Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[6]
```
# Amazon sales analysis

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[9]
```
import numpy as np
import pandas as pd # Ensures Pandas is imported and aliased as 'pd'
import matplotlib.pyplot as plt
import seaborn as sns

# import dataset
df = pd.read_csv('/content/amazon_sales.csv - Sheet1.csv') # Now 'pd' is recognized
df
```

	index	Order ID	Date	Status	Fulfilment	Sales Channel	ship-service-	Category	Size	Courier Status	Qty	currency	Amount	ship-city	ship-state	ship-postal-	co

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[9]

		1101146		to Buyer												
2	2	404-0687676-7273146	4/30/2022	Shipped	Amazon	Amazon.in	Expedited	Shirt	XL	Shipped	1	INR	329.00	NAVI MUMBAI	MAHARASHTRA	410210.0
3	3	403-9615377-8133951	4/30/2022	Cancelled	Merchant	Amazon.in	Standard	Blazzer	L	On the Way	0	INR	753.33	PUDUCHERRY	PUDUCHERRY	605008.0
4	4	407-1069790-7240320	4/30/2022	Shipped	Amazon	Amazon.in	Expedited	Trousers	3XL	Shipped	1	INR	574.00	CHENNAI	TAMIL NADU	600073.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
128971	128970	406-6001380-7673107	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	Shirt	XL	Shipped	1	INR	517.00	HYDERABAD	TELANGANA	500013.0
128972	128971	402-9551604-7544318	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	T-shirt	M	Shipped	1	INR	999.00	GURUGRAM	HARYANA	122004.0
128973	128972	407-9547469-3152358	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	Blazzer	XXL	Shipped	1	INR	690.00	HYDERABAD	TELANGANA	500049.0
		402-														

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[21]
```
# Basic operations
df.shape
df.head()
```

	index	Order ID	Date	Status	Fulfilment	Sales Channel	ship-service-level	Category	Size	Courier Status	Qty	currency	Amount	ship-city	ship-state	ship-postal-code	ship-country
0	0	405-8078784-5731545	4/30/2022	Cancelled	Merchant	Amazon.in	Standard	T-shirt	S	On the Way	0	INR	647.62	MUMBAI	MAHARASHTRA	400081.0	IN
1	1	171-9198151-1101146	4/30/2022	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	Shirt	3XL	Shipped	1	INR	406.00	BENGALURU	KARNATAKA	560085.0	IN
2	2	404-0687676-7273146	4/30/2022	Shipped	Amazon	Amazon.in	Expedited	Shirt	XL	Shipped	1	INR	329.00	NAVI MUMBAI	MAHARASHTRA	410210.0	IN
3	3	403-9615377-	4/30/2022	Cancelled	Merchant	Amazon.in	Standard	Blazzer	L	On the Way	0	INR	753.33	PUDUCHERRY	PUDUCHERRY	605008.0	IN

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[11] df.tail()

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

	index	Order ID	Date	Status	Fulfilment	Sales Channel	ship-service-level	Category	Size	Courier Status	Qty	currency	Amount	ship-city	ship-state	ship-postal-code	ship countr
128971	128970	406-6001380-7673107	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	Shirt	XL	Shipped	1	INR	517.0	HYDERABAD	TELANGANA	500013.0	I
128972	128971	402-9551604-7544318	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	T-shirt	M	Shipped	1	INR	999.0	GURUGRAM	HARYANA	122004.0	I
128973	128972	407-9547469-3152358	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	Blazzer	XXL	Shipped	1	INR	690.0	HYDERABAD	TELANGANA	500049.0	I
128974	128973	402-6184140-0545956	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	T-shirt	XS	Shipped	1	INR	1199.0	Halol	Gujarat	389350.0	I
128975	128974	406-7436540-8728312	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	T-shirt	S	Shipped	1	INR	696.0	Raipur	CHHATTISGARH	492014.0	I

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[13] ---  ------              --------------  -----
 0   Index                 128976 non-null  int64
 1   Order ID              128976 non-null  object
 2   Date                  128976 non-null  object
 3   Status                128976 non-null  object
 4   Fulfilment            128976 non-null  object
 5   Sales Channel         128976 non-null  object
 6   ship-service-level    128976 non-null  object
 7   Category              128976 non-null  object
 8   Size                  128976 non-null  object
 9   Courier Status        128976 non-null  object
 10  Qty                   128976 non-null  int64
 11  currency              121176 non-null  object
 12  Amount                121176 non-null  float64
 13  ship-city             128941 non-null  object
 14  ship-state            128941 non-null  object
 15  ship-postal-code      128941 non-null  float64
 16  ship-country          128941 non-null  object
 17  B2B                   128976 non-null  bool
 18  fulfilled-by          39263 non-null   object
dtypes: bool(1), float64(2), int64(2), object(14)
memory usage: 17.8+ MB
```

```
# cheaking Null values
pd.isnull(df)
```

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

2	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True
3	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
128971	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True
128972	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True
128973	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True
128974	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True
128975	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	False	True

128976 rows × 19 columns

```
[20] # total null values
pd.isnull(df).sum()
```

	0
index	0

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[20]
```

Sales Channel	0
ship-service-level	0
Category	0
Size	0

Courier Status	0
Qty	0
currency	7800
Amount	7800
ship-city	35
ship-state	35
ship-postal-code	35
ship-country	35
B2B	0
fulfilled-by	89713

dtype: int64

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[22]

	index	Order ID	Date	Status	Fulfilment	Sales Channel	service-level	Category	Size	Courier Status	Qty	currency	Amount	ship-city	ship-state	postal-code	cou
0	0	405-8078784-5731545	4/30/2022	Cancelled	Merchant	Amazon.in	Standard	T-shirt	S	On the Way	0	INR	647.62	MUMBAI	MAHARASHTRA	400081.0	
1	1	171-9198151-1101146	4/30/2022	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	Shirt	3XL	Shipped	1	INR	406.00	BENGALURU	KARNATAKA	560085.0	
3	3	403-9615377-8133951	4/30/2022	Cancelled	Merchant	Amazon.in	Standard	Blazzer	L	On the Way	0	INR	753.33	PUDUCHERRY	PUDUCHERRY	605008.0	
7	7	406-7807733-3785945	4/30/2022	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	Shirt	S	Shipped	1	INR	399.00	HYDERABAD	TELANGANA	500032.0	
12	12	405-5513694-8146768	4/30/2022	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	Shirt	XS	Shipped	1	INR	399.00	Amravati.	MAHARASHTRA	444606.0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
128875	128874	405-4724097-1016369	6/1/2022	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	T-shirt	S	Shipped	1	INR	854.00	ALLUR	ANDHRA PRADESH	524315.0	

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

128891	128890	407-0116398-1810752	5/31/2022	Cancelled	Merchant	Amazon.in	Standard	Wallet	Free	On the Way	0	INR	398.10	MADURAI	TAMIL NADU	625007.0
128892	128891	403-0317423-9322704	5/31/2022	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	Blazzer	M	Shipped	1	INR	721.00	UTTAR BAGDOGRA	WEST BENGAL	734014.0

37514 rows × 19 columns

[23]
```
df.isnull().sum()
```

	0
index	0
Order ID	0
Date	0
Status	0
Fulfilment	0
Sales Channel	0

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[23]

Qty	0
currency	7800
Amount	7800
ship-city	35
ship-state	35
ship-postal-code	35
ship-country	35
B2B	0
fulfilled-by	89713

dtype: int64

[32] Start coding or generate with AI.

[33]
```
# change data type
df['ship-postal-code'] = pd.to_numeric(df['ship-postal-code'], errors='coerce').astype('Int64')
df
```

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[33]

1	1	171-9198151-1101146	4/30/2022	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	Shirt	3XL	Shipped	1	INR	406.00	BENGALURU	KARNATAKA	560085
2	2	404-0687676-7273146	4/30/2022	Shipped	Amazon	Amazon.in	Expedited	Shirt	XL	Shipped	1	INR	329.00	NAVI MUMBAI	MAHARASHTRA	410210
3	3	403-9515377-8133951	4/30/2022	Cancelled	Merchant	Amazon.in	Standard	Blazzer	L	On the Way	0	INR	753.33	PUDUCHERRY	PUDUCHERRY	605008
4	4	407-1069790-7240320	4/30/2022	Shipped	Amazon	Amazon.in	Expedited	Trousers	3XL	Shipped	1	INR	574.00	CHENNAI	TAMIL NADU	600073
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
128971	128970	406-6001380-7673107	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	Shirt	XL	Shipped	1	INR	517.00	HYDERABAD	TELANGANA	500013
128972	128971	402-9551604-7544318	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	T-shirt	M	Shipped	1	INR	999.00	GURUGRAM	HARYANA	122004
128973	128972	407-9547469-	5/31/2022	Shipped	Amazon	Amazon.in	Expedited	Blazzer	XXL	Shipped	1	INR	690.00	HYDERABAD	TELANGANA	500049

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[33]

		8728312

128976 rows x 19 columns

[34]
```
#cheak data type
df['ship-postal-code'].dtype
```

```
Int64Dtype()
```

[39]
```
df['Date'] = pd.to_datetime(df['Date'])
df
```

	index	Order ID	Date	Status	Fulfilment	Sales Channel	ship-service-level	Category	Size	Courier Status	Qty	currency	Amount	ship-city	ship-state	ship-postal-code	ship-country
0	0	405-8078784-5731545	2022-04-30	Cancelled	Merchant	Amazon.in	Standard	T-shirt	S	On the Way	0	INR	647.62	MUMBAI	MAHARASHTRA	400081	IN
1	1	171-9198151-1101146	2022-04-30	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	Shirt	3XL	Shipped	1	INR	406.00	BENGALURU	KARNATAKA	560085	IN

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[39]

4	4	407-1069790-7240320	2022-04-30	Shipped	Amazon	Amazon.in	Expedited	Trousers	3XL	Shipped	1	INR	574.00	CHENNAI	TAMIL NADU	600073	IN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
128971	128970	406-6001380-7673107	2022-05-31	Shipped	Amazon	Amazon.in	Expedited	Shirt	XL	Shipped	1	INR	517.00	HYDERABAD	TELANGANA	500013	IN
128972	128971	402-9551604-7544318	2022-05-31	Shipped	Amazon	Amazon.in	Expedited	T-shirt	M	Shipped	1	INR	999.00	GURUGRAM	HARYANA	122004	IN
128973	128972	407-9547469-3152358	2022-05-31	Shipped	Amazon	Amazon.in	Expedited	Blazzer	XXL	Shipped	1	INR	690.00	HYDERABAD	TELANGANA	500049	IN
128974	128973	402-6184140-0545956	2022-05-31	Shipped	Amazon	Amazon.in	Expedited	T-shirt	XS	Shipped	1	INR	1199.00	Halol	Gujarat	389350	IN

128975	128974	408-7436540-8728312	2022-05-31	Shipped	Amazon	Amazon.in	Expedited	T-shirt	S	Shipped	1	INR	696.00	Raipur	CHHATTISGARH	492014	IN

128976 rows × 19 columns

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[40]

count	128976.000000	128976	128976.000000	121176.000000	128941.0
mean	64486.130427	2022-05-12 11:49:26.951991040	0.904401	648.562176	463945.677744
min	0.000000	2022-03-31 00:00:00	0.000000	0.000000	110001.0
25%	32242.750000	2022-04-20 00:00:00	1.000000	449.000000	382421.0
50%	64486.500000	2022-05-10 00:00:00	1.000000	605.000000	500033.0
75%	96730.250000	2022-06-04 00:00:00	1.000000	788.000000	600024.0
max	128974.000000	2022-06-29 00:00:00	15.000000	5584.000000	989898.0
std	37232.897832	NaN	0.313368	281.185041	191458.488954

[41]
```
# describe data with object datatype
df.describe(include='object')
```

	Order ID	Status	Fulfilment	Sales Channel	ship-service-level	Category	Size	Courier Status	currency	ship-city	ship-state	ship-country	fulfilled-by
count	128976	128976	128976	128976	128976	128976	128976	128976	121176	128941	128941	128941	39263
unique	120229	13	2	2	2	9	11	4	1	8948	69	1	1

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
#describe for specific column
df[['Qty','Amount']].describe()
```

