

Neeman's Analysis Report

```
import pandas as pd
import numpy as np
import warnings
warnings.simplefilter('ignore')
```

```
data = pd.read_csv('D:/Datasets/Assignment_Revenue&Exchanges (3).csv', encoding_errors= 'replace', low_memory= False)
```

```
data.head(20)
```

```
data['Order Amount'].value_counts().head(50)
```

```
data.dtypes
```

```
data['Order Amount'] = data['Order Amount'].replace(' - ', '0')
```

```
data['Order Amount'].value_counts().head(10)
```

```
def convert_string(str):
    num = str.replace(',', '')
    return int(num)
```

Created a function to convert string data type to integer data type.

```
convert_string(data['Order Amount'][0])
```

```
Order_Amount = [convert_string(i) for i in data['Order Amount']]
```

Using list comprehensions created a list of order price amounts

```
data['Order Amount'] = pd.Series(Order_Amount)
```

```
data
```

```
data.dtypes
```

```
data['Ship PinCode'][0]
```

```
data['Ship PinCode'].value_counts().head(80)
```

```
data['Ship PinCode'] = pd.to_numeric(data['Ship PinCode'], errors= 'coerce')
```

Converted ship pincode object datatype to numeric datatype.

```
data.dtypes
```

```
data['Unit Price'] = data['Unit Price'].replace(' - ', '0')
```

```
unit_price = [convert_string(i) for i in data['Unit Price']]
```

Using list comprehensions created a list of unit price amounts

```
data['Unit Price'] = pd.Series(unit_price)
```

```
data.dtypes
```

```
data['Discount'] = data['Discount'].replace(' - ', '0')
```

```
Discount = [float(i.replace(',','')) for i in data['Discount']]
```

```
data['Discount'] = pd.Series(Discount)
```

```
data.dtypes
```

```
data.head(15)
```

```
pd.to_numeric(data['Ship PinCode'],downcast= 'signed')
```

```
def convet_integer(float):  
    num = int(float)  
    return num
```

```
data['Ship PinCode'] = data['Ship PinCode'].replace(np.nan, 0)
```

```
data['Ship PinCode'] = data['Ship PinCode'].apply(convet_integer)
```

Using apply function to convert the data type of feature column.

```
data.dtypes
```

```

data['Size'] = data['Size'].replace(np.nan, 0)
data['Size'] = data['Size'].apply(convet_integer)
data['Order Qty'] = data['Order Qty'].replace(np.nan, 0)
data['Shipped Qty'] = data['Shipped Qty'].replace(np.nan, 0)
data['Order Qty'] = data['Order Qty'].apply(convet_integer)
data['Shipped Qty'] = data['Shipped Qty'].apply(convet_integer)

data.head(15)

```

```
data.to_excel('Neeman_data.xlsx')
```

Created a excel with the clean and transformed data.

```
data['Ship State'].value_counts().head(10)
```

Checking the top 10 shipping locations of the products.

```

data['Discount'].describe()
discount = data['Discount'].values
max(discount)
np.where(discount == max(discount))
data['Discount'][160154]

data = data.drop(index= [160145, 160146, 160147, 160148, 160149, 160150, 1601
51, 160152,
    160153, 160154, 160155, 160156, 160157, 160158, 160159, 160160,
    160161, 160162, 160163, 160164, 160165, 160166, 160167, 160168,
    224575, 224576, 224577, 224578, 224579, 224580, 224581, 224582,
    224583, 224584, 224585, 224586, 224587, 224588, 224589, 224590,
    224591, 224592, 224593, 224594, 224595, 224596, 224597, 224598], axis
= 0)

```

As the discount for the above records is 145154 for one product and that cann
ot be possible so dropped those records.

```
data
```

```
data['City'].value_counts().head(20)
```

Top 10 shipping cities

```
data['City'] = data['City'].replace('MUMBAI', 'Mumbai')
```

As the dataset having two same cities with a different spelling so replaced with the single one.

```
data['City'] = data['City'].replace('PUNE', 'Pune')
```

```
data['City'].value_counts().head(20)
```

```
filter_maharashtra = data[data['Ship State'] == 'Maharashtra']
```

Created a data frame where the shipping location is only Maharashtra.

```
filter_maharashtra.describe()
```

Checked the statistical analysis of data frame.

```
filter_maharashtra['City'].value_counts().head(30)
```

