

Shape of Dataset

Rows	Columns
588,101	5

Column Info

Index	Column	Blank Data Count	NA Data Count	All Cell Count	Data Type
B	user id	0	0	588,101	int64
C	test group	0	0	588,101	object
D	converted	0	0	588,101	object
E	total ads	0	0	588,101	float64
F	most ads day	0	0	588,101	object
G	most ads hour	0	0	588,101	float64

Statistics of Dataset (With Outliers)

Statistic	total ads
Count	588,101
Mean	24.82
Standard Deviation	43.72
Minimum	1
Maximum	2,065
25th Percentile	4
Median	13
75th Percentile	27
Lower Bound	-30.5
Upper Bound	61.5

Outlier Analysis

Column	Unique Value	Count	Normalized Count	Sum
outlier flag	outlier	52,057	0.09	
	not outlier	536,044	0.91	1.00

Distribution of Unique Values in "Object" Columns

Column	Unique Value	Normalized Count	Sum
test group	ad	0.96	
	psa	0.04	1.00
converted	FALSE	0.97	
	TRUE	0.03	1.00
most ads day	Monday	0.15	
	Tuesday	0.13	
	Friday	0.16	
	Saturday	0.14	
	Wednesday	0.14	
	Sunday	0.15	
	Thursday	0.14	1.00

Statistics of Dataset (After Removing Outliers)

Statistic	total ads
Count	536,044
Mean	15.13
Standard Deviation	14.06
Minimum	1
Maximum	61
25th Percentile	4
Median	11
75th Percentile	22
Lower Bound	-23
Upper Bound	49

Data Transformation

Max. and Min. Values of "total ads" Column	
Max. Value	61
Min. Value	1

"total ads group" Column is added

Value	Meaning
Low exposure	< 6
Moderate low exposure	6-15
Moderate exposure	16-30
High exposure	31-45
Very high exposure	> 45

"most ads hour group" Column is added

Value	Meaning
Midnight to Early Morning	00:00 - 05:59
Morning	06:00 - 11:59
Afternoon	12:00 - 17:59
Evening	18:00 - 21:59
Late Night	22:00 - 23:59