

# Nashville Housing Data Cleaning

Imported Data to Microsoft SQL

Objective: Data Cleaning

SQLQuery5.sql - LA...V36\Aerial Uy (55)) SQLQuery4.sql - LA...V36\Aerial Uy (52))\*

```
/****** Script for SelectTopNRows command from SSMS *****/  
SELECT TOP (1000) [UniqueID]  
,[ParcelID]  
,[LandUse]  
,[PropertyAddress]  
,[SaleDate]  
,[SalePrice]  
,[LegalReference]  
,[SoldAsVacant]  
,[OwnerName]  
,[OwnerAddress]  
,[Acreage]  
,[TaxDistrict]  
,[LandValue]  
,[BuildingValue]  
,[TotalValue]  
,[YearBuilt]  
,[Bedrooms]  
,[FullBath]  
,[HalfBath]
```

70 %

Results Messages

	LandUse	PropertyAddress	SaleDate	SalePrice	LegalReference	SoldAsVacant	OwnerName
108	SINGLE FAMILY	300 FANNIN DR, GOODLETTSVILLE	2016-01-28 00:00:00.000	220000	20160201-0009137	No	MCCOMB, DAVINIA L. & JOSEF
109	SINGLE FAMILY	105 JUDITH CT, GOODLETTSVILLE	2013-04-30 00:00:00.000	124900	20130502-0043992	No	NULL
110	SINGLE FAMILY	108 JUDITH CT, GOODLETTSVILLE	2013-12-18 00:00:00.000	144500	20131226-0129701	No	NULL
111	SINGLE FAMILY	336 DORR DR, GOODLETTSVILLE	2016-07-20 00:00:00.000	175900	20160726-0076602	No	NULL
112	SINGLE FAMILY	105 MAVELLA CT, GOODLETTSVILLE	2014-03-31 00:00:00.000	110000	20140402-0027284	No	NULL
113	SINGLE FAMILY	105 MAVELLA CT, GOODLETTSVILLE	2015-01-13 00:00:00.000	158000	20150116-0004659	No	NULL
114	SINGLE FAMILY	111 MAVELLA CT, GOODLETTSVILLE	2014-04-29 00:00:00.000	159900	20140502-0037305	No	NULL
115	SINGLE FAMILY	112 MAVELLA CT, GOODLETTSVILLE	2013-09-20 00:00:00.000	164800	20130923-0099794	No	NULL
116	SINGLE FAMILY	104 MAVELLA CT, GOODLETTSVILLE	2013-07-19 00:00:00.000	140000	20130722-0075620	No	NULL
117	SINGLE FAMILY	102 MAVELLA CT, GOODLETTSVILLE	2015-09-17 00:00:00.000	170000	20150925-0097548	No	NULL
118	SINGLE FAMILY	318 DORR DR, GOODLETTSVILLE	2013-06-06 00:00:00.000	149000	20130610-0058845	No	NULL
119	SINGLE FAMILY	331 DORR DR, GOODLETTSVILLE	2014-02-10 00:00:00.000	132900	20140213-0012707	No	NULL

Query executed successfully. LAPTOP-KRK6RV36 (16.0 RTM) LAPTOP-KRK6RV36\Aerial... PortfolioProject 00:00:00 1,000 rows

