

t4bj2hngj

November 18, 2025

```
[ ]: pip install pandas
```

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

```
[ ]: df=pd.read_csv(r'C:\Users\HP\Downloads\sales_performance_dataset.csv')
df
```

```
[ ]:      Employee_ID Employee_Name Department  Experience_Years  Monthly_Sales \
0            1001        Ravi    Finance             15          41834
1            1002       Sneha  Marketing            13          38047
2            1003       Amit       IT                1          46105
3            1004       Priya  Marketing            9          95766
4            1005       Karan  Marketing            7          35707
..           ...
95           1096      Megha    Finance            12          93656
96           1097     Pritam    Finance            12          59384
97           1098     Ramesh  Marketing            4          67254
98           1099    Shivani     Sales              14          41918
99           1100      Niraj  Marketing            14         105981

      Customer_Satisfaction
0                      10
1                      3
2                      7
3                     10
4                     9
..                     ...
95                     6
96                     7
97                     2
98                     10
99                     2

[100 rows x 6 columns]
```

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

```
[ ]: Employee_ID Employee_Name Department Experience_Years Monthly_Sales \
0 1001 Ravi Finance 15 41834
1 1002 Sneha Marketing 13 38047
2 1003 Amit IT 1 46105
3 1004 Priya Marketing 9 95766
4 1005 Karan Marketing 7 35707
```

Customer_Satisfaction

```
0 10
1 3
2 7
3 10
4 9
```

```
[ ]: df.tail()
```

```
[ ]: Employee_ID Employee_Name Department Experience_Years Monthly_Sales \
95 1096 Megha Finance 12 93656
96 1097 Pritam Finance 12 59384
97 1098 Ramesh Marketing 4 67254
98 1099 Shivani Sales 14 41918
99 1100 Niraj Marketing 14 105981
```

Customer_Satisfaction

```
95 6
96 7
97 2
98 10
99 2
```

```
[ ]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Employee_ID      100 non-null    int64  
 1   Employee_Name    100 non-null    object  
 2   Department       100 non-null    object  
 3   Experience_Years 100 non-null    int64  
 4   Monthly_Sales    100 non-null    int64  
 5   Customer_Satisfaction 100 non-null    int64  
dtypes: int64(4), object(2)
memory usage: 4.8+ KB
```

```
[ ]: df.describe()
```

```
[ ]: df.describe()
```

	Employee_ID	Experience_Years	Monthly_Sales	Customer_Satisfaction
count	100.000000	100.000000	100.000000	100.000000
mean	1050.500000	8.450000	72030.750000	4.840000
std	29.011492	4.349329	30111.719996	2.805910
min	1001.000000	1.000000	20854.000000	1.000000
25%	1025.750000	5.000000	44832.000000	3.000000
50%	1050.500000	8.500000	72697.500000	4.500000
75%	1075.250000	12.000000	101270.000000	7.000000
max	1100.000000	15.000000	118506.000000	10.000000

```
[ ]: df.shape
```

```
[ ]: (100, 6)
```

```
[ ]: df.columns
```

```
[ ]: Index(['Employee_ID', 'Employee_Name', 'Department', 'Experience_Years',  
          'Monthly_Sales', 'Customer_Satisfaction'],  
          dtype='object')
```

```
[ ]: df.dtypes
```

```
[ ]: Employee_ID           int64  
Employee_Name          object  
Department            object  
Experience_Years        int64  
Monthly_Sales          int64  
Customer_Satisfaction  int64  
dtype: object
```

```
[ ]: df_drop=df.drop(columns=['Experience_Years'])  
df_drop
```

```
[ ]: df
```

	Employee_ID	Employee_Name	Department	Monthly_Sales	Customer_Satisfaction
0	1001	Ravi	Finance	41834	10
1	1002	Sneha	Marketing	38047	3
2	1003	Amit	IT	46105	7
3	1004	Priya	Marketing	95766	10
4	1005	Karan	Marketing	35707	9
..
95	1096	Megha	Finance	93656	6
96	1097	Pritam	Finance	59384	7
97	1098	Ramesh	Marketing	67254	2
98	1099	Shivani	Sales	41918	10
99	1100	Niraj	Marketing	105981	2

```
[100 rows x 5 columns]
```

```
[ ]: df_renamed=df.rename(columns={'Employee_Name':'Employee_FullName'})
df_renamed
```

