

```

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

order_report=pd.read_excel(r'C:\Users\RAHUL\Desktop\Cointab Data Analyst - Challenge\Company X - Order Report.xlsx')
pin_cod_zon=pd.read_excel(r'C:\Users\RAHUL\Desktop\Cointab Data Analyst - Challenge\Company X - Pincode Zones.xlsx')
suv_master=pd.read_excel(r'C:\Users\RAHUL\Desktop\Cointab Data Analyst - Challenge\Company X - SKU Master.xlsx')
courier_invoice=pd.read_excel(r'C:\Users\RAHUL\Desktop\Cointab Data Analyst - Challenge\Courier Company - Invoice.xlsx')
courier_Rate=pd.read_excel(r'C:\Users\RAHUL\Desktop\Cointab Data Analyst - Challenge\Courier Company - Rates.xlsx')
Expected_Result=pd.read_excel(r'C:\Users\RAHUL\Desktop\Cointab Data Analyst - Challenge\Expected_Result.xlsx')

```

```
order_report.sample(5)
```

	ExternOrderNo	SKU	Order Qty
330	2001806735	8904223818850	1.0
249	2001807970	8904223819321	1.0
32	2001821190	8904223819338	1.0
397	2001806226	8904223818850	2.0
73	2001813009	8904223818706	1.0

```
order_report.isna().sum()
```

```

ExternOrderNo    0
SKU              0
Order Qty        0
dtype: int64

```

```
courier_invoice.sample(5)
```

	AWB Code	Order ID	Charged Weight	Warehouse Pincode \
30	1091121981575	2001825261	1.60	121003
75	1091119367193	2001816131	0.70	121003
8	1091117323812	2001807186	0.50	121003
94	1091117229183	2001806968	0.68	121003
19	1091117436383	2001808507	0.79	121003

	Customer Pincode Zone	Type of Shipment	Billing Amount (Rs.)
30	517128 d	Forward and RT0 charges	345.0
75	302017 d	Forward charges	90.2
8	396001 d	Forward charges	45.4

94	305801	d	Forward charges
90.2			
19	208002	b	Forward charges
61.3			

```
suv_master.sample(5)
```

	SKU	Weight (g)
45	8904223819239	290
7	8904223817334	170
59	8904223819321	600
21	8904223815804	160
31	8904223818980	110

```
Total_weight_KG=suv_master['Weight (g)']/1000
```

```
Total_weight_KG.head()
```

0	0.210
1	0.165
2	0.113
3	0.065
4	0.120

```
Name: Weight (g), dtype: float64
```

```
#Total weight as per Courier Company (KG)
```

```
Total_weight_Courier_Company_kg=Total_weight_KG
```

```
Total_weight_KG.isna().sum()
```

```
0
```

```
suv_master['Weight (g)']=suv_master['Weight (g)'].apply(lambda x: 0.5
if x<=500 else 1)
```

```
# Create a new column for the weight slab
```

```
suv_master['weight_slab'] = suv_master['Weight (g)'].replace('Weight
(g)', 'weight_slab')
```

```
suv_master_new=suv_master.drop(['Weight (g)'],axis=1)
```

```
suv_master_new.head()
```

	SKU	weight_slab
0	8904223815682	0.5
1	8904223815859	0.5
2	8904223815866	0.5
3	8904223815873	0.5
4	8904223816214	0.5

```
pin_cod_zon.head()
```

	Warehouse	Pincode	Customer	Pincode	Zone
0		121003		507101	d
1		121003		486886	d
2		121003		532484	d
3		121003		143001	b
4		121003		515591	d

company Zone

```
pin_cod_zon['Zone'].unique()
```

```
array(['d', 'b', 'e'], dtype=object)
```

#courier Zone

```
courier_invoice['Zone'].unique()
```

```
array(['d', 'b', 'e'], dtype=object)
```

courier_Rate

	fwd_a_fixed	fwd_a_additional	fwd_b_fixed	fwd_b_additional
fwd_c_fixed \				
0	29.5	23.6	33	28.3
40.1				

	fwd_c_additional	fwd_d_fixed	fwd_d_additional	fwd_e_fixed	\
0	38.9	45.4	44.8	56.6	

	fwd_e_additional	rto_a_fixed	rto_a_additional	rto_b_fixed	\
0	55.5	13.6	23.6	20.5	

	rto_b_additional	rto_c_fixed	rto_c_additional	rto_d_fixed	\
0	28.3	31.9	38.9	41.3	

	rto_d_additional	rto_e_fixed	rto_e_additional
0	44.8	50.7	55.5

```
Total_charge_a=courier_Rate['fwd_a_fixed']
+courier_Rate['fwd_a_additional']
```

Total_charge_a

```
0    53.1
dtype: float64
```

```
Total_charge_b=courier_Rate['fwd_b_fixed']
+courier_Rate['fwd_b_additional']
```

```
print(Total_charge_b)
```

```
0    61.3
dtype: float64
```

```
Total_charge_c=courier_Rate['fwd_c_fixed']  
+courier_Rate['fwd_c_additional']  
print(Total_charge_c)
```

```
0    79.0  
dtype: float64
```

```
Total_charge_d=courier_Rate['fwd_d_fixed']  
+courier_Rate['fwd_d_additional']  
print(Total_charge_d)
```

```
0    90.2  
dtype: float64
```

```
Total_rto_a=courier_Rate['rto_a_fixed']  
+courier_Rate['rto_a_additional']  
print(Total_rto_a)
```

```
0    37.2  
dtype: float64
```

```
Total_rto_b=courier_Rate['rto_b_fixed']  
+courier_Rate['rto_b_additional']  
print(Total_rto_b)
```

```
0    48.8  
dtype: float64
```

```
Total_rto_c=courier_Rate['rto_c_fixed']  
+courier_Rate['rto_c_additional']  
print(Total_rto_c)
```

```
0    70.8  
dtype: float64
```

```
Total_rto_d=courier_Rate['rto_d_fixed']  
+courier_Rate['rto_d_additional']  
print(Total_rto_d)
```

```
0    86.1  
dtype: float64
```

```
Total_rto_e=courier_Rate['rto_e_fixed']  
+courier_Rate['rto_e_additional']  
print(Total_rto_e)
```

```
0    106.2  
dtype: float64
```

```
courier_invoice['Zone'].unique()
```

```
array(['d', 'b', 'e'], dtype=object)
```

```
courier_invoice.shape
```

```
(124, 8)
```

```
order_report.rename(columns={'ExternOrderNo': 'Order ID'}, inplace=True)
```

```
order_report.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 400 entries, 0 to 399
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype
---  -
 0   Order ID    400 non-null   int64
 1   SKU         400 non-null   object
 2   Order Qty   400 non-null   float64
dtypes: float64(1), int64(1), object(1)
memory usage: 9.5+ KB
```

```
order_report.describe()
```

	Order ID	Order Qty
count	4.000000e+02	400.000000
mean	2.001811e+09	1.297500
std	5.187698e+03	0.919842
min	2.001806e+09	1.000000
25%	2.001807e+09	1.000000
50%	2.001809e+09	1.000000
75%	2.001812e+09	1.000000
max	2.001827e+09	8.000000

```
order_report.isnull().sum()
```

```
Order ID    0
SKU         0
Order Qty   0
dtype: int64
```

```
order_report.duplicated().sum()
```

```
2
```

```
order_report.shape
```

```
(400, 3)
```

```
order_report.columns
```

```
Index(['Order ID', 'SKU', 'Order Qty'], dtype='object')
```

```
order_report.dtypes
```

```
Order ID    int64
SKU         object
```

```
Order Qty    float64
dtype: object
```

```
order_report.nunique()
```

```
Order ID      124
SKU           65
Order Qty      6
dtype: int64
```

```
order_report.groupby('Order ID').size()
```

```
Order ID
2001806210    2
2001806226    1
2001806229    3
2001806232    2
2001806233    2
..
2001821995    2
2001822466    2
2001823564    3
2001825261    5
2001827036    8
Length: 124, dtype: int64
```

```
order_report.groupby('Order ID').size().sort_values(ascending=False)
```

```
Order ID
2001809820    14
2001827036     8
2001821185     8
2001817093     8
2001812854     8
..
2001820690     1
2001807012     1
2001806226     1
2001806768     1
2001807328     1
Length: 124, dtype: int64
```

```
order_report.groupby('Order ID').size().sort_values(ascending=False).head(10)
```

```
Order ID
2001809820    14
2001827036     8
2001821185     8
2001817093     8
2001812854     8
2001806471     7
```

