

Data Quality Report

1. Data Description

The dataset (New York Property Valuation and Assessment) was **provided by NYC Department of Finance**. This Dataset has **32 fields** and **1070994 records**. The dataset represents NYC property assessments for the purpose of calculating Property Tax and Grant eligible properties Exemptions and/or Abatements. Data is collected and entered into the system by various City employees, like Property Exemption specialists, Property Assessors, ACRIS reporting, Department of Building reporting, etc.

Numerical Variables:

Field Name	# Records Have Values	% Populated	# Zeros	# Unique Values	Mode	Min Value	Standard Deviation	Max Value
LTFRONT	1,070,994	100.00%	169,108	1297	0	0	9,999	74.03284
LTDEPTH	1,070,994	100.00%	170,128	1370	100	0	9,999	76.39628
STORIES	1,014,730	94.75%	0	111	2	1	119	8.365707
FULLVAL	1,070,994	100.00%	13,007	109,324	0	0	6.15E+09	11582431
AVLAND	1,070,994	100.00%	13,009	70,921	0	0	2.67E+09	4057260
AVTOT	1,070,994	100.00%	13,007	112,914	0	0	4.67E+09	6877529
EXLAND	1,070,994	100.00%	491,699	33,419	0	0	2.67E+09	3981576
EXTOT	1,070,994	100.00%	432,572	64,255	0	0	4.67E+09	6508403
BLDFRONT	1,070,994	100.00%	228,815	612	0	0	7,575	35.5797
BLDDEPTH	1,070,994	100.00%	228,853	621	0	0	9,393	42.70715

AVLAND2	282726	26.40%	0	58591	2408	3	2.37E+09	6178963
AVTOT2	282732	26.40%	0	111360	750	3	4.5E+09	11652529
EXLAND2	87449	8.17%	0	22195	2090	1	2.37E+09	10802213
EXTOT2	130828	12.22%	0	48348	2090	7	4.5E+09	16072510

Categorical Variables:

Field Name	# Records Have Values	% Populated	# Zeros	# Unique Values	Most Common Value
RECORD	1,070,994	100.00%	0	1,070,994	1
BBLE	1,070,994	100.00%	0	1,070,994	1000010101
BORO	1,070,994	100.00%	0	5	4
BLOCK	1,070,994	100.00%	0	13,984	3944
LOT	1,070,994	100.00%	0	6,366	1
EASEMENT	4,636	0.43%	0	12	E
OWNER	1,039,249	97.04%	0	863,347	PARKCHESTER PRESERVAT
BLDGCL	1,070,994	100.00%	0	200	R4
TAXCLASS	1,070,994	100.00%	0	11	1
EXT	354,305	33.08%	0	3	G
EXCD1	638,488	59.62%	0	129	1017
STADDR	1,070,318	99.94%	0	839,280	501 SURF AVENUE
ZIP	1,041,104	97.21%	0	196	10314
EXMPTCL	15,579	1.45%	0	14	X1
EXCD2	92,948	8.68%	0	60	1017
PERIOD	1,070,994	100.00%	0	1	FINAL
YEAR	1,070,994	100.00%	0	1	2010/11
VALTYPE	1,070,994	100.00%	0	1	AC-TR

2. Visualization of Each Field

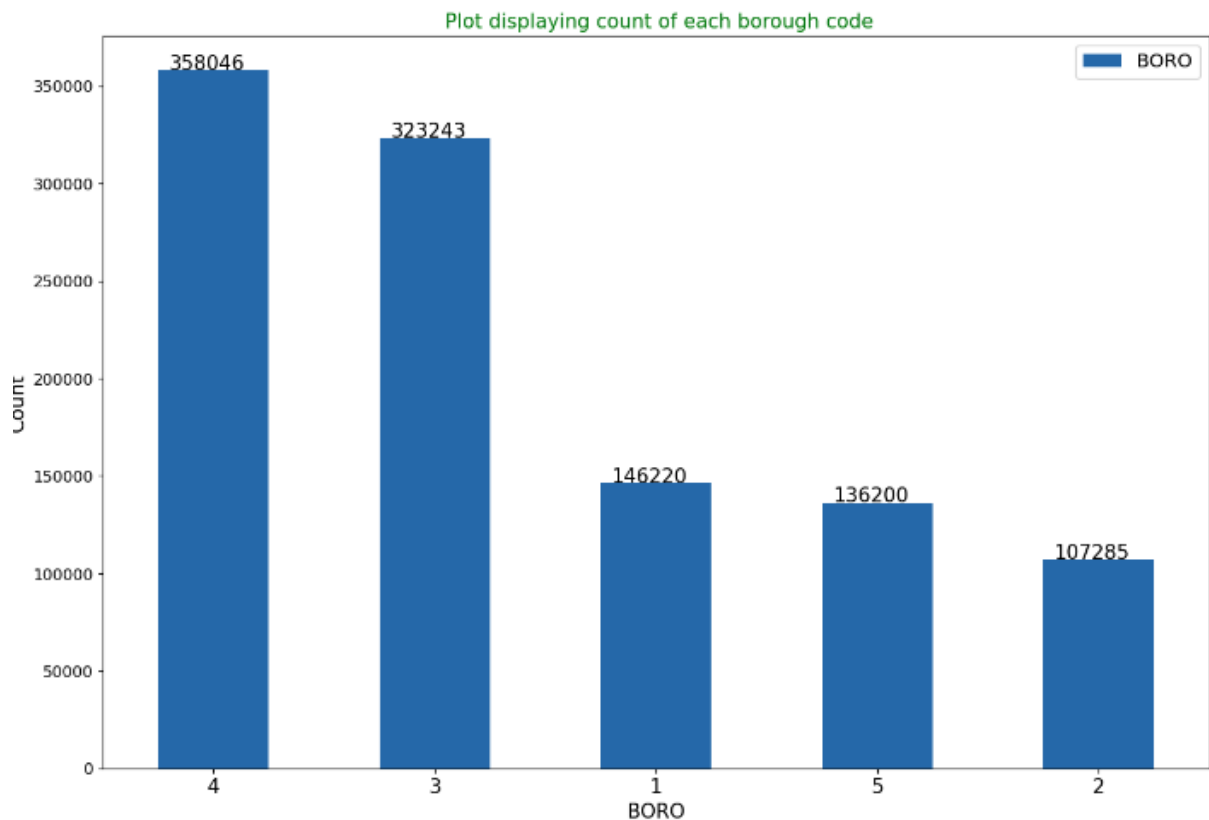
The Line charts, KDE and bar charts represented for each categorical and numerical field give some high-level information of fields of New York Property Valuation and Assessment Data.

1. Field Name: RECORD

Description: A categorical field containing unique positive integers for each record from 1 to 10,70994.

2. Field Name: BBLE

Description: This field is the concatenation of BORO, BLOCK, LOT, and EASEMENT fields, giving a unique number to each record. There are 10,70994 unique values of BBLE which is one unique value for each record.

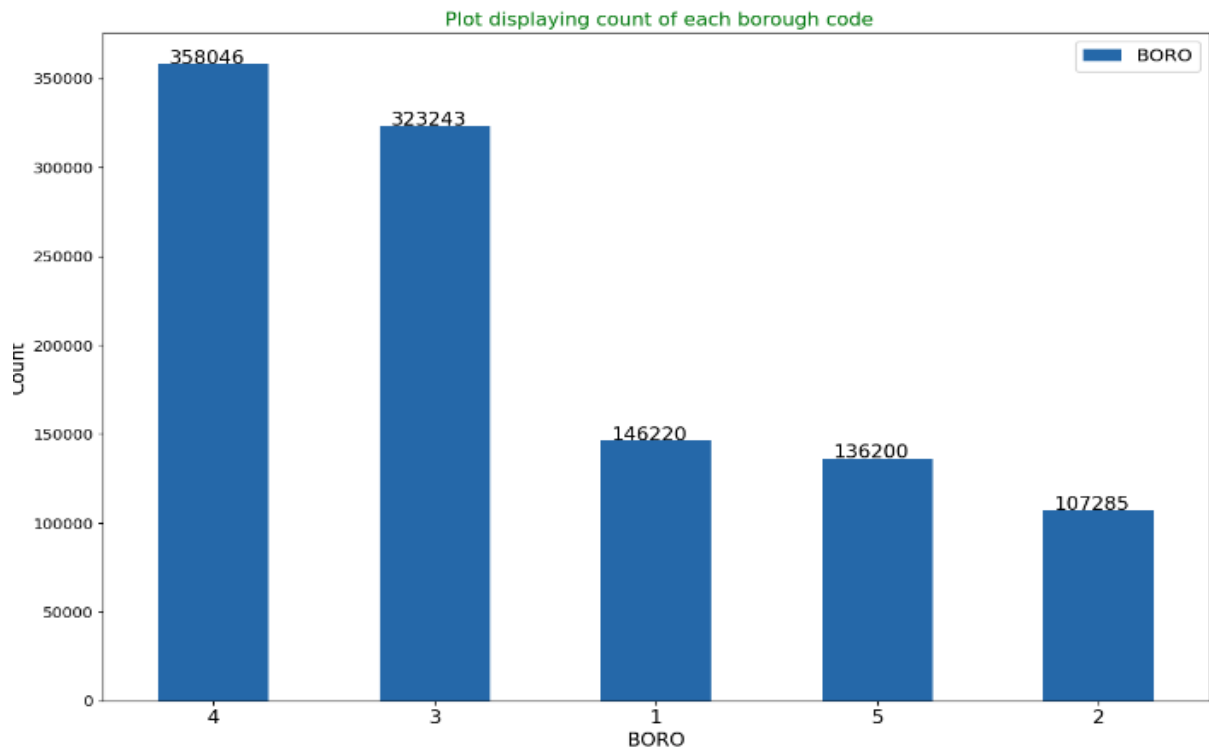


3. Field Name: BORO

Description: This field represents Borough codes which are as follows:

