

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

In [2]: address = (r"Superstore_dataset.csv")
sales = pd.read_csv(address, encoding='latin1')
sales.head()
```

	Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	City	State	...	Product ID	Category	Sub-Category	Product Name	Sales	Quantity	Discount	Profit	Shipping Cost	Order Priority
0	42433	AG-2011-2040	1/1/2011	6/1/2011	Standard Class	TB-11280	Toby Braunhardt	Consumer	Constantine	Constantine	...	OFF-TEN-10000025	Office Supplies	Storage	Tenex Lockers, Blue	408.300	2	0.0	106.140	35.46	Medium
1	22253	IN-2011-47883	1/1/2011	8/1/2011	Standard Class	JH-15985	Joseph Holt	Consumer	Wagga Wagga	New South Wales	...	OFF-SU-10000618	Office Supplies	Supplies	Acme Trimmer, High Speed	120.366	3	0.1	36.036	9.72	Medium
2	48883	HU-2011-1220	1/1/2011	5/1/2011	Second Class	AT-735	Annie Thurman	Consumer	Budapest	Budapest	...	OFF-TEN-10001585	Office Supplies	Storage	Tenex Box, Single Width	66.120	4	0.0	29.640	8.17	High
3	11731	IT-2011-3647632	1/1/2011	5/1/2011	Second Class	EM-14140	Eugene Moren	Home Office	Stockholm	Stockholm	...	OFF-PA-10001492	Office Supplies	Paper	Enermax Note Cards, Premium	44.865	3	0.5	-26.055	4.82	High
4	22255	IN-2011-47883	1/1/2011	8/1/2011	Standard Class	JH-15985	Joseph Holt	Consumer	Wagga Wagga	New South Wales	...	FUR-FU-10003447	Furniture	Furnishings	Eldon Light Bulb, Duo Pack	113.670	5	0.1	37.770	4.70	Medium

5 rows × 24 columns

```
In [3]: #Looking number of rows and columns in the dataset
sales.shape
```

(51290, 24)

```
In [4]: #getting the most sales State
sales['State'].value_counts()
```

California 2001  
England 1499  
New York 1128  
Texas 985  
Ile-de-France 981  
...  
Arad 1  
Edirne 1  
Kastamonu 1  
Gangwon 1  
Inhambane 1  
Name: State, Length: 1094, dtype: int64

```
In [6]: #getting the Top ten sales in city
sales['City'].value_counts()
```

New York City 915  
Los Angeles 747  
Philadelphia 537  
San Francisco 510  
Santo Domingo 443  
...  
Melitopol' 1  
Terneuzen 1  
Athis-Mons 1  
Yuyao 1  
Deer Park 1  
Name: City, Length: 3636, dtype: int64

```
In [7]: #getting the most Purchasing customer
sales['Customer Name'].value_counts()
```

Muhammed Yedwab 108  
Steven Ward 106  
Bill Eplett 102  
Gary Hwang 102  
Patrick O'Brill 102  
...  
Andy Reiter 35  
David Bremer 34  
Darren Budd 31  
Nicole Brennan 31  
Michael Oakman 29  
Name: Customer Name, Length: 795, dtype: int64

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```