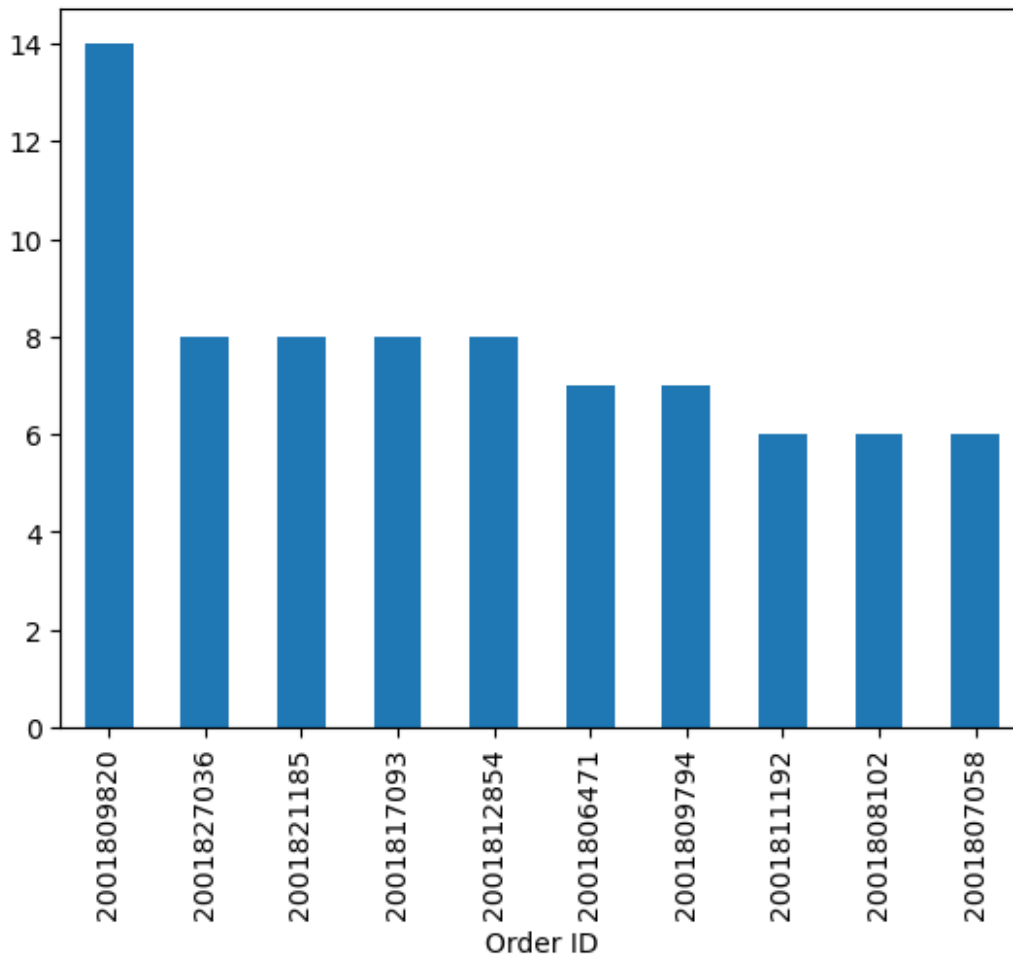
```
2001809794      7
2001811192      6
2001808102      6
2001807058      6
dtype: int64
```

```
order_report.groupby('Order
ID').size().sort_values(ascending=False).tail(10)
```

```
Order ID
2001821742      1
2001814580      1
2001807930      1
2001806823      1
2001806686      1
2001820690      1
2001807012      1
2001806226      1
2001806768      1
2001807328      1
dtype: int64
```

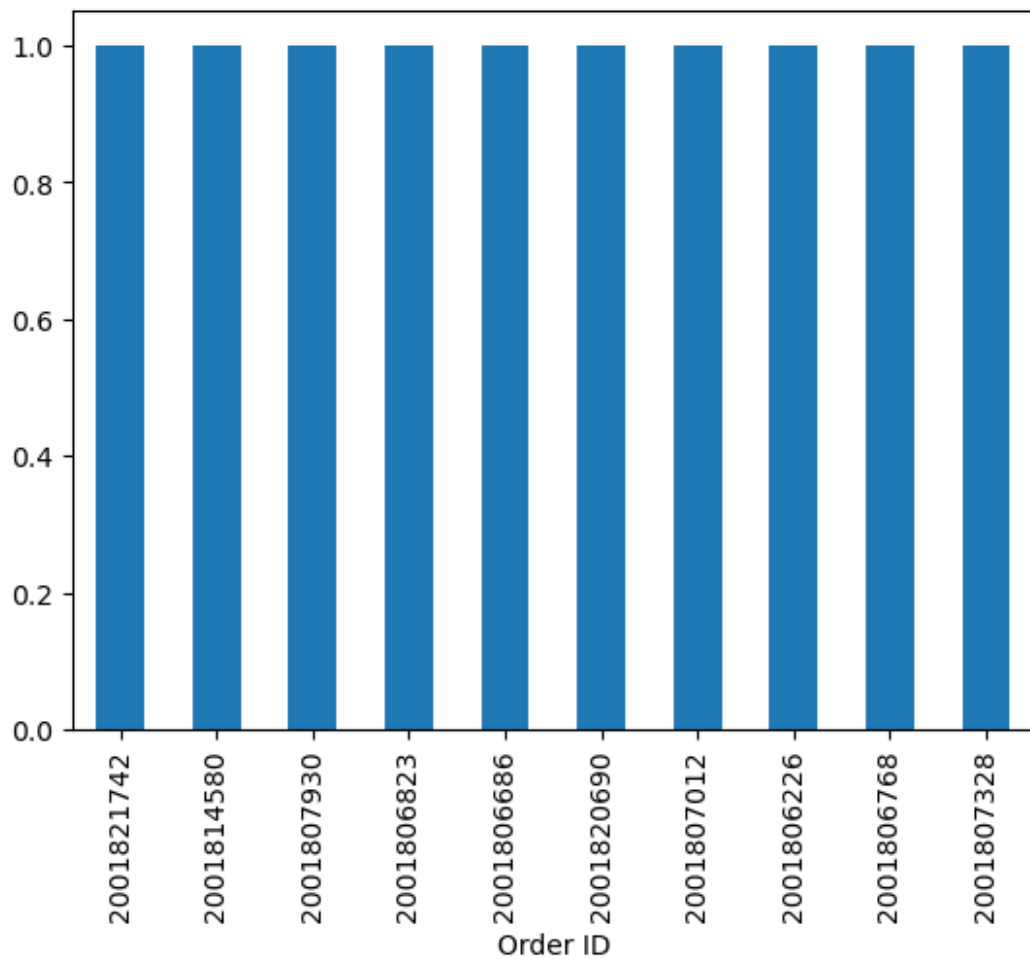
```
order_report.groupby('Order
ID').size().sort_values(ascending=False).head(10).plot(kind='bar')
```

```
<AxesSubplot:xlabel='Order ID'>
```

