

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

my_dataset = pd.read_csv('../csvs/dom_data_ecom.csv')
my_dataset
```

Out [ ]:

	State	City	Category	Customer_Login_type	Delivery_Type	Quantity	Transact S
0	State	New York City	Accessories	Member	one-day deliver	NaN	
1	State	Los Angles	Accessories	Member	one-day deliver	17.0	
2	State	Seattle	Accessories	Member	Normal Delivery	23.0	
3	State	Seattle	Accessories	Member	Normal Delivery	23.0	
4	State	Seattle	Accessories	Member	Normal Delivery	23.0	
...	...	...	...	...	...	...	
97	New York	New York City	Clothing	Member	one-day deliver	16.0	
98	New York	New York City	Clothing	Member	one-day deliver	15.0	
99	New York	New York City	Clothing	Member	one-day deliver	15.0	
100	New York	New York City	Clothing	Member	one-day deliver	15.0	
101	New York	New York City	Clothing	Member	one-day deliver	18.0	

102 rows × 12 columns



```
In [ ]: #2
print(my_dataset.describe())

print("Shape:", my_dataset.shape)

print(my_dataset.info())
```

```

    Quantity  Transaction Start  Transaction_Result
count 101.000000          102.0      102.000000
mean  17.732673           1.0        0.774510
std    3.255430           0.0        0.419968
min   12.000000           1.0        0.000000
25%   15.000000           1.0        1.000000
50%   19.000000           1.0        1.000000
75%   20.000000           1.0        1.000000
max   23.000000           1.0        1.000000
Shape: (102, 12)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 102 entries, 0 to 101
Data columns (total 12 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   State                                102 non-null    object
1   City                                102 non-null    object
2   Category                            102 non-null    object
3   Customer_Login_type                 102 non-null    object
4   Delivery_Type                       102 non-null    object
5   Quantity                            101 non-null    float64
6   Transaction Start                   102 non-null    int64
7   Transaction_Result                  102 non-null    int64
8   Amount US$                          102 non-null    object
9   Individual_Price_US$                102 non-null    object
10  Year_Month                          102 non-null    object
11  Time                                102 non-null    object
dtypes: float64(1), int64(2), object(9)
memory usage: 9.7+ KB
None
```

```
In [ ]: #3
my_dataset_delete = my_dataset.dropna()
my_dataset_delete
```

Out[ ]:

	State	City	Category	Customer_Login_type	Delivery_Type	Quantity	Transact St
1	State	Los Angles	Accessories	Member	one-day deliver	17.0	
2	State	Seattle	Accessories	Member	Normal Delivery	23.0	
3	State	Seattle	Accessories	Member	Normal Delivery	23.0	
4	State	Seattle	Accessories	Member	Normal Delivery	23.0	
5	State	Seattle	Accessories	Member	Normal Delivery	19.0	
...	...	...	...	...	...	...	
97	New York	New York City	Clothing	Member	one-day deliver	16.0	
98	New York	New York City	Clothing	Member	one-day deliver	15.0	
99	New York	New York City	Clothing	Member	one-day deliver	15.0	
100	New York	New York City	Clothing	Member	one-day deliver	15.0	
101	New York	New York City	Clothing	Member	one-day deliver	18.0	

101 rows × 12 columns



In [ ]:

```
my_dataset_replace = my_dataset.fillna(0)
print(my_dataset_replace)
```

	State	City	Category	Customer_Login_type	\
0	State	New York City	Accessories	Member	
1	State	Los Angeles	Accessories	Member	
2	State	Seattle	Accessories	Member	
3	State	Seattle	Accessories	Member	
4	State	Seattle	Accessories	Member	
..	...	...	...	...	
97	New York	New York City	Clothing	Member	
98	New York	New York City	Clothing	Member	
99	New York	New York City	Clothing	Member	
100	New York	New York City	Clothing	Member	
101	New York	New York City	Clothing	Member	

  

	Delivery_Type	Quantity	Transaction_Start	Transaction_Result	\
0	one-day deliver	0.0	1	0	
1	one-day deliver	17.0	1	1	
2	Normal Delivery	23.0	1	0	
3	Normal Delivery	23.0	1	0	
4	Normal Delivery	23.0	1	1	
..	...	...	...	...	
97	one-day deliver	16.0	1	1	
98	one-day deliver	15.0	1	1	
99	one-day deliver	15.0	1	1	
100	one-day deliver	15.0	1	1	
101	one-day deliver	18.0	1	1	

  

	Amount_US\$	Individual_Price_US\$	Year_Month	Time
0	6,910	576	13-Nov	22:35:51
1	1,699	100	13-Nov	6:44:41
2	4,998	217	13-Oct	0:41:24
3	736	32	13-Oct	22:04:03
4	4,389	191	13-Oct	15:00:46
..	...	...	...	...
97	6,915	432	13-Oct	7:04:38
98	4,845	323	13-Oct	9:34:20
99	1,898	127	13-Oct	13:45:21
100	799	53	13-Oct	10:44:01
101	16,563	920	13-Oct	16:57:02

[102 rows x 12 columns]

