SQL Data Cleaning

```
/*
Cleaning Data in SQL Queries
*/


SELECT *
FROM [Portfolio SQL].dbo.NashvilleHousing;
```

	UniqueID	ParcelID	LandUse	PropertyAddress	SaleDate	SalePrice	LegalReference	SoldAsVacant	OwnerName
1	2045	007 00 0 125.00	SINGLE FAMILY	1808 FOX CHASE DR, GOODLETTSVILLE	2013-04-09 00:00:00.000	240000	20130412-0036474	No	FRAZIER, CYRENTHA LYNETTE
2	16918	007 00 0 130.00	SINGLE FAMILY	1832 FOX CHASE DR, GOODLETTSVILLE	2014-06-10 00:00:00.000	366000	20140619-0053768	No	BONER, CHARLES & LESLIE
3	54582	007 00 0 138.00	SINGLE FAMILY	1864 FOX CHASE DR, GOODLETTSVILLE	2016-09-26 00:00:00.000	435000	20160927-0101718	No	WILSON, JAMES E. & JOANNE
4	43070	007 00 0 143.00	SINGLE FAMILY	1853 FOX CHASE DR, GOODLETTSVILLE	2016-01-29 00:00:00.000	255000	20160129-0008913	No	BAKER, JAY K. & SUSAN E.
5	22714	007 00 0 149.00	SINGLE FAMILY	1829 FOX CHASE DR, GOODLETTSVILLE	2014-10-10 00:00:00.000	278000	20141015-0095255	No	POST, CHRISTOPHER M. & SAMANTHA
6	18367	007 00 0 151.00	SINGLE FAMILY	1821 FOX CHASE DR, GOODLETTSVILLE	2014-07-16 00:00:00.000	267000	20140718-0063802	No	FIELDS, KAREN L. & BRENT A.
7	19804	007 14 0 002.00	SINGLE FAMILY	2005 SADIE LN, GOODLETTSVILLE	2014-08-28 00:00:00.000	171000	20140903-0080214	No	HINTON, MICHAEL R. & CYNTHIA M. MC
8	54583	007 14 0 024.00	SINGLE FAMILY	1917 GRACELAND DR, GOODLETTSVILLE	2016-09-27 00:00:00.000	262000	20161005-0105441	No	BAILOR, DARRELL & TAMMY
9	36500	007 14 0 026.00	SINGLE FAMILY	1428 SPRINGFIELD HWY, GOODLETTSVILLE	2015-08-14 00:00:00.000	285000	20150819-0083440	No	ROBERTS, MISTY L. & ROBERT M.
10	19805	007 14 0 034.00	SINGLE FAMILY	1420 SPRINGFIELD HWY, GOODLETTSVILLE	2014-08-29 00:00:00.000	340000	20140909-0082348	No	LEE, JEFFREY & NANCY
11	29467	007 14 0A 024.00	SINGLE FAMILY	2209 KAYLA DR, GOODLETTSVILLE	2015-04-14 00:00:00.000	425000	20150415-0033442	No	NULL
12	10754	007 14 0A 027.00	SINGLE FAMILY	109 BAILEY VIEW CT, GOODLETTSVILLE	2013-12-12 00:00:00.000	585000	20131227-0130352	No	NULL
13	34751	007 14 0B 010.00	RESIDENTIAL CONDO	1900 TINNIN RD, GOODLETTSVILLE	2015-07-13 00:00:00.000	190000	20150717-0069947	No	NULL
14	4512	007 15 0 002.00	SINGLE FAMILY	629 GAYLEMORE DR, GOODLETTSVILLE	2013-06-07 00:00:00.000	189900	20130612-0059715	No	URRUTIA, CARLOS MIGUEL & REBECCA
15	16919	007 15 0 003.00	SINGLE FAMILY	633 GAYLEMORE DR, GOODLETTSVILLE	2014-06-30 00:00:00.000	157500	20140702-0058050	No	SALDANA, ALMA J. & EDWARD A., II
16	16920	007 15 0 004.00	SINGLE FAMILY	637 GAYLEMORE DR, GOODLETTSVILLE	2014-06-30 00:00:00.000	247400	20140630-0057267	No	MCKINNEY, ROBERT K. & BETH E.
17	51967	007 15 0 008.00	SINGLE FAMILY	1976 SADIE LN, GOODLETTSVILLE	2016-07-15 00:00:00.000	211500	20160720-0074793	No	MILLER, JAMES L. & EILEEN T.
18	28155	007 15 0 014.00	SINGLE FAMILY	644 GAYLEMORE DR, GOODLETTSVILLE	2015-03-31 00:00:00.000	185900	20150402-0029022	No	SANDAGE, LEAH S.
19	8899	007 15 0 020.00	SINGLE FAMILY	1921 NORMERLE DR, GOODLETTSVILLE	2013-10-11 00:00:00.000	349900	20131018-0109102	No	VAUGHN, JOHN
20	4513	007 15 0 031.00	SINGLE FAMILY	1916 NORMERLE DR, GOODLETTSVILLE	2013-06-28 00:00:00.000	192500	20130711-0071698	No	PARKS, KRISTEN LEIGH & WILLIAMS, PA
21	27161	007 15 0 044.00	SINGLE FAMILY	2050 GRACELAND DR, GOODLETTSVILLE	2015-02-10 00:00:00.000	279900	20150212-0012993	No	ANGELL, MARGO G. & CHARLES D.
22	46859	007 15 0 048.00	SINGLE FAMILY	2034 GRACELAND DR, GOODLETTSVILLE	2016-04-14 00:00:00.000	383600	20160418-0036715	No	DOSS, GLEN KEVIN & ANGELA DAWN
23	5802	007 15 0 052.00	SINGLE FAMILY	811 BENTON CT, GOODLETTSVILLE	2013-07-08 00:00:00.000	192500	20130712-0072376	No	BRAKE, MICHAEL E. & ERIN L.
24	36501	010 00 0 045.00	SINGLE FAMILY	331 VIEW RIDGE DR, GOODLETTSVILLE	2015-08-31 00:00:00.000	193500	20150903-0090036	No	BENJAMIN, DAVID AND LINDA REVOCAI