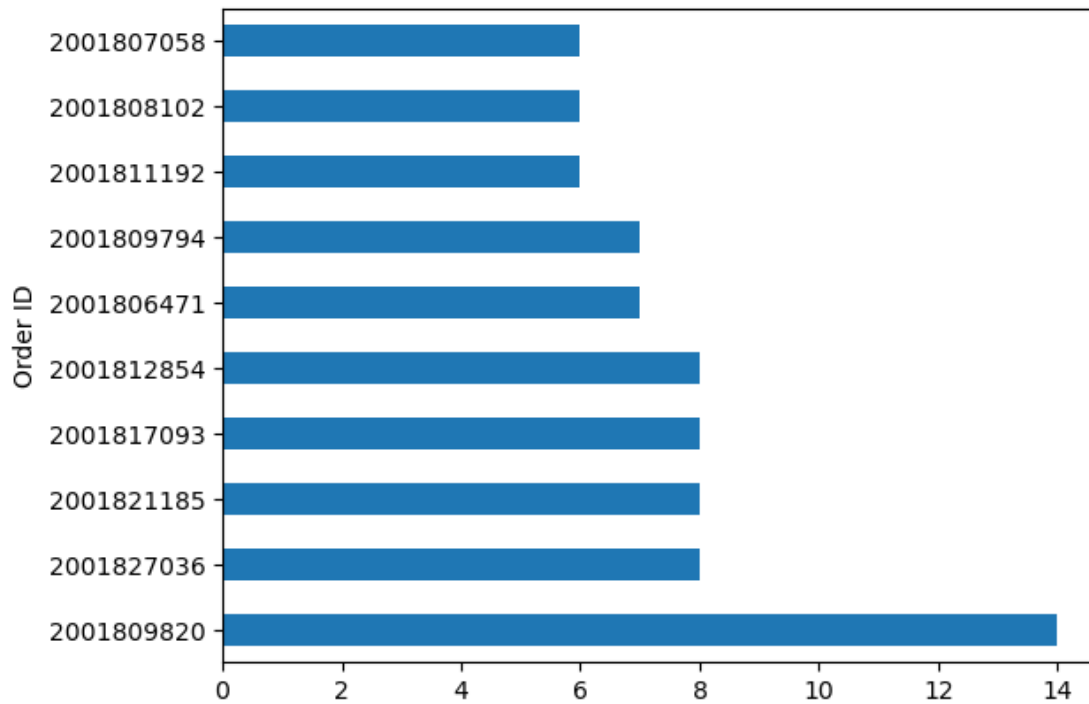


```
order_report.groupby('Order ID').size().sort_values(ascending=False).tail(10).plot(kind='bar')
```

```
<AxesSubplot:xlabel='Order ID'>
```

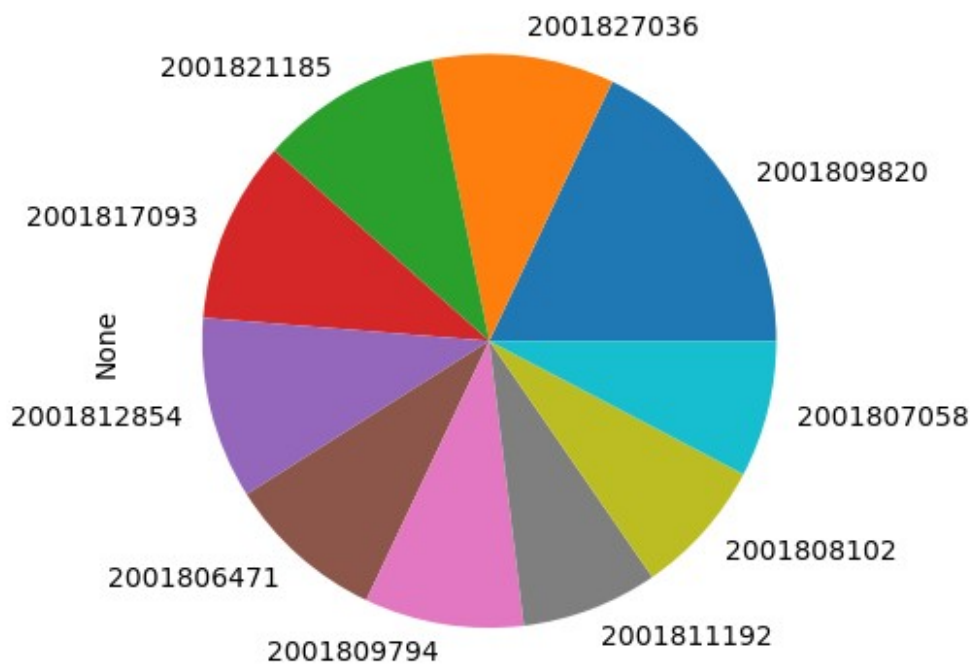



```
order_report.groupby('Order ID').size().sort_values(ascending=False).head(10).plot(kind='barh')  
<AxesSubplot:ylabel='Order ID'>
```



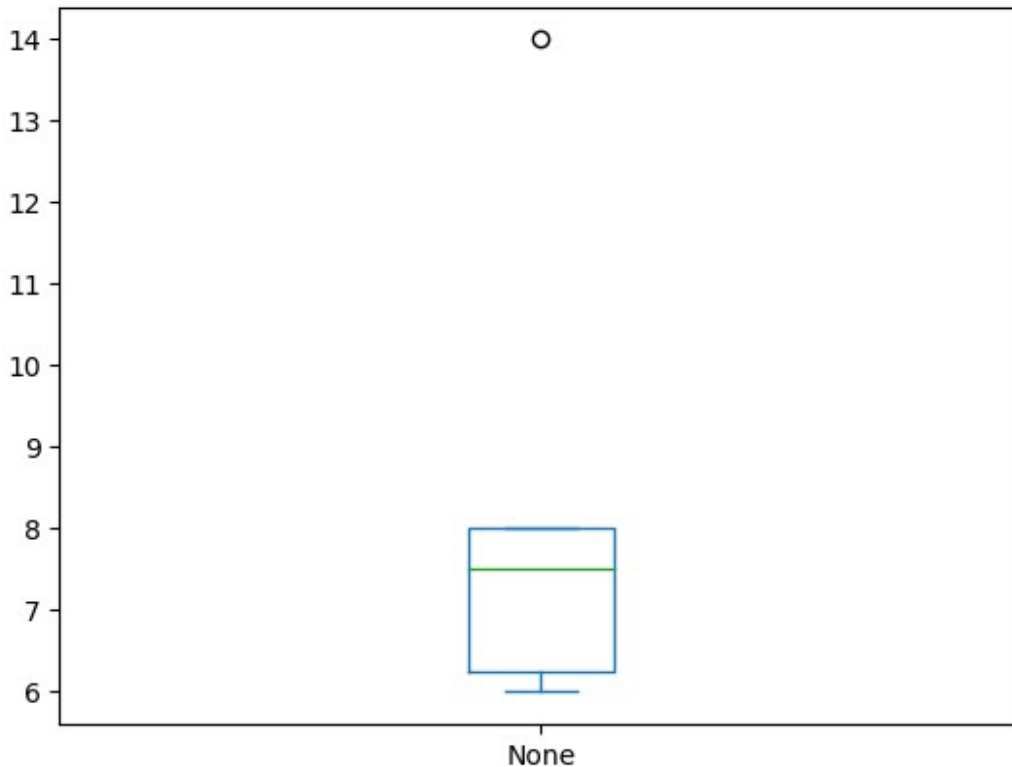
```
order_report.groupby('Order ID').size().sort_values(ascending=False).head(10).plot(kind='pie')
```

<AxesSubplot:ylabel='None'>



```
order_report.groupby('Order ID').size().sort_values(ascending=False).head(10).plot(kind='box')
```

<AxesSubplot:>



```
courier_invoice.sample(5)
```

	AWB Code	Order ID	Charged Weight	Warehouse Pincode \
92	1091117228133	2001806823	0.59	121003
82	1091121034641	2001821766	0.20	121003
90	1091117222360	2001806304	0.71	121003
49	1091117227573	2001806776	2.86	121003
113	1091117806263	2001810549	1.86	121003

	Customer Pincode	Zone	Type of Shipment	Billing Amount (Rs.)
92	314001	d	Forward charges	90.2
82	313001	d	Forward charges	45.4
90	302017	d	Forward charges	90.2
49	226004	b	Forward charges	174.5
113	302017	d	Forward charges	179.8

```
courier_invoice_group=courier_invoice.groupby(['Zone', 'Type of Shipment'])
```

```
courier_invoice_group_weight=courier_invoice_group['Charged Weight'].sum()
```

```
#creat data set for total Charged_weight
courier_invoice_group_weight
```

```
Zone  Type of Shipment
b     Forward and RT0 charges    1.90
      Forward charges           24.60
d     Forward and RT0 charges    8.77
      Forward charges           82.28
e     Forward and RT0 charges    1.00
Name: Charged Weight, dtype: float64
```

```
#creat data set for total price
```

```
courier_invoice_group_price=courier_invoice_group['Charged
Weight'].sum()
```

```
#reset index
```

```
courier_invoice_groupby=courier_invoice.reset_index()
courier_invoice_groupby.head()
```

	index	AWB Code	Order ID	Charged Weight	Warehouse Pincode
\					
0	0	1091117222124	2001806232	1.30	121003
1	1	1091117222194	2001806273	1.00	121003
2	2	1091117222931	2001806408	2.50	121003
3	3	1091117223244	2001806458	1.00	121003
4	4	1091117229345	2001807012	0.15	121003

