

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

```
In [ ]: df = pd.read_csv("employee.csv")
```

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

```
Out[ ]:
```

	Empid	Name	Salary
0	1	Smith	5000
1	2	Jones	6000
2	3	Harry	4000
3	4	Jem	5000
4	5	Sylvia	4000

```
In [ ]: df
```

```
Out[ ]:
```

	Empid	Name	Salary
0	1	Smith	5000
1	2	Jones	6000
2	3	Harry	4000
3	4	Jem	5000
4	5	Sylvia	4000
5	6	Charles	4000
6	7	William	5000
7	7	William	5000

```
In [ ]: # How to remove duplicate rows
```

```
In [ ]: df.drop_duplicates() # Without making changes in dataframe
```

```
Out[ ]:      Empid  Name  Salary
0         1  Smith   5000
1         2   Jones   6000
2         3   Harry   4000
3         4     Jem   5000
4         5  Sylvia   4000
5         6 Charles   4000
6         7 William   5000
```

```
In [ ]: df
```

```
Out[ ]:      Empid  Name  Salary
0         1  Smith   5000
1         2   Jones   6000
2         3   Harry   4000
3         4     Jem   5000
4         5  Sylvia   4000
5         6 Charles   4000
6         7 William   5000
7         7 William   5000
```

```
In [ ]: df.drop_duplicates(inplace = True) # Changes will be made in the dataframe itself.
```

```
In [ ]: df
```

```
Out[ ]:      Empid  Name  Salary
0         1  Smith   5000
1         2   Jones   6000
2         3   Harry   4000
3         4     Jem   5000
4         5  Sylvia   4000
5         6 Charles   4000
6         7 William   5000
```

```
In [ ]: ### How to Drop Column Salary ?
      ## axis specifies row (0) and column(1)
```

```
In [ ]: df.drop(['Salary'], axis = 1, inplace = True )
```

```
In [ ]: df
```

```
Out[ ]:
```

	Empid	Name
0	1	Smith
1	2	Jones
2	3	Harry
3	4	Jem
4	5	Sylvia
5	6	Charles
6	7	William

```
In [ ]: df.drop([4], axis = 0 )
```

```
Out[ ]:
```

	Empid	Name
0	1	Smith
1	2	Jones
2	3	Harry
3	4	Jem
5	6	Charles
6	7	William

```
In [ ]: df
```

```
Out[ ]:
```

	Empid	Name
0	1	Smith
1	2	Jones
2	3	Harry
3	4	Jem
4	5	Sylvia
5	6	Charles
6	7	William

```
In [ ]:
```