Top 30 shipping cities in Maharashtra.

```
filter_maharashtra['City'] = filter_maharashtra['City'].replace('THANE', 'Thane')
```

```
filter_maharashtra['City'] = filter_maharashtra['City'].replace('NAGPUR', 'Nagpur')
```

```
filter_maharashtra['City'] = filter_maharashtra['City'].replace('NAGPUR', 'Nagpur')
```

```
filter_maharashtra['City'] = filter_maharashtra['City'].replace(['NASHIK', 'Nashik'], 'Nashik')
```

```
filter_maharashtra['City'] = filter_maharashtra['City'].replace('mumbai', 'Mumbai')
```

```
filter_maharashtra['City'] = filter_maharashtra['City'].replace('pune', 'Pune')
```

```
filter_maharashtra['City'].value_counts()
```

```
data['Size'].value_counts()
```

Variants of Product Distribution.

```
data['Status'].value_counts()
```

Status of the orders distribution

```
filtered_delivered = data[data['Status'] == 'Delivered']
```

Created a data frame where the status of the orders is delivered.

```
filtered_delivered
```

```
filtered_delivered['Size'].value_counts().
```

```
filtered_delivered['Ship State'].value_counts()
```

```
data.head(10)
```

```
data['OrderDate'] = pd.to_datetime(data['OrderDate'], format='%d-%m-%Y')
```

Changed the order date into above specified format.

```
data.dtypes
```

Checked the data types of the features.

```
data['OrderMonth'] = data['OrderDate'].dt.month
```

Separated the order date and created an order month.

```
data['OrderYear'] = data['OrderDate'].dt.year
```

Separated the order date and created an order year.

```
data.head(10)
```

```
data['OrderMonth'].value_counts()
```

```
filtered_delivered['OrderDate'] = pd.to_datetime(filtered_delivered['OrderDate'], format='%d-%m-%Y')
```

```
filtered_delivered['OrderMonth'] = filtered_delivered['OrderDate'].dt.month
```

```
filtered_delivered['OrderYear'] = filtered_delivered['OrderDate'].dt.year
```

```
filtered_delivered['OrderMonth'].value_counts()
```

Monthly wise delivered products.

```
filtered_delivered
```

```
data.groupby('Status')['Discount'].sum()
```

```
data.groupby('Order Amount')['Discount'].count().sort_values(ascending= False
)
```

Grouped the order amount data with the counts of discount

```
data['Order Amount'].value_counts()
```

```
data[data['Order Amount'] == 2]
```

```
filtered_delivered.groupby('Order Amount')['Discount'].sum().sort_values(ascending= False).head(30)
```

Grouped the order amount data with the sum of discount amounts.

```
filtered_delivered[filtered_delivered['Order Amount'] == 2]
```

```
filtered_delivered
```

```
filtered_delivered = filtered_delivered.drop(filtered_delivered[filtered_delivered['Order Amount'] == 2].index)
```

I see that there is a large sum of discount amounts for the order amount 2 so I have gone ahead and dropped those records.

```
filtered_delivered.groupby('Order Amount')['Discount'].sum().sort_values(ascending= False).head(30)
```

Post - Analysis

By checking the shipping state, we can say that offline expansion should be in Maharashtra.

And the Store should be in Mumbai because we have so many customers from Mumbai and next second is Pune which is very near to Mumbai and it can be useful for shipping purposes too.

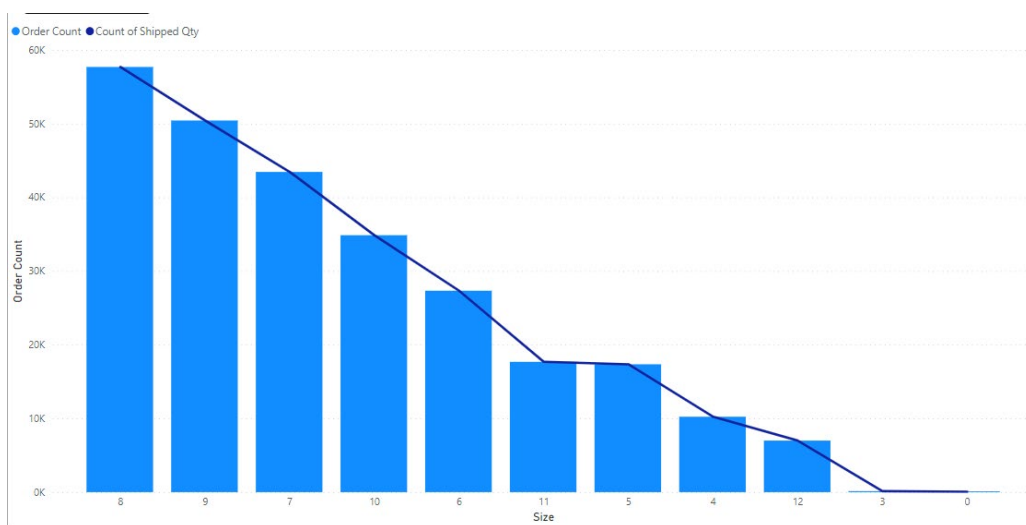
As per Analysis, the variants which are sold more than others are size 8, size 9, and size 7. So, it is best to have a greater number of these types of variants to run the business smoothly and increase revenue.

