Practical Assignment

September 17, 2022

C.K.Pithawala College of Engineering and Technology, Surat Group Member's name and enrollment no. : Kunal Santoshkumar Maurya (200090107084) Melvin Divyeshkumar Jariwala (200090107131) Div. : B Subject : Python for Data Science Subject Code : 3150713 Data Set : Netflix Movies and TV Shows Data Set Link : https://www.kaggle.com/datasets/shivamb/netflix-shows Academic Year : 2022-2023 Subject Faculty : Dr. Ami Tusharkant Choksi

1 CO-1 Assignment:

1.1 Practical-1 :

Write a program to implement the following using operators in Python. start no=110, result=550. Store operators(+,-,*,/,1/x,%,// etc.) in a list. Take a random operator from the list, apply the operation on the start no and result. Store answer in result variable. Take input for a random operator till it is valid input. [Use import random print(random.randint(3, 9))]

```python
[1]: import random

startno,result=4404,8807
op=['+','-','*','/','%','//']#list of operators

while (True):
    r=random.randint(0,9)
    if(r>=len(op)):
        print("Invalid Operator.")
        break
    else:
        print("Start No. :",  startno,"\nOperator :",op[r],"\nResult(Initial) :
 ", result)
        rand=str(startno)+op[r]+str(result)
        result=eval(rand)
        print("Result =",result,"\n")
```

```
Start No. : 4404
Operator : /
Result(Initial) : 8807
Result = 0.5000567730214602

Start No. : 4404
```

```
Operator : -
Result(Initial) : 0.5000567730214602
Result = 4403.499943226979

Invalid Operator.
```

1.2 Practical-2 :

Write a program to create the following pattern. Take input for n. for n=5. [ord()-for character to ascii, chr()-for ascii to char] A AB ABC ABCDEF ABCDEFGHIJKL

```
[2]: n=int(input("Enter value of n : "))
     l=[1,2]

     for i in range(n):
         l.append(sum(l))
         num=65
         for j in range(l[i]):
             print(chr(num),end="")
             num+=1
         print("\r")
```

```
Enter value of n : 5
A
AB
ABC
ABCDEF
ABCDEFGHIJKL
```

1.3 Practical-3:

Use list comprehension to create the following output for your dataset items. E.g.my dataset is of car car=["Swift", "BMW", "Skoda5", "i10", "Lamborghini"] Create a new list of indices where the car name has a number in the name. Output for the example: [2, 3]

```
[3]: import re

     ##list of shows
     shows=["Kota Factory", "Jaws 2", "Squid Game", "Jaws 3", "Schumacher", "6␣
      ↪Bullets"]

     print("List of shows : " + str(shows))

     #list comprehension + enumerate()
     indices=[i for i,x in enumerate(shows) if re.search("[0-9]+",x)]

     print("Indices where shows have a number in the name : " + str(indices))
```

```
List of shows : ['Kota Factory', 'Jaws 2', 'Squid Game', 'Jaws 3', 'Schumacher',
'6 Bullets']
Indices where shows have a number in the name : [1, 3, 5]
```

1.4 Practical-4:

Write a program to make a module. Implement (a)method overloading(b)method overriding(c)subclass(d)multilevel inheritance(e)multiple inheritance. Use the classes, properties and methods according to your data set. Import it in other program

(a) Method Overloading:

```python
[4]: class netflix:
    def favourite_shows(self,*args):
        print("Favourite Shows : ")
        for i in args:
            print(i)
        print()

n=netflix()
n.favourite_shows("Kota Factory")
n.favourite_shows("Grown Ups","Schumacher")
n.favourite_shows("Clickbait","Ant-Man and the Wasp","Marvel's Iron Fist")
```

```
Favourite Shows :
Kota Factory

Favourite Shows :
Grown Ups
Schumacher

Favourite Shows :
Clickbait
Ant-Man and the Wasp
Marvel's Iron Fist
```

(b) Method Overriding:

```python
[5]: class netflix:
    def watching(self):
        print("I am watching a Netflix Show.")

class shows:
    def watching(self):
        print("I am watching Kota Factory on Netflix.")

a=netflix()
b=shows()
```

```
a.watching()
b.watching()
```

```
I am watching a Netflix Show.
I am watching Kota Factory on Netflix.
```

(c) Subclass:

```
[6]: class netflix_shows():
         def __init__(self,show_id,show_name):
             self.show_id=show_id
             self.show_name=show_name
         def display1(self):
             print("Netflix Show :\nShow ID : "+str(self.show_id)+"\nShow Name :␣
      ↪"+str(self.show_name)+"\n")
     class tv_shows(netflix_shows):
         def display2(self):
             print("TV Show :\nShow ID : "+str(self.show_id)+"\nShow Name :␣
      ↪"+str(self.show_name)+"\n")

     show=tv_shows("s5","Kota Factory")

     show.display1()
     show.display2()
```

```
Netflix Show :
Show ID : s5
Show Name : Kota Factory

TV Show :
Show ID : s5
Show Name : Kota Factory
```

(d) Multilevel Inheritance:

```
[7]: class netflix_titles:
         def __init__(self,show_id,show_name):
             self.show_id=show_id
             self.show_name=show_name
         def title_info(self):
             print("Netflix Show :\nShow ID : "+str(self.show_id)+"\nShow Name :␣
      ↪"+str(self.show_name)+"\n")

     class movies(netflix_titles):
         def movie_info(self):
             print("Movie :\nMovie ID : "+str(self.show_id)+"\nMovie Name :␣
      ↪"+str(self.show_name)+"\n")
```

```python
class tv_shows(movies):
    def tv_show_info(self):
        print("TV Show :\nShow ID : "+str(self.show_id)+"\nShow Name :␣
↪"+str(self.show_name)+"\n")

show=tv_shows("s5","Kota Factory")
show.title_info()
show.movie_info()
show.tv_show_info()
```

```
Netflix Show :
Show ID : s5
Show Name : Kota Factory

Movie :
Movie ID : s5
Movie Name : Kota Factory

TV Show :
Show ID : s5
Show Name : Kota Factory
```

(e) Multiple Inheritance:

```python
[8]: class tv_shows:
    def __init__(self,show_id,show_name):
        self.show_id=show_id
        self.show_name=show_name
    def tv_show_info(self):
        print("TV Show :\nShow ID : "+str(self.show_id)+"\nShow Name :␣
↪"+str(self.show_name)+"\n")

class movies:
    def __init__(self,show_id,show_name):
        self.show_id=show_id
        self.show_name=show_name
    def movie_info(self):
        print("Movie :\nMovie ID : "+str(self.show_id)+"\nMovie Name :␣
↪"+str(self.show_name)+"\n")

class shows(tv_shows,movies):
    pass

s1=shows("s5","Kota Factory")
s1.tv_show_info()
```

```
s1.movie_info()

s2=shows("s28","Grown Ups")
s2.tv_show_info()
s2.movie_info()
```

```
TV Show :
Show ID : s5
Show Name : Kota Factory

Movie :
Movie ID : s5
Movie Name : Kota Factory

TV Show :
Show ID : s28
Show Name : Grown Ups

Movie :
Movie ID : s28
Movie Name : Grown Ups
```

1.5 Practical-5:

Write a program to insert, delete, update, retrieve, indexing, slicing, concatenation, join etc. for string(s) from data of your dataset.

```python
[9]: #function to insert string at desired position/index
def insert():
    global str1
    print("String before insertion : " + str1)

    sub=input("Enter the string for insertion : ")#string to be inserted
    pos=int(input("Enter position for insertion: "))#position/index for string␣
  ↪to be inserted

    temp=str1[:pos]+sub+str1[pos:]#insertion of string at the position
    str1=temp

    print("String after insertion : " + str1)

#function to delete string
def delete():
    global str1
    print("String before delete : " + str1)

    str1=""#string is now empty
```

```python
    print("String after delete : " + str1)

#function to update the string
def update():
    global str1
    print("String before updating : " + str1)
    str1=input("Enter the updated string : ")#input of new to string to update⎵
 ↪the existing string
    print("String after updating : " + str1)

def retrieve():
    global str1
    print("String retrieved : " + str1)

#function to get index of desired string/element
def indexing():
    global str1
    ch=input("Enter the string to get its index : ")
    print("Index of "+ch+" is : "+str(str1.index(ch)))

#function for string sliccing
def slicing():
    global str1
    start=int(input("Enter start point : "))
    end=int(input("Enter end point : "))
    step=int(input("Enter step point : "))
    print(str1[start:end:step])#string slicing

#function to concatenate a string with existing string
def concatenation():
    global str1
    print("String before concatenation : " + str1)

    temp=input("Enter a string for concatenation : ")#input of new string to be⎵
 ↪concatenated
    str1+=temp
    print("String after concatenation : " + str1)

#function to perfrom join operation on string
def join_str():
    global str1
    print("String before join : " + str1)
    str2=input("Enter join character for string : ")
    str1=str2.join(str1)
    print("String after join : " + str1)
```

```python
#switcher to select operation
def switch(argument):
    switcher={
        "1":insert,
        "2":delete,
        "3":update,
        "4":retrieve,
        "5":indexing,
        "6":slicing,
        "7":concatenation,
        "8":join_str,
    }
    r=switcher.get(argument)()

#initial string
str1="Kota Factory"
print("String : " + str1)
print()

print("""Select the operation from the following:
1. Insert
2. Delete
3. Update
4. Retrieve
5. Indexing
6. Slicing
7. Concatenation
8. Join
""")

arg=input("Enter numbers to select operation : ")#input to select operation
print()

switch(arg)
```

