Tut 1: Getting started with pandas

Name - Tanishq Thuse

Year - SY

Branch - CSE(AI)

Div - B

Roll no. - 60

Submitted to : S M Jaybhaye Ma'am

Importing Libraries

```
import pandas as pd
import numpy as np
```

Object Creation

```
# df = pd.Series()
# print(df)

a = [1,2,3,4,5,6,7,8,9]
DF = pd.Series(a)
print(DF)
```

```
0    1
1    2
2    3
3    4
4    5
5    6
6    7
7    8
8    9
dtype: int64
```

```
dataFrame = {'Roll no': [1,2,3,4,5] ,
             'Name': ['Tanishq', 'Ayush', 'Aditya', 'Shantanu', 'Yash'],
             'DV': ['100', '98', '97', '91','95'],
            'OS' : ['98', '95', '90', '89', '83'],
             'AI' : ['98', '95', '90', '89', '83'],
             'ADS' : ['98', '95', '90', '89', '83']
}

df = pd.DataFrame(dataFrame)
print(df)

# calories = {"day1": 420, "day2": 380, "day3": 390}

# myvar = pd.Series(calories, index = ["day1", "day2"])

# print(myvar)
```

```
   Roll no      Name   DV  OS  AI ADS
0        1   Tanishq  100  98  98  98
1        2     Ayush   98  95  95  95
2        3    Aditya   97  90  90  90
3        4  Shantanu   91  89  89  89
4        5      Yash   95  83  83  83
```

```
df.dtypes
```

```
Roll no     int64
Name       object
DV         object
OS         object
```

```
AI         object
ADS        object
dtype: object
```

Viewing Data

```
df.head()
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98
1	2	Ayush	98	95	95	95
2	3	Aditya	97	90	90	90
3	4	Shantanu	91	89	89	89
4	5	Yash	95	83	83	83

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
print(df.loc[0])
```

```
Roll no          1
Name       Tanishq
DV             100
OS              98
AI              98
ADS             98
Name: 0, dtype: object
```

```
print(df.iloc[ 0 : 1])
dataFrame = df
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98

```
# print(df.loc['0'])
```

```
dataFrame.isnull()
```

	Roll no	Name	DV	OS	AI	ADS
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False

```
dataFrame.dropna()
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98
1	2	Ayush	98	95	95	95
2	3	Aditya	97	90	90	90
3	4	Shantanu	91	89	89	89
4	5	Yash	95	83	83	83

```
print(dataFrame.loc[0])
```

```
Roll no          1
Name       Tanishq
```

```
DV             100
OS              98
AI              98
ADS             98
Name: 0, dtype: object
```

```
dataFrame
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98
1	2	Ayush	98	95	95	95
2	3	Aditya	97	90	90	90
3	4	Shantanu	91	89	89	89
4	5	Yash	95	83	83	83

Next steps: Generate code with dataFrame | View recommended plots | New interactive sheet

```
dates = pd.date_range("20130101", periods=6)
dates
```

```
DatetimeIndex(['2013-01-01', '2013-01-02', '2013-01-03', '2013-01-04',
               '2013-01-05', '2013-01-06'],
              dtype='datetime64[ns]', freq='D')
```

```
dataFrame
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98
1	2	Ayush	98	95	95	95
2	3	Aditya	97	90	90	90
3	4	Shantanu	91	89	89	89
4	5	Yash	95	83	83	83

Next steps: Generate code with dataFrame | View recommended plots | New interactive sheet

```
dataFrame.dtypes
```

```
Roll no     int64
Name       object
DV         object
OS         object
AI         object
ADS        object
dtype: object
```

```
df = dataFrame
df.dtypes
```

```
Roll no     int64
Name       object
DV         object
OS         object
AI         object
ADS        object
dtype: object
```

```
df.T
```

	0	1	2	3	4
Roll no	1	2	3	4	5
Name	Tanishq	Ayush	Aditya	Shantanu	Yash
DV	100	98	97	91	95
OS	98	95	90	89	83
AI	98	95	90	89	83
ADS	98	95	90	89	83

```
df.describe()
```