Results Messages								
	UniqueID	ParcelID	LandUse	PropertyAddress	SaleDate	SalePrice	LegalReference	SoldAsVacant
1	52229	075 12 0E 018.00	SINGLE FAMILY	1213 LILAC DR, HERMITAGE	2016-07-16 00:00:00.000	312000	20160719-0074160	No
2	29812	075 12 0E 021.00	SINGLE FAMILY	1313 WEEPING CHERRY LN, HERMITAGE	2015-04-20 00:00:00.000	268000	20150428-0037869	No
3	13595	075 12 0E 023.00	SINGLE FAMILY	1321 WEEPING CHERRY LN, HERMITAGE	2014-03-19 00:00:00.000	235000	20140321-0023388	No
4	17225	075 12 0E 024.00	SINGLE FAMILY	1325 WEEPING CHERRY LN, HERMITAGE	2014-06-19 00:00:00.000	310000	20140626-0056087	No
5	21634	075 12 0E 030.00	SINGLE FAMILY	1304 WEEPING CHERRY LN, HERMITAGE	2014-09-22 00:00:00.000	244500	20140924-0087876	No
6	43340	075 12 0E 031.00	SINGLE FAMILY	1300 WEEPING CHERRY LN, HERMITAGE	2016-01-20 00:00:00.000	250000	20160125-0006594	No
7	27319	075 12 0E 036.00	SINGLE FAMILY	1220 LILAC DR, HERMITAGE	2015-02-09 00:00:00.000	284900	20150211-0012533	No
8	8175	075 12 0E 037.00	SINGLE FAMILY	1216 LILAC DR, HERMITAGE	2013-09-27 00:00:00.000	295000	20131001-0102885	No
9	8176	075 12 0E 039.00	SINGLE FAMILY	1208 LILAC DR, HERMITAGE	2013-09-17 00:00:00.000	280000	20130923-0099837	No
10	56358	075 12 0F 001.00	RESIDENTIAL CONDO	735 TULIP GROVE RD, HERMITAGE	2016-10-28 00:00:00.000	185500	20161107-0117405	No
11	35201	075 12 0F 002.00	RESIDENTIAL CONDO	735 TULIP GROVE RD, HERMITAGE	2015-07-24 00:00:00.000	159900	20150803-0076366	No
12	18702	075 12 0F 005.00	RESIDENTIAL CONDO	735 TULIP GROVE RD, HERMITAGE	2014-07-30 00:00:00.000	131000	20140807-0071173	No

Results Messages											
	ate	SalePrice	LegalReference	SoldAsVacant	OwnerName	OwnerAddress	Acreage	TaxDistrict	LandValue	BuildingValue	TotalValue
1	07-16 00:00:00.000	312000	20160719-0074160	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
2	04-20 00:00:00.000	268000	20150428-0037869	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3	03-19 00:00:00.000	235000	20140321-0023388	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
4	06-19 00:00:00.000	310000	20140626-0056087	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
5	09-22 00:00:00.000	244500	20140924-0087876	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
6	01-20 00:00:00.000	250000	20160125-0006594	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
7	02-09 00:00:00.000	284900	20150211-0012533	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
8	09-27 00:00:00.000	295000	20131001-0102885	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
9	09-17 00:00:00.000	280000	20130923-0099837	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
10	10-28 00:00:00.000	185500	20161107-0117405	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
11	07-24 00:00:00.000	159900	20150803-0076366	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL
12	07-30 00:00:00.000	131000	20140807-0071173	No	NULL	NULL	NULL	NULL	NULL	NULL	NULL

## Tasks

### First Task

- Standardized Date Format on SaleDate Column showing the date and time

```
Select SaleDate, CONVERT(Date,SaleDate)
From PortfolioProject..NashvilleHousing
```

	Results	Messages
	SaleDate	
1	2013-04-09 00:00:00.000	
2	2014-06-10 00:00:00.000	
3	2016-09-26 00:00:00.000	
4	2016-01-29 00:00:00.000	
5	2014-10-10 00:00:00.000	
6	2014-07-16 00:00:00.000	
7	2014-08-28 00:00:00.000	
8	2016-09-27 00:00:00.000	
9	2015-08-14 00:00:00.000	
10	2014-08-29 00:00:00.000	
11	2015-04-14 00:00:00.000	
12	2013-12-12 00:00:00.000	
13	2015-07-13 00:00:00.000	



	Results	Messages
	SaleDate	(No column name)
1	2013-04-09 00:00:00.000	2013-04-09
2	2014-06-10 00:00:00.000	2014-06-10
3	2016-09-26 00:00:00.000	2016-09-26
4	2016-01-29 00:00:00.000	2016-01-29
5	2014-10-10 00:00:00.000	2014-10-10
6	2014-07-16 00:00:00.000	2014-07-16
7	2014-08-28 00:00:00.000	2014-08-28
8	2016-09-27 00:00:00.000	2016-09-27
9	2015-08-14 00:00:00.000	2015-08-14
10	2014-08-29 00:00:00.000	2014-08-29
11	2015-04-14 00:00:00.000	2015-04-14
12	2013-12-12 00:00:00.000	2013-12-12
13	2015-07-13 00:00:00.000	2015-07-13