(1 = Manhattan, 2 = Bronx, 3 = Brooklyn, 4 = Queens, 5 = Staten Island)

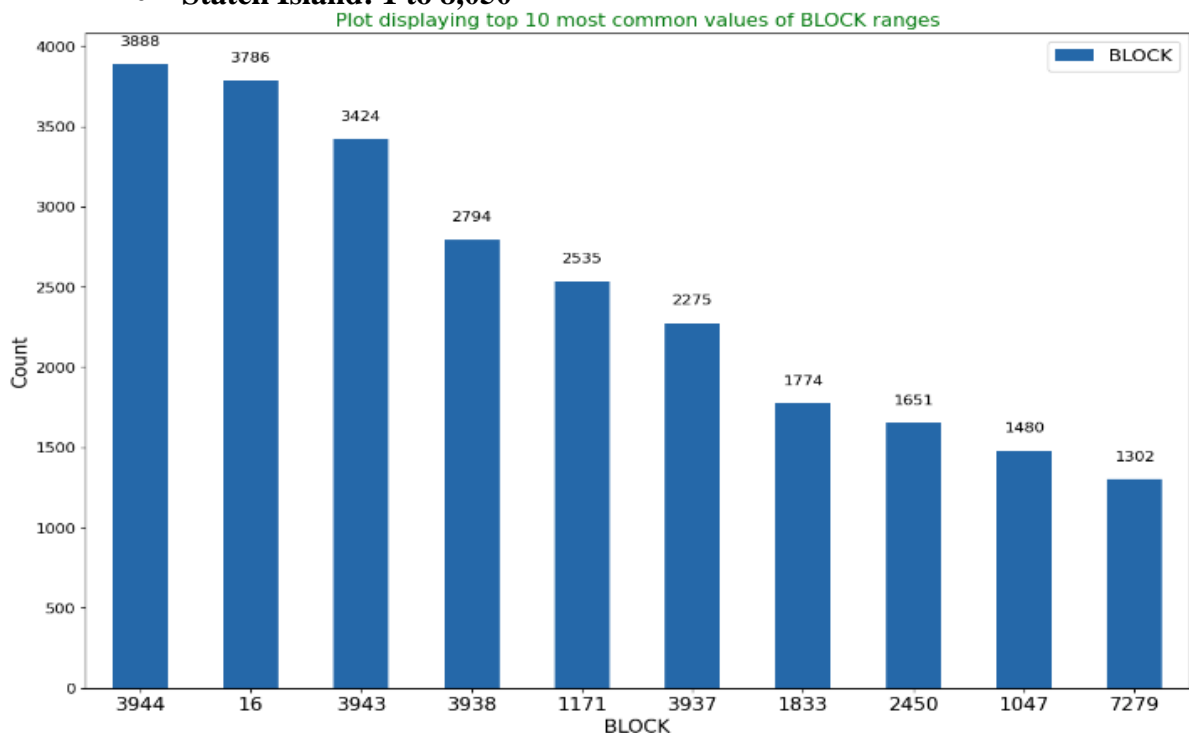
The bar chart below shows the count for each borough code in descending order of their count.



4. Field Name: BLOCK

Description: This field represents the valid block ranges by borough codes. This histogram plot shows the count of 10 most common values for Block ranges.

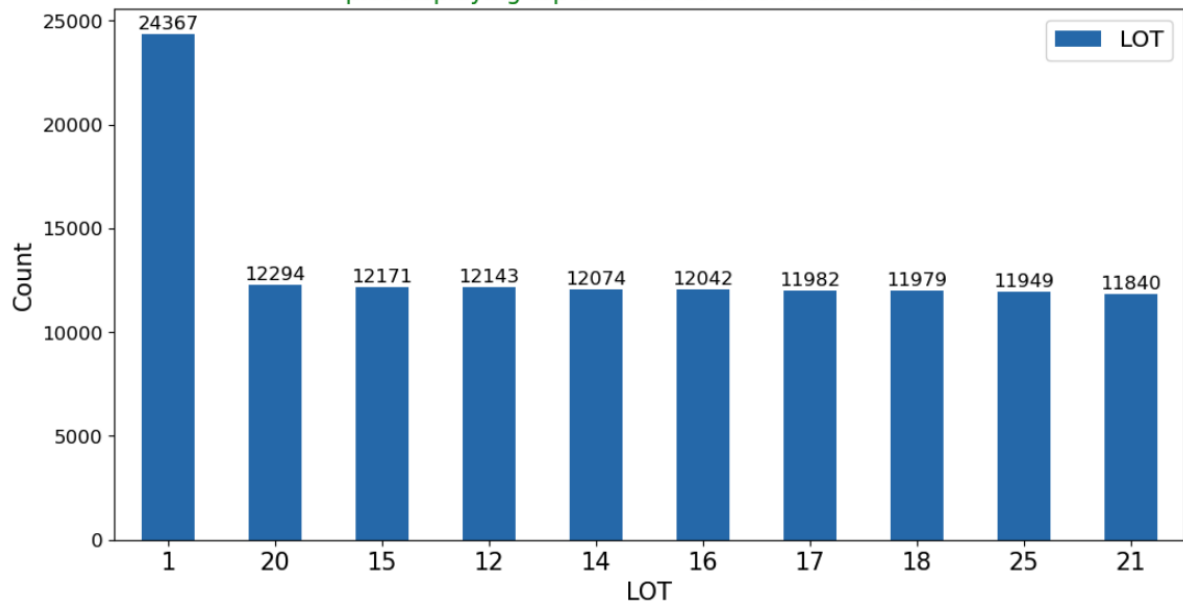
- **Manhattan:** 1 to 2,255
- **Bronx:** 2,260 to 5,958
- **Brooklyn:** 1 to 8,955
- **Queens:** 1 to 16,350
- **Staten Island:** 1 to 8,050



5. Field Name: LOT

Description: This field represents unique codes withing borough codes and ranges.

plot displaying top 10 most common values in LOT

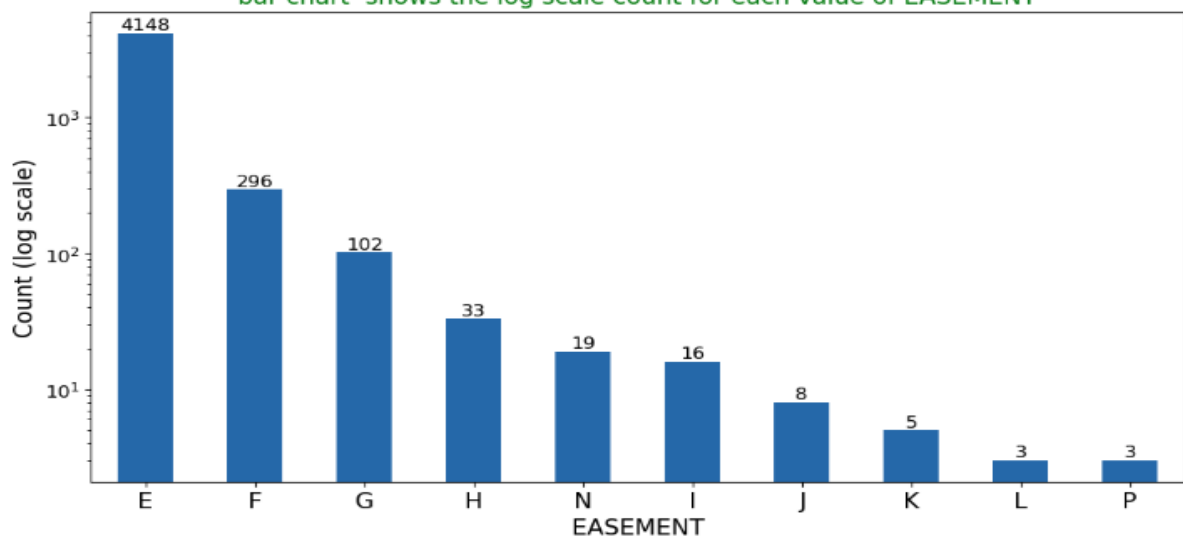


6. Field Name: EASEMENT

Description: This field represents a legal right to use another's land for a specific purpose.