	Customer Pincode	Zone	Type of Shipment	Billing Amount (Rs.)
0	507101	d	Forward charges	135.0
1	486886	d	Forward charges	90.2
2	532484	d	Forward charges	224.6
3	143001	b	Forward charges	61.3
4	515591	d	Forward charges	45.4

```
courier_invoice.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 124 entries, 0 to 123
```

```
Data columns (total 8 columns):
```

#	Column	Non-Null Count	Dtype
0	AWB Code	124 non-null	int64
1	Order ID	124 non-null	int64
2	Charged Weight	124 non-null	float64
3	Warehouse Pincode	124 non-null	int64

```

4    Customer Pincode      124 non-null    int64
5    Zone                  124 non-null    object
6    Type of Shipment      124 non-null    object
7    Billing Amount (Rs.)   124 non-null    float64
dtypes: float64(2), int64(4), object(2)
memory usage: 7.9+ KB

```

```
courier_invoice.describe()
```

	AWB Code	Order ID	Charged Weight	Warehouse
Pincode \				
count	1.240000e+02	1.240000e+02	124.000000	124.0
mean	1.091118e+12	2.001811e+09	0.956048	121003.0
std	1.473661e+06	5.167329e+03	0.662815	0.0
min	1.091117e+12	2.001806e+09	0.150000	121003.0
25%	1.091117e+12	2.001807e+09	0.667500	121003.0
50%	1.091117e+12	2.001809e+09	0.725000	121003.0
75%	1.091119e+12	2.001812e+09	1.100000	121003.0
max	1.091122e+12	2.001827e+09	4.130000	121003.0

	Customer Pincode	Billing Amount (Rs.)
count	124.000000	124.000000
mean	365488.072581	110.066129
std	152156.322130	64.060832
min	140301.000000	33.000000
25%	302017.000000	86.700000
50%	321304.500000	90.200000
75%	405102.250000	135.000000
max	845438.000000	403.800000

```
courier_invoice.isnull().sum()
```

```

AWB Code      0
Order ID      0
Charged Weight 0
Warehouse Pincode 0
Customer Pincode 0
Zone          0
Type of Shipment 0
Billing Amount (Rs.) 0
dtype: int64

```

```

courier_invoice.columns
Index(['AWB Code', 'Order ID', 'Charged Weight', 'Warehouse Pincode',
      'Customer Pincode', 'Zone', 'Type of Shipment', 'Billing Amount
      (Rs.)'],
      dtype='object')

courier_invoice.shape
(124, 8)

courier_invoice.dtypes
AWB Code                int64
Order ID                int64
Charged Weight          float64
Warehouse Pincode       int64
Customer Pincode        int64
Zone                    object
Type of Shipment        object
Billing Amount (Rs.)    float64
dtype: object

courier_invoice.nunique()
AWB Code                124
Order ID                124
Charged Weight          54
Warehouse Pincode       1
Customer Pincode        108
Zone                    3
Type of Shipment        2
Billing Amount (Rs.)    20
dtype: int64

courier_invoice.groupby('Type of Shipment').count()

Type of Shipment      AWB Code  Order ID  Charged Weight  \
Forward and RT0 charges      15        15             15
Forward charges             109        109             109

Type of Shipment      Warehouse Pincode  Customer Pincode  Zone  \
Forward and RT0 charges              15             15      15
Forward charges                    109             109     109

Type of Shipment      Billing Amount (Rs.)
Forward and RT0 charges              15
Forward charges                    109

```

```
courier_invoice.groupby('Zone').count()
```

AWB Code	Order ID	Charged Weight	Warehouse Pincode	Customer Pincode \ Zone
----------	----------	----------------	-------------------	-------------------------

b	24	24	24	24
d	98	98	98	98
e	2	2	2	2

Type of Shipment	Billing Amount (Rs.)
Zone	
b	24
d	98
e	2

```
courier_invoice.groupby('Type of Shipment').sum()
```

AWB Code	Order ID	Charged Weight
\ Type of Shipment		
Forward and RT0 charges	16366792436826	30027225332
Forward charges	118931867272952	218197307648

Warehouse Pincode	Customer Pincode \
Type of Shipment	
Forward and RT0 charges	1815045
Forward charges	13189327

Billing Amount (Rs.)
Type of Shipment
Forward and RT0 charges
Forward charges

```
courier_invoice.groupby('Zone').sum()
```

AWB Code	Order ID	Charged Weight	Warehouse Pincode
\ Zone			
b	26186834344767	48043432119	26.50
d	106929587914889	196177474176	91.05

e	2182237450122	4003626685	1.00	242006
---	---------------	------------	------	--------

	Customer Pincode	Billing Amount (Rs.)
Zone		
b	4978452	1969.9
d	38885766	11357.5
e	1456303	320.8

courier_invoice.groupby('Type of Shipment').mean()

	AWB Code	Order ID	Charged Weight \
Type of Shipment			
Forward and RT0 charges	1.091119e+12	2.001815e+09	0.77800
Forward charges	1.091118e+12	2.001810e+09	0.98055

	Warehouse Pincode	Customer Pincode \
Type of Shipment		
Forward and RT0 charges	121003.0	466300.800000
Forward charges	121003.0	351614.761468

	Billing Amount (Rs.)
Type of Shipment	
Forward and RT0 charges	170.740000
Forward charges	101.716514

courier_invoice.groupby('Zone').mean()

	AWB Code	Order ID	Charged Weight	Warehouse Pincode \
Zone				
b	1.091118e+12	2.001810e+09	1.104167	121003.0
d	1.091118e+12	2.001811e+09	0.929082	121003.0
e	1.091119e+12	2.001813e+09	0.500000	121003.0

	Customer Pincode	Billing Amount (Rs.)
Zone		
b	207435.500000	82.079167
d	396793.530612	115.892857
e	728151.500000	160.400000

courier_invoice.groupby('Type of Shipment').median()

	AWB Code	Order ID	Charged Weight \
Type of Shipment			
Forward and RT0 charges	1.091119e+12	2.001815e+09	0.70
Forward charges	1.091117e+12	2.001808e+09	0.73

	Warehouse Pincode	Customer Pincode \
Type of Shipment		
Forward and RT0 charges	121003.0	411014.0
Forward charges	121003.0	313333.0

	Billing Amount (Rs.)
Type of Shipment	
Forward and RT0 charges	172.8
Forward charges	90.2

```
courier_invoice.groupby('Type of Shipment').min()
```

Type of Shipment	AWB Code	Order ID	Charged Weight \
Forward and RT0 charges	1091117327496	2001807976	0.15
Forward charges	1091117221940	2001806210	0.15

Type of Shipment	Warehouse Pincode	Customer Pincode	Zone \
Forward and RT0 charges	121003	244001	b
Forward charges	121003	140301	b

	Billing Amount (Rs.)
Type of Shipment	
Forward and RT0 charges	86.7
Forward charges	33.0

#Merging the dataframes

```
mergel=pd.merge(order_report,courier_invoice,how='left',on='Order ID')
```

```
mergel
```

	Order ID	SKU	Order Qty	AWB Code	Charged Weight \
0	2001827036	8904223818706	1.0	1091122418320	1.60
1	2001827036	8904223819093	1.0	1091122418320	1.60
2	2001827036	8904223819109	1.0	1091122418320	1.60
3	2001827036	8904223818430	1.0	1091122418320	1.60
4	2001827036	8904223819277	1.0	1091122418320	1.60
..
..					
395	2001806229	8904223818942	1.0	1091117222080	0.71
396	2001806229	8904223818850	1.0	1091117222080	0.71
397	2001806226	8904223818850	2.0	1091117222065	0.68
398	2001806210	8904223816214	1.0	1091117221940	2.92
399	2001806210	8904223818874	1.0	1091117221940	