- Update the data with New Column named 'SaleDateConverted'

**ALTER TABLE** was used as the previous **UPDATE** did not work and changed the column name to SaleDateConverted to be the new column. SaleDate column to be deleted.

```

Update NashvilleHousing
SET SaleDate = CONVERT(Date, SaleDate)

ALTER TABLE NashvilleHousing
Add SaleDateConverted Date;

Update NashvilleHousing
SET SaleDateConverted = CONVERT(Date, SaleDate)

```

	SaleDateConverted	(No column name)
4	2016-01-29	2016-01-29
5	2014-10-10	2014-10-10
6	2014-07-16	2014-07-16
7	2014-08-28	2014-08-28
8	2016-09-27	2016-09-27
9	2015-08-14	2015-08-14
10	2014-08-29	2014-08-29
11	2015-04-14	2015-04-14
12	2013-12-12	2013-12-12
13	2015-07-13	2015-07-13
14	2013-06-07	2013-06-07
15	2014-06-30	2014-06-30
16	2014-06-30	2014-06-30

## Second Task

- Populate Property Address Data

```

Select PropertyAddress
From PortfolioProject.dbo.NashvilleHousing
Where PropertyAddress is null

```

	PropertyAddress
1	NULL
2	NULL
3	NULL
4	NULL
5	NULL
6	NULL
7	NULL
8	NULL

To populate the Null values, the ParcelID was used as it shows that if it is the same ParcelID then the PropertyAddress is the same.

	UniqueID	ParcelID	LandUse	PropertyAddress
44	6993	015 14 0 060.00	SINGLE FAMILY	3113 MILLIKEN DR, JOELTON
45	43071	015 14 0 060.00	SINGLE FAMILY	3113 MILLIKEN DR, JOELTON
46	45286	015 14 0 061.00	SINGLE FAMILY	3117 MILLIKEN DR, JOELTON
47	3267	015 14 0 071.00	SINGLE FAMILY	3107 MARGIE DR, JOELTON
48	34753	015 14 0 075.00	SINGLE FAMILY	3124 MARGIE DR, JOELTON

```

Select a.ParcelID, a. PropertyAddress, b.ParcelID, b.PropertyAddress
From PortfolioProject..NashvilleHousing a
JOIN PortfolioProject..NashvilleHousing b
    ON a.ParcelID = b.ParcelID
    AND a.[UniqueID] <> b. [UniqueID]
Where PropertyAddress is null

```

We use JOIN to populate the property address. Alias were used such as 'a' and 'b' for the table name. Query shows to populate the address with the same parcel ID but not the same unique ID.

Results Messages				
	ParcelID	PropertyAddress	ParcelID	PropertyAddress
1	025 07 0 031.00	NULL	025 07 0 031.00	410 ROSEHILL CT, GOODLETTSVILLE
2	026 01 0 069.00	NULL	026 01 0 069.00	141 TWO MILE PIKE, GOODLETTSVILLE
3	026 05 0 017.00	NULL	026 05 0 017.00	208 EAST AVE, GOODLETTSVILLE
4	026 06 0A 038.00	NULL	026 06 0A 038.00	109 CANTON CT, GOODLETTSVILLE
5	033 06 0 041.00	NULL	033 06 0 041.00	1129 CAMPBELL RD, GOODLETTSVILLE
6	033 06 0A 002.00	NULL	033 06 0A 002.00	1116 CAMPBELL RD, GOODLETTSVILLE
7	033 15 0 123.00	NULL	033 15 0 123.00	438 W CAMPBELL RD, GOODLETTSVILLE
8	034 03 0 059.00	NULL	034 03 0 059.00	2117 PAULA DR, MADISON
9	034 03 0 059.00	NULL	034 03 0 059.00	2117 PAULA DR, MADISON
10	034 07 0B 015.00	NULL	034 07 0B 015.00	2524 VAL MARIE DR, MADISON
11	034 07 0B 015.00	NULL	034 07 0B 015.00	2524 VAL MARIE DR, MADISON
12	034 07 0B 015.00	NULL	034 07 0B 015.00	2524 VAL MARIE DR, MADISON
13	034 16 0A 004.00	NULL	034 16 0A 004.00	213 WARREN CT, OLD HICKORY



70 % Results Messages					
	ParcelID	PropertyAddress	ParcelID	PropertyAddress	(No column name)
1	025 07 0 031.00	NULL	025 07 0 031.00	410 ROSEHILL CT, GOODLETTSVILLE	410 ROSEHILL CT, GOODLETTSVILLE
2	026 01 0 069.00	NULL	026 01 0 069.00	141 TWO MILE PIKE, GOODLETTSVILLE	141 TWO MILE PIKE, GOODLETTSVILLE
3	026 05 0 017.00	NULL	026 05 0 017.00	208 EAST AVE, GOODLETTSVILLE	208 EAST AVE, GOODLETTSVILLE
4	026 06 0A 038.00	NULL	026 06 0A 038.00	109 CANTON CT, GOODLETTSVILLE	109 CANTON CT, GOODLETTSVILLE
5	033 06 0 041.00	NULL	033 06 0 041.00	1129 CAMPBELL RD, GOODLETTSVILLE	1129 CAMPBELL RD, GOODLETTSVILLE
6	033 06 0A 002.00	NULL	033 06 0A 002.00	1116 CAMPBELL RD, GOODLETTSVILLE	1116 CAMPBELL RD, GOODLETTSVILLE
7	033 15 0 123.00	NULL	033 15 0 123.00	438 W CAMPBELL RD, GOODLETTSVILLE	438 W CAMPBELL RD, GOODLETTSVILLE
8	034 03 0 059.00	NULL	034 03 0 059.00	2117 PAULA DR, MADISON	2117 PAULA DR, MADISON
9	034 03 0 059.00	NULL	034 03 0 059.00	2117 PAULA DR, MADISON	2117 PAULA DR, MADISON
10	034 07 0B 015.00	NULL	034 07 0B 015.00	2524 VAL MARIE DR, MADISON	2524 VAL MARIE DR, MADISON
11	034 07 0B 015.00	NULL	034 07 0B 015.00	2524 VAL MARIE DR, MADISON	2524 VAL MARIE DR, MADISON
12	034 07 0B 015.00	NULL	034 07 0B 015.00	2524 VAL MARIE DR, MADISON	2524 VAL MARIE DR, MADISON
13	034 16 0A 004.00	NULL	034 16 0A 004.00	213 WARREN CT, OLD HICKORY	213 WARREN CT, OLD HICKORY

Update a

```

SET PropertyAddress = ISNULL( a.PropertyAddress,b.PropertyAddress)
From PortfolioProject..NashvilleHousing a
JOIN PortfolioProject..NashvilleHousing b
    ON a.ParcelID = b.ParcelID
    AND a.[UniqueID] <> b.[UniqueID]
Where a.PropertyAddress is null

```

### Third Task

- Separate the addresses to individual columns (Address, City, State)

```

Select PropertyAddress
From PortfolioProject.dbo.NashvilleHousing
--Where PropertyAddress is null
--order by ParcelID

```

Results Messages	
	PropertyAddress
1	1808 FOX CHASE DR, GOODLETTSVILLE
2	1832 FOX CHASE DR, GOODLETTSVILLE
3	1864 FOX CHASE DR, GOODLETTSVILLE
4	1853 FOX CHASE DR, GOODLETTSVILLE
5	1829 FOX CHASE DR, GOODLETTSVILLE
6	1821 FOX CHASE DR, GOODLETTSVILLE
7	2005 SADIE LN, GOODLETTSVILLE
8	1917 GRACELAND DR, GOODLETTSVILLE
9	1428 SPRINGFIELD HWY, GOODLETTSVILLE
10	1420 SPRINGFIELD HWY, GOODLETTSVILLE
11	2209 KAYLA DR, GOODLETTSVILLE
12	109 BAILEY VIEW CT, GOODLETTSVILLE
13	1900 TINNIN RD, GOODLETTSVILLE

```
SELECT
SUBSTRING(PropertyAddress, 1, CHARINDEX(',', PropertyAddress)) as Address
From PortfolioProject..NashvilleHousing
```

```
SELECT
SUBSTRING(PropertyAddress, 1, CHARINDEX(',', PropertyAddress)-1) as Address
From PortfolioProject..NashvilleHousing
```