As our sales are great in the second half of the year so it is better to increase our inventory in the last 6 months and plan according to regarding the variants of sizes 9,8,7 to make an impact on the business

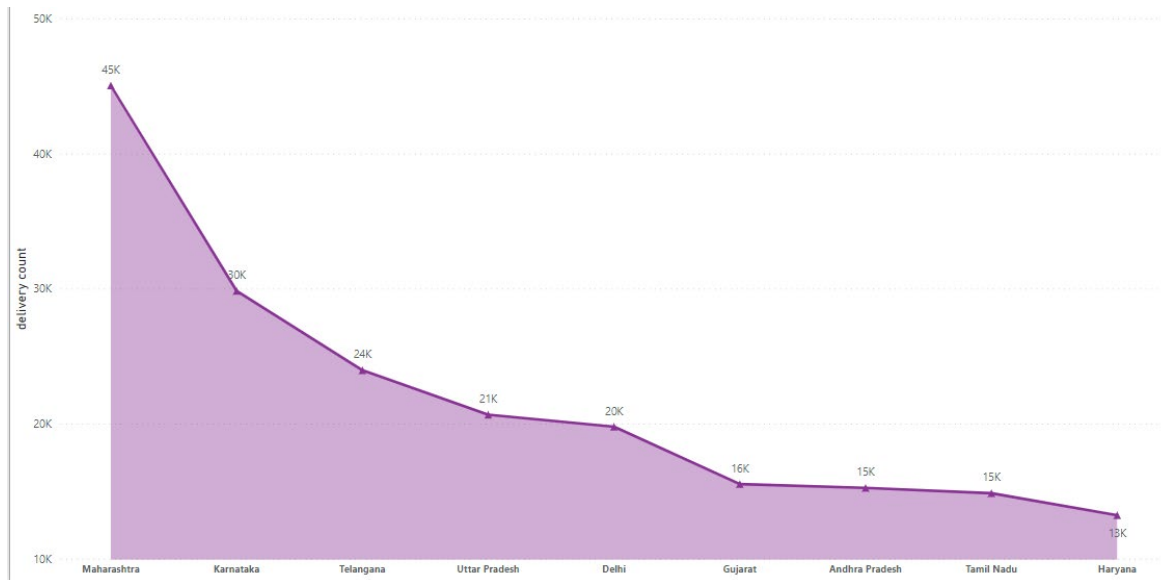
From Analysis Flat Discount is not needed for the products. But a Partial/Upto discount is needed to boost the sales. Most Probably a discount is needed for the order amount which is more than 2500 and for those below the 2500 amount products we do not need discounts because we already have existing customers. Sometimes customers expect discounts for those products too, in that case, we can put discounts when there are any festivals and public holidays for boosting the brand.



