E-commerce_Cancellation_Analysis

May 18, 2024

[3]:
```
import pandas as pd
```

[4]:
```
df = pd.read_csv('E-commerce-data.csv')
```

[5]:
```
df
```

[5]:
```
      Period Payment Channel Dest Zone          Remarks  Count        State  \
0     Feb W1     COD     APP      EAST  Verified Orders  41387       ORISSA
1     Feb W1     COD     APP      EAST  Verified Orders  42652       ORISSA
2     Feb W1     COD     APP      EAST  Verified Orders  38427  WEST BENGAL
3     Feb W1     COD     APP      EAST  Verified Orders  37933  WEST BENGAL
4     Feb W1     COD     APP     METRO  Verified Orders  59200        DELHI
…       …      …     …       …              …    …          …
2011  Mar W4     STD      PC     METRO        Post-Ship    206      HARYANA
2012  Mar W4     STD      PC     METRO        Post-Ship     20      HARYANA
2013  Mar W4     STD      PC     METRO        Post-Ship    114    KARNATAKA
2014  Mar W4     STD      PC     METRO        Post-Ship    107    KARNATAKA
2015  Mar W4     STD      PC     METRO        Post-Ship    166    KARNATAKA

     Country
0      India
1      India
2      India
3      India
4      India
…      …
2011   India
2012   India
2013   India
2014   India
2015   India

[2016 rows x 8 columns]
```

[6]:
```
df.shape
```

[6]:
```
(2016, 8)
```

[7]:
```
df['Remarks'].value_counts()
```

[7]:
```
Remarks
Verified Orders    504
Pre-Manifest       504
M2S                504
Post-Ship          504
Name: count, dtype: int64
```

[8]:
```
df_copy = df[df['Remarks'] != 'Verified Orders']
df_copy
```

[8]:

	Period	Payment	Channel	Dest Zone	Remarks	Count	State	\
504	Feb W1	COD	APP	TIER 2	Pre-Manifest	3179	KARNATAKA	
505	Feb W1	COD	APP	TIER 2	Pre-Manifest	2912	KERALA	
506	Feb W1	COD	APP	TIER 2	Pre-Manifest	2037	MADHYA PRADESH	
507	Feb W1	COD	APP	TIER 2	Pre-Manifest	2657	MAHARASHTRA	
508	Feb W1	COD	APP	TIER 2	Pre-Manifest	3962	TAMIL NADU	
…	…	…	…	…	…	…	…	
2011	Mar W4	STD	PC	METRO	Post-Ship	206	HARYANA	
2012	Mar W4	STD	PC	METRO	Post-Ship	20	HARYANA	
2013	Mar W4	STD	PC	METRO	Post-Ship	114	KARNATAKA	
2014	Mar W4	STD	PC	METRO	Post-Ship	107	KARNATAKA	
2015	Mar W4	STD	PC	METRO	Post-Ship	166	KARNATAKA	

	Country
504	India
505	India
506	India
507	India
508	India
…	…
2011	India
2012	India
2013	India
2014	India
2015	India

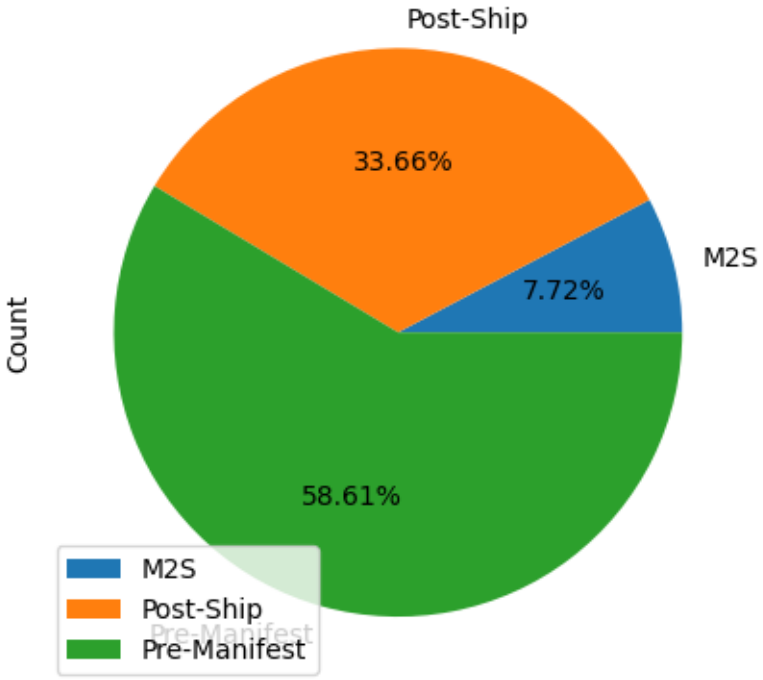
[1512 rows x 8 columns]