	Employee_ID	Employee_FullName	Department	Experience_Years	Monthly_Sales
0	1001	Ravi	Finance	15	41834
1	1002	Sneha	Marketing	13	38047
2	1003	Amit	IT	1	46105
3	1004	Priya	Marketing	9	95766
4	1005	Karan	Marketing	7	35707
..
95	1096	Megha	Finance	12	93656
96	1097	Pritam	Finance	12	59384
97	1098	Ramesh	Marketing	4	67254
98	1099	Shivani	Sales	14	41918
99	1100	Niraj	Marketing	14	105981
		Customer_Satisfaction			
0		10			
1		3			
2		7			
3		10			
4		9			
..		...			
95		6			
96		7			
97		2			
98		10			
99		2			

[100 rows x 6 columns]

```
[ ]: df_sorted=df.sort_values(by='Experience_Years')
df_sorted
```

	Employee_ID	Employee_Name	Department	Experience_Years	Monthly_Sales
2	1003	Amit	IT	1	46105
7	1008	Arjun	IT	1	43776
14	1015	Anjali	HR	1	29474
18	1019	Suman	HR	1	114856
31	1032	Monika	HR	1	71005
..
46	1047	Shweta	HR	15	110084
58	1059	Sameer	Finance	15	118506
59	1060	Sonia	Finance	15	32688
65	1066	Rehan	Finance	15	109045
76	1077	Prakash	HR	15	69811

```
Customer_Satisfaction  
2 7  
7 2  
14 3  
18 8  
31 9  
.. ...  
46 8  
58 3  
59 6  
65 5  
76 8
```

[100 rows x 6 columns]

```
[ ]: df_fillna=df.fillna(0)  
df_fillna
```

```
[ ]: Employee_ID Employee_Name Department Experience_Years Monthly_Sales \\  
0 1001 Ravi Finance 15 41834  
1 1002 Sneha Marketing 13 38047  
2 1003 Amit IT 1 46105  
3 1004 Priya Marketing 9 95766  
4 1005 Karan Marketing 7 35707  
.. ... ... ... ... ...  
95 1096 Megha Finance 12 93656  
96 1097 Pritam Finance 12 59384  
97 1098 Ramesh Marketing 4 67254  
98 1099 Shivani Sales 14 41918  
99 1100 Niraj Marketing 14 105981
```

```
Customer_Satisfaction  
0 10  
1 3  
2 7  
3 10  
4 9  
.. ...  
95 6  
96 7  
97 2  
98 10  
99 2
```

[100 rows x 6 columns]

```
[ ]: drop_uniques=df.drop_duplicates()  
drop_uniques
```

```
[ ]: Employee_ID Employee_Name Department Experience_Years Monthly_Sales \  
0 1001 Ravi Finance 15 41834  
1 1002 Sneha Marketing 13 38047  
2 1003 Amit IT 1 46105  
3 1004 Priya Marketing 9 95766  
4 1005 Karan Marketing 7 35707  
.. ... ... ... ... ...  
95 1096 Megha Finance 12 93656  
96 1097 Pritam Finance 12 59384  
97 1098 Ramesh Marketing 4 67254  
98 1099 Shivani Sales 14 41918  
99 1100 Niraj Marketing 14 105981  
  
Customer_Satisfaction  
0 10  
1 3  
2 7  
3 10  
4 9  
.. ...  
95 6  
96 7  
97 2  
98 10  
99 2
```

[100 rows x 6 columns]

```
[ ]: grouped_df=df.groupby('Employee_Name').sum()  
grouped_df
```

```
[ ]: Employee_ID Department Experience_Years Monthly_Sales \  
Employee_Name  
Aarav 1069 HR 4 42671  
Aditya 1041 Sales 6 87172  
Alok 1064 Support 13 72083  
Amit 1003 IT 1 46105  
Anaya 1090 Marketing 14 91295  
... ... ... ... ...  
Varun 1048 Finance 14 58623  
Vikas 1031 IT 2 74384  
Vinay 1095 Finance 12 28155  
Vivek 1012 IT 15 26776  
Yash 1035 Support 14 82003
```

```

Customer_Satisfaction
Employee_Name
Aarav                      3
Aditya                     4
Alok                       1
Amit                       7
Anaya                      6
...
Varun                      8
Vikas                      3
Vinay                      2
Vivek                      7
Yash                       2