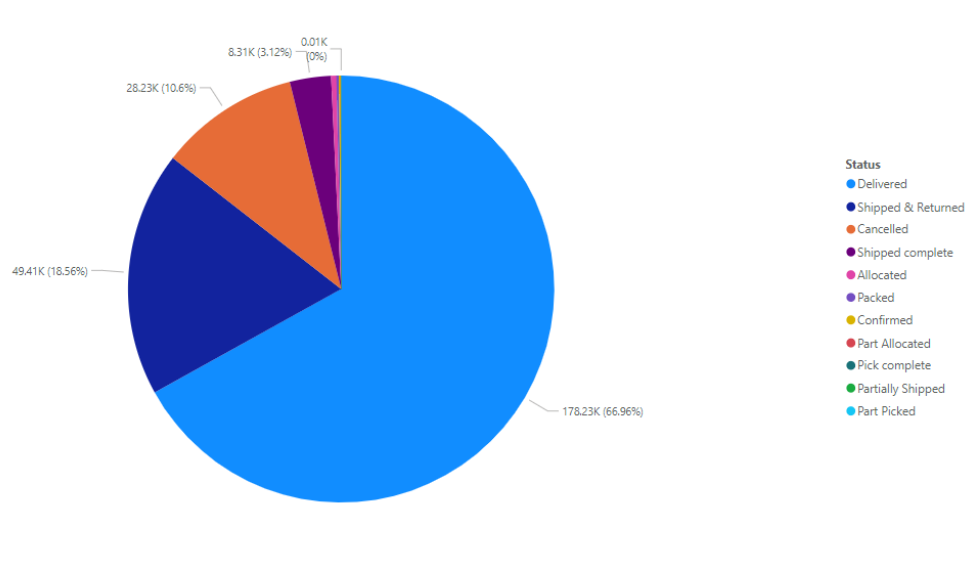
As per the above visual, the visual represents the Total shipping cities of the orders.



Order count and Shipped Quantity according to the variants of the products.



Top cities which have more delivered orders.



Status of the order wise distribution

Created a Matrix type visual for the Count of discounts within all variants of products according to order amounts.

Order Amount	0	10	11	12	3	4	5	6	7	8	9	Total
69900000									1		1	1
983102									1			1
366189				1								1
209930								1				1
183014											1	1
131199			1									1
117939			1				2	1	3	3	5	18
114037				2			1					3
110963									1			1
89970					1							1
77589										1		1
63054										4		4
52784											2	2
50328										1		1
49990										2		2
46186							1				1	2
45675											2	2
44991										1		1
39992											2	12
39144									1	10		1
36289			2									2
35068										1		1
34993										2		2
32990					1						4	5
32970			2							1		3
30756											1	1
30069			2									2
29990										1	1	2
29691				2								8
Total	56	34859	17681	6997	138	10233	17350	27330	43445	57685	50422	266196

Order Amount	0	10	11	12	3	4	5	6	7	8	9	Total
29691				2							1	8
28092										2		2
27069										2	3	5
26392											1	7
25980			3	1								4
24995											10	10
24728			20							4		24
24366							1			1	1	3
24070			2									2
23393										4		4
23093										2	16	18
21938			2						2			16
21413			2						2		5	9
21293			10							4		14
21113			2								2	4
21070										8		8
21068						1					1	2
20993											1	1
20543			14							4		18
20071											2	2
19794							2		1	5	4	12
19572										3		3
19485											2	2
19374			4			2	2				4	12
18594										2		2
18474										2		2
18069							1			5	1	7
17994										1		1
17665												7
Total	56	34859	17681	6997	138	10233	17350	27330	43445	57685	50422	266196

Order Amount	0	10	11	12	3	4	5	6	7	8	9	Total
16995					3						10	27
16889											2	2
16596											2	4
16495			6		1			3	1	2	34	49
16100											14	14
16088												7
15835							2					2
15797									2			2
15771												7
15575			4			2	2			3	4	12
15424										2	4	12
15395			3							2		5
15378					1							1
15356			6		6							12
15171					2				6	2		14
15120		2			6					4		10
15072								2		4	1	8
14997					1						2	3
14995			5			2				4		11
14946								2	6	2		10
14896											4	4
14876					2	10						12
14836									3		3	6
14754										1		1
14715						5				1		6
14400									1			1
14274										10	8	18
14245			3					1			2	11
13980										1		1
Total	56	34859	17681	6997	138	10233	17350	27330	43445	57685	50422	266196

Order Amount	0	10	11	12	3	4	5	6	7	8	9	Total
10987			4									4
10969			1							1	2	4
10956			10									10
10922								2	1		1	7
10891							1	1	1	3	2	8
10800			6						2	2		10
10794			1					1	2		2	6
10772									1		2	3
10706										2		6
10697											2	2
10636										4	2	8
10597				2	2							4
10560											1	1
10556					1				3		4	9
10498										1		1
10497											3	3
10476									5			5
10472					2				1			3
10427			2					2			4	8
10399			4						2		1	7
10398											6	10
10397			2	5	2				1			10
10272			5									5
10260	1								1	2		4
10257							1	2		8		8
10238												3
10198			3									3
10097				3								3
10076								2		2		4
Total	56	34859	17681	6997	138	10233	17350	27330	43445	57685	50422	266196