String : Kota Factory

Select the operation from the following:
1. Insert
2. Delete
3. Update
4. Retrieve
5. Indexing
6. Slicing
7. Concatenation
8. Join

```
Enter numbers to select operation : 7

String before concatenation : Kota Factory
Enter a string for concatenation :   is a good show
String after concatenation : Kota Factory is a good show
```

1.6   Practical-6:

Write a program to insert, delete, update, retrieve, indexing, slicing, concatenation, join etc. for list(s) from data of your dataset.

```python
[10]: #funtion to insert desired elements at desired position/index in the list
      def insert():
          global list1
          print("List before insertion : " + str(list1))

          sub=input("Enter the string for insertion : ")#input of desired string for␣
       ↪insertion

          pos =int(input("Enter the position for insertion : "))#insertion of desired␣
       ↪position/index for insertion
          list1.insert(pos,sub)

          print("List after insertion : " + str(list1))

      #funciton to delete desired element from the list
      def delete():
          global list1
          print("List before deletion : " + str(list1))
          list1.remove(input("Enter the element to delete : "))#input of element to␣
       ↪be deleted and done so with remove function
          print("List after deletion : " + str(list1))

      #function to update the existing element of the list
      def update():
          global list1
          print("List before updating : " + str(list1))

          idx=list1.index(input("Enter the element to be replaced : "))#index of␣
       ↪element to be updated

          list1[idx]=input("Enter the new element : ")#input of new element
          print("List after updating : " + str(list1))

      def retrieve():
          global list1
          print("List retrieved : " + str(list1))
```

```python
#function to retrieve the index of desired element
def indexing():
    global list1
    print(list1.index(input("Enter the string to get its index : ")))

#function slicing of list
def slicing():
    global list1
    start=int(input("Enter the start position : "))
    end=int(input("Enter the end position : "))
    step=int(input("Enter step : "))
    print(list1[start:end:step])#list slicing

#function for concatenation of desired elements to the list
def concatenation():
    global list1
    print("List before concatenation : " + str(list1))
    list1.append(input("Enter element to be concatenated : "))
    print("List after concatenation : " + str(list1))

#funciton to perform join operation on list
def join_list():
    global list1
    print("List before concatenation : " + str(list1))
    list1=input("Enter the element for join : ").join(list1)
    print("List after join : "+str(list1))

#switcher to select operation
def switch(argument):
    switcher={
        "1":insert,
        "2":delete,
        "3":update,
        "4":retrieve,
        "5":indexing,
        "6":slicing,
        "7":concatenation,
        "8":join_list,
    }
    r=switcher.get(argument)()

#initial list
list1=["6","s5","TV Show","Kota Factory"]
print("List : " + str(list1) + "\n")
```

```python
print("""Select the operation from the following:
1. Insert
2. Delete
3. Update
4. Retrieve
5. Indexing
6. Slicing
7. Concatenation
8. Join
""")

arg=input("Enter numbers to select operation : ")#input to select operation
print()

switch(arg)
```

```
List : ['6', 's5', 'TV Show', 'Kota Factory']

Select the operation from the following:
1. Insert
2. Delete
3. Update
4. Retrieve
5. Indexing
6. Slicing
7. Concatenation
8. Join

Enter numbers to select operation : 8

List before concatenation : ['6', 's5', 'TV Show', 'Kota Factory']
Enter the element for join : :
List after join : 6:s5:TV Show:Kota Factory
```

1.7 Practical-7:

Write a program to insert, delete, update, retrieve, indexing, slicing, concatenation etc. for tuple(s) from data of your dataset.

```python
#function to insert desired element at desired position/index
def tup_insert():
    global shows_tup
    print("Tuple before insertion : " + str(shows_tup))

    shows_tup=list(shows_tup)#convert tuple to list as tuple is immutable

    shows_tup.insert(int(input("Enter position for insertion")),input("Enter⎵
    ↪the string for insertion : "))
```

11

```python
    shows_tup=tuple(shows_tup)#convert list back to tuple
    print("Tuple before insertion : " + str(shows_tup))


#function to delete desired elements at desired position/index
def tup_delete():
    global shows_tup

    print("Tuple before deletion : " + str(shows_tup))

    #convert tuple to list as tuple is immutable
    shows_tup=list(shows_tup)

    #delete element with remove function
    shows_tup.remove(input("Enter the element to be deleted : "))

    #convert list back to tuple
    shows_tup=tuple(shows_tup)

    print("Tuple before deletion : " + str(shows_tup))

#function to update desired elements in the tuple
def tup_update():
    global shows_tup

    print("Tuple before updating : " + str(shows_tup))

    #converting tuple to list
    shows_tup=list(shows_tup)

    #index of element to be updated
    idx=shows_tup.index(input("Enter the element to be replaced : "))

    #input of the new element
    shows_tup[idx]=input("Enter the new element : ")

    #convert list back to tuple
    shows_tup=tuple(shows_tup)

    print("Tuple after updating : " + str(shows_tup))

def tup_retrieve():
    global shows_tup
    print("Tuple retrieved : " + str(shows_tup))

#funciton to retrieve index of desired element
```

```python
def tup_indexing():
    global shows_tup
    print(shows_tup.index(input("Enter the element to get its index : ")))

#function for slicing of tuple
def tup_slicing():
    global shows_tup
    start=int(input("Enter the start position : "))
    end=int(input("Enter the end position : "))
    step=int(input("Enter step : "))
    print(shows_tup[start:end:step])#tuple slicing

#function for concatenating elements to the tuple
def tup_concatenation():
    global shows_tup
    tup1=shows_tup
    print("Tuple before concatenation : " + str(shows_tup))

    #convert tuple to list as tuple is immutable
    tup1=list(tup1)

    #concatenation with the help of append function
    tup1.append(input("Enter the element to be concatenated : "))

    #convert list back to tuple
    shows_tup=tuple(tup1)
    print("Tuple after concatenation : " + str(shows_tup))

#switcher to select operation to perform on tuple
def switch(argument):
    switcher={
        "1":tup_insert,
        "2":tup_delete,
        "3":tup_update,
        "4":tup_retrieve,
        "5":tup_indexing,
        "6":tup_slicing,
        "7":tup_concatenation,
    }
    r=switcher.get(argument)()


#initial tuple
shows_tup=("6","s5","TV Show","Kota Factory")
print("Tuple : " + str(shows_tup) + "\n")
```

```
print("""Select the operation from the following:
1. Insert
2. Delete
3. Update
4. Retrieve
5. Indexing
6. Slicing
7. Concatenation
""")
arg=input("Enter numbers to select operation : ")#input to select operation
print()

switch(arg)
```

```
Tuple : ('6', 's5', 'TV Show', 'Kota Factory')

Select the operation from the following:
1. Insert
2. Delete
3. Update
4. Retrieve
5. Indexing
6. Slicing
7. Concatenation

Enter numbers to select operation : 7

Tuple before concatenation : ('6', 's5', 'TV Show', 'Kota Factory')
Enter the element to be concatenated : Grown Ups
Tuple after concatenation : ('6', 's5', 'TV Show', 'Kota Factory', 'Grown Ups')
```