```

[100 rows x 5 columns]

```
[ ]: agg_df=df.groupby('Experience_Years').agg({'Customer_Satisfaction':'sum'})
agg_df
```

```

Customer_Satisfaction
Experience_Years
1                         37
2                          5
3                         46
4                          9
5                         24
6                         13
7                         61
8                         33
9                         37
10                        25
11                        28
12                        47
13                        26
14                        40
15                        53

```

```
[ ]: df_count=df.groupby('Employee_Name').count()
df_count
```

```

Employee_ID  Department  Experience_Years  Monthly_Sales \
Employee_Name
Aarav          1           1                  1                 1
Aditya         1           1                  1                 1
Alok           1           1                  1                 1
Amit           1           1                  1                 1

```

Anaya	1	1	1	1
...
Varun	1	1	1	1
Vikas	1	1	1	1
Vinay	1	1	1	1
Vivek	1	1	1	1
Yash	1	1	1	1

Customer_Satisfaction

Employee_Name	
Aarav	1
Aditya	1
Alok	1
Amit	1
Anaya	1
...	...
Varun	1
Vikas	1
Vinay	1
Vivek	1
Yash	1

[100 rows x 5 columns]

```
[ ]: df_sum=df.groupby('Experience_Years').sum()
df_sum
```

```
[ ]: Employee_ID \
Experience_Years
1           8239
2           2118
3           9373
4           4318
5           6320
6           3219
7          11529
8           7149
9           6212
10          6309
11          6293
12         10635
13          7461
14          7499
15          8376
```

```
Employee_Name \
Experience_Years
```

1	Amit	Arjun	Anjali	Suman	Monika	Arnav	Manish	Sudhir				
2							Vikas	Tanvi				
3	Nitin	Rakesh	Deepa	Shreya	Geeta	Payal	Juhি	Namita	Mohan			
4							Aarav	Tushar	JayRamesh			
5					Isha	Ankit	Dev	Ritu	Sandeep	Deepak		
6								Aditya	Nisha	Shalini		
7	Karan	Gaurav	Sahil	Rajesh	Sheetal	Ira	Lakshmi	Niharik...				
8					Neha	Rohit	Meena	Harsh	Kabir	Poonam	Komal	
9						Priya	Divya	Rekha	Nidhi	Bhavana	Sujit	
10							Tanya	Manoj	Rohan	Tejas	Diya	Hemant
11							Pooja	Rahul	Chirag	Sunny	Mitali	Kavita
12	Simran	Tanuj	Preeti	Tarun	Naveen	Raghu	Omkar	Vinay	Meg...			
13					Sneha	Alok	Lavanya	Snehal	Harini	Vandana	Usha	
14						Yash	Varun	Krishna	Rajni	Anaya	Shivani	Niraj
15					Ravi	Vivek	Shweta	Ria	Sameer	Sonia	Rehan	Prakash

Experience_Years	Department	\
1	IT	IT
2	IT	HR
3	Marketing	Support
4	Support	IT
5	Marketing	Support
6	Sales	Support
7	Marketing	Finance
8	IT	Marketing
9	Marketing	Finance
10	Sales	HR
11	HR	Finance
12	IT	Marketing
13	Marketing	Support
14	Support	Finance
15	Finance	IT

Experience_Years	Monthly_Sales	Customer_Satisfaction
1	528255	37
2	103064	5
3	591332	46
4	315391	9
5	415121	24
6	231946	13
7	741142	61
8	573244	33
9	454871	37
10	518784	25
11	386840	28

```
12          736547        47
13          502103        26
14          549382        40
15          555053        53
```

```
[ ]: df_mean=df.groupby('Experience_Years').mean()
df_mean
```

```
[ ]: df_max=df.groupby('Employee_Name').max()
df_max
```

```
[ ]:           Employee_ID Department  Experience_Years  Monthly_Sales \
Employee_Name
Aarav            1069        HR             4        42671
Aditya           1041      Sales            6        87172
Alok             1064    Support           13        72083
Amit              1003        IT             1        46105
Anaya            1090  Marketing          14        91295
...
Varun            1048      Finance          14        58623
Vikas            1031        IT             2        74384
Vinay            1095      Finance          12        28155
Vivek            1012        IT             15        26776
Yash              1035    Support           14        82003
```

