

Data Cleaning

CustomerId	StreetNo	Streetname	OwnerOccupied	NoBedRooms	NoBathRooms	AreaSqFt	Color
234	104	MG Road	Y	3	1	1000	White
344	197	Shivaji Road	N	3	1.5	--	Red
245		Netaji Subash Road	N	n/a	1	850	pink
523	201	Shivaji Road	12	1	NaN	700	Yello
123	203	Azad Road	Y	3	2	1600	Green
235	207	Azad Road	Y	NA	1	800	WHite
255	NA	Pragati Marg		2	Delhi	950	RED
367	213	Gole Market	Y	1	1		Green
347	215	Gole Market	Y	na	2	1800	pink
523	201	Shivaji Road	12	1	NaN	700	Yello
255	NA	Pragati Marg		2	Delhi	950	RED

1. Count the number of missing values per column
2. Count total missing values
3. How many observations remain if we drop rows with missing values?
4. Include "n/a","na" and "--" as missing values
5. Rename the column 'AreaSqFt' with 'Area'
6. Count number of duplicate rows and remove them.
7. Remove column 'color'
8. Make CustomerId column index
9. In the "OwnerOccupied" column, ensure that there are no numeric values.
10. For customer id=245, fill the "Streetno" with value "125"
11. Fill the "NoBedRooms" with median value of the column