[9]:
```
total_count = df_copy['Count'].sum()
total_count
```

[9]:
```
1196517
```

[10]:
```
df_copy.groupby(['Remarks']).sum().plot(kind='pie', y='Count',autopct='%1.2f%%')
```

[10]:
```
<Axes: ylabel='Count'>
```



[11]:
```
df2 = df_copy.groupby(['Period','Remarks']).agg({'Count': ['sum']})
df2
```

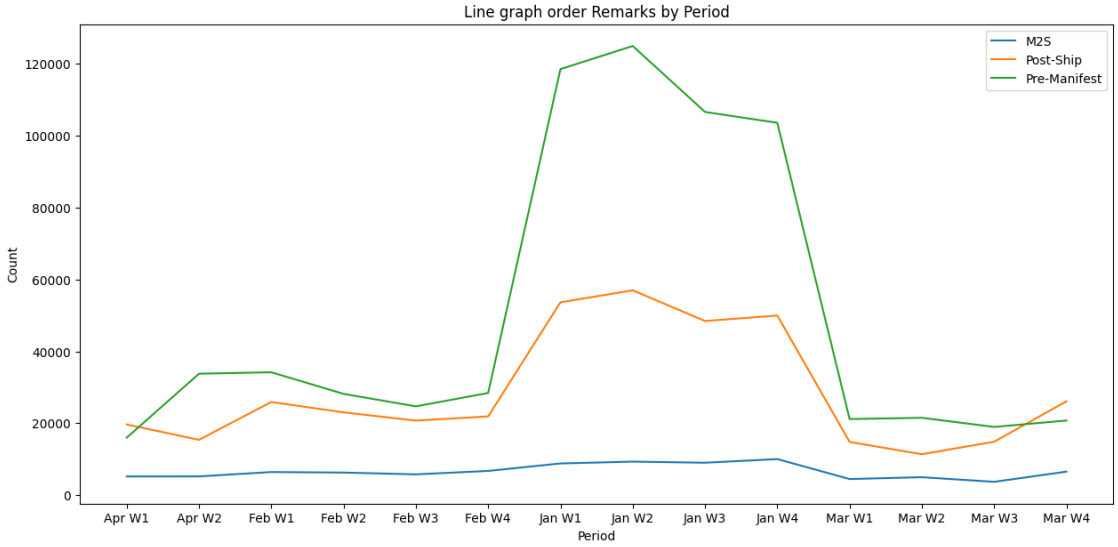
[11]:

Period	Remarks	Count sum
Apr W1	M2S	5189
	Post-Ship	19658
	Pre-Manifest	15977
Apr W2	M2S	5205
	Post-Ship	15381
	Pre-Manifest	33776
Feb W1	M2S	6418
	Post-Ship	25911
	Pre-Manifest	34185
Feb W2	M2S	6270
	Post-Ship	23025
	Pre-Manifest	28171
Feb W3	M2S	5775
	Post-Ship	20733
	Pre-Manifest	24699
Feb W4	M2S	6728
	Post-Ship	21896

```
       Pre-Manifest   28411
Jan W1 M2S             8809
       Post-Ship      53645
       Pre-Manifest  118531
Jan W2 M2S             9315
       Post-Ship      56986
       Pre-Manifest  124964
Jan W3 M2S             9008
       Post-Ship      48444
       Pre-Manifest  106612
Jan W4 M2S            10026
       Post-Ship      49972
       Pre-Manifest  103599
Mar W1 M2S             4470
       Post-Ship      14790
       Pre-Manifest   21163
Mar W2 M2S             4982
       Post-Ship      11373
       Pre-Manifest   21512
Mar W3 M2S             3696
       Post-Ship      14858
       Pre-Manifest   18964
Mar W4 M2S             6530
       Post-Ship      26104
       Pre-Manifest   20756
```

[22]:
```
import matplotlib.pyplot as plt

plt.figure(figsize=(15,7))
for name, group in df2.groupby('Remarks'):
    plt.plot(group['Period'],group['Count']['sum'], label=name)
plt.title('Line graph order Remarks by Period')
plt.xlabel("Period")
plt.ylabel("Count")
plt.legend()
plt.show()
```



[25]:
```
df3 = df_copy.groupby(['Channel', 'Remarks']).agg({'Count': ['sum']})
df3
```

[25]:

		Count
		sum
Channel	Remarks	
APP	M2S	73446
	Post-Ship	293317
	Pre-Manifest	508629
M-Site	M2S	9518
	Post-Ship	57976
	Pre-Manifest	112780
PC	M2S	9457
	Post-Ship	51483
	Pre-Manifest	79911

df3.unstack().plot(kind='bar') plt.title('Bar graph order Remarks by Channel') plt.xlabel("Channel") plt.ylabel("Count") plt.legend() plt.show()