	Qty	Amount
count	128976.000000	121176.000000
mean	0.904401	648.562176
std	0.313368	281.185041
min	0.000000	0.000000
25%	1.000000	449.000000
50%	1.000000	605.000000
75%	1.000000	788.000000
max	15.000000	5584.000000

[ ] Start coding or generate with AI.

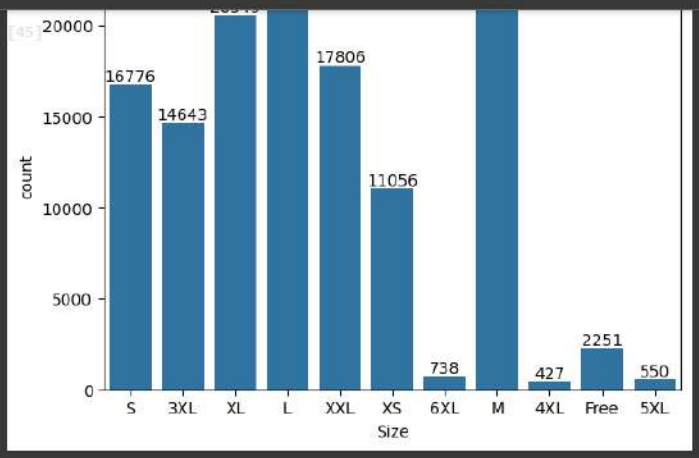
[45]
```
#cheaking for Data labels
ax=sns.countplot(x='Size',data=df)
```

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text



[ ] Start coding or generate with AI.

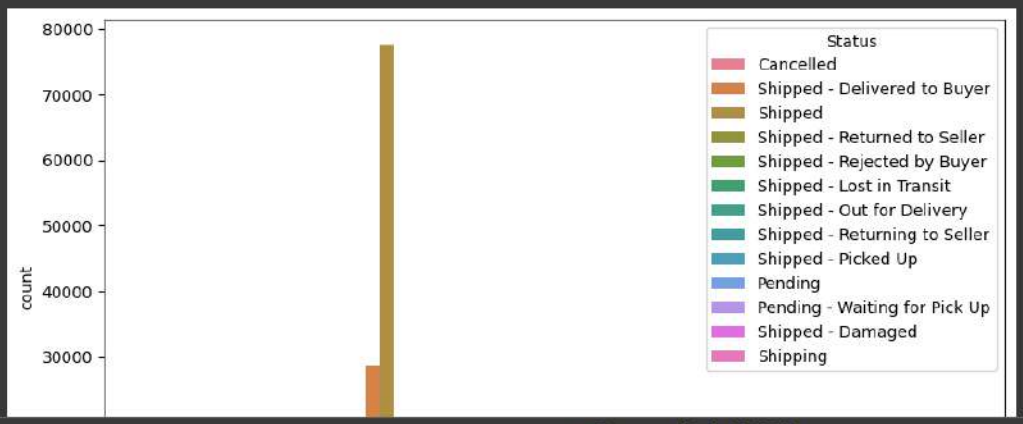
✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[52] ax=sns.countplot(x='Courier Status',hue='Status',data=df)
     plt.show(
     )
```



✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Majority of the order ship through the courier

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

	index	Order ID	Date	Status	Fulfilment	Sales Channel	ship-service-level	Category	Size	Courier Status	Qty	currency	Amount	ship-city	ship-state	ship-postal-code	ship-country	B2B
0	0	405-8078784-5731545	2022-04-30	Cancelled	Merchant	Amazon.in	Standard	T-shirt	S	On the Way	0	INR	647.62	MUMBAI	MAHARASHTRA	400081	IN	False
1	1	171-9198151-1101146	2022-04-30	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard	Shirt	3XL	Shipped	1	INR	406.00	BENGALURU	KARNATAKA	560085	IN	False
2	2	404-0687676-7273146	2022-04-30	Shipped	Amazon	Amazon.in	Expedited	Shirt	XL	Shipped	1	INR	329.00	NAVI MUMBAI	MAHARASHTRA	410210	IN	True

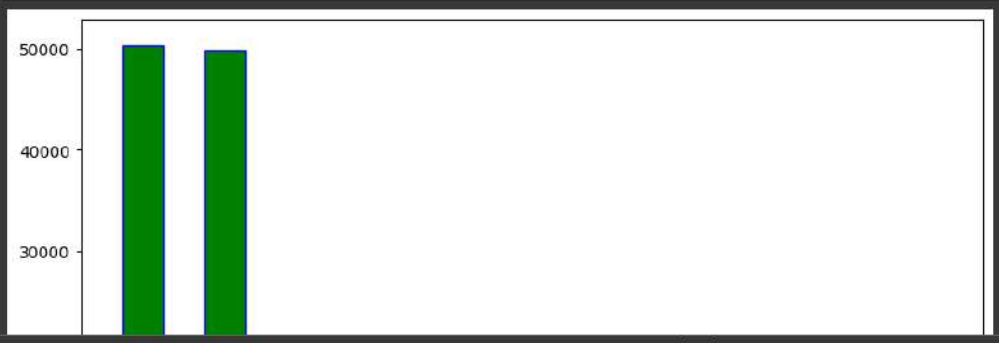
✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[64] df['category'] = df['category'].astype('str')
     c_d=df['Category']
     plt.figure(figsize=(10,6))
     plt.hist(c_d,bins=20,edgecolor='blue',color='g')
     plt.xticks(rotation=90)
     plt.show()
```