Order Amount	0	10	11	12	3	4	5	6	7	8	9	Total
9655			9	3								18
9598			3									3
9597				3	3			4				10
9572							4					4
9522			3							2		5
9518				2							2	9
9499			1						3	2		6
9498											2	2
9403											3	3
9402											3	3
9307				2					2	1		5
9297										3		3
9278											6	6
9277			3									9
9237			1					1		1	7	13
9213			1								2	3
9148			2							2		8
9117			2					1				3
9116			4								4	8
9095			3	4			2					9
9093										2		2
9073										1	1	2
9072			10	5	3		6		2	1	8	45
9047				3								3
9037									9			9
9000									1	1		2
8998				2						2	1	9
8997			11	4				1	6	2	11	49
Total	56	34859	17681	6997	138	10233	17350	27330	43445	57685	50422	266196

Order Amount	0	10	11	12	3	4	5	6	7	8	9	Total
1394			11	1			3	9	8	11	8	57
1366			53	30	2		22	65	40	75	79	422
1359											1	1
1348			49	22			15	23	21	23	19	198
1336			2					1			2	5
1328			25	11	11		16	66	69	45	44	322
1299			1002	700			75	168	431	1013	1448	6330
1289				1								2
1281			27	23			5	20	36	40	38	216
1278							2				2	4
1258			18	12	7		28	37	45	49	51	309
1234			1	2				1	1	4	2	13
1198			4									4
1195			22	11	2		19	30	26	40	30	228
1169			3	2				1	1	4	4	17
1158				1					1		1	3
1148				1					1		1	3
1138			2									2
1118			403	188	117		286	514	643	723	787	4226
1100										4		4
1099			1965	1201			267	388	908	2107	2987	12609
1079											2	2
1076				1								1
1062			87	51	32		85	176	134	205	169	1070
1048			144	51	32		99	155	163	176	208	1206
1044			237	132			38	56	105	187	293	1351
1039			479	284			44	96	176	345	641	2714
1011											1	1
1006											1	1
Total	56	34859	17681	6997	138	10233	17350	27330	43445	57685	50422	266196

Order Amount	0	10	11	12	3	4	5	6	7	8	9	Total	
900									4	2		2	8
899								2			2		4
894									1				1
884			1							1	1	2	5
879			1300	721			192	306	501	1097	1817	1820	7754
878			2								2		4
863												2	2
854											2		2
842									1				1
835			696	383			99	167	284	601	937	930	4097
824			289	156			47	56	97	220	361	375	1601
800							1		1				2
798			1	1							1	2	5
783			267	138			42	56	98	197	360	384	1542
782												1	1
779			3									1	4
774										1			1
769			255	170			29	82	158	309	486	478	1967
758				2	1		2			1			6
751												2	2
750										1			1
749												2	2
748										2	2		4
740										1			1
736											1	1	2
730			195	125			33	57	109	218	311	269	1317
724					1						1	2	4
717										3			3
713									1				1
Total		56	34859	17681	6997	138	10233	17350	27330	43445	57685	50422	266196

Order Amount	0	10	11	12	3	4	5	6	7	8	9	Total	
524			622	311	129		319	543	737	850	1065	892	5468
522			1	2							1		4
509								1					1
499			11	3	1				4	7	7	12	45
498			203	91	60		135	188	287	291	314	260	1829
490			81	65						2	117	24	289
489			510	241	133		244	459	520	732	836	797	4472
474			1	1	1				1			2	6
464			181	118	45		124	235	283	349	362	331	2028
450												4	4
424			5	4	4			3	10	6	13	11	56
419									1				1
409			4								2		6
402											2	2	4
400							1			1		1	3
399			7						3	1	7	3	21
379			1	2	1				1	2		2	9
374					1					1			2
355											1		1
349			8		1		2	1	4	2	5	1	24
332			1	1	1		2				2	3	10
331			1							2	5	2	10
199			2	1					1			1	5
99			1										1
6								3					3
4											2		2
2			119	42	8	5	94	115	157	176	181	189	1086
1			2				5	2	2		10		21
0		2	1990	1190	400	16	520	1125	1584	2086	3026	2877	14816
Total		56	34859	17681	6997	138	10233	17350	27330	43445	57685	50422	266196