1.8   Practical-8:

Write a program to do set operations from data of your dataset: (i)intersection (ii)union
(iii)difference (iv)symmetric difference (v)check s1 is a subset of s2(vi)check if s1 is a superset of
s2(vii)find whether two sets are disjoint or not(viii)find all subsets of a set without using itertools

```python
[12]: #function to perform set intersection
      def set_intersection():
          global s1,s2

          #intersection of sets s1 and s2 with the help of intersection function
          print("Intersection of s1 and s2 : "+str(s1.intersection(s2)))

      #function for union of two sets
      def set_union():
          global s1,s2
```

```python
    #union of sets s1 and s2 with the help of union funciton
    print("Union of s1 and s2 : " + str(s1.union(s2)))

#function for set difference
def set_difference():
    global s1,s2

    #Set difference of sets s1 and s2 (i.e. s1-s2) with the help of difference
↪function
    print("Set Difference : " + str(s1.difference(s2)))

#funciton for set symmetric difference
def set_symmetric_difference():
    global s1,s2

    #Set symmetric difference of sets s1 and s2 with the help of
↪symmetric_difference function
    print("Symmetric Difference : " + str(s1.symmetric_difference(s2)))

#function to check if a set is subset of another
def set_subset():
    global s1,s2

    #check if set s1 is subset of s2
    #returns True if s1 is subset of s2 else False
    print("s1 subset of s2 : " + str(s1.issubset(s2)))

#function to check if a set is superset of another
def set_superset():
    global s1,s2

    #check if set s1 is superset of s2
    #returns True if s1 is superset of s2 else False
    print("s1 superset of s2 : "+str(s1.issuperset(s2)))

def set_disjoint():
    global s1,s2

    #check if sets s1 and s2 are disjoint sets
    print("s1 and s2 are disjoint : "+str(s1.isdisjoint(s2)))

#funciton to find all the subsets of a set
def set_subsets():
    global s1
    output=[[]]
    for i in s1:
        output+=[lst+[i] for lst in output]
```

```python
        print("All subsets of s1 :\n"+str(output))

#switcher to select operation
def switch(argument):
    switcher={
        "1":set_intersection,
        "2":set_union,
        "3":set_difference,
        "4":set_symmetric_difference,
        "5":set_subset,
        "6":set_superset,
        "7":set_disjoint,
        "8":set_subsets,
    }
    r=switcher.get(argument)()

#initial set s1
s1={"6","s5","TV Show","Kota Factory"}

#initial set s2
s2={"29","s28"," Movie","Grown Ups"}

print("s1 : " + str(s1) + "\n")
print("s2 : " + str(s2) + "\n")

print("""Select the operation from the following:
1. Intersection
2. Union
3. Difference
4. Symmetric difference
5. Check s1 is a subset of s2
6. Check if s1 is a superset of s2
7. Find whether two sets are disjoint or not
8. Find all subsets of a set without using itertools
""")
arg=input("Enter numbers to select operation : ")#input to select operation
print()

switch(arg)
```

s1 : {'6', 's5', 'TV Show', 'Kota Factory'}

s2 : {'s28', 'Grown Ups', '29', ' Movie'}

Select the operation from the following:
1. Intersection
2. Union

```
3. Difference
4. Symmetric difference
5. Check s1 is a subset of s2
6. Check if s1 is a superset of s2
7. Find whether two sets are disjoint or not
8. Find all subsets of a set without using itertools

Enter numbers to select operation : 5

s1 subset of s2 : False
```

1.9   Practical-9:

Write a program to do operations on a dictionary from data of your dataset.

```python
[13]: #function to insert desired key-value pair at desired position/index
      def dict_insert():
          global dict1
          print("Dictionary before insertion : " + str(dict1))
          x=list(dict1.items())
          x.insert(int(input("Enter position for insertion : ")), (input("Enter Key :␣
       ↪"),input("Enter value : ")))
          dict1=dict(x)
          print("Dictionary after insertion : " + str(dict1))

      #function to delete key-value pair in a dictionary
      def dict_delete():
          global dict1
          print("Dictionary before deletion : " + str(dict1))

          key_to_delete=(input("Enter key to be deleted : "))#input of key to be␣
       ↪deleted
          i=0

          #traverse through the dictionary to find the desired key
          #Once found, it is deleted
          for key in dict1.keys():
              if key_to_delete in dict1:
                  del dict1[key_to_delete]
                  i=1
                  break
          if i==0:
              print("Key not found!")
          print("Dictionary after deletion : " + str(dict1))

      #funciton to update dictionary
      def dict_update():
          global dict1
```

17

```python
    print("Tuple before updating : " + str(dict1))

    key_to_replace=(input("Enter the key to be replaced : "))#input of key for
 ↪updating the dictionary
    value_to_replace=input("Enter the value : ")#input of value for updating
 ↪the dictionary

    #update dictionary with the help of 'update' function
    dict1.update({key_to_replace:value_to_replace})

    print("Tuple after updating : " + str(dict1))

def dict_retrieve():
    global dict1
    print("Dictionary retrieved : " + str(dict1))

#funciton to get index of key
def dict_indexing():
    global dict1
    k=input("Enter the key to get its value : ")#input of key

    #using list + keys() + index() we get the index of key
    res=list(dict1.keys()).index(k)
    print("Index of key : " + str(res))

#function for dictionary slicing
def dict_slicing():
    global dict1
    start=int(input("Enter the start position : "))
    end=int(input("Enter the end position : "))
    res=dict()

    #loop to traverse through dictionary keys
    for i in dict1:
        res[i]=dict1[i][start:end]
    print(res)

#function to concatenate key-value pair to the existing dictionary
def dict_concatenation():
    global dict1
    print("Tuple before concatenation : " + str(dict1))

    #input of key to be concatenated
    key_to_concatenate=(input("Enter the key to be concatenated : "))

    #input of value to be concatenated
    value_to_concatenate=input("Enter the value to be concatenated : ")
```

```python
    #concatenation using update()
    dict1.update({key_to_concatenate:value_to_concatenate})
    print("Tuple after concatenation : " + str(dict1))

#switcher to select operation
def switch(argument):
    switcher={
        "1":dict_insert,
        "2":dict_delete,
        "3":dict_update,
        "4":dict_retrieve,
        "5":dict_indexing,
        "6":dict_slicing,
        "7":dict_concatenation,
    }
    r=switcher.get(argument)()

#initial dictionary
dict1={"6":"s5","TV Show":"Kota Factory"}
print("Dictionary : " + str(dict1) + "\n")


print("""Select the operation from the following:
1. Insert
2. Delete
3. Update
4. Retrieve
5. Indexing
6. Slicing
7. Concatenation
""")
arg=input("Enter numbers to select operation : ")#input to select operation
print()

switch(arg)
```

Dictionary : {'6': 's5', 'TV Show': 'Kota Factory'}

Select the operation from the following:
1. Insert
2. Delete
3. Update
4. Retrieve
5. Indexing
6. Slicing
7. Concatenation

```
Enter numbers to select operation : 7

Tuple before concatenation : {'6': 's5', 'TV Show': 'Kota Factory'}
Enter the key to be concatenated : Movie
Enter the value to be concatenated : Grown Ups
Tuple after concatenation : {'6': 's5', 'TV Show': 'Kota Factory', 'Movie':
'Grown Ups'}
```

2 CO-2 Assignment:

2.1 Practical-10:

Read and analyze your data set for the following: a. Number of rows b. Number of attributes
c. Number of missing values for each attribute

```python
import pandas as pd

#read dataset panda and read_csv()
file=pd.read_csv('netflix_titles.csv')

#Number of rows using len()
print("Number of rows :" + str(len(file)) + "\n")

#Number of attributes using len()
print("Number of attributes :" + str(len(file.dtypes)) + "\n")

#No. of missing values for each attribute using isnull() and sum()
print("Number of missing values for each attribute :\n" + str(file.isnull().
 ↪sum()) + "\n")

#Total no. of missing values in the entire dataset using isnull() and sum()
print("Total number of missing values :" + str(file.isnull().sum().sum()) +␣
 ↪"\n")
```

```
Number of rows :8807

Number of attributes :12

Number of missing values for each attribute :
show_id            0
type               0
title              0
director        2634
cast             825
country          831
date_added        10
release_year       0
```

20

```
rating           4
duration         3
listed_in        0
description      0
dtype: int64


Total number of missing values :4307
```

2.2 Practical-11:

Write a program to parse HTML documents w.r.to your dataset using Beautiful Soup.

```
[16]: import requests
      import lxml
      from bs4 import BeautifulSoup

      #url according to dataset
      URL="https://www.netflix.com/"

      #response from url using requests()
      r=requests.get(URL)

      #Parsing
      soup=BeautifulSoup(r.content,'lxml')
      print(f'{soup.h1.name} : {soup.h1.text}')#h1
      print(f'{soup.h2.name} : {soup.h2.text}')#h2
      print(f'{soup.h3.name} : {soup.h3.text}')#h3
      print(f'{soup.li.name} : {soup.li.text}')#li
```

```
h1 : Unlimited movies, TV shows and more.
h2 : Watch anywhere. Cancel anytime.
h3 : Ready to watch? Enter your email to create or restart your membership.
li : Email address
```

