Basic web scraping

In [139…

```
#!pip install gazpacho
from gazpacho import Soup
from requests import get
```

In [140…
```
url = "https://datarockie.com"
```

In [141…
```
web = get(url)

datarockie = Soup(web.text)

print(type(datarockie))
```

```
<class 'gazpacho.soup.Soup'>
```

In [144…
```
##find information we want in this soup

datarockie.find("h2",mode="first").strip()
```

Out[144…
```
'Data Analysis Made Simple'
```

In [147…
```
for h2 in datarockie.find("h2"):
    print(h2.strip())
```

```
Data Analysis Made Simple
Recent Posts
Career Guide
```

In [148…
```
for h4 in datarockie.find("h4"):
    print(h4.strip())
```

```
เปิดโลกวิชา Digital Marketing 101 สำหรับผู้เริ่มต้น
สูตรหารค่าข้าวด้วย Google Sheets ง่ายๆ ทำตามได้ทันที
ดาวน์โหลดฟรี หนังสือ DuckDB in Action ของดีที่ควรมีติดเชลฟ์
Suddenly Talented วิธีพัฒนาตัวเองแบบเก่งปุ๋บปับ
Data Analyst Career Guide แนะนำงานสาย Data ฉบับสมบูรณ์
เริ่มต้นใช้งาน Bard AI ตัวใหม่จาก Google ออกมาสู้กับ ChatGPT เต็มๆ
```

SQLite

In [150…
```
import sqlite3
import pandas as pd
```

In [151…
```
##create connection
con = sqlite3.connect("chinook.db")
```

In [153…
```
df = pd.read_sql("select * from customers limit 5", con)
```

In [154…
```
df
```

Out[154…

	CustomerId	FirstName	LastName	Company	Address	City	State	Country	PostalCode
0	1	Luís	Gonçalves	Embraer - Empresa Brasileira de Aeronáutica S.A.	Av. Brigadeiro Faria Lima, 2170	São José dos Campos	SP	Brazil	12227-000
1	2	Leonie	Köhler	None	Theodor-Heuss-Straße 34	Stuttgart	None	Germany	70174
2	3	François	Tremblay	None	1498 rue Bélanger	Montréal	QC	Canada	H2G 1A7
3	4	Bjørn	Hansen	None	Ullevålsveien 14	Oslo	None	Norway	0171
4	5	František	Wichterlová	JetBrains s.r.o.	Klanova 9/506	Prague	None	Czech Republic	14700

In [155…

```
df[["FirstName","City"]]
```

Out[155…

	FirstName	City
0	Luís	São José dos Campos
1	Leonie	Stuttgart
2	François	Montréal
3	Bjørn	Oslo
4	František	Prague

In [156…

```
con.close()
```

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: