**Field 1**

**Field Name:** BBLE

**Description:**

BBLE is a categorical variable representing the concatenation of AV\_BORO (boro code), AV\_BLOCK (block code), AV\_LOT (lot code), AV\_EASEMENT (describes easement). It has 11 length digits.

**Unique Values:**

BBLE has 1070994 unique values ranging from 1000010101 to 5080500094, meaning each value in BBLE is unique.

**Field 2**

**Field Name:** B

**Description:**

B is a categorical variable representing boro code.

**Unique Values:**

B has 5 unique values ranging from 1 to 5. There are 0 properties with the B of 0 in the dataset. No missing values exist. The statistics and distribution excluding 0 records are shown below:

|  |  |
| --- | --- |
| 1 | 146220 |
| 2 | 107285 |
| 3 | 323243 |
| 4 \* most common | 358046 |
| 5 | 136200 |

**Field 3**

**Field Name:** BLOCK

**Description:**

BLOCK is a categorical variable representing block code.

**Unique Values:**

BLOCK has 13984 unique values. There are 0 properties with the BLOCK of 0 in the dataset. No missing values exist.

**Field 4**

**Field Name:** LOT

**Description:**

BLOCK is a categorical variable representing lot code.

**Unique Values:**

LOT has 6366 unique values. There are 0 properties with the LOT of 0 in the dataset. No missing values exist.

**Field 5**

**Field Name:** EASEMENT

**Description:**

EASEMENT is a categorical variable describing easement.

**Unique Values:**

LOT has 13 unique values. There are 0 properties with the LOT of 0 in the dataset. There are 1066358 missing values. The statistics and distribution excluding 0 records are shown below:

|  |  |
| --- | --- |
| E \* most common | 4148 |
| F | 296 |
| G | 102 |
| H | 33 |
| N | 19 |
| I | 16 |
| J | 8 |
| K | 5 |
| P | 3 |
| L | 3 |
| M | 2 |
| U | 1 |
| I | 16 |

**Field 6**

**Field Name:** OWNER

**Description:**

OWNER is a categorical variable describing name of the owner

**Unique Values:**

OWNER has 863348 unique values. There are 31745 missing values. The top 10 categories are:

A picture containing bottle, text

Description generated with high confidence

**Field 7**

**Field Name:** BLDGCL

**Description:**

BLDGCL is a categorical variable describing building class. This variable is corelated to TAXCLASS.

**Unique Values:**

BLDGCL has 200 unique values. There are no missing values. The top 10 categories are:

|  |  |
| --- | --- |
| R4 | 139879 |
| A1 | 123369 |
| A5 | 96984 |
| B1 | 84208 |
| B2 | 77598 |
| C0 | 73111 |
| B3 | 59240 |
| A2 | 51130 |
| A9 | 26177 |
| B9 | 26133 |

**Field 8**

**Field Name:** TAXCLASS

**Description:**

TAXCLASS is a categorical variable describing Current Property Tax Class Code

**Unique Values:**

TAXCLASS has 11 unique values. There are 0 properties with the TAXCLASS of 0 in the dataset. There are 0 missing values. The statistics and distribution excluding 0 records are shown below:

|  |  |
| --- | --- |
| 1 | 660721 |
| 2 | 188612 |
| 4 | 104310 |
| 2A | 40574 |
| 1B | 24738 |
| 1A | 21667 |
| 2B | 13964 |
| 2C | 10795 |
| 3 | 4638 |
| 1C | 946 |
| 1D | 29 |

**Field 9**

**Field Name:** LTFRONT

**Description:**

LTFRONT is a numerical variable describing lot frontage in feet.

**Unique Values:**

LTFRONT has 1297 unique values ranging from 0 to 9999. There are 169108 properties with the LTFRONT of 0 in the dataset. No missing values exist. The statistics and distribution excluding 0 records are shown below:

|  |  |
| --- | --- |
| Max | 9999 |
| Min | 1 |
| Medium | 25 |
| Mean | 43.5 |
| SD | 78.8 |

**Field 10**

**Field Name:** LTDEPTH

**Description:**

LTDEPTH is a numerical variable describing lot depth in feet.

**Unique Values:**

LTDEPTH has 1370 unique values ranging from 0 to 9999. There are 170128 properties with the LTDEPTH of 0 in the dataset. No missing values exist. The statistics and distribution excluding 0 records are shown below:

|  |  |
| --- | --- |
| Max | 9999 |
| Min | 1 |
| Medium | 100 |
| Mean | 105.64 |
| SD | 71.87 |

**Field 11**

**Field Name:** EXT

**Description:**

EXT is a categorical variable describing extension

**Unique Values:**

EXT has 3 unique values. There are 716689 missing values. The distribution excluding is shown below:

|  |  |
| --- | --- |
| G | 266970 |
| E | 49442 |
| EG | 37893 |

**Field 12**

**Field Name:** STORIES

**Description:**

STORIES is a numerical value describing the number of floors of the building.

**Unique Values:**

STORIES has 112 unique values ranging from 1 to 119. There are 0 properties with the STORIES of 0 in the dataset. 56264 missing values exist. The statistics and distribution excluding 0 records are shown below:

|  |  |
| --- | --- |
| Max | 119 |
| Min | 1 |
| Medium | 2 |
| Mean | 5.006918e+00 |
| SD | 8.365707e+00 |

**Field 20**

**Field Name:** ZIP

**Description:**

ZIP is a categorical variable describing Postal Zip code of the property.

**Unique Values:**

ZIP has 197 unique values. There are 29890 missing values. The top 10 categories are:

|  |  |
| --- | --- |
| 10314 | 24606 |
| 11234 | 20001 |
| 10312 | 18127 |
| 10462 | 16905 |
| 10306 | 16578 |
| 11236 | 15678 |
| 11385 | 14921 |
| 11229 | 12793 |
| 11211 | 12710 |
| 11207 | 12293 |

**Field 21**

**Field Name:** EXMPTCL

**Description:**

EXMPTCL is a categorical variable describing Exempt Class used for fully exempt properties only.

**Unique Values:**

EXMPTCL has 15 unique values. There are 1055415 missing values. The distribution are:

|  |  |
| --- | --- |
| X1 | 6912 |
| X5 | 5208 |
| X7 | 820 |
| X2 | 770 |
| X6 | 764 |
| X4 | 441 |
| X8 | 292 |
| X3 | 259 |
| X9 | 108 |
| A9 | 1 |
| R4 | 1 |
| 5 | 1 |
| KI | 1 |
| VI | 1 |

**Field 29**

**Field Name:** PERIOD

**Description:**

PERIOD is a categorical variable describing ?

**Unique Values:**

PERIOD has 1 unique value, ‘FINAL’. There are 0 missing values.

**Field 30**

**Field Name:** YEAR

**Description:**

YEAR is a categorical variable describing ?

**Unique Values:**

YEAR has 1 unique value, ‘2010/11’. There are 0 missing values.

**Field 31**

**Field Name:** VALTYPE

**Description:**

VALTYPE is a categorical variable describing ?

**Unique Values:**

VALTYPE has 1 unique value, ‘AC-TR’. There are 0 missing values.