3 CO-3 Assignment:

3.1 Practical-12:

Display graphics and multimedia video related to your data set in Jupyter notebook.

```
[45]: from IPython.display import YouTubeVideo
      from IPython.display import Image

      #URL of image of Kota Factory season 2 poster
      image_url="https://m.media-amazon.com/images/M/
       ↪MV5BOGVmMGYwZTEtOGFlYS00ODRhLTg3M2MtMzc1OGNhNzU3N2IzXkEyXkFqcGdeQXVyNDY4NjAxNTc@.
       ↪_V1_.jpg"
```

```python
#id of video of official trailer of Kota Factory season 2
video_id=YouTubeVideo("aG8KWYko8VY")

print("Kota Factory Season 2 poster : ")

display(Image(image_url,width=300,height=300))#display image of poster

print("Official trailer of Kota Factory Season 2 : ")

display(video_id)#display video of trailer
```

Kota Factory Season 2 poster :

A NETFLIX SERIES

KOTA FACTORY

SEASON 2

A TVF CREATION

NOW STREAMING | NETFLIX

Official trailer of Kota Factory Season 2 :

3.2 Practical-13:

Read your data set and do the following: (a) Validating Your Data, Figuring out what's in your data, (b) Removing duplicates, Creating a data map and data plan, Manipulating (c) Categorical Variables, Creating categorical variables, Renaming levels, (d) Combining levels, Dealing with Dates in Your Data, Formatting date and time values, Using the right time transformation, (e) Dealing with Missing Data, Finding the missing data, (f) Encoding missingness, Imputing missing data, Slicing and Dicing: i. Filtering and Selecting Data, Slicing rows, Slicing columns, Dicing, ii. Concatenating and Transforming, Adding new cases and variables, Removing data iii.Sorting and shuffling, Aggregating Data at Any Level.

3.2.1 (a) Validating Your Data, Figuring out what's in your data:

```
[35]: import pandas as pd

df=pd.read_csv('netflix_titles.csv')
```

```
[33]: def validate_df(df):
          if(df.empty):
              print("File is empty")
          else:
              print("File isn't empty.\nBelow is DataFrame:")
              return df
```

```
[34]: validate_df(df)
```

File isn't empty.
Below is DataFrame:

```
[34]:      show_id     type                    title           director  \
      0          s1    Movie    Dick Johnson Is Dead   Kirsten Johnson
      1          s2  TV Show          Blood & Water               NaN
      2          s3  TV Show              Ganglands   Julien Leclercq
      3          s4  TV Show    Jailbirds New Orleans             NaN
      4          s5  TV Show           Kota Factory              NaN
      ...       ...      ...                    ...               ...
      8802    s8803    Movie                 Zodiac     David Fincher
      8803    s8804  TV Show            Zombie Dumb              NaN
      8804    s8805    Movie             Zombieland   Ruben Fleischer
      8805    s8806    Movie                   Zoom      Peter Hewitt
      8806    s8807    Movie                 Zubaan       Mozez Singh

                                                     cast         country  \
      0                                               NaN   United States
      1     Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…    South Africa
      2     Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…             NaN
      3                                               NaN             NaN
      4     Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…           India
      ...                                               …               …
      8802  Mark Ruffalo, Jake Gyllenhaal, Robert Downey J…  United States
      8803                                              NaN             NaN
      8804  Jesse Eisenberg, Woody Harrelson, Emma Stone, …  United States
      8805  Tim Allen, Courteney Cox, Chevy Chase, Kate Ma…  United States
      8806  Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan…           India

                    date_added  release_year rating   duration  \
      0     September 25, 2021          2020  PG-13     90 min
      1     September 24, 2021          2021  TV-MA   2 Seasons
      2     September 24, 2021          2021  TV-MA    1 Season
      3     September 24, 2021          2021  TV-MA    1 Season
      4     September 24, 2021          2021  TV-MA   2 Seasons
      ...                  …            …     …          …
      8802   November 20, 2019          2007      R    158 min
      8803        July 1, 2019          2018  TV-Y7   2 Seasons
```

```
8804         November 1, 2019         2009      R       88 min
8805         January 11, 2020         2006     PG       88 min
8806           March 2, 2019          2015  TV-14      111 min

                                              listed_in  \
0                                          Documentaries
1             International TV Shows, TV Dramas, TV Mysteries
2        Crime TV Shows, International TV Shows, TV Act…
3                                   Docuseries, Reality TV
4        International TV Shows, Romantic TV Shows, TV …
…                                                       …
8802                      Cult Movies, Dramas, Thrillers
8803             Kids' TV, Korean TV Shows, TV Comedies
8804                             Comedies, Horror Movies
8805                  Children & Family Movies, Comedies
8806     Dramas, International Movies, Music & Musicals

                                            description
0        As her father nears the end of his life, filmm…
1        After crossing paths at a party, a Cape Town t…
2        To protect his family from a powerful drug lor…
3        Feuds, flirtations and toilet talk go down amo…
4        In a city of coaching centers known to train I…
…                                                       …
8802    A political cartoonist, a crime reporter and a…
8803    While living alone in a spooky town, a young g…
8804    Looking to survive in a world taken over by zo…
8805    Dragged from civilian life, a former superhero…
8806    A scrappy but poor boy worms his way into a ty…

[8807 rows x 12 columns]
```

Figuring out what's in data

```
[37]: df.head()
```

```
[37]:   show_id     type                      title          director  \
      0     s1    Movie   Dick Johnson Is Dead  Kirsten Johnson
      1     s2  TV Show          Blood & Water              NaN
      2     s3  TV Show              Ganglands  Julien Leclercq
      3     s4  TV Show  Jailbirds New Orleans              NaN
      4     s5  TV Show           Kota Factory              NaN


                                                cast         country  \
      0                                          NaN   United States
      1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
      2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…             NaN
```

```
3                                                    NaN                NaN
4  Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…             India


         date_added  release_year rating   duration  \
0  September 25, 2021          2020  PG-13     90 min
1  September 24, 2021          2021  TV-MA   2 Seasons
2  September 24, 2021          2021  TV-MA    1 Season
3  September 24, 2021          2021  TV-MA    1 Season
4  September 24, 2021          2021  TV-MA   2 Seasons


                                          listed_in  \
0                                      Documentaries
1      International TV Shows, TV Dramas, TV Mysteries
2    Crime TV Shows, International TV Shows, TV Act…
3                              Docuseries, Reality TV
4    International TV Shows, Romantic TV Shows, TV …


                                        description
0  As her father nears the end of his life, filmm…
1  After crossing paths at a party, a Cape Town t…
2  To protect his family from a powerful drug lor…
3  Feuds, flirtations and toilet talk go down amo…
4  In a city of coaching centers known to train I…
```

[39]: `df.index`

[39]: RangeIndex(start=0, stop=8807, step=1)

[3]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 12 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   show_id       8807 non-null   object
 1   type          8807 non-null   object
 2   title         8807 non-null   object
 3   director      6173 non-null   object
 4   cast          7982 non-null   object
 5   country       7976 non-null   object
 6   date_added    8797 non-null   object
 7   release_year  8807 non-null   int64
 8   rating        8803 non-null   object
 9   duration      8804 non-null   object
 10  listed_in     8807 non-null   object
 11  description   8807 non-null   object
```

```
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```

[4]: `df.describe()`

[4]:
```
       release_year
count  8807.000000
mean   2014.180198
std       8.819312
min    1925.000000
25%    2013.000000
50%    2017.000000
75%    2019.000000
max    2021.000000
```

