Table of Contents

Basic Description1

Assessment2

ImprovementDwelling4

ImprovementInterior7

Property9

PropertyAssessmentHistory16

RealEstateAssessmentPaymentHistory22

SalesHistory24

**Basic Description**

Assessment: Historical assessment information for a parcel (identified by RealEstatePropertyCode) in Arlington County.

ImprovementDwelling: Information, such as year built and heating type, about each dwelling on a parcel (identified by RealEstatePropertyCode) in Arlington County.

ImprovementInterior: Information about the interior of a dwelling, such as number of bedrooms or bathrooms, on a parcel (identified by RealEstatePropertyCode) in Arlington County.

Property: The basic information, primary address, owners, etc., about a parcel (identified by RealEstatePropertyCode) in Arlington County.

PropertyAssessmentHistory: Information about the assessment history for a parcel

RealEstateAssessmentPaymentHistory: Information about the assessment and tax payment history of a parcel (identified by RealEstatePropertyCode)

Sales History: Contains one instance for each time the ownership status of a parcel has changed.

|  |  |  |
| --- | --- | --- |
| **Table** | **Observations** | **Years** |
| Assessment | 410,570 | 2009-July 2015 |
| ImprovementDwelling | 60,730 | All records |
| ImprovementInterior | 134,698 | All records |
| Property  | 68,949 | All records |
| PropertyAssessmentHistory | 1,891,156 | All records |
| RealEstateAssessmentPaymentHistory | 5,000,549 | All records |
| SalesHistory | 42,552 | 2009-July 2015 |

**Assessment**

**\*\*\* Needs Cleaning**

Unit of observation: A property's assessment event

8,5047 (21%) are duplicated rows

"Assessment.Key”

* The ID for that assessment
* Factor
* 85,047 are duplicated (21%)
* None missing
	+ 2,662 of properties listed in Properties Table have no assessment history

"provalLrsnId"

* Proval land record serial number: identification for all properties Factor
	+ Follows "Real.Estate.Property.Code”
* None Missing

"Real.Estate.Property.Code”

* Property Code (parcel) for the Property assessed.
* Factor
* 83% are duplication
	+ 99% of the properties appear more than once
* None Missing

"Assessment.Change.Reason.Type.Dsc"

* Explanation of why assessment was changed
* Levels: "01- Annual”,"02- Permit”,"03- Board of Equalization","04- Court Order","05- Review”,"09- New Construction","16- Tax to Exempt”,"18- Exempt to Tax"
* 199 missing

"Assessment.Date"

* Date that assessment took place
* YYYY-MM-DD
* Invalid entries
	+ YYYY-01-01 for all years have high frequencies as compared to other days
* None missing

"Improvement.Value.Amount"

* Value of the improvements -- an "improvement' usually means a home/house/structure of some type.
* Numeric
* 5% coded as 0 (No units on land?)
* None missing

"Land.Value.Amount”

* Assessed value of the land
* Numeric
* 12% as 0
* None missing

"Total.Value.Amount"

* Total Assessed Value of the property
* Numeric
* Invalid entries: 331 are listed at 0
	+ In active properties

**ImprovementDwelling**

**\*\*\* Needs Cleaning**

Unit of observation: dwelling

66 are duplicated rows

"propertyKey”

* Unknown Definition
	+ Follows "realEstatePropertyCode”
* 167 are duplicated
* None Missing
	+ 8,386 (12%) of properties listed in Properties Table have no Improvement Dwelling information

"realEstatePropertyCode”

* Property Code (parcel) for the Property assessed.
* Factor
* 167 are duplicated
	+ 160 properties appear more than once (multiple dwellings on same property)
* None missing

"provalLrsnId”

* Proval land record serial number: identification for all properties
	+ Follows "Real.Estate.Property.Code”
* None Missing

"dwellingKey”

* ID for the dwelling
* 66 are duplicates
* These have all the same values

"extensionNbr”

* Number of bedrooms in the improvement
* Levels: "R01","R02","R03","R04","R05",”R06",”R07","R10"
* None missing

"basementFinishedRecRoomSquareFeetQty"

* Area of Finished basement Rec Room
* Numeric
* 60% Missing
* Invalid entry: 1 -240 value

"basementRecRoomTypeDsc”

* Description of basement Rec Room
* Levels: "Flr,ceil fin”,"Flr,ceil,wall fin”,"Flr,ceil,wall fin,ptn”,"Flr,ceil,wall fin,ptn,blt-in”,"None"
* 164 are missing

"coolingTypeDsc”

* Type of cooling system
* Levels: "Central air”, "Forced Cool Air", "None”, "Package unit”, "Window units"
* 35 missing

"dwellingTypeDsc”

* Description of unit in regards to unit type
* 24 different levels
* 105 are missing

"heatingTypeDsc”

* Type of heating system
* 24 different levels
* 13 missing

"convertedExtraLivingUnitCnt”

* Unknown definition
* Levels: 0,1,2,3,4
	+ 97% 0’s

"designedExtraLivingUnitCnt”

* Unknown definition
* Levels: “0,1,2,3,”
	+ 99.6% 0’s

dwellingYearBuiltDate”

* The year the unit was build
* Invalid entries: 68 are older than 1800 (to include 0s)
* None missing
	+ 78 "realEstatePropertyCode" have different "propertyYearBuilt" and “dwellingYearBuilt"
	+ 1,865 "realEstatePropertyCode" with "propertyYearBuilt" have no “dwellingYearBuilt”
	+ 0 "realEstatePropertyCode" with a "dwellingYearbuilt" has no “propertyYearBuilt”

"extraFixtureCnt”

* Unknown definition
* Levels:”0,1,2,3,4,5,6,7”
	+ 98% are 0’s.

"storiesQuantityDsc”

* Number of stories in the unit
* Inconsistent coding (e.g. 1.5, 1.5+, 1.50, 1.50+)
* Invalid entries: <1
* 28 missing

"finishedStorageAreaSquareFeetQty”

* Area of finished storage area
* All missing

"storageAreaSquareFeetQty”

* Area of storage in the unit
* 56% missing

“fireplaceCnt"

* Number of fireplace in the unit
* Factor: 0 - 10
* None missing

**ImprovementInterior**

**\*\*\* Needs Cleaning**

Unit of observation: individual floors of the building on the property

19402 (14%) are duplicated rows

"propertyKey”

* Unknown Definition
	+ Follows "realEstatePropertyCode”
* 55% are duplicated
* None Missing

"realEstatePropertyCode”

* Property Code (parcel) for the Property assessed.
* Factor
* 55% are duplicated
	+ 65% properties appear more than once
		- The key is FloorKey or a combination of RealEstatePropertyCode, DwellingKey, ExtensionNbr, FloorNbr (note this can have values like B for basement and 2.5 for a second floor with sloped ceiling).
* None missing
	+ 8,321 (12%) of properties listed in Properties Table have no Improvement Interior information

"provalLrsnId”

* Proval land record serial number: identification for all properties
	+ Follows "Real.Estate.Property.Code”
* None Missing

"dwellingKey”

* ID for the dwelling
* None missing

"extensionNbr”

* Number of bedrooms in the improvement
* Levels: "R01","R02","R03","R04","R05",”R06",”R07","R10"
* None missing

"baseAreaSquareFeetQty"

* Square footage of housing unit (?)
* Numeric
* Validity depends on definition (ranges from 1 to 41000)

"bedroomCnt”

* Number of bedrooms
* Numeric (ranges tom 1 to 10)
* None missing

"finishedAreaSquareFeetQty”

* Square feet of finished area in the housing unit ???
* None missing
* Invalid entries: 25% listed as 0, one listed at -213
	+ Need to check definition

"twoFixtureBathroomCnt”

* Number of two fixture bathrooms in the housing unit
* Numeric (ranges from 0 to 10)
* None missing

"threeFixtureBathroomCnt"

* Number of three fixture bathrooms in the housing unit
* Numeric (ranges from 0 to 7)
* None missing

"fourFixtureBathroomCnt”

* Number of four fixture bathrooms in the housing unit
* Numeric (ranges from 0 to 4)
* None missing

"fiveFixtureBathroomCnt”

* Number of five fixture bathrooms in the housing unit
* Numeric (ranges from 0 to 5)
* None missing

"floorNbr”

* Number of floors in the unit
* Levels: "1.0","1.5","1.75","2.0","2.5","2.75,”,"3.0","3.5","3.75", "4.0" “A”, “B”, “L”
* None missing

"floorKey"

* The unique identifier of that floor
* Numeric (ranges from 1 to 134700)
* None missing

**Property**

**\*\*\* Needs Cleaning**

Unit of observation: property

No duplicated rows found

"provalLrsnId”

* Proval land record serial number: identification for all properties
* Factor (no duplications)
* None Missing

"hubEffectiveDtm”

* UKNOWN Date
* YYYY-MM-DD HH:DD.SS

"hubExpirationDtm”

* UKNOWN Date
* YYYY-MM-DD HH:DD.SS

"realEstatePropertyCode"

* Property Code (parcel) for the Property assessed.
* Factor
* 0 are duplicated
* None missing

"reasPropertyStatusCode”

* What is the status of the property?
* Levels: “A”, “T”
	+ Active, inactive
* None missing

"propertyClassTypeCode”

* Type of property classification
* 58 levels
	+ Definitions are in “propertyClassTypeDsc”
* None missing

"propertyClassTypeDsc”

* Description of type of property classification
* 58 levels
* None missing

"legalDsc”

* Description of the property
* String
* None missing

"lotSizeQty"

* Lot size of the property
	+ Unknown units
* Numeric (range from 0 to 37480000)
	+ Validity depends on units

“mapBookPageNbr”

* Page number of the map book that the property is located
* XXX-XX
* None missing

"neighborhoodNbr”

* ID of the neighborhood the property is location
	+ Unknown definition of levels
* XXXXXX
* 46 coded as 0

"polygonId"

* ID of the polygon that property is location (for GIS Mapping Center data)’
* 2 missing
	+ NA codded as “”
* 380 appear more than 1

"propertyStreetNbrNameText”

* Number and Street that the property is located
* 753 are missing

"propertyStreetNbr”

* Street number the property is located
* 6% missing

"propertyStreetNbrSuffixCode”

* Number suffix for the property’s location
* Levels: "A" "B" "BK" "C" "D" "E" "F" "G" "H" "I" "J" "K" "L" "N" "S"

"propertyStreetDirectionPrefixCode"

* Street direction for the property’s location
* Levels: “N”, “S”

"propertyStreetName”

* Name of street property is located
* Inconsistent levels (12th, 12TH, 19th, 19TH, 20th, 20TH,8th, 8TH)
* 753 missing

"propertyStreetTypeCode”

* Code for the type o street property is located on
* Levels: "AVE","BLVD","CIR", "CT" ,"DR" , "HWY", "LN", "PIKE”, "PKWY", "PL", "RD", "ST", "TER"
* 753 missing

"propertyStreetDirectionSuffixCode"

* Direction of street that property is located
* Levels: “N”, “S”
* Inconsistent NA coding: NA, “”

"propertyUnitNbr"

* Unit number of that property
* 6,501 levels

"propertyCityName”

* City the property is located in
* Inconsistent coding: Arlington spelled in many different ways
* Invalid entry: “Clarendon”
* 753 NA’s

"propertyZipCode”

* Zip code the property is located in
* Invalid entries: 970000, 12201, 20003, 20009, 20011, 22003, 22004, 2207, 222004, 22208, 2221, 2224, 222O2
* 754 Missing

"gisStreetCode”

* Code the street is on for GIS data
* 515 levels
* 2% missing

"physicalAddressPrimeInd"

* UNKNOWN DEFINITION
* “TRUE”
* 753 missing

"zoningDescListText”

* Zoning description the property is located in
* 168 levels
* 393 Missing
* Invalid entries: see <http://buildingarlington.s3.amazonaws.com/wp-content/uploads/2013/06/ACZO.pdf>

"firstOwnerName”

* Name of the first listed owner of the property
* String
* None missing

"secondOwnerName”

* Name of the second listed owner of the property
* String

"additionalOwnerText"

* Additional information on the owners of the property
* String

"tradeName”

* Name of the business that operates in the property

"ownerStreetText”

* Address of the owner
* 602 missing

"ownerCityName”

* City of the address for the owner
* 603 missing

"ownerStateCode"

* State of the address for the owner
* 903 Missing

"ownerZipCode”

* Zipcode of the address for the owner
* 860 Missing

"compressedOwnerText”

* Owner information
* String

"propertyYearBuilt”

* Year the property was built
* Invalid entries: “0”, “1000”
* 9% Missing
	+ 78 "realEstatePropertyCode" have different "propertyYearBuilt" and “dwellingYearBuilt"
	+ 1,865 "realEstatePropertyCode" with "propertyYearBuilt" have no “dwellingYearBuilt”
	+ 0 "realEstatePropertyCode" with a "dwellingYearbuilt" has no “propertyYearBuilt”

"grossFloorAreaSquareFeetQty

* Total area of floor space
* 98% Missing
* Invalid entries: Range 0-1303000)

"effectiveAgeYearDate”

* Year the property was ready to be used
* Invalid entry “1”
* 98% missing

"numberOfUnitsCnt”

* Number of units in the property
* Numeric (range: 1 to 1295)
* Inconsistent NA coding: “NA”, “0”
* 99% Missing

"storyHeightCnt”

* Number of stories of the property
* Numeric (range: 1-32)
* 98% missing

"valuationYearDate"

* Year the valuation was taken
* “2015”
* 97 % missing

"commercialPropertyTypeDsc”

* Type of commercial property
* Levels: "Apartment”, "General Commercial", "Hotel”, "Office"

"economicUnitNbr”

* UNKNOWN DEFITION
* 99% Missing

"condoModelName”

* Description of the Condo Model
* String

"condoStyleName"

* Description of the condo
* 29 levels
* Inconsistent coding (e.g. High rise, High rise)

"finishedStorageAreaSquareFeetQty"’

* Square footage of finished area in the property
* ALL 0’S or NA

"storageAreaSquareFeetQty”

* Square footage of storage area in the property
* 62% missing
* 226 with value greater than 0

"unitNbr”

* Unit number of that property
* 6,699 levels

 ------------------------------

NEW TABLE?!

"propertyKey”

* Unknown Definition
* None Missing

"reasPropertyOwnerKey

* UNKNOWN DEFITION
* 0% duplicated

"arlingtonStreetKey”

* Street key for the property
	+ Unknown definition of levels
* 2% missing

"propertyExpiredInd”

* UNKNOWN
* ALL “FALSE”

"mixedUseInd"

* Is the property for mixed used?
* “TRUE/FALSE”
* None missing

"commercialInd”

* Is the property for commercial used?
* “TRUE/FALSE”
* None missing

"districtNbr”

* District number that property is located in
	+ Unknown description of levels
* 18 levels
	+ Potentially inconsistent coding: 0C, OC
* 3 missing

"taxExemptionTypeDsc”

* Description of tax exemption
* 21 Levels

"condominiumProjectName"

* Name of condo complex the property is a part of
* Factor

"masterRealEstatePropertyCode”

* Property Code (parcel) for the Property assessed: master real is specific for condo building only
* Factor
* 0 are duplicated
* None missing

"streetNbrOrder”

* Unknown definition
* Levels: “1”,”0”
* None missing

"resourceProtectionAreaInd"

* Is the property in a recourse protected area
* “TRUE/FALSE”
* None missing

**PropertyAssessmentHistory**

**\*\*\* Needs Cleaning**

Unit of observation: a property's assessment event

1,822,174 (96%) duplicated rows

"provalLrsnId”

* Proval land record serial number: identification for all properties
* Factor
* 96% duplicated
	+ 4% of ID’s appear more than once
* None missing

" realEstatePropertyCode "

* Property Code (parcel) for the Property assessed.
* Factor
* 96% are duplicated
	+ 98% of ID’s appear more than once
* None missing

"masterRealEstatePropertyCode”

* Property Code (parcel) for the Property assessed: master real is specific for condo building only
* Factor
* 98% are duplicated
	+ 99% of ID’s appear more than once None missing

"reasPropertyStatusCode”

* What is the status of the property?
* Levels: “A”, “T”
	+ Active, inactive
* None missing

"propertyClassTypeCode”

* Type of property classification
* 58 levels
	+ Definitions are in “propertyClassTypeDsc”
* None missing

"propertyClassTypeDsc”

* Description of type of property classification
* 58 levels
* None missing

"propertyStreetNbrNameText”

* Number and Street that the property is located
* 757 are missing

"propertyStreetNbr”

* Street number the property is located
* 4% missing

"propertyStreetNbrSuffixCode”

* Number suffix for the property’s location
* Levels: "A" "B" "BK" "C" "D" "E" "F" "G" "H" "I" "J" "K" "L" "N" "S"

"propertyStreetDirectionPrefixCode"

* Street direction for the property’s location
* Levels: “N”, “S”

"propertyStreetName”

* Name of street property is located
* Inconsistent levels (12th, 12TH, 19th, 19TH, 20th, 20TH,8th, 8TH)
* 757 missing

"propertyStreetTypeCode”

* Code for the type o street property is located on
* Levels: "AVE","BLVD","CIR", "CT","DR" , "HWY", "LN", "PIKE”, "PKWY", "PL", "RD", "ST", "TER"
* 757 missing

"propertyStreetDirectionSuffixCode"

* Direction of street that property is located
* Levels: “N”, “S”
* Inconsistent NA coding: NA, “”

"propertyUnitNbr"

* Unit number of that property
* 6,501 levels

"propertyCityName”

* City the property is located in
* Inconsistent coding: Arlington spelled in many different ways
* Invalid entry: “Clarendon”
* 757 NA’s

"propertyZipCode”

* Zip code the property is located in
* Invalid entries: 970000, 12201, 20003, 20009, 20011, 22003, 22004, 2207, 222004, 22208, 2221, 2224, 222O2
* 757 Missing

"zoningDescListText”

* Zoning description the property is located in
* 168 levels
* 487 Missing
* Invalid entries: see <http://buildingarlington.s3.amazonaws.com/wp-content/uploads/2013/06/ACZO.pdf>

"currentfirstOwnerName”

* Name of the first listed owner of the property
* String
* None missing

"currentsecondOwnerName”

* Name of the second listed owner of the property
* String

"currentadditionalOwnerText"

* Additional information on the owners of the property
* String

"tradeName”

* Name of the business that operates in the property
* String

"propertyYearBuilt”

* Year the property was built
* Invalid entries: “0”, “1000”
* 6% Missing

"grossFloorAreaSquareFeetQty

* Total area of floor space
* 98% Missing
* Invalid entries: Range 0-1303000)

"effectiveAgeYearDate”

* Year the property was ready to be used
* Invalid entry “1”
* 98% missing

"numberOfUnitsCnt”

* Number of units in the property
* Numeric (range: 1 to 1295)
* Inconsistent NA coding: “NA”, “0”
* 99% missing

"storyHeightCnt”

* Number of stories of the property
* Numeric (range: 1-32)
* 97% missing

"valuationYearDate"

* Year the valuation was taken
* “2015”
* 96 % missing

"commercialPropertyTypeDsc”

* Type of commercial property
* Levels: "Apartment”, "General Commercial", "Hotel”, "Office"

"condoModelName”

* Description of the Condo Model
* String

"condoStyleName"

* Description of the condo
* 29 levels
* Inconsistent coding (e.g. High rise, High rise)

"finishedStorageAreaSquareFeetQty"’

* Square footage of finished area in the property
* ALL 0’S or NA

"storageAreaSquareFeetQty”

* Square footage of storage area in the property
* 65% missing
* 6830 with value greater than 0

"unitNbr”

* Unit number of that property
* 6,699 levels

"propertyKey”

* Unknown Defintion
* None Missing
* 96% duplicated
	+ 97% of property key’s appear more than once

------------------------------

NEW TABLE?!

"reasPropertyOwnerKey

* UNKNOWN DEFINITION
* 0% duplicated

"arlingtonStreetKey”

* Street key for the property
	+ Unknown definition of levels
* 1654 missing

"propertyExpiredInd”

* UNKNOWN DEFINITION
* ALL “FALSE”

"commercialInd”

* Is the property for commercial used?
* “TRUE/FALSE”
* None missing

"districtNbr”

* District number that property is located in
	+ Unknown description of levels
* 18 levels
	+ Potentially inconsistent coding: 0C, OC
* 3 missing

"taxExemptionTypeDsc”

* Description of tax exemption
* 21 Levels

"condominiumProjectName"

* Name of condo complex the property is a part of
* Factor

The rest of the variables do not vary across time for each property. Each property only has values for one assessment date, which is then repeated for the rest of that property’s observations.

"taxBalanceAmt”

* Unknown definition (Balance of tax account?)
* Numeric (range: -74570 to 214000)
	+ Validity depends on definition
* 31 % missing

"assessedYearDate”

* Year the assessment was conducted for that property
* Factor
* 1471 Missing

"totalAssessedAmt”

* Total assed value of that property
* Numeric: range: 0 to 405600000)
	+ 1% listed at 0
* 1471 missing

"assessmentDate"

* Date that the assessment took place
* YYY-MM-DD HH:MM.SS
* 1471 missing

"assessmentChangeReasonTypeDsc"

* Reason the assessment was changed
* Levels: "01- Annual”, "02- Permit”, "03- Board of Equalization", "04- Court Order", "05- Review", “09- New Construction”, "16- Tax to Exempt”, "18- Exempt to Tax"
* 47% missing

**RealEstateAssessmentPaymentHistory**

**\*\*\* Needs Cleaning**

Unit of observation an action on a tax case

1774 are duplicated rows

"realEstatePropertyCode”

* Property Code (parcel) for the Property assessed.
* Factor
* 36% are duplicated
	+ 100% properties appear more than once
* None missing
	+ 4,522 (7%) of properties listed in Properties Table have no Real Estate Assessment Payment information

“propertyKey”

* Unknown Definition
	+ Follows "realEstatePropertyCode” ??
* 99% are duplicated
	+ 100% appear more than once
* None Missing

“taxYearDate”

* Tax year
* XXXX the assessment payment is made
* None missing

“taxInstallmentName”

* Name of tax installment
* Levels: "1st", "2nd", "New"
* None missing

“displayLabelSequenceNbr”

* Code for the Type of Payment
* Levels: "10" "20" "30" "40" "50"
	+ See “displayLabelTypeName” for description
* None missing

“displayLabelTypeName”

* Type of Payment
* Levels: "Adjustment", "Deferral", "Levy", "Payment", "Relief"
	+ Fir every levy, there needs to be an adjustment
* None missing

“taxDueDate”

* Due date taxes are due for that property
* YYYY-MM-DD HH:MM:ss
* None missing

“assessedTaxAmt”

* Amount of assessed taxes for that property
* Numeric
* 50% are below 0
* None missing

“penaltyAmt”

* Amount of any penalties for that property
* Numeric
* 4% do not equal 0
* None are missing

“interestAmt”

* Amount due in interest for that property
* Numeric
* 2% do not equal 0
* None missing

“feeAmt”

* Amount due in fees for that property
* Numeric
* 13,752 do not equal 0
* None missing

“totalAmt”

* Total amount due for that property
* Numeric
* 1 equals 0
* None missing

"provalLrsnId"

* Proval land record serial number: identification for all properties
* Factor (Do not know if duplicates are valid as no definition)
	+ Follows "Real.Estate.Property.Code” ??
* None Missing

**SalesHistory**

**\*\*\* Needs Cleaning**

Unit of observation: a property's sales transaction

473 (1%) of rows duplicated

“propertyKey”

* Unknown Definition
	+ Follows "realEstatePropertyCode”
* 36% are duplicated
* None Missing

"realEstatePropertyCode”

* Property Code (parcel) for the Property assessed.
* Factor
* 36% are duplicated
	+ 36% properties appear more than once
* None missing
	+ 41,581 (60%) of properties listed in Properties Table have no Real Estate Assessment Payment information

"provalLrsnId"

* Proval land record serial number: identification for all properties Factor (Do not know if duplicates are valid as no definition)
	+ Follows "Real.Estate.Property.Code”
* None Missing

"neighborhoodNbr”

* ID of the neighborhood the property is location
	+ Unknown definition of levels
* XXXXXX
* 1 coded as 0

"propertyStreetNbrNameText”

* Number and Street that the property is located
* String
* 1 is missing

“salesTypeCode”

* Code for the type of sale the transaction was
* 25 levels
	+ For levels definition, see “SalesTypeDsc”
* 41% missing

“salesTypeDsc”

* Description of the type of sale the transaction was
* 25 levels
* 41% missing

“deedBookNbr”

* The book number the property’s deed is located
* Some have letters in front of number
* NA coded as “”
* 5% are missing

“deedPageNbr”

* The page number of the deed book the property is located
* 5% missing
* 2 listed as page 0

“granteeName”

* Name of grantee
* String
* None missing

“saleAmt”

* The amount the property was sold for
* Numeric
* None missing
* 46% listed as 0

“saleDate”

* The date the transaction took place
* YYYY-MM-DD HH:MM.ss
* None missing

“salesHistoryKey”

* Unknown ID
* 0 are duplicated
* None missing

LandRecordDocumentNbr

* Number of Land Record
* 95% missing
* Invalid entries: dates, 0, ‘
* Inconsistent lengths