	Roll no
count	5.000000
mean	3.000000
std	1.581139
min	1.000000
25%	2.000000
50%	3.000000
75%	4.000000
max	5.000000

```
df.isnull()
```

	Roll no	Name	DV	OS	AI	ADS
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False

Selection

```
#TO sort by values in DV Subject
df.sort_values(by='DV')
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98
3	4	Shantanu	91	89	89	89
4	5	Yash	95	83	83	83
2	3	Aditya	97	90	90	90
1	2	Ayush	98	95	95	95

```
df['DV']
```

```
0    100
1     98
2     97
3     91
4     95
Name: DV, dtype: object
```

```
df[0:3]
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98
1	2	Ayush	98	95	95	95
2	3	Aditya	97	90	90	90

Selection by label

```
df.loc[:, ['Name','DV' ]]
```

	Name	DV
0	Tanishq	100
1	Ayush	98
2	Aditya	97
3	Shantanu	91
4	Yash	95

```
df.loc[0:2, ['Name','DV', 'ADS']]
```

	Name	DV	ADS
0	Tanishq	100	98
1	Ayush	98	95
2	Aditya	97	90

Selection by position

```
df.iloc[3]
```

```
Roll no          4
Name       Shantanu
DV               91
OS               89
AI               89
ADS              89
Name: 3, dtype: object
```

```
df.iloc[0:1,:]
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98

```
df
```

	Roll no	Name	DV	OS	AI	ADS
0	1	Tanishq	100	98	98	98
1	2	Ayush	98	95	95	95
2	3	Aditya	97	90	90	90
3	4	Shantanu	91	89	89	89
4	5	Yash	95	83	83	83

Next steps: Generate code with dataFrame | View recommended plots | New interactive sheet

```
#A python function to assign DV Subject grade
def grade_dv(marks):
    marks = int(marks)
    if marks > 95:
        return 'A+'
    elif marks > 90 :
        return 'A'
    else:
        return 'B'

# Apply the grading function to the DV marks
df['DV Grade'] = df['DV'].apply(grade_dv)

# Display the updated DataFrame
print(df)
```

```
   Roll no      Name   DV  OS  AI ADS DV Grade
0        1   Tanishq  100  98  98  98        A+
1        2     Ayush   98  95  95  95        A+
2        3    Aditya   97  90  90  90        A+
3        4  Shantanu   91  89  89  89         A
4        5      Yash   95  83  83  83         A
```

Getting

```
df["Name"]
```

```
0     Tanishq
1       Ayush
2      Aditya
3    Shantanu
4        Yash
Name: Name, dtype: object
```

```
df[0:3]
```

	Roll no	Name	DV	OS	AI	ADS	DV Grade
0	1	Tanishq	100	98	98	98	A+
1	2	Ayush	98	95	95	95	A+
2	3	Aditya	97	90	90	90	A+

```
df[0:4]
```

	Roll no	Name	DV	OS	AI	ADS	DV Grade
0	1	Tanishq	100	98	98	98	A+
1	2	Ayush	98	95	95	95	A+
2	3	Aditya	97	90	90	90	A+
3	4	Shantanu	91	89	89	89	A

Selection by label

```
df.head() #having a look at the dataframe
```

	Roll no	Name	DV	OS	AI	ADS	DV Grade
0	1	Tanishq	100	98	98	98	A+
1	2	Ayush	98	95	95	95	A+
2	3	Aditya	97	90	90	90	A+
3	4	Shantanu	91	89	89	89	A
4	5	Yash	95	83	83	83	A

Next steps: Generate code with df | View recommended plots | New interactive sheet

```
df.loc[:, ["Name", "DV"]]
```

	Name	DV
0	Tanishq	100
1	Ayush	98
2	Aditya	97
3	Shantanu	91
4	Yash	95

```
df.loc[0:2, ["Name", "DV"]]
```

	Name	DV
0	Tanishq	100
1	Ayush	98
2	Aditya	97

Selection by position