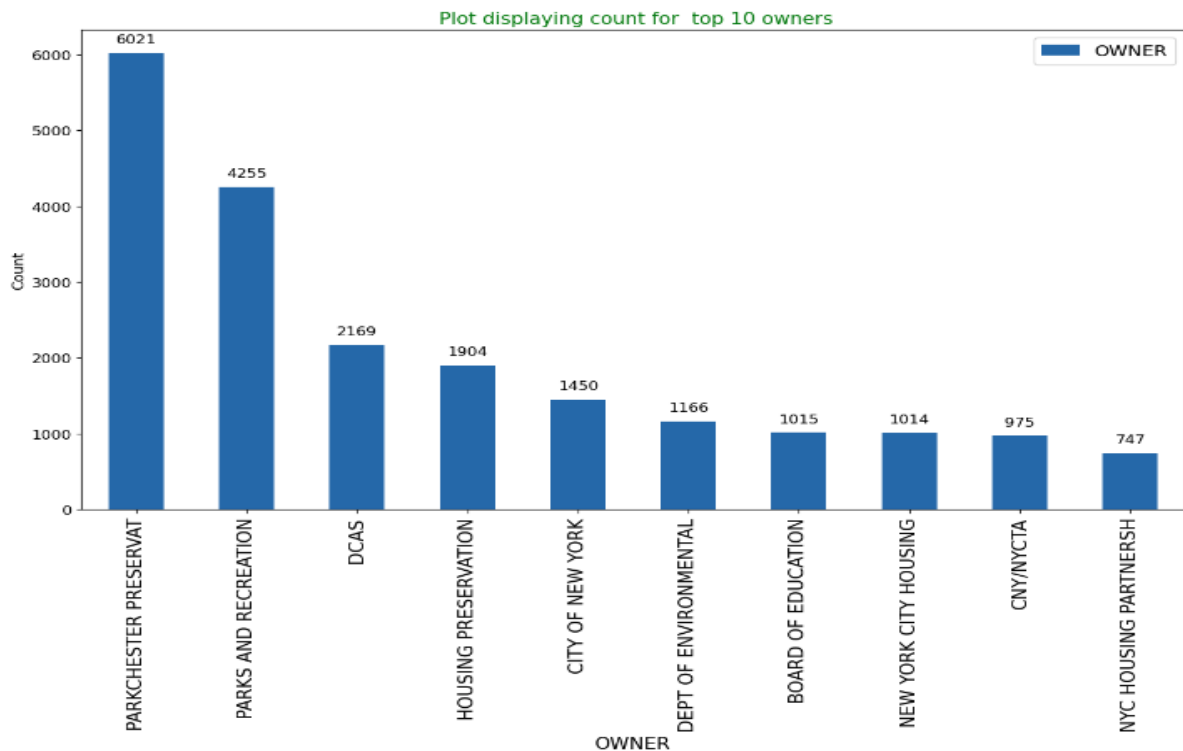
- **A:** Indicates the portion of the Lot that has an Air Easement
- **B:** Indicates Non-Air Rights
- **E:** Indicates the portion of the lot that has a Land Easement
- **F through M** Are duplicates of 'E'**N:** Indicates Non-Transit Easement
- **P:** Indicates Piers
- **R:** Indicates Railroads
- **S:** Indicates Street
- **U:** Indicates U.S. Government

bar chart shows the log scale count for each value of EASEMENT



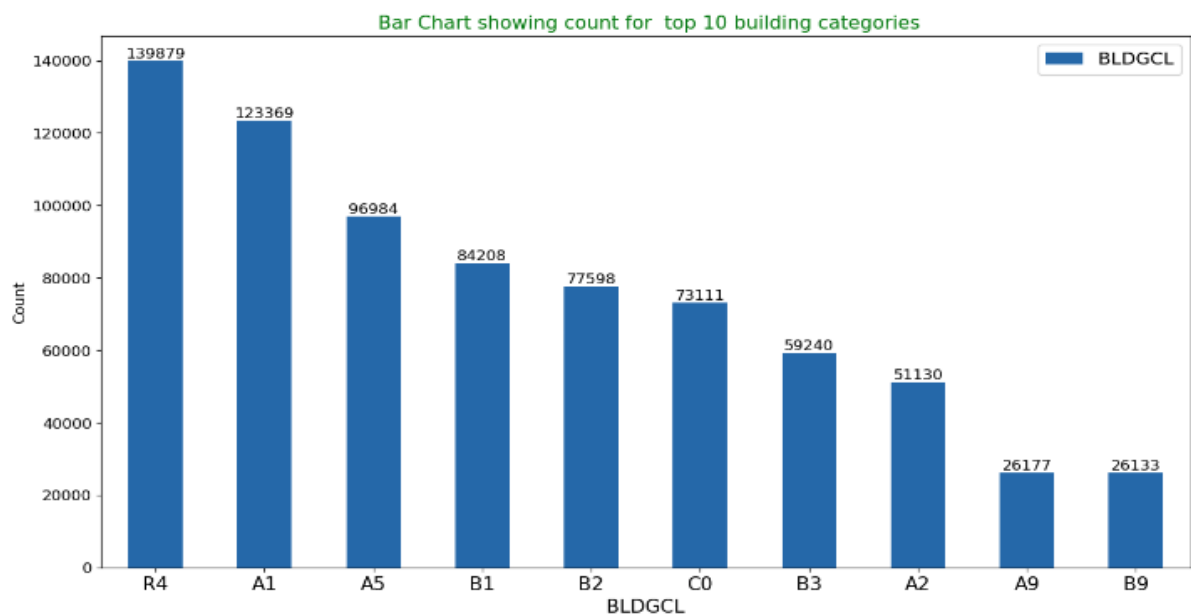
7. Field Name: OWNER

Description: This field represents the property owner's name, plot displaying count for top 10 owners.



8. Field Name: BLDGCL

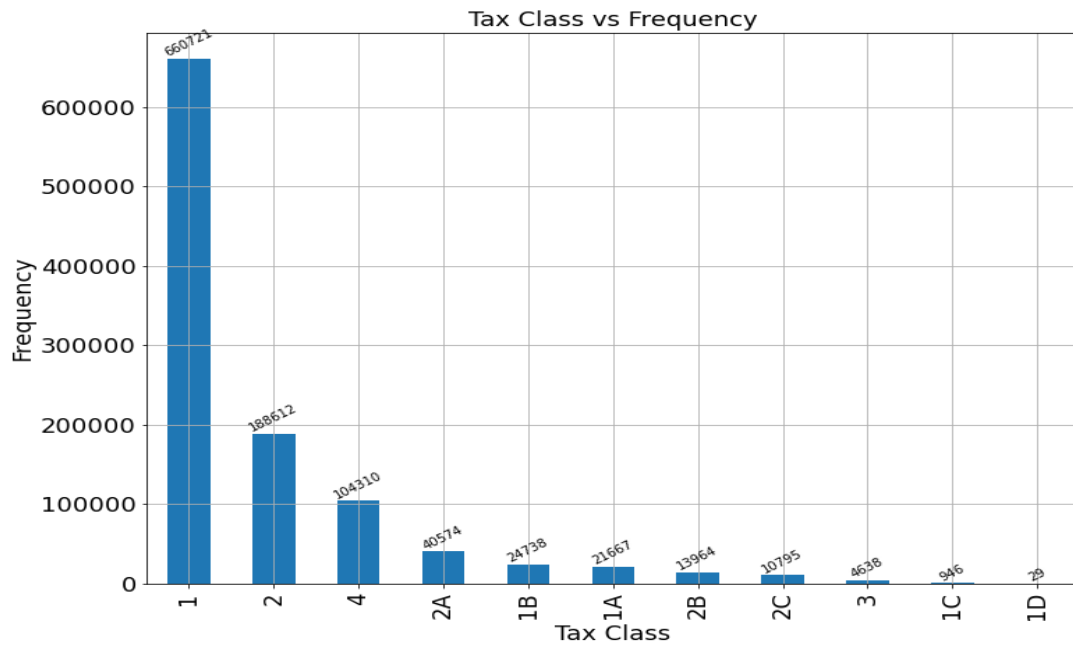
Description: This field indicates the building class, bar chart shows the count of the top 10 building class



