

Pandasjupyter

January 20, 2022

0.1 Introduction

```
[ ]: import pandas as pd
import numpy as np
```

0.2 Create database

```
[ ]: dict = {"Name":['Aryan' , 'Vishesh'],
            "Stream" : ['PCMC' , 'PCBPE']}
```

```
[ ]: df = pd.DataFrame(dict)
```

0.3 Converting into Excel File and other function of to_csv()

```
[ ]: #to convert it into CSV file
#Don't want index
df.to_csv('friends.csv' , index = False)
```

0.4 Extracting data from table

```
[ ]: # one from top
df.head(1)
```

```
[ ]:      Name Stream
0  Aryan   PCMC
```

```
[ ]: #one from bottom
df.tail(1)
```

```
[ ]:      Name Stream
1  Vishesh  PCBPE
```

```
[ ]: #Get all information about in numerical column (max , min in column percentile,
↵, count , mean )
df.describe()
```

```
[ ]:      Name Stream
count      2      2
```

```
unique      2      2
top      Aryan  PCMC
freq        1      1
```

```
[ ]: # Get table from another excel sheet
t = pd.read_csv('test.csv')
```

```
[ ]: t
```

```
[ ]:      Train No  Speed  City
0      123458     34   JBL
1      123456     56   MU
2      123457     88  KOTA
```

```
[ ]: t['Speed']
```

```
[ ]: 0      34
1      56
2      88
Name: Speed, dtype: int64
```

```
[ ]: #Copy it into another csv file
t.to_csv('train.csv')
```

0.5 Changing index in excel file

```
[ ]: t.index = [1,2,3]
t
```

```
[ ]:      Train No  Speed  City
1      123458     34   JBL
2      123456     56   MU
3      123457     88  KOTA
```

```
[ ]: t.sort_index(axis = 0 , ascending= False)
```

```
[ ]:      Train No  Speed  City
3      123457     88  KOTA
2      123456     56   MU
1      123458     34   JBL
```

1 Loc

it is use to change the value in the dataframe without showing error eg:-

```
[ ]: #it is use to change the value in the dataframe without showing error eg:-
t['Train No'][1] = 123490
```

```
t.head()
```

C:\Users\Aryan\AppData\Local\Temp\ipykernel_12616\2691491249.py:2:

SettingWithCopyWarning:

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
t['Train No'][1] = 123490
```

```
[ ]:   Train No  Speed  City
      1    123490    34   JBL
      2    123456    56   MU
      3    123457    88  KOTA
```

```
[ ]: #Don't Show error
      t.loc[1, 'Train No'] = 123499 #[row , column]
      t.head()
```

```
[ ]:   Train No  Speed  City
      1    123499    34   JBL
      2    123456    56   MU
      3    123457    88  KOTA
```

- Now you can see that we by mistake created 2 columns (2,0) and (1,0) So we can remove it by using Drop
 - `t = t.drop((2,0), axis = 1)`
 - `t = t.drop((1,0), axis = 1)` python or instead of using `t = t.drop((2,0), axis = 1)` we can write `t.drop((2,0), axis = 1, inplace = true)`

```
[ ]: t.head()
```

```
[ ]:   Train No  Speed  City
      1    123499    34   JBL
      2    123456    56   MU
      3    123457    88  KOTA
```

```
[ ]: #to get particular value with row and column
      t.loc[[2,3], ['Train No', 'City']]
#to get all value of particular column syntax =[:,[Speed]] , for row = [[1,2],:]
      ↪]
```

```
[ ]:   Train No  City
      2    123456   MU
      3    123457  KOTA
```

2 To Run complex query in table

```
[ ]: t.loc[(t['Speed'] > 58)]
```

```
[ ]:   Train No  Speed  City  
      3    123457     88  KOTA
```

```
[ ]: t.loc[(t['Speed'] < 58)]
```

```
[ ]:   Train No  Speed  City  
      1    123499     34  JBL  
      2    123456     56   MU
```

```
[ ]: t.loc[(t['Speed'] < 58) & (t['Train No'] < 123456)]
```

```
[ ]: Empty DataFrame  
      Columns: [Train No, Speed, City]  
      Index: []
```

```
[ ]: t.loc[(t['Speed'] < 58) & (t['Train No'] <=123456)]
```

```
[ ]:   Train No  Speed  City  
      2    123456     56   MU
```

```
[ ]: t.loc[(t['Speed'] < 58) & (t['Train No'] <=123459)]
```

```
[ ]:   Train No  Speed  City  
      2    123456     56   MU
```

3 Iloc

It is use to count from index like in loc we have to mention particular row and column name but in iloc we can just write only index number of row and column

3.1 Some more comman command we will use

To change all value of a particular row or column with same values
syntax:

```
table_name loc[[row_name , :]] = value or table_name[: , column_name] = value
```