```
df.iloc[0]
```

```
Roll no          1
Name          Tanishq
DV              100
OS               98
AI               98
ADS              98
DV Grade         A+
Name: 0, dtype: object
```

```
df.iloc[0:2, :]
```

	Roll no	Name	DV	OS	AI	ADS	DV Grade
0	1	Tanishq	100	98	98	98	A+
1	2	Ayush	98	95	95	95	A+

```
df.iloc[:, 1:3]
```

	Name	DV
0	Tanishq	100
1	Ayush	98
2	Aditya	97
3	Shantanu	91
4	Yash	95

Practise of some python functions

Hello world practice

```
def say_hello():
  print("Hello, World!")
```

```
say_hello()
```

```
Hello, World!
```

```
say_hello()
say_hello()
say_hello()
say_hello()
say_hello()
```

```
Hello, World!
Hello, World!
Hello, World!
Hello, World!
Hello, World!
```

Square of a user entered number

```
def square(i):
  print(i*i)
```

```
n=int(input("Enter number:"))
square(n)
```

```
Enter number:6
36
```

Practice of methods

append, copy, clear, count, extend, insert, pop, remove, reverse, sort, min, max

```
colours=['Red','Blue','Green']
capacities=[100,200,300]
```

append adds to the list (at last)

```
capacities.append(400)
```

```
print(capacities)
```

```
[100, 200, 300, 400]
```

```
colours.append('Yellow')
```

```
print(colours)
```

```
['Red', 'Blue', 'Green', 'Yellow']
```

insert() takes 2 input values first => position on which value is to be inserted, second => value to be inserted

```
capacities.insert(1,200)
print(capacities)
```

```
[100, 200, 200, 300, 400]
```

what you want to remove

```
colours.remove('Red')
```

```
print(colours)
```

```
['Blue', 'Green', 'Yellow']
```

gives the postion of the input value

```
capacities.index(400)
```

```
4
```

if there are 2 same values then it returns the position of the first one

```
capacities.index(200)
```

```
1
```

Copies the content of one list to another

```
a=colours.copy()
print(a)
```

```
['Blue', 'Green', 'Yellow']
```

Clears the given list completely

```
a.clear()
print(a)
```

```
[]
```

Counts how many times the given input is repeated in the list

```
capacities.count(200)
```

```
2
```

Removes the last item from the list

```
print(colours)
colours.pop()
print(colours)
```

```
['Blue', 'Green', 'Yellow']
['Blue', 'Green']
```

Sorts the list in ascending order

```
capacities.sort()
print(capacities)
```

```
[100, 200, 200, 300, 400]
```

Maximum and minimum

```
print(min(capacities))
print(max(capacities))
```

```
100
400
```

Extend merges two lists

```
print(colours)
print(capacities)

colours.extend(capacities)

print(colours)
```

```
['Blue', 'Green']
[100, 200, 200, 300, 400]
['Blue', 'Green', 100, 200, 200, 300, 400]
```

Type of data

```
a=-1
b='hello'
c='g'
d=10.345
print(type(a))
print(type(b))
print(type(c))
print(type(d))
```

```
<class 'int'>
<class 'str'>
<class 'str'>
<class 'float'>
```

if-else

```
num1=eval(input('Enter a number:'))
num2=eval(input('Enter another number:'))

if num1==num2:
  print('both are same')
elif num1>num2:
  print('1st is greater than 2nd')
elif num2>num1:
  print('2nd is greater than 1st')
```

```
Enter a number:9
Enter another number:10
2nd is greater than 1st
```

Slicing practice

```
print(capacities)
```

```
[100, 200, 200, 300, 400]
```

```
print(capacities[:])
```

```
[100, 200, 200, 300, 400]
```

```
print(capacities[2:])
```

```
[200, 300, 400]
```

```
print(capacities[:3])
```

```
[100, 200, 200]
```

```
print(capacities[-4:])
```

```
[200, 200, 300, 400]
```