

# E-commerce Analytics: Understanding Pakistan Market

From the word **E-commerce**; it is the marketing of goods and services. Every transaction done online involves a seller and a buyer. This report provides an analysis of the current state of e-commerce, key trends, challenges, and future endeavors. The objective is to offer insights into the Pakistan E-commerce dataset using a spreadsheet tool i.e., Excel. The dataset set consisted of 584525 rows and 21 columns but due to the size of the dataset, a sample of 5000 rows was extracted to Python for easy analysis.

1. **Data Exploration:** Data Exploration: This dataset structure was examined in which missing values and blanks were dropped due to the low percentage of missing values and the data type was also corrected. After this, summary statistics was carried out to understand key variable which is: Price, Quantity ordered, Grand total, and MV.

[illegible]

2. Descriptive Statistics: Trends in the dataset were found after getting the total unit and average unit for each category.

[illegible]

- | INFERENTIAL STATISTICS       |             |                |   |              |             |
|------------------------------|-------------|----------------|---|--------------|-------------|
| z-Test: Two Sample for Means |             |                | t-Test: Two-Sample Assuming Unequal Variances |              |             |
|                              | Orders      | Payment Method |   | Variable 1   | Variable 2  |
| Mean                         | 1.1434      | 555.5555556    | Mean  | 1.1434       | 555.5555556 |
| Known Variance               | 0.912218884 | 1734085.778    | Variance                                      | 0.912218884  | 1734085.778 |
| Observations                 | 5000        | 9              | Observations                                  | 5000         | 9           |
| Hypothesized Mean            | 0           |                | Hypothesized Mean Difference                  | 0            |             |
| z                            | -1.26304467 |                | df  | 8            |             |
| P(Z<=z) one-tail             | 0.103286564 |                | t Stat  | -1.263044669 |             |
| z Critical one-tail          | 1.644853627 |                | P(T<=t) one-tail                              | 0.12107007   |             |
| P(Z<=z) two-tail             | 0.206573127 |                | t Critical one-tail                           | 1.859548038  |             |
| z Critical two-tail          | 1.959963985 |                | P(T<=t) two-tail                              | 0.24214014   |             |
|                              |             |                | t Critical two-tail                           | 2.306004135  |             |

- [illegible]

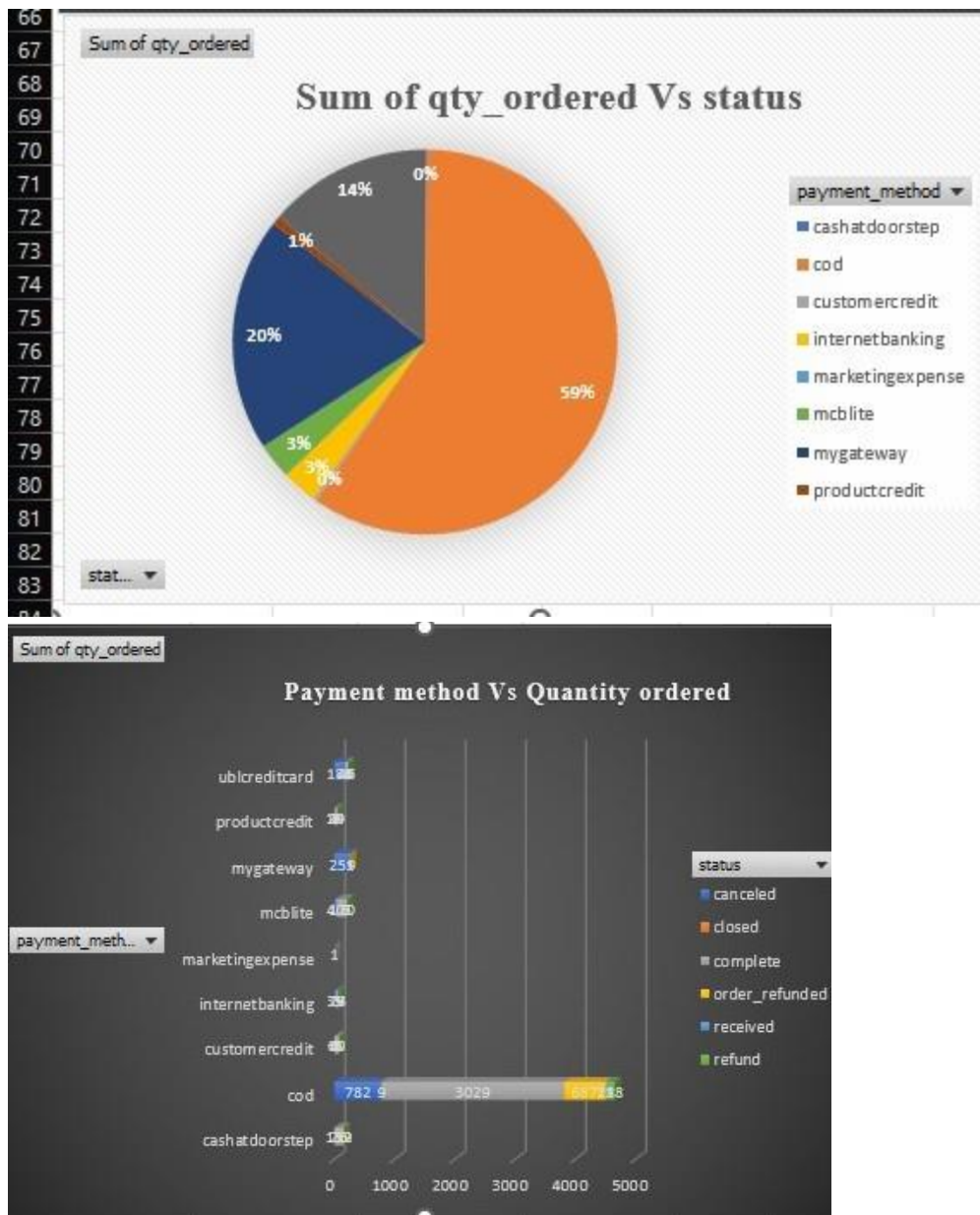
[illegible]

