14. write a program to find the replace the missing values in given Dataframes which do not have any Valuable information

Aim:-

To find the replace the missing values in the given Dataframes which do not have any Valuable information.

Pseudo Code:-

→ Import pandas and numpy

→ Generate a Dataframe with random values and introduce some NaN

→ choose a replacement strategy and apply it using fillna()

→ Display the Dataframe.

Sample input:-

df = pd.Dataframe (np.random.rand (10,4). Columns = ['A', 'B', 'c', 'D'])

Sample output:-

	A	B	C	D	
0	0.79	0.34	0.65	0.12	0
1	0.32	0.45	0.57	0.89	1
2	0.51	0.14	0.76	0.37	2
..	...	...	...	..	
9	0.29	0.43	0.69	0.46	

Result:-

this Code executed Sucassfully and we got the output

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

# Sample DataFrame based on your screenshot
data = {
    'ord_no': [70001, np.nan, 70002, 70004, np.nan, 70005, '--', 70010, 70003, 70012, np.nan, 70013],
    'purch_amt': [150.5, 270.65, 65.26, 110.5, 948.5, 2400.6, 5760, '?', 12.43, 2480.4, 250.45, 3045.6],
    'ord_date': ['?', '2012-09-10', np.nan, '2012-08-17', '2012-09-10', '2012-07-27', '2012-09-10', '2012-10-10', '2012-10-10', '2012-06-27', '2012-08-17', '2012-04-25'],
    'customer_id': [3002, 3001, 3001, 3003, 3002, 3001, 3001, 3004, '--', 3002, 3001, 3001],
    'salesman_id': [5002, 5003, '?', 5001, np.nan, 5001, 5001, '?', 5003, 5002, 5003, '--']
}

df = pd.DataFrame(data)

# Replace placeholders ('?', '--') with NaN
df.replace(['?', '--'], np.nan, inplace=True)

# Display the DataFrame after replacement
print("DataFrame after replacing placeholders with NaN:\n", df)

# Optionally, you can fill NaN values with suitable replacements, like 0 for numerical columns o
df_filled = df.fillna({
    'ord_no': 0,            # Fill missing order numbers with 0
    'purch_amt': df['purch_amt'].mean(),  # Fill missing purchase amounts with the column mean
    'ord_date': 'Unknown',  # Fill missing order dates with 'Unknown'
    'customer_id': 0,       # Fill missing customer IDs with 0
    'salesman_id': 0        # Fill missing salesman IDs with 0
})

# Display the DataFrame after filling NaN values
print("\nDataFrame after filling NaN values:\n", df_filled)
```

IDLE Shell 3.12.4

File Edit Shell Debug Options Window Help

```
DataFrame after replacing placeholders with NaN:
     ord_no  purch_amt    ord_date  customer_id  salesman_id
0   70001.0     150.50         NaN       3002.0       5002.0
1       NaN     270.65  2012-09-10       3001.0       5003.0
2   70002.0      65.26         NaN       3001.0          NaN
3   70004.0     110.50  2012-08-17       3003.0       5001.0
4       NaN     948.50  2012-09-10       3002.0          NaN
5   70005.0    2400.60  2012-07-27       3001.0       5001.0
6       NaN    5760.00  2012-09-10       3001.0       5001.0
7   70010.0        NaN  2012-10-10       3004.0          NaN
8   70003.0      12.43  2012-10-10          NaN       5003.0
9   70012.0    2480.40  2012-06-27       3002.0       5002.0
10      NaN     250.45  2012-08-17       3001.0       5003.0
11  70013.0    3045.60  2012-04-25       3001.0          NaN

DataFrame after filling NaN values:
     ord_no    purch_amt    ord_date  customer_id  salesman_id
0   70001.0   150.500000     Unknown       3002.0       5002.0
1       0.0   270.650000  2012-09-10       3001.0       5003.0
2   70002.0    65.260000     Unknown       3001.0          0.0
3   70004.0   110.500000  2012-08-17       3003.0       5001.0
4       0.0   948.500000  2012-09-10       3002.0          0.0
5   70005.0  2400.600000  2012-07-27       3001.0       5001.0
6       0.0  5760.000000  2012-09-10       3001.0       5001.0
7   70010.0  1408.626364  2012-10-10       3004.0          0.0
8   70003.0    12.430000  2012-10-10          0.0       5003.0
9   70012.0  2480.400000  2012-06-27       3002.0       5002.0
10      0.0   250.450000  2012-08-17       3001.0       5003.0
11  70013.0  3045.600000  2012-04-25       3001.0          0.0
>>>
```

Ln: 54 Col: 0