9. Field name: TAXCLASS

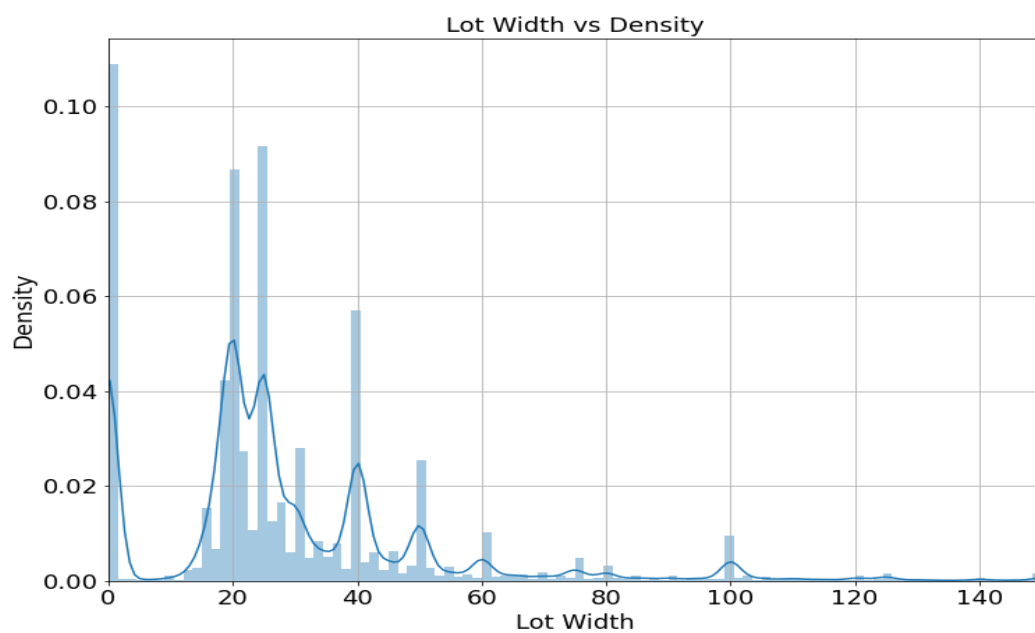
Description: This field represents the tax class of the property

1. 1 - 3 Unit Residence
2. Apartments, 2A = 4, 5, or 6 Units
3. Utilities
4. All Others



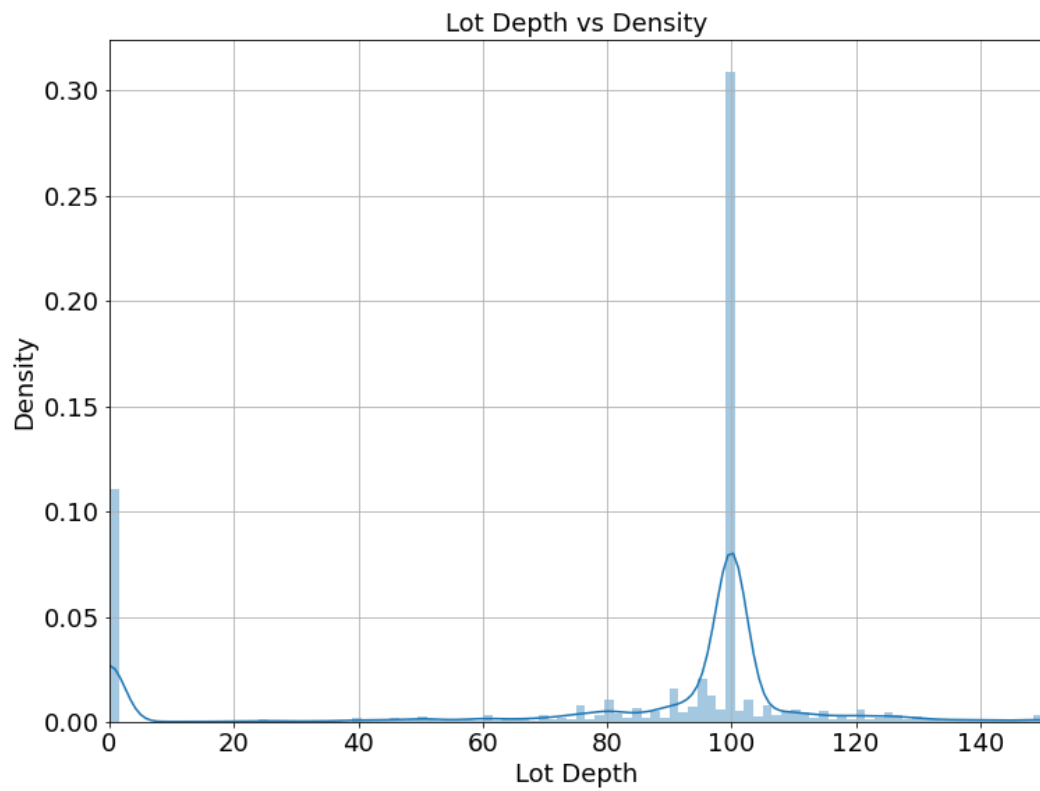
10. Field Name: LTFRONT

Description: It is the lot width. The below figure uses a logarithmic scale, and the lot width is limited to 150..



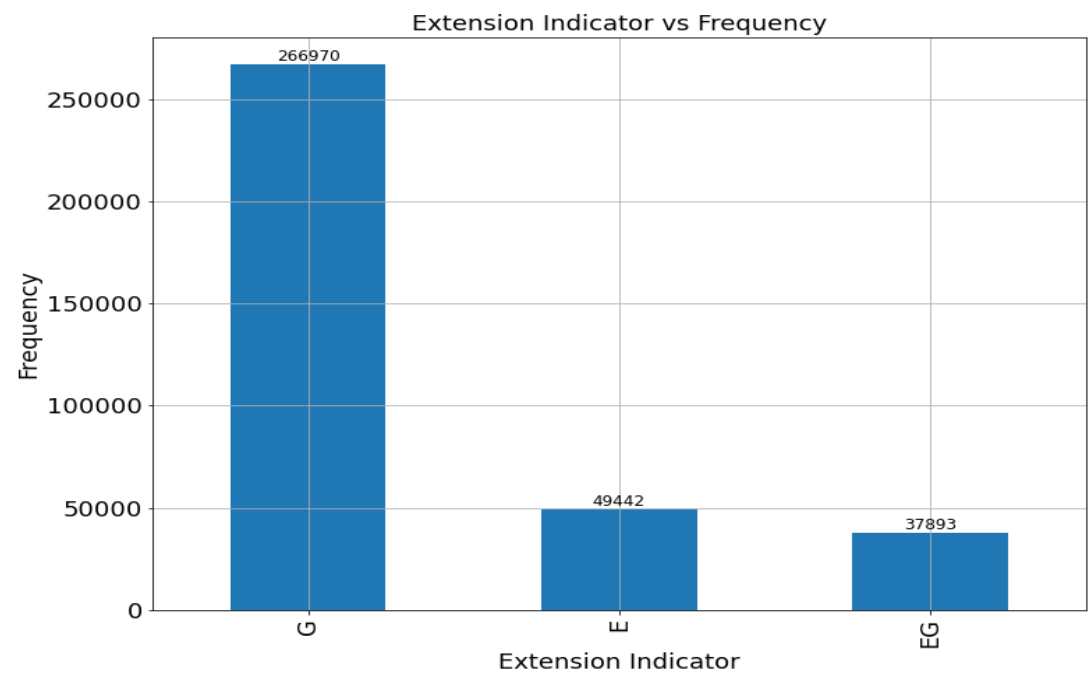
11.Field name: LTDEPTH: It is the Lot Depth.

Description: The below figure uses a logarithmic scale and the Lot Depth is limited to 150.