```
In [ ]: my_dataset_fill_forward = my_dataset.ffmpeg()
        print(my_dataset_fill_forward)
```

	State	City	Category	Customer_Login_type	\
0	State	New York City	Accessories	Member	
1	State	Los Angeles	Accessories	Member	
2	State	Seattle	Accessories	Member	
3	State	Seattle	Accessories	Member	
4	State	Seattle	Accessories	Member	
..	...	...	...	...	
97	New York	New York City	Clothing	Member	
98	New York	New York City	Clothing	Member	
99	New York	New York City	Clothing	Member	
100	New York	New York City	Clothing	Member	
101	New York	New York City	Clothing	Member	

	Delivery_Type	Quantity	Transaction_Start	Transaction_Result	\
0	one-day deliver	NaN	1	0	
1	one-day deliver	17.0	1	1	
2	Normal Delivery	23.0	1	0	
3	Normal Delivery	23.0	1	0	
4	Normal Delivery	23.0	1	1	
..	...	...	...	...	
97	one-day deliver	16.0	1	1	
98	one-day deliver	15.0	1	1	
99	one-day deliver	15.0	1	1	
100	one-day deliver	15.0	1	1	
101	one-day deliver	18.0	1	1	

	Amount_US\$	Individual_Price_US\$	Year_Month	Time
0	6,910	576	13-Nov	22:35:51
1	1,699	100	13-Nov	6:44:41
2	4,998	217	13-Oct	0:41:24
3	736	32	13-Oct	22:04:03
4	4,389	191	13-Oct	15:00:46
..	...	...	...	...
97	6,915	432	13-Oct	7:04:38
98	4,845	323	13-Oct	9:34:20
99	1,898	127	13-Oct	13:45:21
100	799	53	13-Oct	10:44:01
101	16,563	920	13-Oct	16:57:02

[102 rows x 12 columns]

```
In [ ]: my_dataset_fill_backward = my_dataset.bfill()
        print(my_dataset_fill_backward)
```

	State	City	Category	Customer_Login_type	\
0	State	New York City	Accessories	Member	
1	State	Los Angeles	Accessories	Member	
2	State	Seattle	Accessories	Member	
3	State	Seattle	Accessories	Member	
4	State	Seattle	Accessories	Member	
..	...	...	...	...	
97	New York	New York City	Clothing	Member	
98	New York	New York City	Clothing	Member	
99	New York	New York City	Clothing	Member	
100	New York	New York City	Clothing	Member	
101	New York	New York City	Clothing	Member	

  

	Delivery_Type	Quantity	Transaction_Start	Transaction_Result	\
0	one-day deliver	17.0	1	0	
1	one-day deliver	17.0	1	1	
2	Normal Delivery	23.0	1	0	
3	Normal Delivery	23.0	1	0	
4	Normal Delivery	23.0	1	1	
..	...	...	...	...	
97	one-day deliver	16.0	1	1	
98	one-day deliver	15.0	1	1	
99	one-day deliver	15.0	1	1	
100	one-day deliver	15.0	1	1	
101	one-day deliver	18.0	1	1	

  

	Amount_US\$	Individual_Price_US\$	Year_Month	Time
0	6,910	576	13-Nov	22:35:51
1	1,699	100	13-Nov	6:44:41
2	4,998	217	13-Oct	0:41:24
3	736	32	13-Oct	22:04:03
4	4,389	191	13-Oct	15:00:46
..	...	...	...	...
97	6,915	432	13-Oct	7:04:38
98	4,845	323	13-Oct	9:34:20
99	1,898	127	13-Oct	13:45:21
100	799	53	13-Oct	10:44:01
101	16,563	920	13-Oct	16:57:02

[102 rows x 12 columns]