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1		Sum of qty_ordered	Column Labels																									
2		Row Labels		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	27	34	Grand Total						
3		canceled		920	156	36	20	10	6	14	8	9	30	11				15	20	27	34	1316						
4		closed		15																		15						
5		complete		2938	266	84	48	45	18	21		9	30		12	13	14					3498						
6		order_refunded		640	58	12							10									720						
7		received		24		4																28						
8		refund		109	26		5															140						
9		Grand Total		4646	506	132	72	60	24	35	8	18	70	11	12	13	14	15	20	27	34	5717						
10																												
11	Observed	Row Labels		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	27	34	Row Total						
12		canceled		920	156	36	20	10	6	14	8	9	30	11				15	20	27	34	1316						
13		closed		15																		15						
14		complete		2938	266	84	48	45	18	21		9	30		12	13	14					3498						
15		order_refunded		640	58	12							10									720						
16		received		24		4																28						
17		refund		109	26		5															140						
18		Column Total		4646	506	132	72	60	24	35	8	18	70	11	12	13	14	15	20	27	34	5717	Grand Total					
19																												
20	Expected	Row Labels		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	27	34							
21		canceled		1069.465804	116	30	17	14	6	8.06	1.8415	4	16	3	3	3	3	5	6	8								
22		closed		12.18995977	1.3	0.3	0	0	0	0.09	0.021	0	0	0	0	0	0	0	0	0								

C41																											
21		canceled		1069.465804	116	30	17	14	6	8.06	1.8415	4	16	3	3	3	3	5	6	8							
22		closed		12.18995977	1.3	0.3	0	0	0	0.09	0.021	0	0	0	0	0	0	0	0	0							
23		complete		2842.698618	310	81	44	37	15	21.4	4.8949	11	43	7	7	8	9	9	12	17	21						
24		order_refunded		585.1180689	64	17	9	8	3	4.41	1.0075	2	9	1	2	2	2	2	3	3	4						
25		received		22.75459157	2.5	0.6	0	0	0	0.17	0.0392	0	0	0	0	0	0	0	0	0	0						
26		refund		113.7729578	12	3.2	2	1	1	0.86	0.1959	0	2	0	0	0	0	0	0	1	1						
27																											
28																											
29	(O-E)^2/E	Row Labels		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	27	34						
30		canceled		20.88895822	13	1	1	1	0	4.38	20.595	6	12	28	3	3	3	3	51	70	88						
31		closed		0.647772942	1.3	0.3	0	0	0	0.09	0.021	0	0	0	0	0	0	0	0	0	0						
32		complete		3.194975832	6.1	0.1	0	2	1	0.01	4.8949	0	4	7	3	3	3	9	12	17	21						
33		order_refunded		5.147724057	0.5	1.3	9	8	3	4.41	1.0075	2	0	1	2	2	2	2	3	3	4						
34		received		0.06816392	2.5	0.6	38	0	0	0.17	0.0392	0	0	0	0	0	0	0	0	0	0						
35		refund		0.200233228	15	3.2	2	8	1	0.86	0.1959	0	2	0	0	0	0	0	0	1	1						
36																											
37																											
38		Xsquare		595.8082195																							
39		df		85																							
40		p-value		1.06214E-77																							
41																											
42																											

5. Visualization: This section consists of plots carried out in the analysis.





In conclusion, the e-commerce industry is poised for continued growth driven by technological innovation and evolving consumer behaviors. This analysis serves as a snapshot of the e-commerce landscape as of July 2016 and the industry is expected to undergo further transformation in the years to come. Staying informed about market dynamics and consumer preferences will be essential for success in these dynamic sectors.

It is observed that consumers are more comfortable with the cash-on-delivery. From the hypotheses carried out, it shows that there is an association between orders and payment methods.