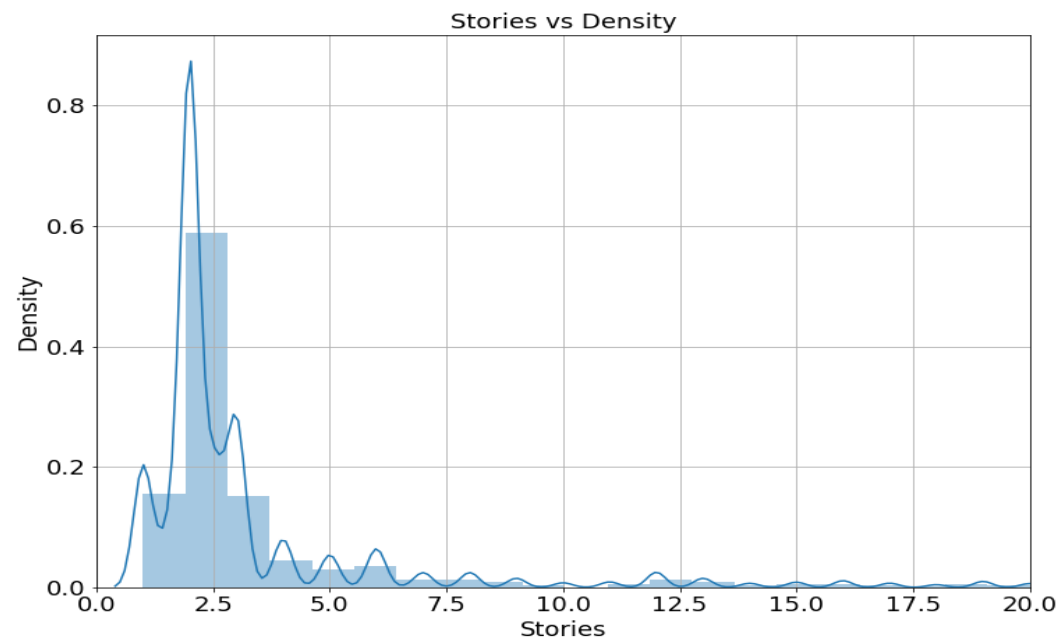
12. Field Name: EXT

Description: Extension Indicator



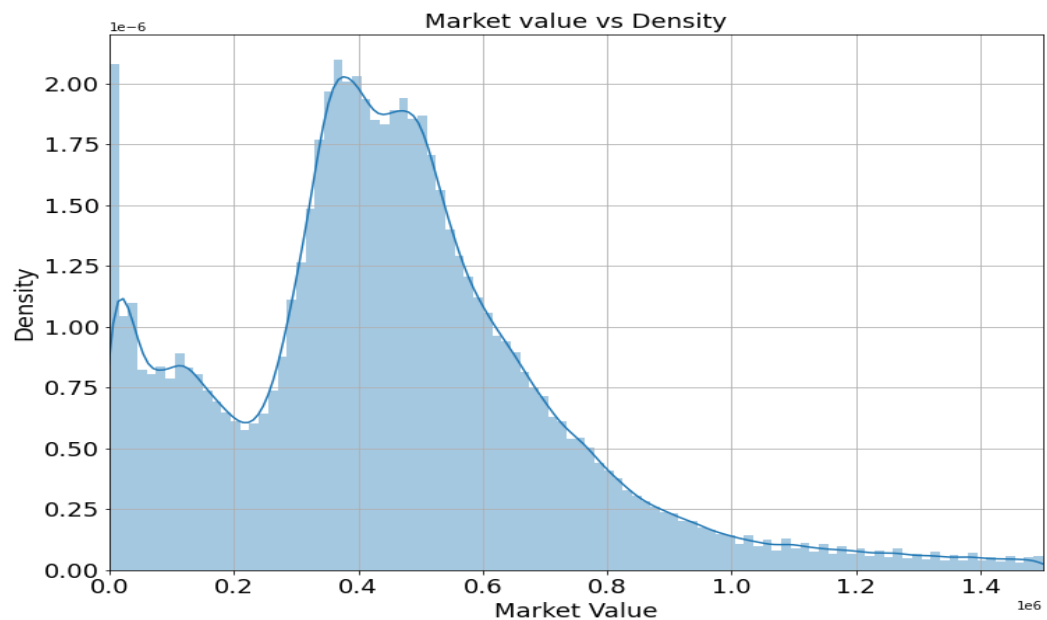
13. Field Name: Stories

Description: Number of Stories in Building.



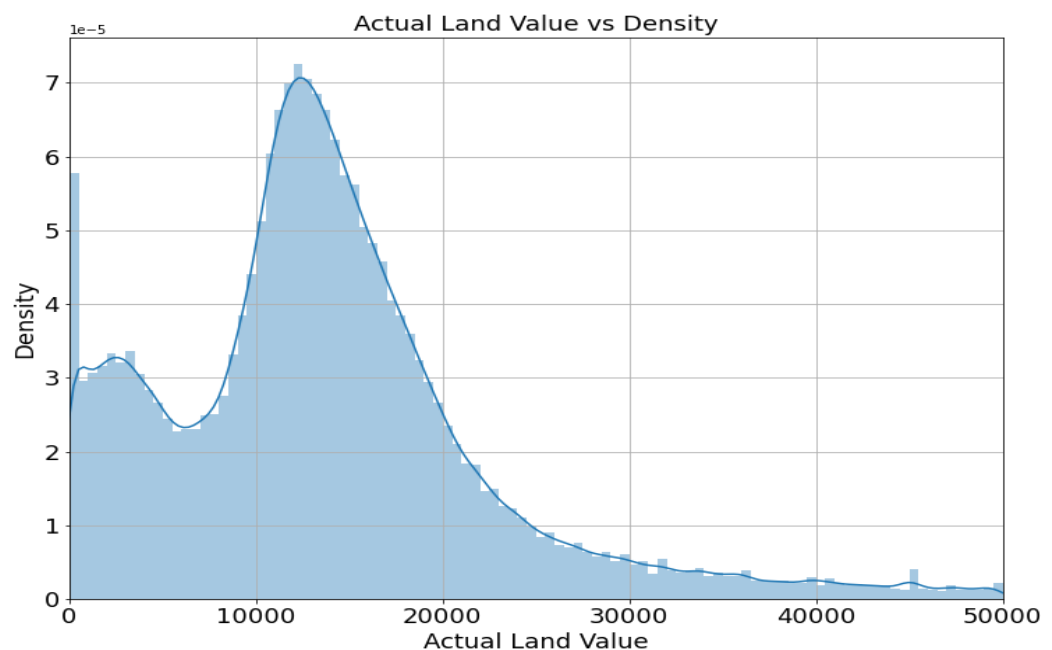
14. Field Name: FULLVAL

Description: The below figure uses a logarithmic scale and displays market value.



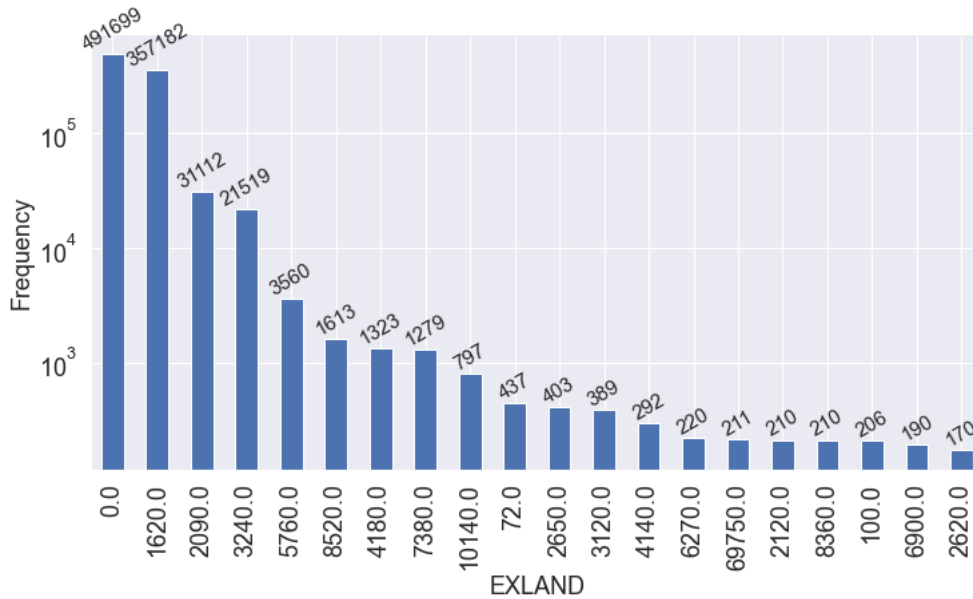
15. Field Name: AVLAND Actual Land Value

Description: This field represents the actual land value
The distribution displays up to a value of 50,000.



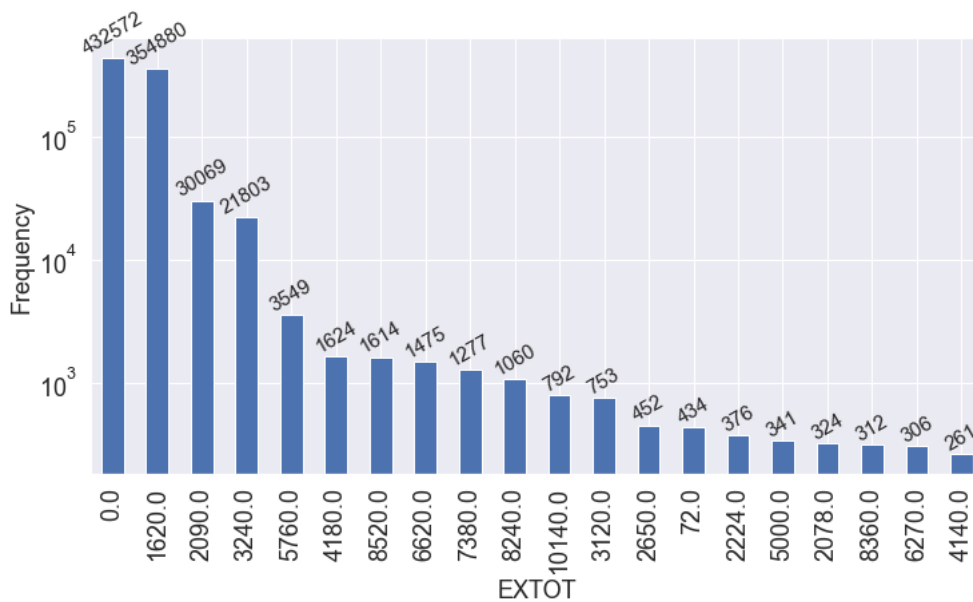
16.Field Name: EXLAND

Description: Merch Zip: The field “EXLAND” refers to the Actual Exempt Land Value. The distribution indicates the top 20 most common Actual Exempt Land Value.The most common value of the EXTOT is 0.0, the count is 491699.