Results Messages	
	Address
1	1808 FOX CHASE DR,
2	1832 FOX CHASE DR,
3	1864 FOX CHASE DR,
4	1853 FOX CHASE DR,
5	1829 FOX CHASE DR,
6	1821 FOX CHASE DR,
7	2005 SADIE LN,
8	1917 GRACELAND DR,
9	1428 SPRINGFIELD HWY,
10	1420 SPRINGFIELD HWY,
11	2209 KAYLA DR,
12	109 BAILEY VIEW CT,
13	1900 TINNIN RD,



Results Messages	
	Address
1	1808 FOX CHASE DR
2	1832 FOX CHASE DR
3	1864 FOX CHASE DR
4	1853 FOX CHASE DR
5	1829 FOX CHASE DR
6	1821 FOX CHASE DR
7	2005 SADIE LN
8	1917 GRACELAND DR
9	1428 SPRINGFIELD HWY
10	1420 SPRINGFIELD HWY
11	2209 KAYLA DR
12	109 BAILEY VIEW CT
13	1900 TINNIN RD

```
SELECT
SUBSTRING(PropertyAddress, 1, CHARINDEX(',', PropertyAddress) -1 ) as Address
, SUBSTRING(PropertyAddress, CHARINDEX(',', PropertyAddress) + 1 , LEN(PropertyAddress))
as Address
```

```
From PortfolioProject.dbo.NashvilleHousing
```

+1 was removed to remove the comma from the result and 'LEN' to separate the address even with different lengths

Results Messages		
	Address	Address
1	1808 FOX CHASE DR	GOODLETTSVILLE
2	1832 FOX CHASE DR	GOODLETTSVILLE
3	1864 FOX CHASE DR	GOODLETTSVILLE
4	1853 FOX CHASE DR	GOODLETTSVILLE
5	1829 FOX CHASE DR	GOODLETTSVILLE
6	1821 FOX CHASE DR	GOODLETTSVILLE
7	2005 SADIE LN	GOODLETTSVILLE
8	1917 GRACELAND DR	GOODLETTSVILLE
9	1428 SPRINGFIELD HWY	GOODLETTSVILLE
10	1420 SPRINGFIELD HWY	GOODLETTSVILLE
11	2209 KAYLA DR	GOODLETTSVILLE
12	109 BAILEY VIEW CT	GOODLETTSVILLE
13	1900 TINNIN RD	GOODLETTSVILLE

Select

```
PARSENAME(REPLACE(OwnerAddress, ',', '.'), 1)
From PortfolioProject.dbo.NashvilleHousing
```

--Another easier way of splitting the addresses. Commas were replaced with periods as PARSENAME is useful only for periods.

Results Messages	
	(No column name)
1	TN
2	TN
3	TN
4	TN
5	TN
6	TN
7	TN
8	TN
9	TN
10	TN
11	NULL
12	NULL
13	NULL

Select

```
PARSENAME(REPLACE(OwnerAddress, ',', '.'), 1)
, PARSENAME(REPLACE(OwnerAddress, ',', '.'), 2)
, PARSENAME(REPLACE(OwnerAddress, ',', '.'), 3)
From PortfolioProject.dbo.NashvilleHousing
```

Select

```
PARSENAME(REPLACE(OwnerAddress, ',', '.'), 1)
, PARSENAME(REPLACE(OwnerAddress, ',', '.'), 2)
, PARSENAME(REPLACE(OwnerAddress, ',', '.'), 3)
From PortfolioProject.dbo.NashvilleHousing
```

	Results	Messages
	(No column name)	(No column name)
1	TN	GOODLETTSVILLE
2	TN	GOODLETTSVILLE
3	TN	GOODLETTSVILLE
4	TN	GOODLETTSVILLE
5	TN	GOODLETTSVILLE
6	TN	GOODLETTSVILLE
7	TN	GOODLETTSVILLE
8	TN	GOODLETTSVILLE
9	TN	GOODLETTSVILLE
10	TN	GOODLETTSVILLE
11	NULL	NULL
12	NULL	NULL
13	NULL	NULL