2.92

	Warehouse	Pincode	Customer	Pincode	Zone	Type of Shipment	\
0		121003		173213	b	Forward charges	
1		121003		173213	b	Forward charges	
2		121003		173213	b	Forward charges	
3		121003		173213	b	Forward charges	
4		121003		173213	b	Forward charges	
...		
395		121003		421204	d	Forward charges	
396		121003		421204	d	Forward charges	
397		121003		723146	d	Forward charges	
398		121003		140604	b	Forward charges	
399		121003		140604	b	Forward charges	

	Billing Amount (Rs.)
0	117.9
1	117.9
2	117.9
3	117.9
4	117.9
...	...
395	90.2
396	90.2
397	90.2
398	174.5
399	174.5

[400 rows x 10 columns]

#Merging the dataframes

```
merge2=pd.merge(pin_cod_zon,courier_invoice,how='left',on='Zone')
```

merge2

	Warehouse	Pincode_x	Customer	Pincode_x	Zone	AWB Code
Order ID \						
0		121003		507101	d	1091117222124
2001806232						
1		121003		507101	d	1091117222194
2001806273						
2		121003		507101	d	1091117222931
2001806408						
3		121003		507101	d	1091117229345
2001807012						
4		121003		507101	d	1091117229555
2001806686						
...	
...						
5629		121003		302020	b	1091118004245

2001811604				
5630	121003	302020	b	1091120352712
2001819252				
5631	121003	302020	b	1091122418320
2001827036				
5632	121003	302020	b	1091117436652
2001808585				
5633	121003	302020	b	1091121844806
2001811475				

	Charged Weight	Warehouse Pincode_y	Customer Pincode_y	\
0	1.30	121003	507101	
1	1.00	121003	486886	
2	2.50	121003	532484	
3	0.15	121003	515591	
4	0.15	121003	326502	
...	
5629	0.80	121003	173212	
5630	0.30	121003	174101	
5631	1.60	121003	173213	
5632	0.72	121003	175101	
5633	0.50	121003	173212	

	Type of Shipment	Billing Amount (Rs.)
0	Forward charges	135.0
1	Forward charges	90.2
2	Forward charges	224.6
3	Forward charges	45.4
4	Forward charges	45.4
...
5629	Forward charges	61.3
5630	Forward charges	33.0
5631	Forward charges	117.9
5632	Forward charges	61.3
5633	Forward charges	33.0

[5634 rows x 10 columns]

#Merging the dataframes

```
merge3=pd.merge(order_report,suv_master,how='left',on='SKU')
```

merge3

	Order ID	SKU	Order Qty	Weight (g)	weight_slab
0	2001827036	8904223818706	1.0	0.5	0.5
1	2001827036	8904223819093	1.0	0.5	0.5
2	2001827036	8904223819109	1.0	0.5	0.5
3	2001827036	8904223818430	1.0	0.5	0.5
4	2001827036	8904223819277	1.0	0.5	0.5
..

396	2001806229	8904223818942	1.0	0.5	0.5
397	2001806229	8904223818850	1.0	0.5	0.5
398	2001806226	8904223818850	2.0	0.5	0.5
399	2001806210	8904223816214	1.0	0.5	0.5
400	2001806210	8904223818874	1.0	0.5	0.5

[401 rows x 5 columns]

courier_invoice.sample(2)

	AWB Code	Order ID	Charged Weight	Warehouse	Pincode \
115	1091117904860	2001811039	0.68		121003
8	1091117323812	2001807186	0.50		121003

	Customer Pincode	Zone	Type of Shipment	Billing Amount (Rs.)
115	302020	d	Forward charges	90.2
8	396001	d	Forward charges	45.4

courier_Rate

	fwd_a_fixed	fwd_a_additional	fwd_b_fixed	fwd_b_additional
fwd_c_fixed \				
0	29.5	23.6	33	28.3
40.1				

	fwd_c_additional	fwd_d_fixed	fwd_d_additional	fwd_e_fixed \
0	38.9	45.4	44.8	56.6

	fwd_e_additional	rto_a_fixed	rto_a_additional	rto_b_fixed \
0	55.5	13.6	23.6	20.5

	rto_b_additional	rto_c_fixed	rto_c_additional	rto_d_fixed \
0	28.3	31.9	38.9	41.3

	rto_d_additional	rto_e_fixed	rto_e_additional
0	44.8	50.7	55.5

courier_invoice['Type of Shipment'].unique()

array(['Forward charges', 'Forward and RT0 charges'], dtype=object)

#creat data set ('Forward charges':29.5,'Forward and RT0 charges':43.1).groupby(courier_invoice['Zone'])

#Expected Charge as per X (Rs.)

courier_invoice['Zone_d'] = courier_invoice['Type of Shipment'].map({'Forward charges':45.4,'Forward and RT0 charges':86.7})

courier_invoice['Zone_b'] = courier_invoice['Type of Shipment'].map({'Forward charges':33,'Forward and RT0 charges':53.5})

courier_invoice['Zone_e'] = courier_invoice['Type of

```
Shipment'].map({'Forward charges':56.6,'Forward and RT0
charges':107.3}))
```

#Delivery Zone charged by Courier Company

```
courier_invoice['Zone_d_Courier'] = courier_invoice['Type of
Shipment'].map({'Forward charges':89.8,'Forward and RT0
charges':134.6})
courier_invoice['Zone_b_Courier'] = courier_invoice['Type of
Shipment'].map({'Forward charges':61.3,'Forward and RT0
charges':89.6})
courier_invoice['Zone_e_Courier'] = courier_invoice['Type of
Shipment'].map({'Forward charges':112.1,'Forward and RT0
charges':167.6})
```

courier_invoice

	AWB Code	Order ID	Charged Weight	Warehouse Pincode \
0	1091117222124	2001806232	1.30	121003
1	1091117222194	2001806273	1.00	121003
2	1091117222931	2001806408	2.50	121003
3	1091117223244	2001806458	1.00	121003
4	1091117229345	2001807012	0.15	121003
...
119	1091118551656	2001812941	0.73	121003
120	1091117614452	2001809383	0.50	121003
121	1091120922803	2001820978	0.50	121003
122	1091121844806	2001811475	0.50	121003
123	1091121846136	2001811305	0.50	121003

(Rs.) \	Customer Pincode	Zone	Type of Shipment	Billing Amount
0	507101	d	Forward charges	
135.0				
1	486886	d	Forward charges	
90.2				
2	532484	d	Forward charges	
224.6				
3	143001	b	Forward charges	
61.3				
4	515591	d	Forward charges	
45.4				
...	
...				
119	325207	d	Forward charges	
90.2				
120	303702	d	Forward and RT0 charges	
86.7				
121	313301	d	Forward charges	
45.4				
122	173212	b	Forward charges	
33.0				

```
123          302020      d          Forward charges
45.4
```

	Zone_d	Zone_b	Zone_e	Zone_d_Courier	Zone_b_Courier
Zone_e_Courier					
0	45.4	33.0	56.6	89.8	61.3
112.1					
1	45.4	33.0	56.6	89.8	61.3
112.1					
2	45.4	33.0	56.6	89.8	61.3
112.1					
3	45.4	33.0	56.6	89.8	61.3
112.1					
4	45.4	33.0	56.6	89.8	61.3
112.1					
...
...					
119	45.4	33.0	56.6	89.8	61.3
112.1					
120	86.7	53.5	107.3	134.6	89.6
167.6					
121	45.4	33.0	56.6	89.8	61.3
112.1					
122	45.4	33.0	56.6	89.8	61.3
112.1					
123	45.4	33.0	56.6	89.8	61.3
112.1					

```
[124 rows x 14 columns]
```

```
courier_invoice.isna().sum()
```

```
AWB Code          0
Order ID          0
Charged Weight    0
Warehouse Pincode 0
Customer Pincode  0
Zone              0
Type of Shipment  0
Billing Amount (Rs.) 0
Zone_d            0
Zone_b            0
Zone_e            0
Zone_d_Courier    0
Zone_b_Courier    0
Zone_e_Courier    0
dtype: int64
```

```
courier_invoice.drop(['Zone', 'Type of Shipment'],axis=1,inplace=True)
```

```
courier_invoice.dtypes
```

```

AWB Code                int64
Order ID                int64
Charged Weight          float64
Warehouse Pincode       int64
Customer Pincode        int64
Billing Amount (Rs.)    float64
Zone_d                  float64
Zone_b                  float64
Zone_e                  float64
Zone_d_Courier          float64
Zone_b_Courier          float64
Zone_e_Courier          float64
dtype: object

```

#Expected Charge as per X (Rs.)

```

courier_invoice['Expected_Charge_as_per_X']=courier_invoice['Zone_d']
+courier_invoice['Zone_b']+courier_invoice['Zone_e']

```

courier_invoice

	AWB Code	Order ID	Charged Weight	Warehouse Pincode	\
0	1091117222124	2001806232	1.30	121003	
1	1091117222194	2001806273	1.00	121003	
2	1091117222931	2001806408	2.50	121003	
3	1091117223244	2001806458	1.00	121003	
4	1091117229345	2001807012	0.15	121003	
...
119	1091118551656	2001812941	0.73	121003	
120	1091117614452	2001809383	0.50	121003	
121	1091120922803	2001820978	0.50	121003	
122	1091121844806	2001811475	0.50	121003	
123	1091121846136	2001811305	0.50	121003	

	Customer Pincode	Billing Amount (Rs.)	Zone_d	Zone_b	Zone_e	\
0	507101	135.0	45.4	33.0	56.6	
1	486886	90.2	45.4	33.0	56.6	
2	532484	224.6	45.4	33.0	56.6	
3	143001	61.3	45.4	33.0	56.6	
4	515591	45.4	45.4	33.0	56.6	
...
119	325207	90.2	45.4	33.0	56.6	
120	303702	86.7	86.7	53.5	107.3	
121	313301	45.4	45.4	33.0	56.6	
122	173212	33.0	45.4	33.0	56.6	
123	302020	45.4	45.4	33.0	56.6	

	Zone_d_Courier	Zone_b_Courier	Zone_e_Courier
Expected_Charge_as_per_X			
0	89.8	61.3	112.1
135.0			
1	89.8	61.3	112.1

135.0			
2	89.8	61.3	112.1
135.0			
3	89.8	61.3	112.1
135.0			
4	89.8	61.3	112.1
135.0			
..
...			
119	89.8	61.3	112.1
135.0			
120	134.6	89.6	167.6
247.5			
121	89.8	61.3	112.1
135.0			
122	89.8	61.3	112.1
135.0			
123	89.8	61.3	112.1
135.0			

[124 rows x 13 columns]

#Charges Billed by Courier Company (Rs.)

courier_invoice['Charges Billed by Courier Company

(Rs.)']=courier_invoice['Zone_d_Courier']

+courier_invoice['Zone_b_Courier']+courier_invoice['Zone_e_Courier']

courier_invoice.head()

	AWB Code	Order ID	Charged Weight	Warehouse	Pincode	\
0	1091117222124	2001806232	1.30		121003	
1	1091117222194	2001806273	1.00		121003	
2	1091117222931	2001806408	2.50		121003	
3	1091117223244	2001806458	1.00		121003	
4	1091117229345	2001807012	0.15		121003	

	Customer	Pincode	Billing Amount (Rs.)	Zone_d	Zone_b	Zone_e	\
0		507101	135.0	45.4	33.0	56.6	
1		486886	90.2	45.4	33.0	56.6	
2		532484	224.6	45.4	33.0	56.6	
3		143001	61.3	45.4	33.0	56.6	
4		515591	45.4	45.4	33.0	56.6	

	Zone_d_Courier	Zone_b_Courier	Zone_e_Courier
Expected_Charge_as_per_X \			
0	89.8	61.3	112.1
135.0			
1	89.8	61.3	112.1
135.0			
2	89.8	61.3	112.1

135.0			
3	89.8	61.3	112.1
135.0			
4	89.8	61.3	112.1
135.0			

Charges Billed by Courier Company (Rs.)	
0	263.2
1	263.2
2	263.2
3	263.2
4	263.2

```
#Difference Between Expected Charges and Billed Charges (Rs.)
courier_invoice['Difference Between Expected Charges and Billed
Charges (Rs.)']=courier_invoice['Expected_Charge_as_per_X']-
courier_invoice['Billing Amount (Rs.)']
```

```
courier_invoice.sample(5)
```

	AWB Code	Order ID	Charged Weight	Warehouse Pincode	\
42	1091117224902	2001806567	1.16	121003	
91	1091117227116	2001806768	1.02	121003	
19	1091117436383	2001808507	0.79	121003	
31	1091117957780	2001811192	1.13	121003	
59	1091117223351	2001806471	1.70	121003	

	Customer Pincode	Billing Amount (Rs.)	Zone_d	Zone_b	Zone_e	\
42	370201	135.0	45.4	33.0	56.6	
91	322201	135.0	45.4	33.0	56.6	
19	208002	61.3	45.4	33.0	56.6	
31	562110	258.9	86.7	53.5	107.3	
59	313027	179.8	45.4	33.0	56.6	

	Zone_d_Courier	Zone_b_Courier	Zone_e_Courier
Expected_Charge_as_per_X	\		
42	89.8	61.3	112.1
135.0			
91	89.8	61.3	112.1
135.0			
19	89.8	61.3	112.1
135.0			
31	134.6	89.6	167.6
247.5			
59	89.8	61.3	112.1
135.0			

Charges Billed by Courier Company (Rs.)		\
42	263.2	
91	263.2	

19	263.2
31	391.8
59	263.2

	Difference Between Expected Charges and Billed Charges (Rs.)
42	0.0
91	0.0
19	73.7
31	-11.4
59	-44.8