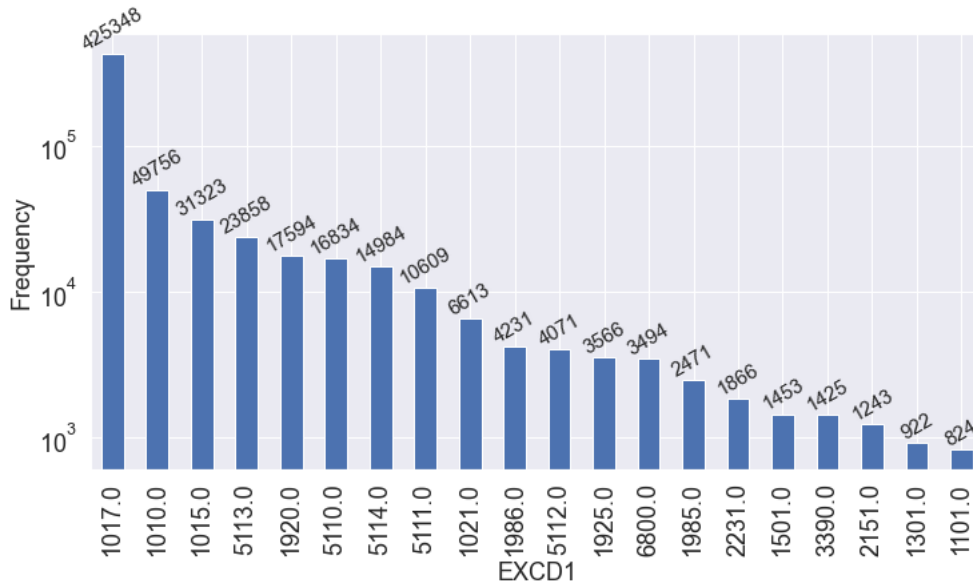
17.Field Name: EXTOT

Description: Merch Zip: The field “EXTOT” refers to the Actual Exempt Land Total. The distribution indicates the top 20 most common Actual Exempt Land Total used.The most common value of the EXTOT is 0.0, the count is 432572.



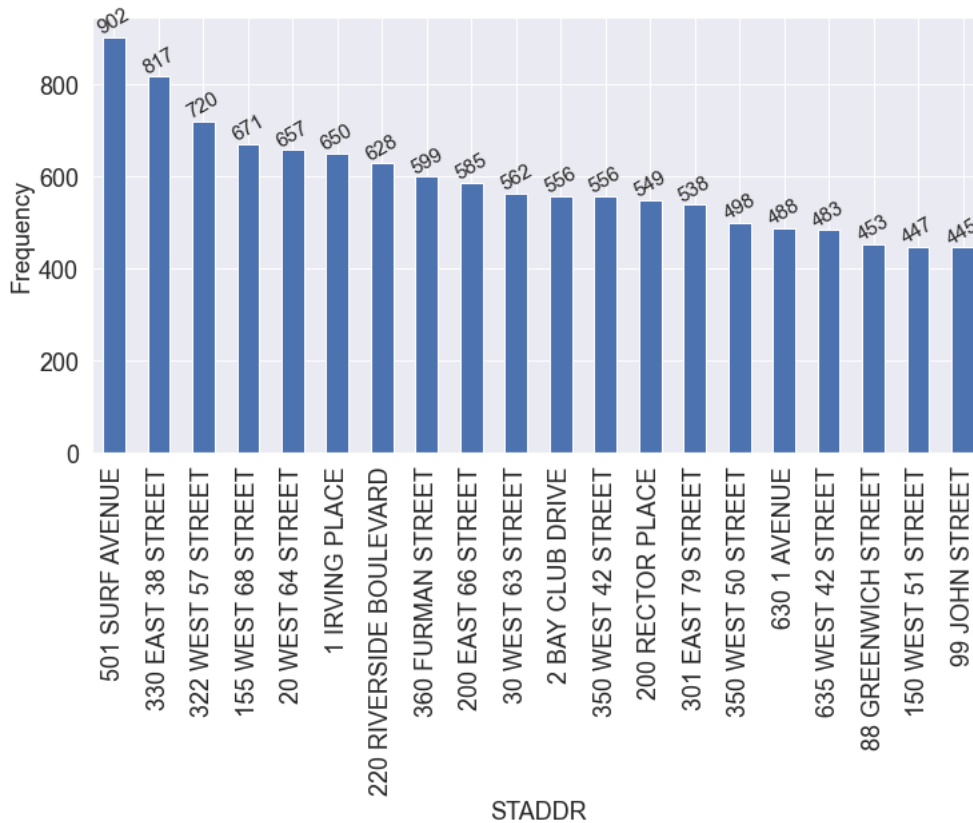
18.Field Name: EXCD1

Description: The field “EXCD1” refers to the Exemption Code 1. The distribution indicates the top 20 most common Exemption Code 1 used.The most common value of the EXCD1 is 1017.0, the count is 425348.



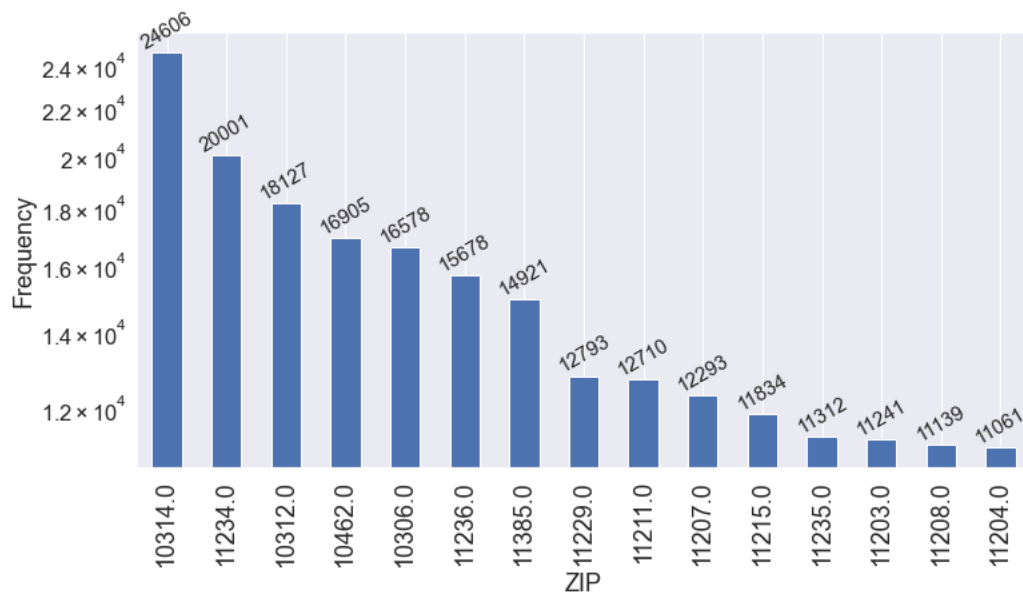
19.Field Name: STADDR

Description: The field “STADDR” refers to the Street Address of the property. The distribution indicates the top 20 most common STADDR used.The most common value of the STADDR is 501 Surf Avenue, the count is 902.



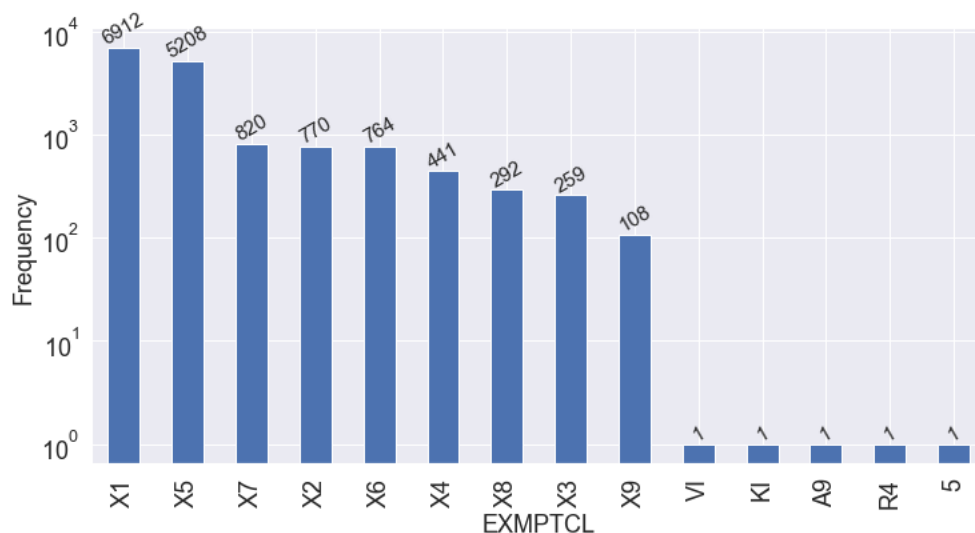
20. Field Name: ZIP

Description: The field “ZIP” refers to the Zip code of the property. The distribution indicates the top 15 most common Zip code used. The most common value of the Zip code is 10314, the count is 24606.