	Results	Messages
	(No column name)	(No column name)
1	1808 FOX CHASE DR	GOODLETTSVILLE
2	1832 FOX CHASE DR	GOODLETTSVILLE
3	1864 FOX CHASE DR	GOODLETTSVILLE
4	1853 FOX CHASE DR	GOODLETTSVILLE
5	1829 FOX CHASE DR	GOODLETTSVILLE
6	1821 FOX CHASE DR	GOODLETTSVILLE
7	2005 SADIE LN	GOODLETTSVILLE
8	1917 GRACELAND DR	GOODLETTSVILLE
9	1428 SPRINGFIELD HWY	GOODLETTSVILLE
10	1420 SPRINGFIELD HWY	GOODLETTSVILLE
11	NULL	NULL
12	NULL	NULL
13	NULL	NULL

```
ALTER TABLE NashvilleHousing
Add OwnerSplitAddress Nvarchar(255);
```

```
Update NashvilleHousing
SET OwnerSplitAddress = PARSENAME(REPLACE(OwnerAddress, ',', '.'), 3)
```

```
ALTER TABLE NashvilleHousing
Add OwnerSplitCity Nvarchar(255);
```

```
Update NashvilleHousing
SET OwnerSplitCity = PARSENAME(REPLACE(OwnerAddress, ',', '.'), 2)
```

```
ALTER TABLE NashvilleHousing
Add OwnerSplitState Nvarchar(255);
```

```
Update NashvilleHousing
SET OwnerSplitState = PARSENAME(REPLACE(OwnerAddress, ',', '.'), 1)
```

OwnerSplitAddress	OwnerSplitCity	OwnerSplitState
1808 FOX CHASE DR	GOODLETTSVILLE	TN
1832 FOX CHASE DR	GOODLETTSVILLE	TN
1864 FOX CHASE DR	GOODLETTSVILLE	TN
1853 FOX CHASE DR	GOODLETTSVILLE	TN
1829 FOX CHASE DR	GOODLETTSVILLE	TN
1821 FOX CHASE DR	GOODLETTSVILLE	TN
2005 SADIE LN	GOODLETTSVILLE	TN
1917 GRACELAND DR	GOODLETTSVILLE	TN
1428 SPRINGFIELD HWY	GOODLETTSVILLE	TN
1420 SPRINGFIELD HWY	GOODLETTSVILLE	TN
NULL	NULL	NULL
NULL	NULL	NULL

#### Fourth Task

- Replacing the Y and N to Yes and No in “Sold As Vacant” field

```
Select Distinct (SoldasVacant), COUNT(SoldasVacant)
From PortfolioProject..NashvilleHousing
Group by SoldasVacant
Order by 2
```

	SoldasVacant	(No column name)
1	Y	52
2	N	399
3	Yes	4623
4	No	51403

```
Select SoldAsVacant
, CASE When SoldAsVacant = 'Y' THEN 'Yes'
      When SoldAsVacant = 'N' THEN 'No'
      ELSE SoldAsVacant
      END
From PortfolioProject.dbo.NashvilleHousing
```

	SoldAsVacant	(No column name)
163	No	No
164	No	No
165	No	No
166	No	No
167	No	No
168	No	No
169	No	No
170	No	No
171	No	No
172	No	No
173	No	No
174	No	No
175	No	No

#### Fifth Task

--Remove Duplicates by Identifying the Duplicates

```
WITH RowNumCTE AS(
Select *,
      ROW_NUMBER() OVER (
        PARTITION BY ParcelID,
                    PropertyAddress,
                    SalePrice,
                    SaleDate,
                    LegalReference
        Order by
                    UniqueID
        ) as row_num
From PortfolioProject.dbo.NashvilleHousing
--order by ParcelID
)
```

```
Select *
From RowNumCTE
Where row_num > 1
order by PropertyAddress
```