[5]: `df.tail()`

[5]:
```
     show_id     type         title          director  \
8802  s8803    Movie        Zodiac     David Fincher
8803  s8804  TV Show   Zombie Dumb               NaN
8804  s8805    Movie    Zombieland   Ruben Fleischer
8805  s8806    Movie          Zoom      Peter Hewitt
8806  s8807    Movie        Zubaan       Mozez Singh

                                            cast        country  \
8802  Mark Ruffalo, Jake Gyllenhaal, Robert Downey J…  United States
8803                                             NaN            NaN
8804  Jesse Eisenberg, Woody Harrelson, Emma Stone, …  United States
8805  Tim Allen, Courteney Cox, Chevy Chase, Kate Ma…  United States
8806  Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan…          India

           date_added  release_year rating   duration  \
8802  November 20, 2019         2007      R    158 min
8803       July 1, 2019         2018  TV-Y7  2 Seasons
8804   November 1, 2019         2009      R     88 min
8805   January 11, 2020         2006     PG     88 min
8806      March 2, 2019         2015  TV-14    111 min

                                       listed_in  \
8802               Cult Movies, Dramas, Thrillers
8803          Kids' TV, Korean TV Shows, TV Comedies
8804                       Comedies, Horror Movies
8805             Children & Family Movies, Comedies
8806  Dramas, International Movies, Music & Musicals

                                       description
8802  A political cartoonist, a crime reporter and a…
```

```
8803  While living alone in a spooky town, a young g…
8804  Looking to survive in a world taken over by zo…
8805  Dragged from civilian life, a former superhero…
8806  A scrappy but poor boy worms his way into a ty…
```

[6]: `df.dtypes`

[6]:
```
show_id         object
type            object
title           object
director        object
cast            object
country         object
date_added      object
release_year     int64
rating          object
duration        object
listed_in       object
description     object
dtype: object
```

[7]: `df.shape`

[7]: `(8807, 12)`

[8]: `df.size`

[8]: `105684`

[9]: `df['release_year'].value_counts()`

[9]:
```
2018    1147
2017    1032
2019    1030
2020     953
2016     902
        …
1959       1
1925       1
1961       1
1947       1
1966       1
Name: release_year, Length: 74, dtype: int64
```

3.2.2 (b) Removing duplicates, Creating a data map and data plan, Manipulating:

```
[10]: print("Data frame before removing duplicates : ")
      df
```

Data frame before removing duplicates :

```
[10]:       show_id     type                      title           director  \
      0          s1    Movie   Dick Johnson Is Dead    Kirsten Johnson
      1          s2  TV Show           Blood & Water                NaN
      2          s3  TV Show               Ganglands    Julien Leclercq
      3          s4  TV Show    Jailbirds New Orleans              NaN
      4          s5  TV Show            Kota Factory                NaN
      ...       ...      ...                     ...                ...
      8802    s8803    Movie                  Zodiac      David Fincher
      8803    s8804  TV Show             Zombie Dumb                NaN
      8804    s8805    Movie              Zombieland    Ruben Fleischer
      8805    s8806    Movie                    Zoom        Peter Hewitt
      8806    s8807    Movie                  Zubaan        Mozez Singh

                                                    cast         country  \
      0                                              NaN   United States
      1     Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
      2     Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…            NaN
      3                                              NaN             NaN
      4     Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…           India
      ...                                             …              …
      8802  Mark Ruffalo, Jake Gyllenhaal, Robert Downey J…   United States
      8803                                           NaN             NaN
      8804  Jesse Eisenberg, Woody Harrelson, Emma Stone, …   United States
      8805  Tim Allen, Courteney Cox, Chevy Chase, Kate Ma…   United States
      8806  Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan…           India

                  date_added  release_year rating   duration  \
      0      September 25, 2021          2020  PG-13     90 min
      1      September 24, 2021          2021  TV-MA   2 Seasons
      2      September 24, 2021          2021  TV-MA    1 Season
      3      September 24, 2021          2021  TV-MA    1 Season
      4      September 24, 2021          2021  TV-MA   2 Seasons
      ...                   …           …       …         …
      8802   November 20, 2019          2007      R    158 min
      8803         July 1, 2019          2018  TV-Y7  2 Seasons
      8804    November 1, 2019          2009      R     88 min
      8805    January 11, 2020          2006     PG     88 min
      8806       March 2, 2019          2015  TV-14    111 min

                                           listed_in  \
      0                                 Documentaries
```

```
1        International TV Shows, TV Dramas, TV Mysteries
2        Crime TV Shows, International TV Shows, TV Act…
3                              Docuseries, Reality TV
4        International TV Shows, Romantic TV Shows, TV …
…                                                    …
8802                      Cult Movies, Dramas, Thrillers
8803             Kids' TV, Korean TV Shows, TV Comedies
8804                            Comedies, Horror Movies
8805                 Children & Family Movies, Comedies
8806      Dramas, International Movies, Music & Musicals

                                           description
0     As her father nears the end of his life, filmm…
1     After crossing paths at a party, a Cape Town t…
2     To protect his family from a powerful drug lor…
3     Feuds, flirtations and toilet talk go down amo…
4     In a city of coaching centers known to train I…
…                                                    …
8802  A political cartoonist, a crime reporter and a…
8803  While living alone in a spooky town, a young g…
8804  Looking to survive in a world taken over by zo…
8805  Dragged from civilian life, a former superhero…
8806  A scrappy but poor boy worms his way into a ty…

[8807 rows x 12 columns]
```

[12]: `df.duplicated()`

```
[12]: 0        False
      1        False
      2        False
      3        False
      4        False
               …
      8802     False
      8803     False
      8804     False
      8805     False
      8806     False
      Length: 8807, dtype: bool
```

[13]:
```
df.drop_duplicates()
print("Data frame after removing duplicates : ")
df
```

Data frame after removing duplicates :

```
[13]:        show_id     type                   title            director  \
        0          s1    Movie   Dick Johnson Is Dead    Kirsten Johnson
        1          s2  TV Show          Blood & Water                 NaN
        2          s3  TV Show              Ganglands    Julien Leclercq
        3          s4  TV Show   Jailbirds New Orleans                NaN
        4          s5  TV Show            Kota Factory                 NaN
        …          …       …                      …                   …
        8802     s8803    Movie                 Zodiac      David Fincher
        8803     s8804  TV Show            Zombie Dumb                 NaN
        8804     s8805    Movie             Zombieland    Ruben Fleischer
        8805     s8806    Movie                   Zoom        Peter Hewitt
        8806     s8807    Movie                 Zubaan        Mozez Singh


                                                  cast          country  \
        0                                          NaN    United States
        1      Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…     South Africa
        2      Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…              NaN
        3                                          NaN              NaN
        4      Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…            India
        …                                            …                …
        8802   Mark Ruffalo, Jake Gyllenhaal, Robert Downey J…  United States
        8803                                         NaN              NaN
        8804   Jesse Eisenberg, Woody Harrelson, Emma Stone, …  United States
        8805   Tim Allen, Courteney Cox, Chevy Chase, Kate Ma…  United States
        8806   Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan…          India


                      date_added   release_year  rating   duration  \
        0     September 25, 2021           2020   PG-13     90 min
        1     September 24, 2021           2021   TV-MA   2 Seasons
        2     September 24, 2021           2021   TV-MA    1 Season
        3     September 24, 2021           2021   TV-MA    1 Season
        4     September 24, 2021           2021   TV-MA   2 Seasons
        …                      …              …       …          …
        8802   November 20, 2019           2007       R    158 min
        8803        July 1, 2019           2018   TV-Y7   2 Seasons
        8804   November 1, 2019            2009       R     88 min
        8805   January 11, 2020            2006      PG     88 min
        8806       March 2, 2019           2015   TV-14    111 min


                                              listed_in  \
        0                                 Documentaries
        1       International TV Shows, TV Dramas, TV Mysteries
        2     Crime TV Shows, International TV Shows, TV Act…
        3                         Docuseries, Reality TV
        4     International TV Shows, Romantic TV Shows, TV …
        …                                             …
        8802              Cult Movies, Dramas, Thrillers
```

```
8803                    Kids' TV, Korean TV Shows, TV Comedies
8804                              Comedies, Horror Movies
8805                    Children & Family Movies, Comedies
8806       Dramas, International Movies, Music & Musicals

                                                  description
0       As her father nears the end of his life, filmm…
1       After crossing paths at a party, a Cape Town t…
2       To protect his family from a powerful drug lor…
3       Feuds, flirtations and toilet talk go down amo…
4       In a city of coaching centers known to train I…
…                                                         …
8802    A political cartoonist, a crime reporter and a…
8803    While living alone in a spooky town, a young g…
8804    Looking to survive in a world taken over by zo…
8805    Dragged from civilian life, a former superhero…
8806    A scrappy but poor boy worms his way into a ty…

[8807 rows x 12 columns]
```