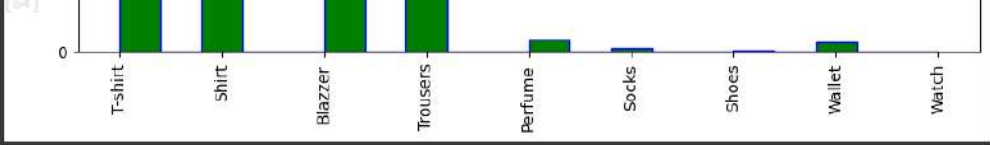


✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text




```
[ ] Start coding or generate with AI.
```

Most of the buyers are t-shirt

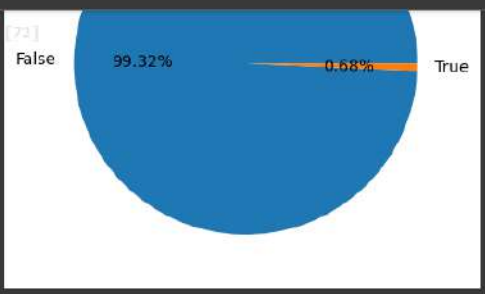
```
[72] # cheak b2b data
     B2b=df['B2B'].value_counts()
     B2b

     # create pie chart
     plt.pie(B2b,labels=B2b.index,autopct='%1.2f%%')
     plt.show()
```

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text



max 99.32% buyers are retailers and 0.68% are wholesalers

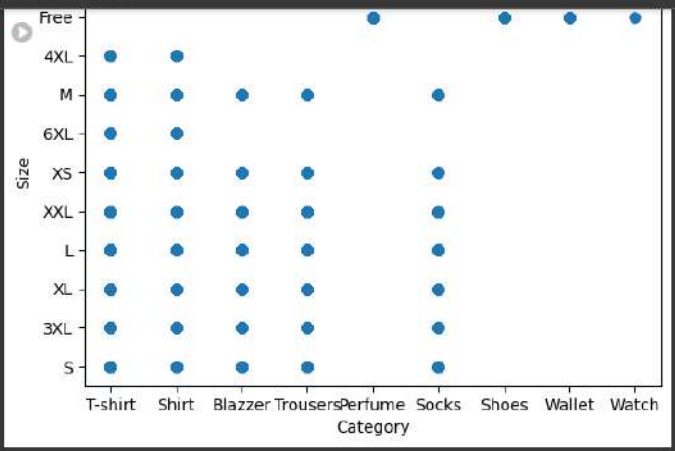
```
# scatter plot
x_data=df['Category']
y_data=df['Size']
plt.scatter(x_data,y_data)
plt.xlabel('Category')
plt.ylabel('Size')
```