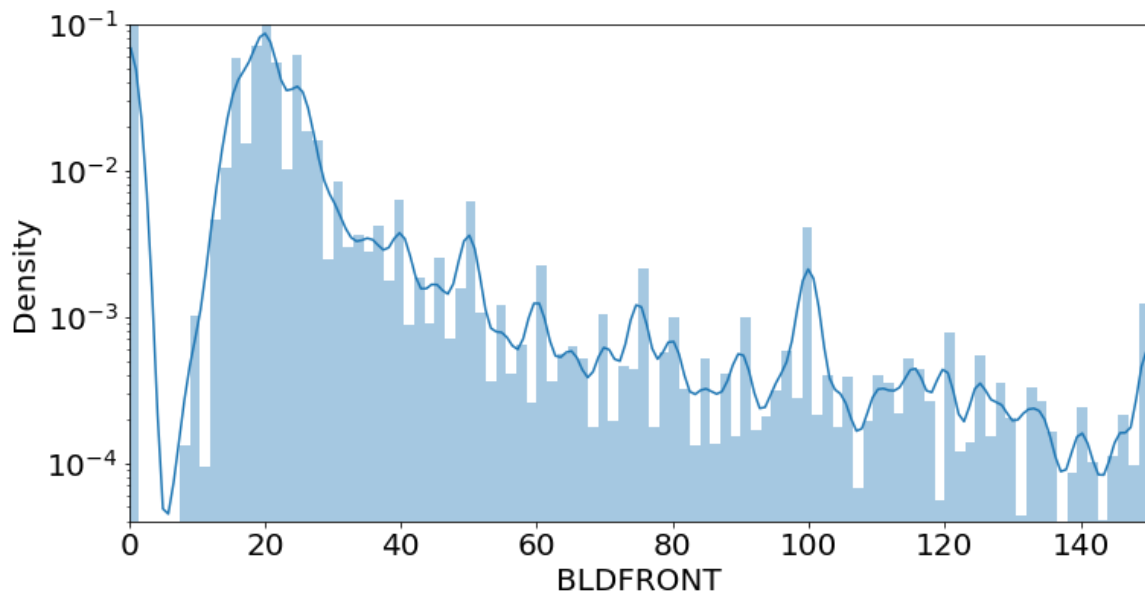
21. Field Name: EXMPTCL

Description: Merch Zip: The field “EXMPTCL” refers to the Exemption Class. The distribution indicates the 14 EXMPTCL used. The most common value of the EXMPTCL is X1, the count is 6912.



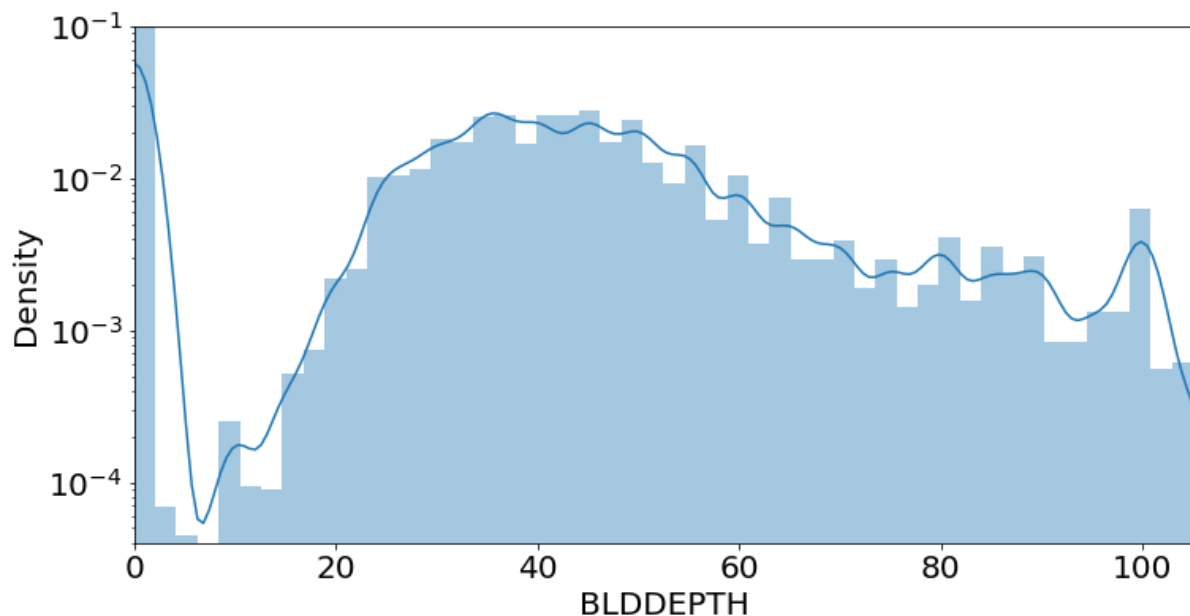
22.Field Name: BLDFRONT

Description: The field “BLDFRONT” refers to the building Width of the property. The distribution displays up to a Building Width of 150.The most common value of the BLDFRONT” is 0, the count is 228815.



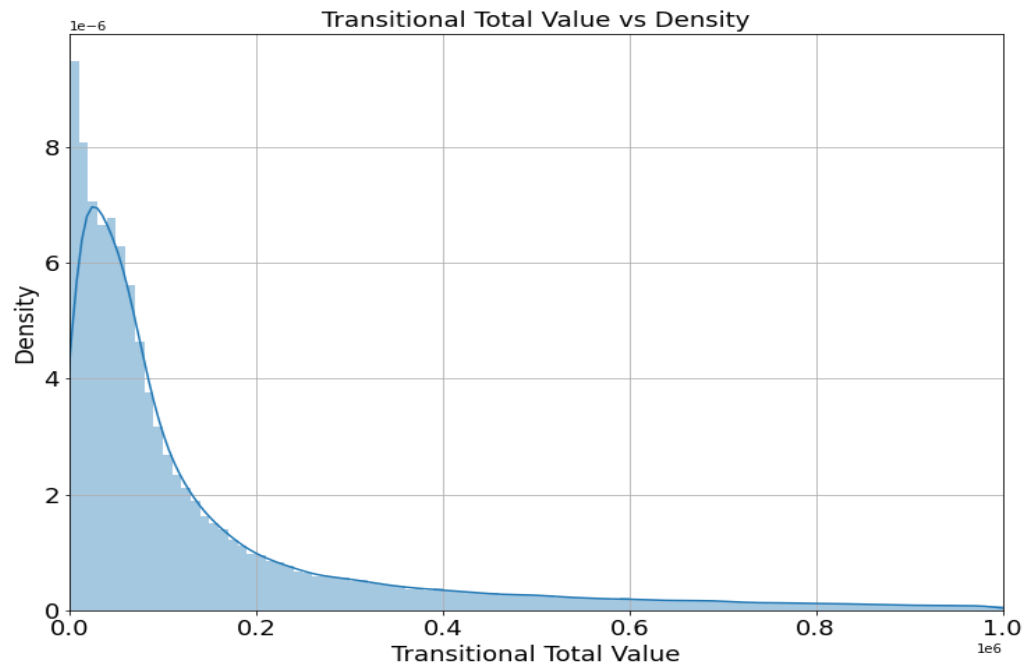
23.Field Name: BLDDEPTH

Description: The field “BLDDEPTH” refers to the building depth. The distribution displays Density up to a Building Depth of 100.The most common value of the Zip code is 0, the count is 228853.