Customer_Satisfaction

```
Employee_Name
Aarav            3
Aditya           4
Alok             1
Amit              7
Anaya            6
...
Varun            8
Vikas            3
Vinay            2
Vivek            7
Yash              2
```

[100 rows x 5 columns]

```
[ ]: df_min=df.groupby('Employee_Name').min()
df_min
```

```
[ ]:           Employee_ID Department  Experience_Years  Monthly_Sales \
Employee_Name
Aarav            1069        HR             4        42671
```

Aditya	1041	Sales	6	87172
Alok	1064	Support	13	72083
Amit	1003	IT	1	46105
Anaya	1090	Marketing	14	91295
...
Varun	1048	Finance	14	58623
Vikas	1031	IT	2	74384
Vinay	1095	Finance	12	28155
Vivek	1012	IT	15	26776
Yash	1035	Support	14	82003

Customer_Satisfaction

Employee_Name	
Aarav	3
Aditya	4
Alok	1
Amit	7
Anaya	6
...	...
Varun	8
Vikas	3
Vinay	2
Vivek	7
Yash	2

[100 rows x 5 columns]

```
[ ]: df_clear=df.dropna()
df_clear
```

```
[ ]: Employee_ID Employee_Name Department Experience_Years Monthly_Sales \
0 1001 Ravi Finance 15 41834
1 1002 Sneha Marketing 13 38047
2 1003 Amit IT 1 46105
3 1004 Priya Marketing 9 95766
4 1005 Karan Marketing 7 35707
...
95 1096 Megha Finance 12 93656
96 1097 Pritam Finance 12 59384
97 1098 Ramesh Marketing 4 67254
98 1099 Shivani Sales 14 41918
99 1100 Niraj Marketing 14 105981
```

Customer_Satisfaction

0	10
1	3
2	7

```
3          10
4          9
..
95         ...
96         6
97         7
98         2
99         10
2
```

[100 rows x 6 columns]

```
[ ]: df_fillna=df.fillna(0)
df_fillna
```

```
[ ]:   Employee_ID Employee_Name Department Experience_Years Monthly_Sales \
0           1001        Ravi     Finance            15       41834
1           1002      Sneha    Marketing           13       38047
2           1003       Amit      IT                1       46105
3           1004      Priya    Marketing           9       95766
4           1005      Karan    Marketing           7       35707
..
95          ...        ...      ...            ...       ...
96          1096      Megha    Finance           12       93656
97          1097     Pritam    Finance           12       59384
98          1098     Ramesh    Marketing          4       67254
99          1100     Shivani   Sales             14       41918
2
```

Customer_Satisfaction

```
0          10
1          3
2          7
3          10
4          9
..
95         ...
96         6
97         7
98         2
99         10
2
```

[100 rows x 6 columns]

```
[ ]: df_replace=df.replace({' ':0})
df_replace
```

```
[ ]:   Employee_ID Employee_Name Department Experience_Years Monthly_Sales \
0           1001        Ravi     Finance            15       41834
2
```

1	1002	Sneha	Marketing	13	38047
2	1003	Amit	IT	1	46105
3	1004	Priya	Marketing	9	95766
4	1005	Karan	Marketing	7	35707
..
95	1096	Megha	Finance	12	93656
96	1097	Pritam	Finance	12	59384
97	1098	Ramesh	Marketing	4	67254
98	1099	Shivani	Sales	14	41918
99	1100	Niraj	Marketing	14	105981

Customer_Satisfaction					
0	10				
1	3				
2	7				
3	10				
4	9				
..	..				
95	6				
96	7				
97	2				
98	10				
99	2				

[100 rows x 6 columns]

```
[ ]: df_contain=df['Employee_Name'].str.contains('o')
df_contain
```

```
[ ]: df['Name']=df['Employee_Name'].str.strip()
df['Name']
```

```
[ ]: count=df['Employee_Name'].values.counts()
count
```