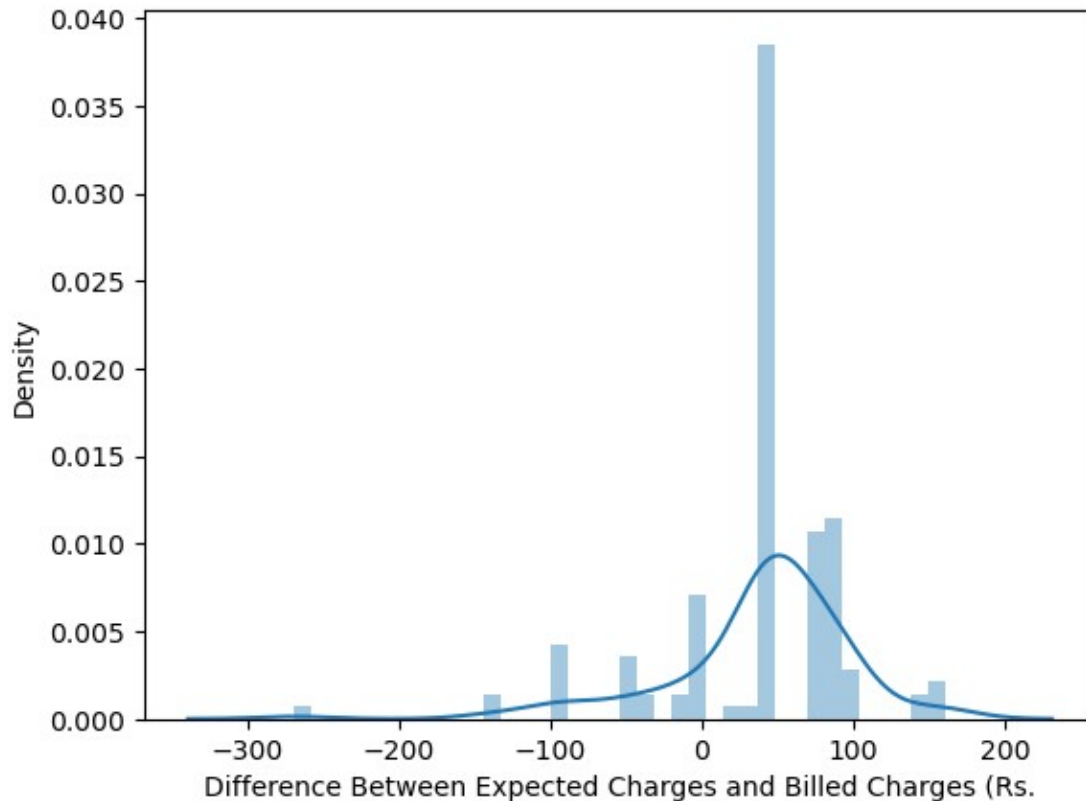
```
courier_invoice.columns
```

```
Index(['AWB Code', 'Order ID', 'Charged Weight', 'Warehouse Pincode',
      'Customer Pincode', 'Billing Amount (Rs.)', 'Zone_d', 'Zone_b',
      'Zone_e', 'Zone_d_Courier', 'Zone_b_Courier', 'Zone_e_Courier',
      'Expected_Charge_as_per_X', 'Charges Billed by Courier Company
      (Rs.)',
      'Difference Between Expected Charges and Billed Charges (Rs.)',
      dtype='object')
```

```
sns.distplot(courier_invoice['Difference Between Expected Charges and
Billed Charges (Rs.)'])
```

```
C:\Users\RAHUL\anaconda3\lib\site-packages\seaborn\
distributions.py:2619: FutureWarning: `distplot` is a deprecated
function and will be removed in a future version. Please adapt your
code to use either `displot` (a figure-level function with similar
flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

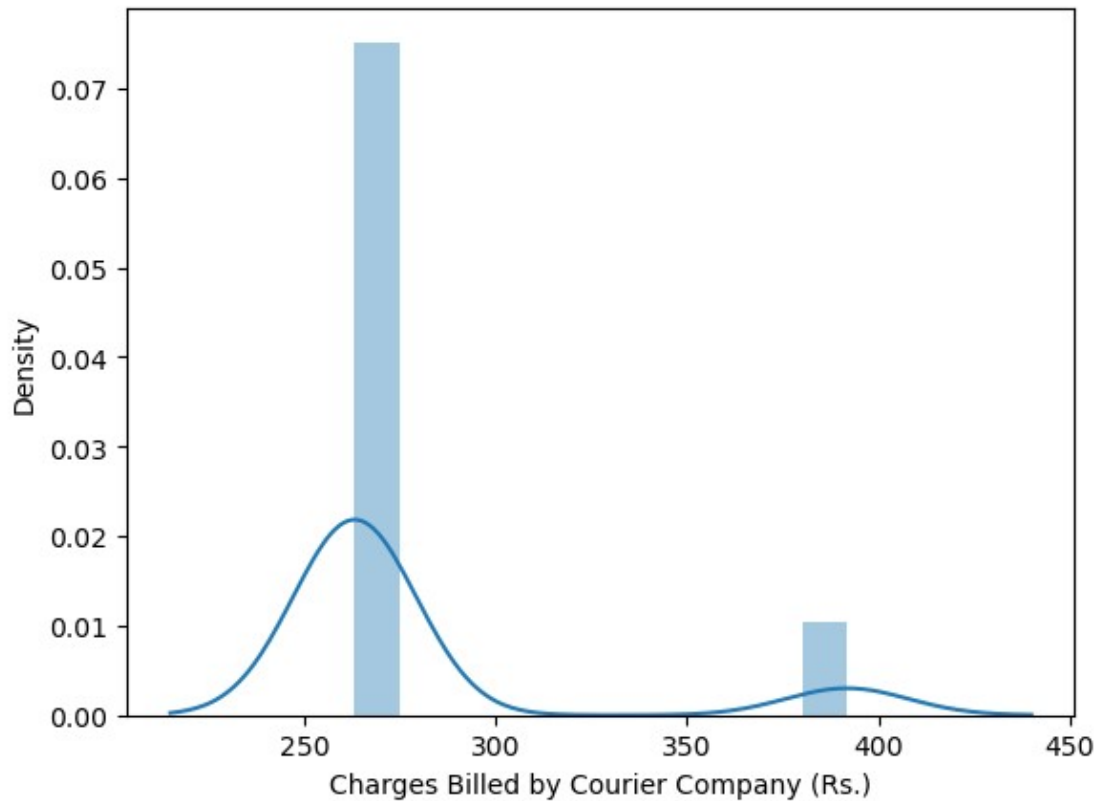
```
<AxesSubplot:xlabel='Difference Between Expected Charges and Billed
Charges (Rs.', ylabel='Density'>
```



```
sns.distplot(courier_invoice['Charges Billed by Courier Company (Rs.)'])
```

```
C:\Users\RAHUL\anaconda3\lib\site-packages\seaborn\
distributions.py:2619: FutureWarning: `distplot` is a deprecated
function and will be removed in a future version. Please adapt your
code to use either `displot` (a figure-level function with similar
flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

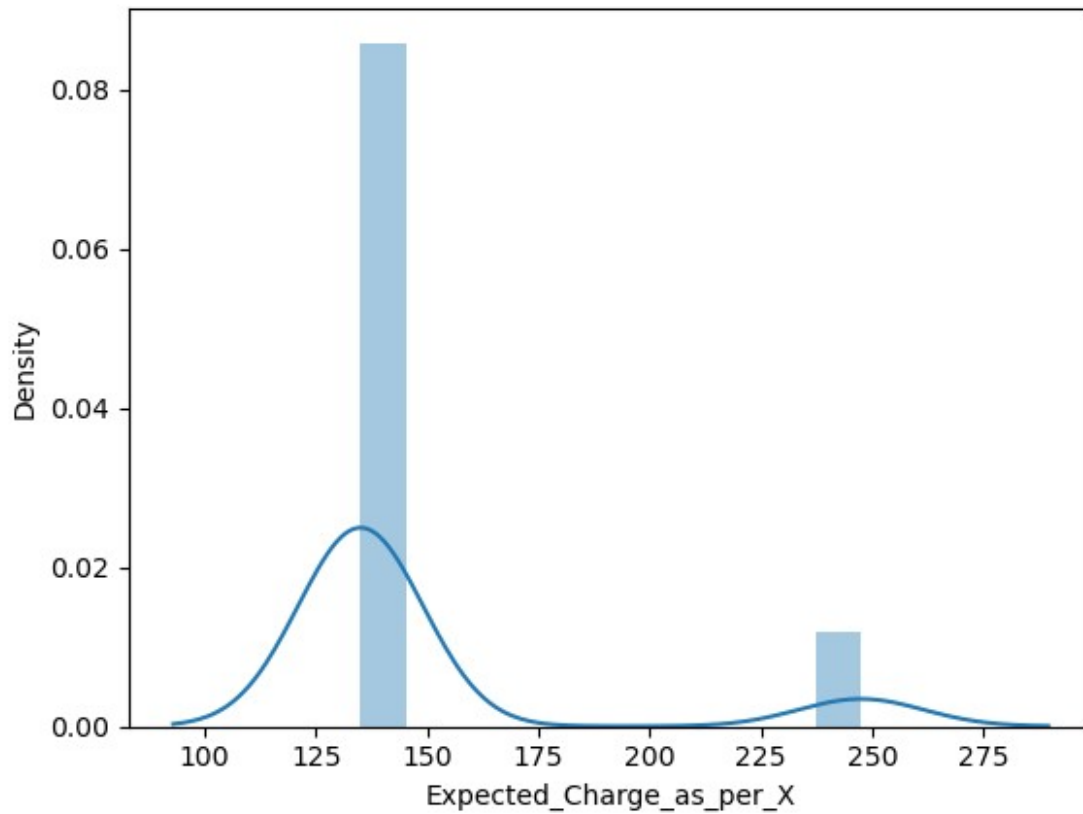
```
<AxesSubplot:xlabel='Charges Billed by Courier Company (Rs.)',
ylabel='Density'>
```



```
sns.distplot(courier_invoice['Expected_Charge_as_per_X'])
```