```
In [ ]: #4
state_groups = my_dataset.groupby('State')['Quantity'].mean()
print(state_groups)
```

```
State
New York    17.259259
State       19.400000
Washington  17.575758
Name: Quantity, dtype: float64
```

```
In [ ]: #5
random_samples = my_dataset.sample(n=20)
hierarchical_df = random_samples.set_index(['State', 'City'])
print(hierarchical_df)
```

State	City	Category	Customer_Login_type	Delivery_Type	\
Washington	Seattle	Accessories	Member	one-day deliver	
New York	New York City	Clothing	Member	one-day deliver	
Washington	Seattle	Accessories	Member	one-day deliver	
New York	New York City	Clothing	Member	one-day deliver	
	New York City	Clothing	Member	one-day deliver	
	New York City	Clothing	Member	Normal Delivery	
	New York City	Clothing	Member	one-day deliver	
Washington	Seattle	Accessories	Member	one-day deliver	
New York	New York City	Clothing	Member	one-day deliver	
	New York City	Clothing	Member	one-day deliver	
	New York City	Clothing	Member	one-day deliver	
	New York City	Clothing	Member	one-day deliver	
	New York City	Clothing	Member	one-day deliver	
State	Seattle	Accessories	Member	Normal Delivery	
	Seattle	Accessories	Member	one-day deliver	
Washington	Seattle	Accessories	Member	Normal Delivery	
	Seattle	Accessories	Member	Normal Delivery	
	Seattle	Accessories	Member	one-day deliver	
New York	New York City	Clothing	Member	one-day deliver	
Washington	Seattle	Accessories	Member	one-day deliver	

State	City	Quantity	Transaction Start	Transaction_Result	\
Washington	Seattle	20.0	1	0	
New York	New York City	15.0	1	1	
Washington	Seattle	12.0	1	0	
New York	New York City	13.0	1	1	
	New York City	22.0	1	1	
	New York City	15.0	1	1	
	New York City	21.0	1	1	
Washington	Seattle	20.0	1	1	
New York	New York City	21.0	1	1	
	New York City	15.0	1	1	
	New York City	16.0	1	0	
	New York City	20.0	1	1	
	New York City	13.0	1	1	
State	Seattle	23.0	1	0	
	Seattle	20.0	1	1	
Washington	Seattle	15.0	1	1	
	Seattle	13.0	1	1	
	Seattle	14.0	1	1	
New York	New York City	16.0	1	1	
Washington	Seattle	20.0	1	0	

State	City	Amount US\$	Individual_Price_US\$	Year_Month	Time
Washington	Seattle	1,563	78	13-Nov	5:19:56
New York	New York City	1,898	127	13-Oct	13:45:21
Washington	Seattle	2,000	167	13-Nov	10:19:10
New York	New York City	300	23	13-Oct	23:39:23
	New York City	800	36	13-Oct	20:01:07
	New York City	6,900	460	13-Oct	10:29:59
	New York City	7,395	352	13-Oct	13:32:56
Washington	Seattle	25,635	1,282	13-Nov	9:17:48
New York	New York City	2,300	110	13-Oct	1:43:58
	New York City	5,391	359	13-Oct	5:12:54
	New York City	699	44	13-Oct	18:21:08
	New York City	584	29	13-Oct	6:05:15

	New York City	3,086	237	13-Oct	18:51:36
State	Seattle	736	32	13-Oct	22:04:03
	Seattle	0	0	13-Oct	13:48:08
Washington	Seattle	3,003	200	13-Nov	11:05:12
	Seattle	1,270	98	13-Nov	3:24:54
	Seattle	2,879	206	13-Nov	17:43:51
New York	New York City	1,999	125	13-Oct	3:50:06
Washington	Seattle	7,926	396	13-Nov	17:03:30