	UniquelD	ParcelID	LandUse	PropertyAddress	SaleDate
46	27380	149 15 0A 418.00	SINGLE FAMILY	3044 BROMLEY WAY, ANTIOCH	2015-02-12 00:00:00.000
47	12692	163 11 0B 081.00	RESIDENTIAL CONDO	305 ASH FORGE DR, ANTIOCH	2014-02-03 00:00:00.000
48	27143	114 14 0A 022.00	SINGLE FAMILY	309 DOGWOOD TRL, NASHVILLE	2015-02-18 00:00:00.000
49	27392	164 07 0A 139.00	SINGLE FAMILY	312 GROVEDALE TRCE, ANTIOCH	2015-02-20 00:00:00.000
50	27355	108 08 0 114.00	SINGLE FAMILY	3121 E LAKE DR, NASHVILLE	2015-02-06 00:00:00.000
51	27356	108 12 0B 031.00	SINGLE FAMILY	3124 HARBORWOOD DR, NASHVILLE	2015-02-13 00:00:00.000
52	27387	150 07 0 181.00	SINGLE FAMILY	3140 COUNTRY MEADOW RD, ANTI...	2015-02-18 00:00:00.000
53	27388	150 09 0A 028.00	SINGLE FAMILY	3313 CALAIS CIR, ANTIOCH	2015-02-13 00:00:00.000
54	27382	150 03 0 442.00	SINGLE FAMILY	3412 WESTCLIFFE CT, ANTIOCH	2015-02-02 00:00:00.000
55	27353	098 13 0C 200.00	SINGLE FAMILY	3420 EARHART RD, MOUNT JULIET	2015-02-06 00:00:00.000
56	27386	150 05 0 353.00	SINGLE FAMILY	365 CEDARCLIFF RD, ANTIOCH	2015-02-10 00:00:00.000
57	27402	176 01 0 017.00	SINGLE FAMILY	3848 ROUND ROCK DR, ANTIOCH	2015-02-13 00:00:00.000

- then Delete the duplicates

```
DELETE
From RowNumCTE
Where row_num > 1
--order by PropertyAddress
```

Messages
(104 rows affected)
Completion time: 2023-02-19T16:10:36.9285450+08:00

	UniquelD	ParcelID	LandUse	PropertyAddress	SaleDate	SalePrice	LegalReference	SoldAsVacant	OwnerName	O
--	----------	----------	---------	-----------------	----------	-----------	----------------	--------------	-----------	---

## Sixth Task

- Delete Unused Columns

```
ALTER TABLE PortfolioProject.dbo.NashvilleHousing
DROP COLUMN OwnerAddress, TaxDistrict, PropertyAddress
```

	UniquelD	ParcelID	LandUse	SaleDate	SalePrice	LegalReference	SoldAsVaca
1	2045	007 00 0 125.00	SINGLE FAMILY	2013-04-09 00:00:00.000	240000	20130412-0036474	No
2	16918	007 00 0 130.00	SINGLE FAMILY	2014-06-10 00:00:00.000	366000	20140619-0053768	No
3	54582	007 00 0 138.00	SINGLE FAMILY	2016-09-26 00:00:00.000	435000	20160927-0101718	No
4	43070	007 00 0 143.00	SINGLE FAMILY	2016-01-29 00:00:00.000	255000	20160129-0008913	No
5	22714	007 00 0 149.00	SINGLE FAMILY	2014-10-10 00:00:00.000	278000	20141015-0095255	No
6	18367	007 00 0 151.00	SINGLE FAMILY	2014-07-16 00:00:00.000	267000	20140718-0063802	No
7	19804	007 14 0 002.00	SINGLE FAMILY	2014-08-28 00:00:00.000	171000	20140903-0080214	No
8	54583	007 14 0 024.00	SINGLE FAMILY	2016-09-27 00:00:00.000	262000	20161005-0105441	No
9	36500	007 14 0 026.00	SINGLE FAMILY	2015-08-14 00:00:00.000	285000	20150819-0083440	No
10	19805	007 14 0 034.00	SINGLE FAMILY	2014-08-29 00:00:00.000	340000	20140909-0082348	No
11	29467	007 14 0A 024.00	SINGLE FAMILY	2015-04-14 00:00:00.000	425000	20150415-0033442	No
12	10754	007 14 0A 027.00	SINGLE FAMILY	2013-12-12 00:00:00.000	585000	20131227-0130352	No

Dataset: Alex the Analyst Guided Project