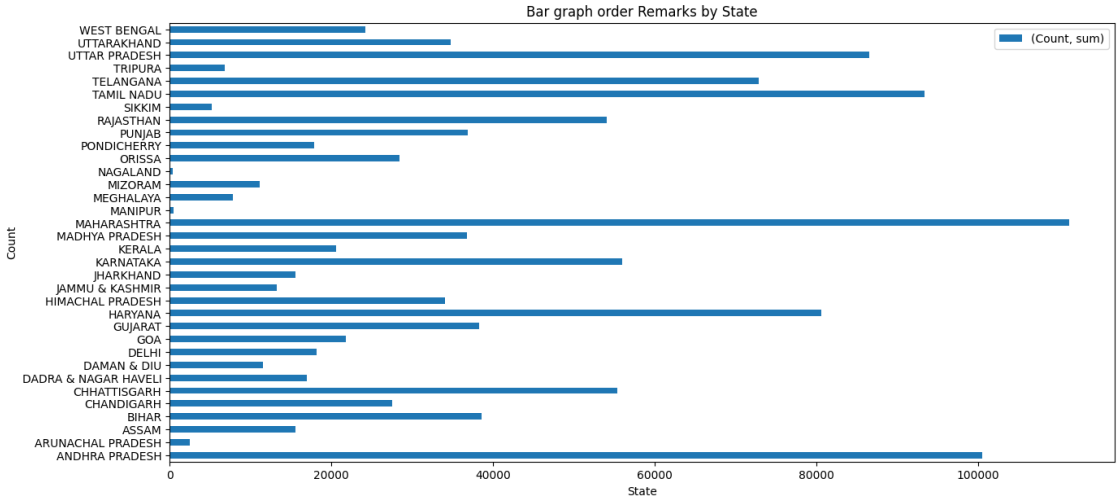
[28]:
```
df4 = df_copy.groupby(['State']).agg({'Count': ['sum']})
df4
```

[28]:

	Count
	sum
State	
ANDHRA PRADESH	100576
ARUNACHAL PRADESH	2559
ASSAM	15558
BIHAR	38584

```
CHANDIGARH             27506
CHHATTISGARH           55386
DADRA & NAGAR HAVELI   17036
DAMAN & DIU            11535
DELHI                  18203
GOA                    21781
GUJARAT                38346
HARYANA                80681
HIMACHAL PRADESH       34122
JAMMU & KASHMIR        13300
JHARKHAND              15623
KARNATAKA              55984
KERALA                 20582
MADHYA PRADESH         36846
MAHARASHTRA           111317
MANIPUR                  455
MEGHALAYA               7795
MIZORAM                11175
NAGALAND                 402
ORISSA                 28464
PONDICHERRY            17872
PUNJAB                 36853
RAJASTHAN              54040
SIKKIM                  5198
TAMIL NADU             93361
TELANGANA              72935
TRIPURA                 6802
UTTAR PRADESH          86594
UTTARAKHAND            34775
WEST BENGAL            24271
```

[30]:
```
df4.plot(kind='barh', figsize=(15,7))
plt.title('Bar graph order Remarks by State')
plt.xlabel("State")
plt.ylabel("Count")
plt.legend()
plt.show()
```