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text



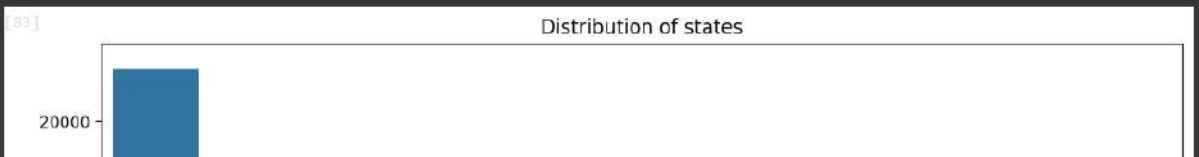

```
[83] top10_states=df['ship-state'].value_counts().head(10)
     top10_states
     plt.figure(figsize=(12,6))
```

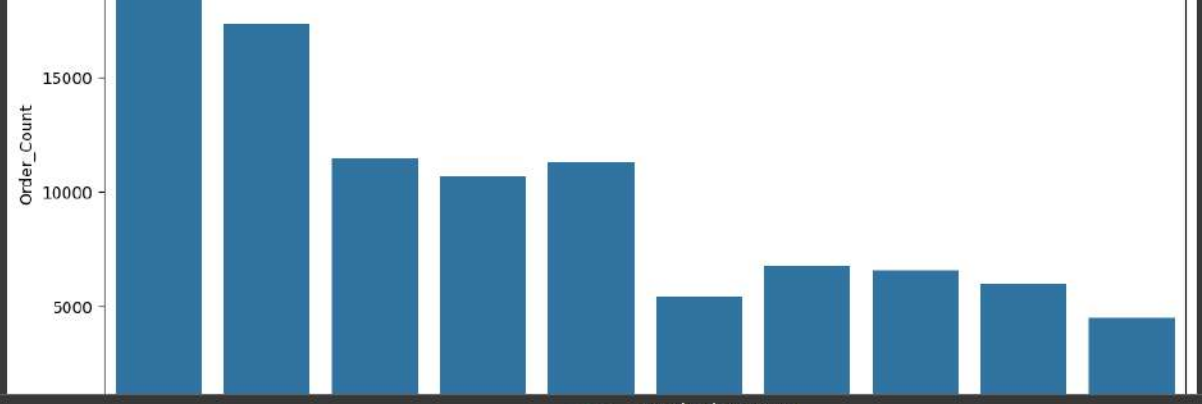
✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text





✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Most of the buyers are Manarastra state

[86]
```
from google.colab import files
files.download("Untitled45.ipynb.png")
```

```
---------------------------------------------------------------------------
FileNotFoundError                         Traceback (most recent call last)
<ipython-input-86-929eb0a00c7d> in <cell line: 2>()
      1 from google.colab import files
----> 2 files.download("Untitled45.ipynb.png")

/usr/local/lib/python3.10/dist-packages/google/colab/files.py in download(filename)
    231   if not _os.path.exists(filename):
    232     msg = 'Cannot find file: {}'.format(filename)
--> 233     raise FileNotFoundError(msg)  # pylint: disable=undefined-variable
    234
    235   comm_manager = _IPython.get_ipython().kernel.comm_manager

FileNotFoundError: Cannot find file: Untitled45.ipynb.png
```

Next steps: Explain error

[87] Start coding or generate with AI.

✓ 0s completed at 5:19 PM

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

[ ] Start coding or generate with AI.

Untitled45.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text