3.2.3 (c)Categorical Variables, Creating categorical variables, Renaming levels:

```
[25]: df_cat=pd.Series(pd.Categorical(df['show_id'],dtype='category'))
      df_cat
```

```
[25]: 0            s1
      1            s2
      2            s3
      3            s4
      4            s5
                  …
      8802      s8803
      8803      s8804
      8804      s8805
      8805      s8806
      8806      s8807
      Length: 8807, dtype: category
      Categories (8807, object): ['s1', 's10', 's100', 's1000', …, 's996', 's997',
      's998', 's999']
```

```
[26]: pd.get_dummies(df_cat)
```

```
[26]:      s1  s10  s100  s1000  s1001  s1002  s1003  s1004  s1005  s1006  …  \
      0      1    0     0      0      0      0      0      0      0      0  …
      1      0    0     0      0      0      0      0      0      0      0  …
      2      0    0     0      0      0      0      0      0      0      0  …
      3      0    0     0      0      0      0      0      0      0      0  …
```

4	0	0	0	0	0	0	0	0	0	0	…
…	..	…	…	…	…	…	…	…	…	…	
8802	0	0	0	0	0	0	0	0	0	0	…
8803	0	0	0	0	0	0	0	0	0	0	…
8804	0	0	0	0	0	0	0	0	0	0	…
8805	0	0	0	0	0	0	0	0	0	0	…
8806	0	0	0	0	0	0	0	0	0	0	…

	s990	s991	s992	s993	s994	s995	s996	s997	s998	s999
0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0
…	…	…	…	…	…	…	…	…	…	…
8802	0	0	0	0	0	0	0	0	0	0
8803	0	0	0	0	0	0	0	0	0	0
8804	0	0	0	0	0	0	0	0	0	0
8805	0	0	0	0	0	0	0	0	0	0
8806	0	0	0	0	0	0	0	0	0	0

[8807 rows x 8807 columns]

```python
[30]: df_data=pd.Series(pd.Categorical(df['show_id'],categories=df_cat))
      df_data
```

```
[30]: 0          s1
      1          s2
      2          s3
      3          s4
      4          s5
               ...
      8802    s8803
      8803    s8804
      8804    s8805
      8805    s8806
      8806    s8807
      Length: 8807, dtype: category
      Categories (8807, object): ['s1', 's10', 's100', 's1000', …, 's996', 's997',
      's998', 's999']
```

```python
[28]: df_cat.cat.categories=df['title']
      df_data.cat.categories=df_cat
      df_data
```

```
[28]: 0       Dick Johnson Is Dead
      1               Blood & Water
```

```
2                   Ganglands
3        Jailbirds New Orleans
4                 Kota Factory
                     …
8802                   Zodiac
8803              Zombie Dumb
8804               Zombieland
8805                     Zoom
8806                   Zubaan
Length: 8807, dtype: category
Categories (8807, object): ['Dick Johnson Is Dead', 'Blood & Water',
'Ganglands', 'Jailbirds New Orleans', …, 'Zombie Dumb', 'Zombieland', 'Zoom',
'Zubaan']
```

3.2.4 (d) Dealing with Dates in Your Data, Formatting date and time values, Using the right time transformation:

```python
from datetime import datetime

df_date=df['date_added'].head()
print("Dates before formatting : \n")
df_date
```

```
Dates before formatting :
```

```
[29]: 0    September 25, 2021
      1    September 24, 2021
      2    September 24, 2021
      3    September 24, 2021
      4    September 24, 2021
      Name: date_added, dtype: object
```

```python
print("\nDates after formatting : \n")
for i in range(len(df_date)):
        d=datetime.strptime(df_date[i], "%B %d, %Y")
        print(d.strftime("%Y-%m-%d"))
```

```
Dates after formatting :

2021-09-25
2021-09-24
2021-09-24
2021-09-24
2021-09-24
```

3.2.5 (e) Dealing with Missing Data, Finding the missing data:

```
[15]: print("'True' values the below matrix shows missing data : ")
      pd.isna(df)
```

'True' values the below matrix shows missing data :

[15]:

	show_id	type	title	director	cast	country	date_added \
0	False	False	False	False	True	False	False
1	False	False	False	True	False	False	False
2	False	False	False	False	False	True	False
3	False	False	False	True	True	True	False
4	False	False	False	True	False	False	False
...	...	...	...	...	...	...	
8802	False	False	False	False	False	False	False
8803	False	False	False	True	True	True	False
8804	False	False	False	False	False	False	False
8805	False	False	False	False	False	False	False
8806	False	False	False	False	False	False	False

	release_year	rating	duration	listed_in	description
0	False	False	False	False	False
1	False	False	False	False	False
2	False	False	False	False	False
3	False	False	False	False	False
4	False	False	False	False	False
...	...	...	...	...	...
8802	False	False	False	False	False
8803	False	False	False	False	False
8804	False	False	False	False	False
8805	False	False	False	False	False
8806	False	False	False	False	False

[8807 rows x 12 columns]

```
[32]: print("\nTotal number of missing values for each attribute' : ")
      pd.isna(df).sum()
```

Total number of missing values for each attribute' :

[32]:
```
show_id         0
type            0
title           0
director     2634
cast          825
country       831
date_added     10
```

36

```
release_year      0
rating            4
duration          3
listed_in         0
description       0
dtype: int64
```

[34]:
```python
print("\nTotal number of missing values in entire data framme : "+str(pd.
↪isna(df).sum().sum()))
```

```
Total number of missing values in entire data framme : 4307
```

3.2.6   (f) Encoding missingness, Imputing missing data, Slicing and Dicing:

(i) Filtering and Selecting Data, Slicing rows, Slicing columns, Dicing:

[35]:
```python
#Selecting Data
# Selecting columns

show_detail=df[["show_id","title"]]
show_detail.head()
```

[35]:
```
  show_id                 title
0      s1  Dick Johnson Is Dead
1      s2          Blood & Water
2      s3             Ganglands
3      s4  Jailbirds New Orleans
4      s5          Kota Factory
```

[36]:
```python
#Selecting rows:
show_detail=show_detail[0:10]
show_detail
```

[36]:
```
  show_id                           title
0      s1            Dick Johnson Is Dead
1      s2                    Blood & Water
2      s3                       Ganglands
3      s4            Jailbirds New Orleans
4      s5                    Kota Factory
5      s6                   Midnight Mass
6      s7  My Little Pony: A New Generation
7      s8                         Sankofa
8      s9      The Great British Baking Show
9     s10                     The Starling
```

[37]:
```python
Movie_df=df[(df["type"]=="Movie")&(df["release_year"]>=2005)]
Movie_df.head()
```

```
[37]:      show_id    type                                    title  \
      0         s1   Movie                    Dick Johnson Is Dead
      6         s7   Movie   My Little Pony: A New Generation
      9        s10   Movie                         The Starling
      12       s13   Movie                          Je Suis Karl
      13       s14   Movie   Confessions of an Invisible Girl

                             director  \
      0               Kirsten Johnson
      6    Robert Cullen, José Luis Ucha
      9                Theodore Melfi
      12          Christian Schwochow
      13                 Bruno Garotti

                                                    cast  \
      0                                              NaN
      6    Vanessa Hudgens, Kimiko Glenn, James Marsden, …
      9    Melissa McCarthy, Chris O'Dowd, Kevin Kline, T…
      12   Luna Wedler, Jannis Niewöhner, Milan Peschel, …
      13   Klara Castanho, Lucca Picon, Júlia Gomes, Marc…

                           country         date_added  release_year rating duration  \
      0             United States  September 25, 2021          2020  PG-13   90 min
      6                       NaN  September 24, 2021          2021     PG   91 min
      9             United States  September 24, 2021          2021  PG-13  104 min
      12  Germany, Czech Republic  September 23, 2021          2021  TV-MA  127 min
      13                      NaN  September 22, 2021          2021  TV-PG   91 min

                             listed_in  \
      0                   Documentaries
      6           Children & Family Movies
      9                 Comedies, Dramas
      12       Dramas, International Movies
      13  Children & Family Movies, Comedies

                                          description
      0    As her father nears the end of his life, filmm…
      6    Equestria's divided. But a bright-eyed hero be…
      9    A woman adjusting to life after a loss contend…
      12   After most of her family is murdered in a terr…
      13   When the clever but socially-awkward Tetê join…
```

Slicing dataFrame using loc and iloc:

Syntax: loc[row labels, columns labels]

```
[38]:  #Slice rows by label.
       df.loc[1:3, :]
```

```
[38]:   show_id     type                title          director  \
     1     s2   TV Show         Blood & Water               NaN
     2     s3   TV Show             Ganglands   Julien Leclercq
     3     s4   TV Show   Jailbirds New Orleans             NaN

                                              cast        country  \
     1   Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
     2   Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…          NaN
     3                                              NaN          NaN

              date_added   release_year  rating   duration  \
     1   September 24, 2021          2021   TV-MA   2 Seasons
     2   September 24, 2021          2021   TV-MA    1 Season
     3   September 24, 2021          2021   TV-MA    1 Season

                                         listed_in  \
     1      International TV Shows, TV Dramas, TV Mysteries
     2   Crime TV Shows, International TV Shows, TV Act…
     3                          Docuseries, Reality TV

                                          description
     1   After crossing paths at a party, a Cape Town t…
     2   To protect his family from a powerful drug lor…
     3   Feuds, flirtations and toilet talk go down amo…

[39]:  #Slice columns by label.
       df.loc[:, "show_id":"country"]

[39]:        show_id     type                title          director  \
     0          s1    Movie   Dick Johnson Is Dead   Kirsten Johnson
     1          s2   TV Show         Blood & Water             NaN
     2          s3   TV Show             Ganglands   Julien Leclercq
     3          s4   TV Show   Jailbirds New Orleans             NaN
     4          s5   TV Show          Kota Factory             NaN
     …          …       …                  …              …
     8802     s8803    Movie               Zodiac    David Fincher
     8803     s8804   TV Show         Zombie Dumb             NaN
     8804     s8805    Movie           Zombieland   Ruben Fleischer
     8805     s8806    Movie                 Zoom     Peter Hewitt
     8806     s8807    Movie               Zubaan      Mozez Singh

                                              cast        country
     0                                              NaN   United States
     1       Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
     2       Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…          NaN
     3                                              NaN          NaN
     4       Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…        India
```

```
...                                          ...              ...
8802  Mark Ruffalo, Jake Gyllenhaal, Robert Downey J…   United States
8803                                               NaN             NaN
8804  Jesse Eisenberg, Woody Harrelson, Emma Stone, …   United States
8805  Tim Allen, Courteney Cox, Chevy Chase, Kate Ma…   United States
8806  Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan…           India

[8807 rows x 6 columns]
```

[40]: #To slice rows by index position.
df.iloc[0:2,:]

```
[40]:   show_id     type                title          director   \
        0      s1    Movie  Dick Johnson Is Dead  Kirsten Johnson
        1      s2  TV Show          Blood & Water              NaN

                                                      cast         country  \
        0                                              NaN   United States
        1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…    South Africa

                  date_added  release_year rating   duration  \
        0  September 25, 2021          2020  PG-13     90 min
        1  September 24, 2021          2021  TV-MA  2 Seasons

                                             listed_in  \
        0                                 Documentaries
        1  International TV Shows, TV Dramas, TV Mysteries

                                           description
        0  As her father nears the end of his life, filmm…
        1  After crossing paths at a party, a Cape Town t…
```

[41]: #To slice columns by index position.
df.iloc[:,1:3]

```
[41]:         type                title
        0      Movie  Dick Johnson Is Dead
        1    TV Show          Blood & Water
        2    TV Show             Ganglands
        3    TV Show  Jailbirds New Orleans
        4    TV Show          Kota Factory
        ...      ...                   ...
        8802   Movie                Zodiac
        8803 TV Show           Zombie Dumb
        8804   Movie            Zombieland
        8805   Movie                  Zoom
        8806   Movie                Zubaan
```

```
[8807 rows x 2 columns]
```

Dicing dataFrame using loc and iloc:

```
[43]: df.loc[0:5, "show_id":"country"]
```

```
[43]:   show_id     type                    title           director  \
        0      s1    Movie    Dick Johnson Is Dead  Kirsten Johnson
        1      s2  TV Show          Blood & Water              NaN
        2      s3  TV Show              Ganglands  Julien Leclercq
        3      s4  TV Show  Jailbirds New Orleans              NaN
        4      s5  TV Show           Kota Factory              NaN
        5      s6  TV Show          Midnight Mass    Mike Flanagan

                                                      cast        country
        0                                              NaN   United States
        1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
        2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…            NaN
        3                                              NaN            NaN
        4  Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…          India
        5  Kate Siegel, Zach Gilford, Hamish Linklater, H…            NaN
```

```
[44]: df.iloc[0:6,1:3]
```

```
[44]:      type                    title
        0    Movie    Dick Johnson Is Dead
        1  TV Show          Blood & Water
        2  TV Show              Ganglands
        3  TV Show  Jailbirds New Orleans
        4  TV Show           Kota Factory
        5  TV Show          Midnight Mass
```

(ii) Concatenating and Transforming, Adding new cases and variables, Removing data:

```
[9]: df1=df.iloc[0:3,0:3]
     df1
```

```
[9]:   show_id     type                 title
       0      s1    Movie  Dick Johnson Is Dead
       1      s2  TV Show         Blood & Water
       2      s3  TV Show             Ganglands
```

```
[10]: df2=df.iloc[3:6,0:3]
      df2
```

```
[10]:    show_id      type                          title
    3        s4   TV Show   Jailbirds New Orleans
    4        s5   TV Show            Kota Factory
    5        s6   TV Show           Midnight Mass
```

```
[11]: df3=df.iloc[6:9,0:3]
      df3
```

```
[11]:    show_id      type                                title
    6        s7     Movie   My Little Pony: A New Generation
    7        s8     Movie                            Sankofa
    8        s9   TV Show       The Great British Baking Show
```

```
[12]: concat_df=pd.concat([df1,df2,df3])
      concat_df
```

```
[12]:    show_id      type                                title
    0        s1     Movie             Dick Johnson Is Dead
    1        s2   TV Show                    Blood & Water
    2        s3   TV Show                        Ganglands
    3        s4   TV Show            Jailbirds New Orleans
    4        s5   TV Show                     Kota Factory
    5        s6   TV Show                    Midnight Mass
    6        s7     Movie   My Little Pony: A New Generation
    7        s8     Movie                          Sankofa
    8        s9   TV Show       The Great British Baking Show
```

```
[13]: df4=df.iloc[0:4,0:5]
      df4
```

```
[13]:    show_id      type                     title            director  \
    0        s1     Movie    Dick Johnson Is Dead   Kirsten Johnson
    1        s2   TV Show           Blood & Water               NaN
    2        s3   TV Show               Ganglands   Julien Leclercq
    3        s4   TV Show   Jailbirds New Orleans               NaN

                                                      cast
    0                                                  NaN
    1   Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…
    2   Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…
    3                                                  NaN
```

```
[14]: pd.concat([df1,df2,df4])
```

```
[14]:    show_id      type                   title      director  \
    0        s1     Movie   Dick Johnson Is Dead           NaN
    1        s2   TV Show          Blood & Water           NaN
```

```
2      s3  TV Show            Ganglands                 NaN
3      s4  TV Show  Jailbirds New Orleans               NaN
4      s5  TV Show           Kota Factory               NaN
5      s6  TV Show          Midnight Mass               NaN
0      s1    Movie   Dick Johnson Is Dead  Kirsten Johnson
1      s2  TV Show          Blood & Water               NaN
2      s3  TV Show            Ganglands  Julien Leclercq
3      s4  TV Show  Jailbirds New Orleans               NaN


                                             cast
0                                             NaN
1                                             NaN
2                                             NaN
3                                             NaN
4                                             NaN
5                                             NaN
0                                             NaN
1   Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…
2   Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…
3                                             NaN
```

(iii) Sorting and shuffling, Aggregating Data at Any Level:

```
[48]: df.head()
```

```
[48]:   show_id     type                  title         director  \
0      s1    Movie   Dick Johnson Is Dead  Kirsten Johnson
1      s2  TV Show          Blood & Water               NaN
2      s3  TV Show            Ganglands  Julien Leclercq
3      s4  TV Show  Jailbirds New Orleans               NaN
4      s5  TV Show           Kota Factory               NaN


                                             cast        country  \
0                                             NaN  United States
1   Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
2   Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…            NaN
3                                             NaN            NaN
4   Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…          India


          date_added  release_year rating   duration  \
0  September 25, 2021          2020  PG-13     90 min
1  September 24, 2021          2021  TV-MA  2 Seasons
2  September 24, 2021          2021  TV-MA   1 Season
3  September 24, 2021          2021  TV-MA   1 Season
4  September 24, 2021          2021  TV-MA  2 Seasons


                                   listed_in  \
```

```
0                               Documentaries
1    International TV Shows, TV Dramas, TV Mysteries
2  Crime TV Shows, International TV Shows, TV Act…
3                          Docuseries, Reality TV
4  International TV Shows, Romantic TV Shows, TV …

                                       description
0  As her father nears the end of his life, filmm…
1  After crossing paths at a party, a Cape Town t…
2  To protect his family from a powerful drug lor…
3  Feuds, flirtations and toilet talk go down amo…
4  In a city of coaching centers known to train I…
```