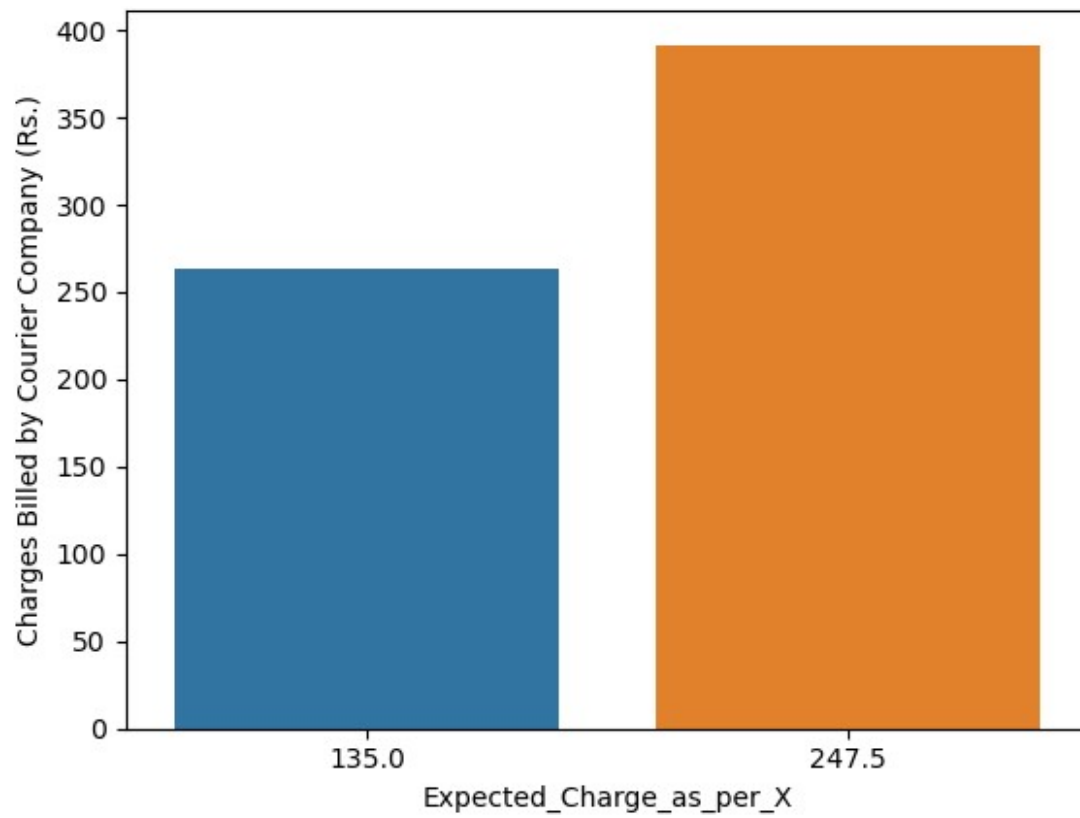
```
C:\Users\RAHUL\anaconda3\lib\site-packages\seaborn\
distributions.py:2619: FutureWarning: `distplot` is a deprecated
function and will be removed in a future version. Please adapt your
code to use either `displot` (a figure-level function with similar
flexibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

```
<AxesSubplot:xlabel='Expected_Charge_as_per_X', ylabel='Density'>
```



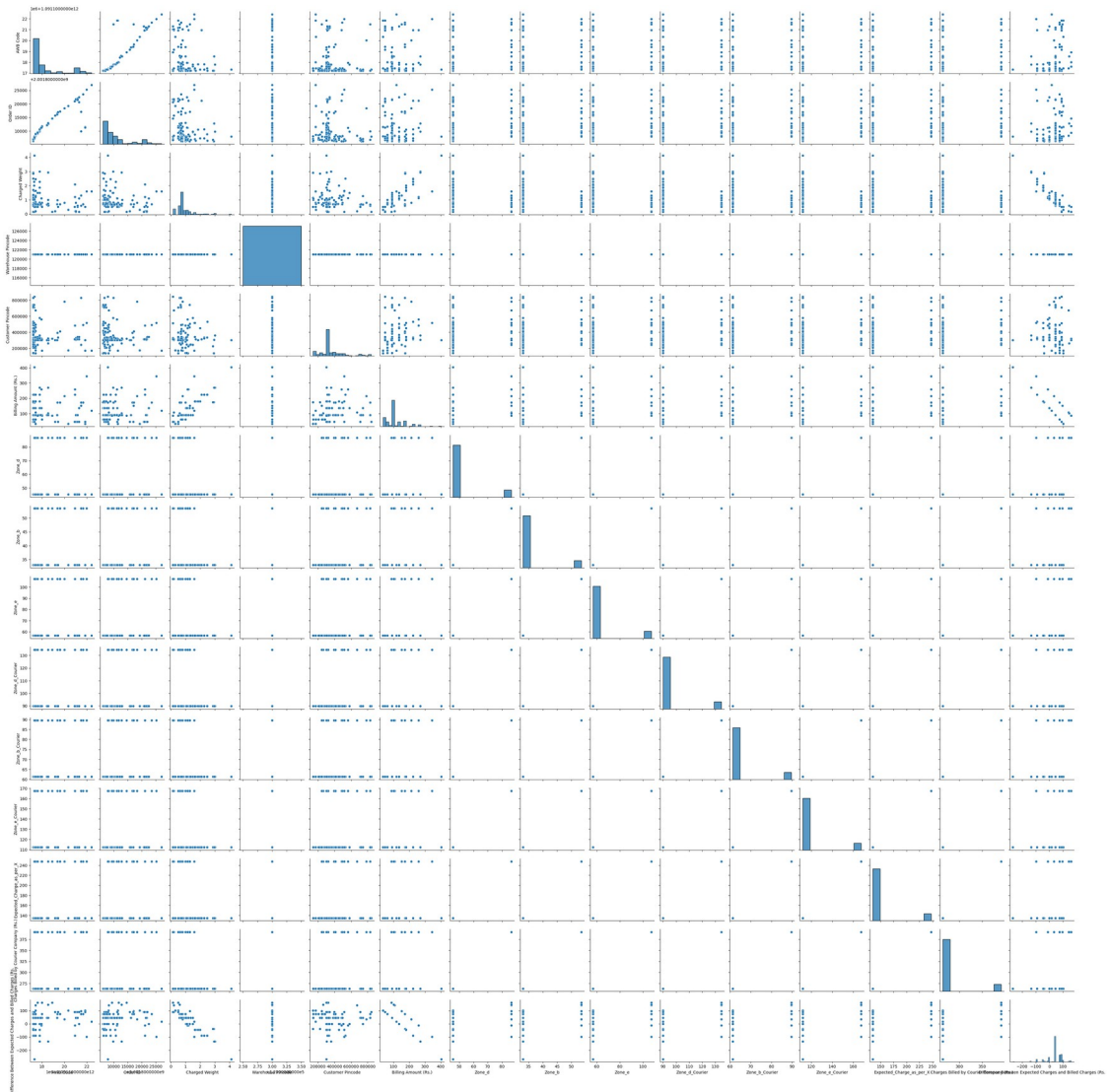
```
sns.barplot(x=courier_invoice['Expected_Charge_as_per_X'],y=courier_invoice['Charges Billed by Courier Company (Rs.)'])
```

```
<AxesSubplot:xlabel='Expected_Charge_as_per_X', ylabel='Charges Billed by Courier Company (Rs.)'>
```



```
sns.pairplot(courier_invoice)
```

```
<seaborn.axisgrid.PairGrid at 0x22a60b433d0>
```



```
e_courier=courier_invoice.to_csv('C://Users//RAHUL//Desktop//Cointab
Data Analyst')
```

#Output Data 2

#Create a summary table

#Count Amount (Rs.)

```
Total_orders_X_charged=courier_invoice['Expected_Charge_as_per_X'].cou
nt()
```

##Count

```
Total_orders_X_charged
```

124

```
#Amount (Rs.)
```

```
Total_orders_X_charged_sum=courier_invoice['Expected_Charge_as_per_X']  
.sum()
```

```
Total_orders_X_charged_sum
```

```
18427.5
```

```
Expected_Result.head()
```

	Unnamed: 0	Count	Amount
0	Total Orders - Correctly Charged	1.0	135.0
1	Total Orders - Over Charged	1.0	-141.5
2	Total Orders - Under Charged	NaN	NaN

```
courier_invoice['']
```