PANDAS PART 2

Working with columns

Creating new columns

```
[1]: import pandas as pd
```

```
[17]: # create a tuple of tuples
      tpl = ((1001, 'Nagesh'), (1002, 'Laxmi'), (1003, 'Ganesh'))
```

```
[18]: df = pd.DataFrame(tpl)
      df
```

[18]:

	0	1
0	1001	Nagesh
1	1002	Laxmi
2	1003	Ganesh

```
[19]: # add column names
      df.columns = ['col1', 'col2']
      df
```

[19]:

	col1	col2
0	1001	Nagesh
1	1002	Laxmi
2	1003	Ganesh

```
[20]: # rename the column names
      df.rename(columns= {'col1':'empid', 'col2': 'name'}, inplace=True)
      df
```

[20]:

	empid	name
0	1001	Nagesh
1	1002	Laxmi
2	1003	Ganesh

Add a new column by typing its name and values

```
[21]:  # add a new column name with its values
       df['dept'] = 'Sales'
       df
```

[21]:

	empid	name	dept
0	1001	Nagesh	Sales
1	1002	Laxmi	Sales
2	1003	Ganesh	Sales

```
[22]:  # same as above with multiple values
       df['dept'] = ['Sales', 'HR', 'Production']
       df
```

[22]:

	empid	name	dept
0	1001	Nagesh	Sales
1	1002	Laxmi	HR
2	1003	Ganesh	Production

Adding a new column in a particular position using insert() method

```
[23]:  # insert() can be used to insert a new col in a particular location
       df.insert(loc=1, column='Salary', value=[35000, 22000, 25000])
       df
```

[23]:

	empid	Salary	name	dept
0	1001	35000	Nagesh	Sales
1	1002	22000	Laxmi	HR
2	1003	25000	Ganesh	Production

Creating new column from existing columns

```
[24]:  # create a new column income tax as 10% of Salary
       df['Tax'] = df['Salary']*10/100
       df
```

[24]:

	empid	Salary	name	dept	Tax
0	1001	35000	Nagesh	Sales	3500.0
1	1002	22000	Laxmi	HR	2200.0
2	1003	25000	Ganesh	Production	2500.0

Creating new column from existing columns using lambda

pyter **Untitled1** Last Checkpoint: 36 minutes ago (autosaved)

Edit   View   Insert   Cell   Kernel   Help

✂ ⎘ 📋 ↑ ↓ ▶ Run ■ C ⏭ | Code ⌄ | ⌨

```
[36]:   1  # create a new column PF as 12.5% of Salary using lambda
        2  df1 = df.assign(PF = lambda x: (x.Salary) * 12.5/100)
        3  df1
```

[36]:

	empid	Salary	name	dept	Tax	PF
0	1001	35000	Nagesh	Sales	3500.0	4375.0
1	1002	22000	Laxmi	HR	2200.0	2750.0
2	1003	25000	Ganesh	Production	2500.0	3125.0

```
[37]:   # create a new column gross as Salary – Tax – PF
        df2 = df1.assign(gross = lambda x: x.Salary – x.Tax – x.PF)
        df2
```

[37]:

	empid	Salary	name	dept	Tax	PF	gross
0	1001	35000	Nagesh	Sales	3500.0	4375.0	27125.0
1	1002	22000	Laxmi	HR	2200.0	2750.0	17050.0
2	1003	25000	Ganesh	Production	2500.0	3125.0	19375.0