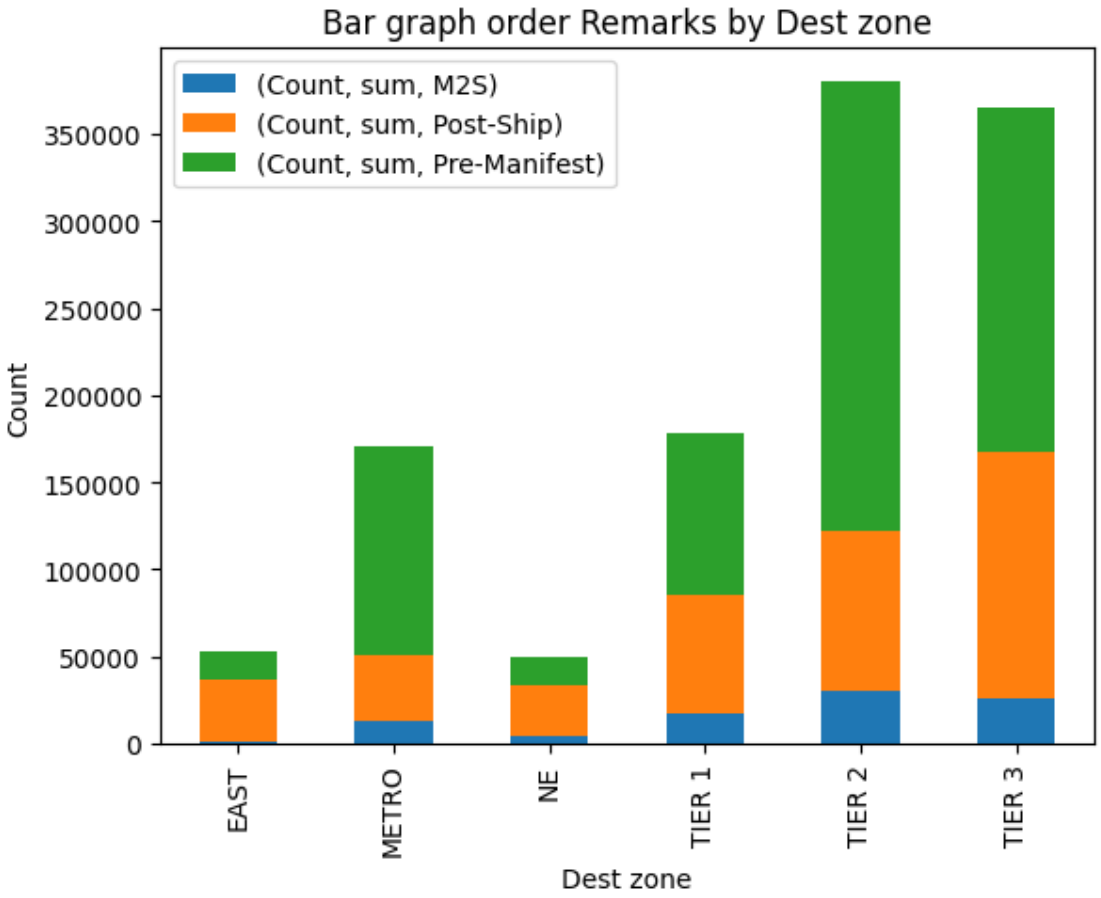
[37]:
```
df5 = df_copy.groupby(['Dest Zone', 'Remarks']).agg({'Count': ['sum']})
df5
```

[37]:

		Count
		sum
Dest Zone	Remarks	
EAST	M2S	958
	Post-Ship	35984
	Pre-Manifest	15793
METRO	M2S	13468
	Post-Ship	37483
	Pre-Manifest	120022
NE	M2S	4231
	Post-Ship	29259
	Pre-Manifest	16454
TIER 1	M2S	17543
	Post-Ship	67653
	Pre-Manifest	92579
TIER 2	M2S	30182
	Post-Ship	91474
	Pre-Manifest	258559
TIER 3	M2S	26039
	Post-Ship	140923
	Pre-Manifest	197913

[39]:
```
df5.unstack().plot(kind='bar', stacked=True)
plt.title('Bar graph order Remarks by Dest zone')
plt.xlabel("Dest zone")
plt.ylabel("Count")
plt.legend()
```

```
plt.show()
```



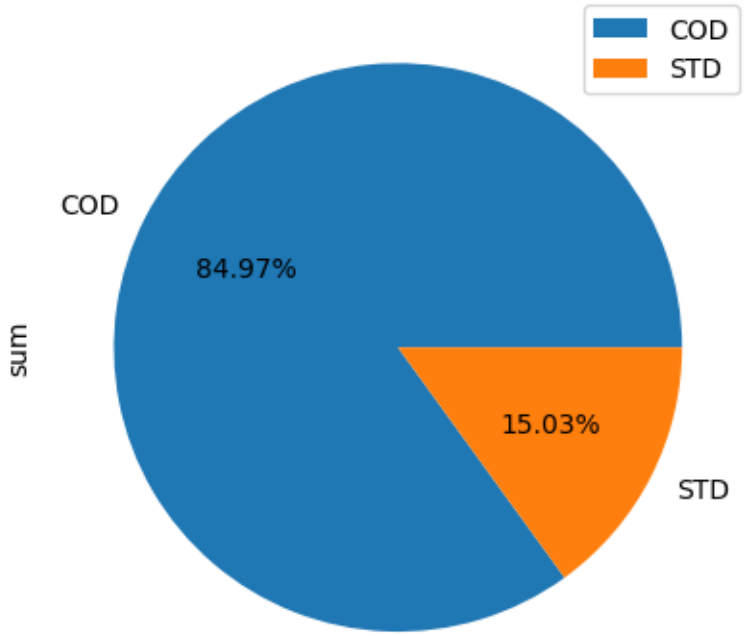
[42]:
```
df6 = df_copy.groupby(['Payment']).agg({'Count': ['sum']})
df6['% Share'] = 100 * df6[('Count', 'sum')] / df6[('Count', 'sum')].sum()
df6
```

[42]:

	Count	% Share
	sum	
Payment		
COD	1016735	84.974555
STD	179782	15.025445

[43]:
```
df6.plot(kind='pie', y='Count', autopct='%1.2f%%')
```

[43]: <Axes: ylabel='sum'>


COD
STD
COD
84.97%
sum
15.03%
STD