[50]:
```python
#soritng

sort_df=df.sort_values(by=['title','country'],ascending=[True,False])
sort_df
```

[50]:
```
      show_id     type                          title  \
2036    s2037    Movie                         #Alive
2304    s2305    Movie  #AnneFrank - Parallel Stories
2481    s2482    Movie              #FriendButMarried
2324    s2325    Movie            #FriendButMarried 2
5973    s5974    Movie                          #Roxy
…          …        …                            …
6177    s6178  TV Show
4914    s4915  TV Show
7101    s7102  TV Show
5022    s5023    Movie
7108    s7109    Movie                            :

                          director  \
2036                        Cho Il
2304  Sabina Fedeli, Anna Migotto
2481                 Rako Prijanto
2324                 Rako Prijanto
5973               Michael Kennedy
…                              …
6177                           NaN
4914                           NaN
7101                           NaN
5022               Hong-seon Kim
7108               Young Jun Lee

                                       cast      country  \
2036               Yoo Ah-in, Park Shin-hye  South Korea
2304           Helen Mirren, Gengher Gatti        Italy
```

```
2481   Adipati Dolken, Vanesha Prescilla, Rendi Jhon,…     Indonesia
2324   Adipati Dolken, Mawar de Jongh, Sari Nila, Von…     Indonesia
5973   Jake Short, Sarah Fisher, Booboo Stewart, Dann…        Canada
…                                                    …          …
6177                                                NaN      Japan
4914   Li Nanxing, Christopher Lee, Jesseca Liu, Appl…        NaN
7101   Si-kyung Sung, Se-yoon Yoo, Dong-yup Shin, Ji-…  South Korea
5022                                     Baek Yoon-sik  South Korea
7108   Um Sang-hyun, Yang Jeong-hwa, Jeon Tae-yeol, S…        NaN

            date_added  release_year rating   duration  \
2036  September 8, 2020          2020  TV-MA     99 min
2304       July 1, 2020          2019  TV-14     95 min
2481       May 21, 2020          2018   TV-G    102 min
2324      June 28, 2020          2020   TV-G    104 min
5973     April 10, 2019          2018  TV-14    105 min
…                    …             …      …          …
6177  December 23, 2018          2012  TV-Y7  2 Seasons
4914     April 27, 2018          2016  TV-14   1 Season
7101  February 19, 2018          2015  TV-MA   1 Season
5022  February 28, 2018          2017  TV-MA    110 min
7108  September 1, 2018          2018  TV-Y7     68 min

                                            listed_in  \
2036      Horror Movies, International Movies, Thrillers
2304               Documentaries, International Movies
2481     Dramas, International Movies, Romantic Movies
2324     Dramas, International Movies, Romantic Movies
5973                      Comedies, Romantic Movies
…                                                   …
6177                            Anime Series, Kids' TV
4914               International TV Shows, TV Dramas
7101  International TV Shows, Korean TV Shows, Stand…
5022         Dramas, International Movies, Thrillers
7108                         Children & Family Movies

                                          description
2036  As a grisly virus rampages a city, a lone man …
2304  Through her diary, Anne Frank's story is retol…
2481  Pining for his high school crush for years, a …
2324  As Ayu and Ditto finally transition from best …
5973  A teenage hacker with a huge nose helps a cool…
…                                                   …
6177  Hailing from the mountains of Iga, Kanzo Hatto…
4914  Two brothers start a new life in Singapore, wh…
7101  Four Korean celebrity men and guest stars of b…
5022  After people in his town start turning up dead…
```

```
7108   Miniforce, a special task force of elite range…

[8807 rows x 12 columns]
```

```
[4]: df.head()
```

```
[4]:   show_id     type                       title          director  \
     0      s1    Movie   Dick Johnson Is Dead  Kirsten Johnson
     1      s2  TV Show          Blood & Water              NaN
     2      s3  TV Show              Ganglands  Julien Leclercq
     3      s4  TV Show  Jailbirds New Orleans              NaN
     4      s5  TV Show           Kota Factory              NaN

                                            cast        country  \
     0                                        NaN   United States
     1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban…   South Africa
     2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi…            NaN
     3                                        NaN            NaN
     4  Mayur More, Jitendra Kumar, Ranjan Raj, Alam K…          India

               date_added  release_year rating   duration  \
     0  September 25, 2021          2020  PG-13     90 min
     1  September 24, 2021          2021  TV-MA  2 Seasons
     2  September 24, 2021          2021  TV-MA   1 Season
     3  September 24, 2021          2021  TV-MA   1 Season
     4  September 24, 2021          2021  TV-MA  2 Seasons

                                       listed_in  \
     0                               Documentaries
     1    International TV Shows, TV Dramas, TV Mysteries
     2  Crime TV Shows, International TV Shows, TV Act…
     3                       Docuseries, Reality TV
     4  International TV Shows, Romantic TV Shows, TV …

                                      description
     0  As her father nears the end of his life, filmm…
     1  After crossing paths at a party, a Cape Town t…
     2  To protect his family from a powerful drug lor…
     3  Feuds, flirtations and toilet talk go down amo…
     4  In a city of coaching centers known to train I…
```

```
[5]: #shuffling

     suff_df = df.sample(frac=1).reset_index(drop=True)
     suff_df.head()
```

```
[5]:   show_id     type                                      title           director  \
    0  s5273  TV Show           VeggieTales in the City                   NaN
    1   s724  TV Show      The American Bible Challenge      Michael Simon
    2  s1299  TV Show                          Sisyphus                   NaN
    3  s4238    Movie  Sam Kinison: Breaking the Rules  Walter C. Miller
    4  s6122    Movie              All the Devil's Men      Matthew Hope

                                          cast          country  \
    0  Phil Vischer, Mike Nawrocki, Rob Paulsen, Tres…    United States
    1                                  Jeff Foxworthy    United States
    2  Cho Seung-woo, Park Shin-hye, Sung Dong-il, Ta…      South Korea
    3                                     Sam Kinison    United States
    4  Milo Gibson, Sylvia Hoeks, Gbenga Akinnagbe, W…  United Kingdom

                date_added  release_year rating   duration  \
    0  September 15, 2017          2017   TV-Y   2 Seasons
    1       June 15, 2021          2014   TV-G    1 Season
    2  February 18, 2021          2021  TV-14    1 Season
    3  December 31, 2018          1987  TV-MA      50 min
    4       May 15, 2019          2018      R     100 min

                                           listed_in  \
    0                                        Kids' TV
    1                                      Reality TV
    2  International TV Shows, TV Action & Adventure,…
    3                                  Stand-Up Comedy
    4                               Action & Adventure

                                         description
    0  With exciting trips to the big city, the ski s…
    1  Join host Jeff Foxworthy as contestants test t…
    2  An unfathomable incident introduces a genius e…
    3  The one-time preacher shakes and shudders as h…
    4  A battle-scarred Special Ops military vet join…
```

Aggregating Data

```
[8]: df.groupby('country').aggregate(['min','sum','mean','max'])
```

/tmp/ipykernel_9859/3496753236.py:1: FutureWarning: ['show_id', 'type', 'title', 'director', 'cast', 'date_added', 'rating', 'duration', 'listed_in', 'description'] did not aggregate successfully. If any error is raised this will raise in a future version of pandas. Drop these columns/ops to avoid this warning.
  df.groupby('country').aggregate(['min','sum','mean','max'])

```
[8]:                                              release_year              \
                                                        min       sum
      country
      , France, Algeria                                2014      2014
      , South Korea                                    2021      2021
      Argentina                                        1985    112896
      Argentina, Brazil, France, Poland, Germany, Den…  2017      2017
      Argentina, Chile                                 2006      4023
      …                                                 …         …
      Venezuela                                        2017      2017
      Venezuela, Colombia                              2007      2007
      Vietnam                                          2014     14124
      West Germany                                     1977      1977
      Zimbabwe                                         2017      2017


                                                        mean      max
      country
      , France, Algeria                            2014.000000   2014
      , South Korea                                2021.000000   2021
      Argentina                                    2016.000000   2021
      Argentina, Brazil, France, Poland, Germany, Den… 2017.000000   2017
      Argentina, Chile                             2011.500000   2017
      …                                                 …         …
      Venezuela                                    2017.000000   2017
      Venezuela, Colombia                          2007.000000   2007
      Vietnam                                      2017.714286   2021
      West Germany                                 1977.000000   1977
      Zimbabwe                                     2017.000000   2017

      [748 rows x 4 columns]
```