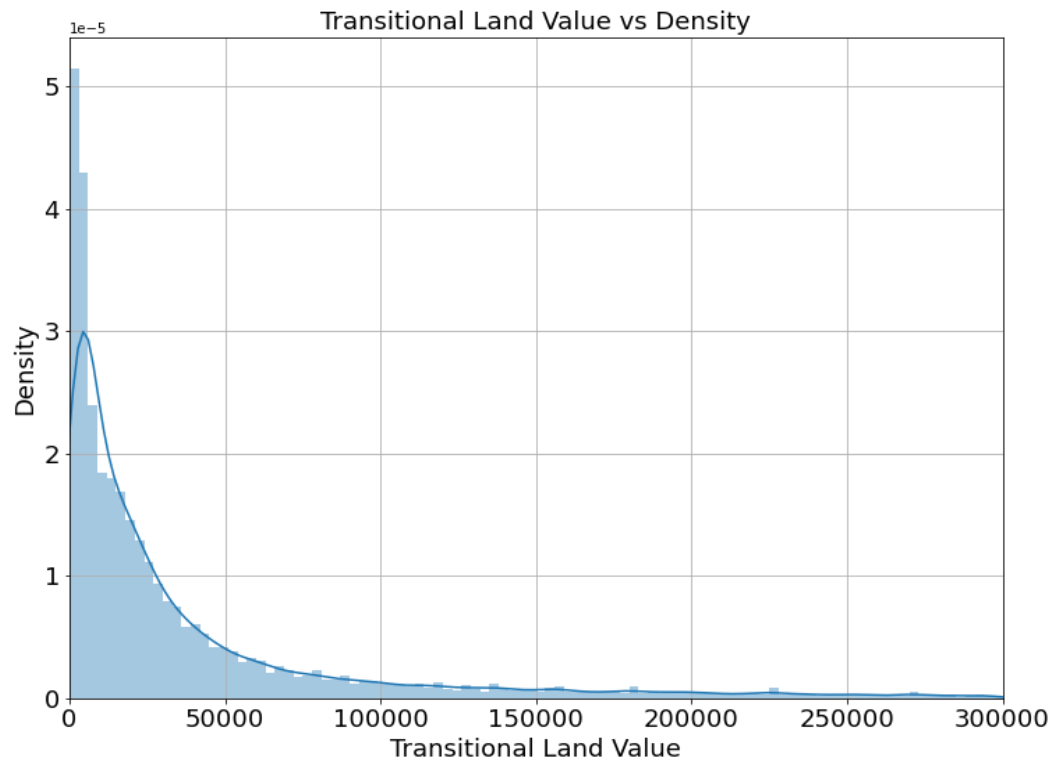
24.Field Name: AVLAND2

Description: This field represents the transactional land value



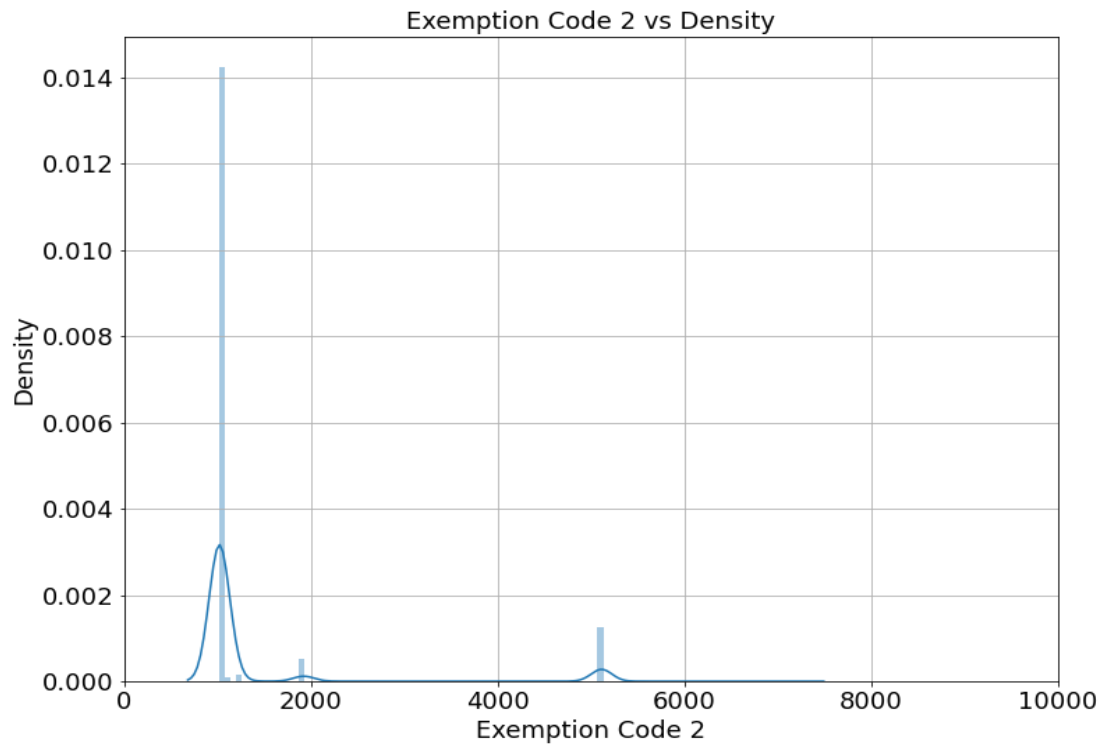
25.Field Name: AVTOT2

Description: This field represents the Transitional Total Value



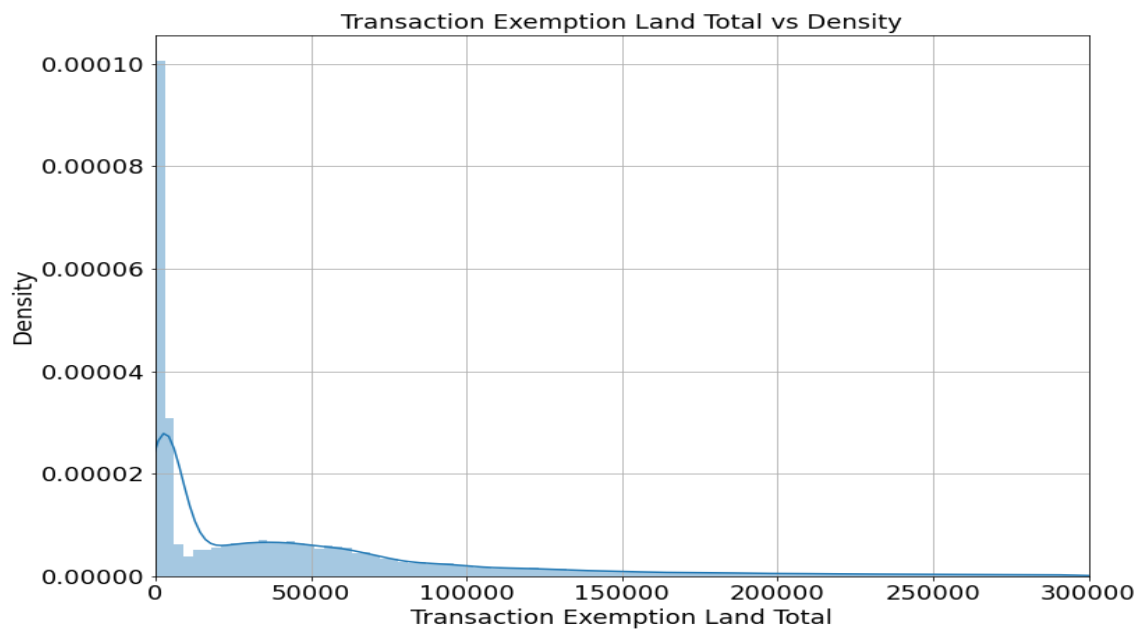
26.Field Name: EXLAND2

Description: This field represents the Transitional Exemption Land Value.



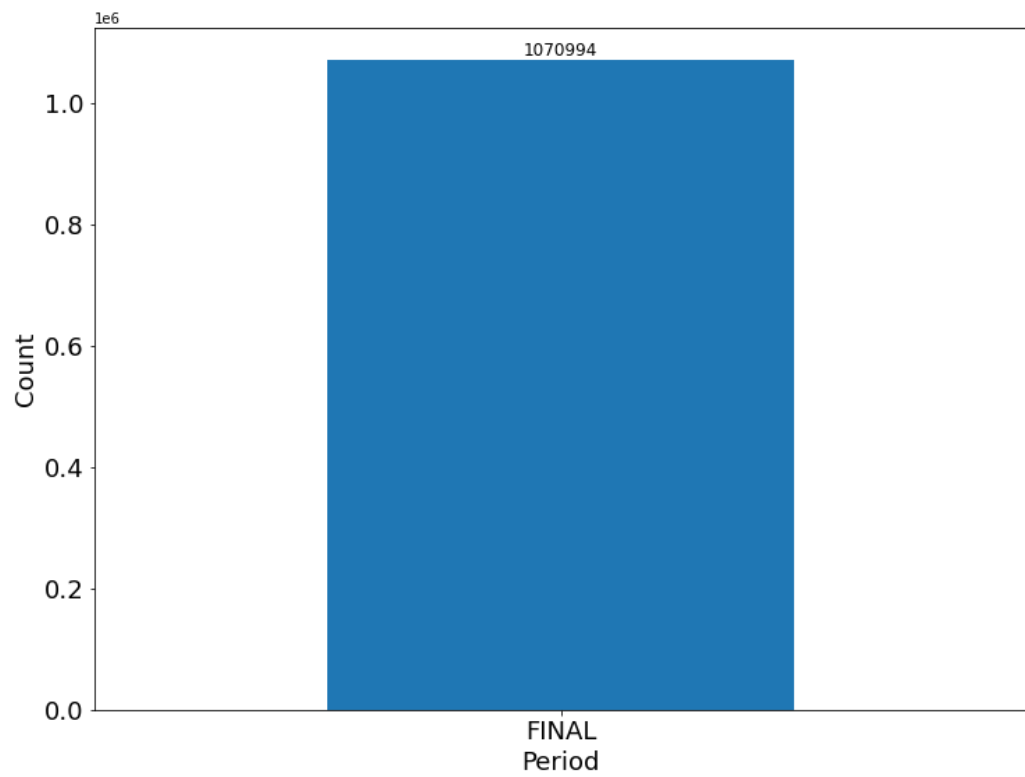
27.Field Name: EXTOT2

Description: This field represents the Transitional Exemption Land Total.



28.Field Name: PERIOD

Description: Has only 1 Value



29.Field Name: YEAR

Description: Has only 1 value.