-- Standardize Date Format

```
ALTER TABLE [Portfolio SQL].dbo.NashvilleHousing
ADD SaleDateConverted Date;


UPDATE [Portfolio SQL].dbo.NashvilleHousing
SET SaleDateConverted = CONVERT(Date, SaleDate);
```

OwnerAddress	Acreage	TaxDistrict	LandValue	BuildingValue	TotalValue	YearBuilt	Bedrooms	FullBath	HalfBath	SaleDateConverted
1808 FOX CHASE DR, GOODLETTSVILLE, TN	2.3	GENERAL SERVICES DISTRICT	50000	168200	235700	1986	3	3	0	2013-04-09
1832 FOX CHASE DR, GOODLETTSVILLE, TN	3.5	GENERAL SERVICES DISTRICT	50000	264100	319000	1998	3	3	2	2014-06-10
1864 FOX CHASE DR, GOODLETTSVILLE, TN	2.9	GENERAL SERVICES DISTRICT	50000	216200	298000	1987	4	3	0	2016-09-26
1853 FOX CHASE DR, GOODLETTSVILLE, TN	2.6	GENERAL SERVICES DISTRICT	50000	147300	197300	1985	3	3	0	2016-01-29
1829 FOX CHASE DR, GOODLETTSVILLE, TN	2	GENERAL SERVICES DISTRICT	50000	152300	202300	1984	4	3	0	2014-10-10
1821 FOX CHASE DR, GOODLETTSVILLE, TN	2	GENERAL SERVICES DISTRICT	50000	190400	259800	1980	3	3	0	2014-07-16
2005 SADIE LN, GOODLETTSVILLE, TN	1.03	GENERAL SERVICES DISTRICT	40000	137900	177900	1976	3	2	0	2014-08-28
1917 GRACELAND DR, GOODLETTSVILLE, TN	1.03	GENERAL SERVICES DISTRICT	40000	157900	197900	1978	3	2	0	2016-09-27
1428 SPRINGFIELD HWY, GOODLETTSVILLE, TN	1.67	GENERAL SERVICES DISTRICT	45400	176900	222300	2000	3	2	1	2015-08-14
1420 SPRINGFIELD HWY, GOODLETTSVILLE, TN	1.3	GENERAL SERVICES DISTRICT	40000	179600	219600	1995	5	3	0	2014-08-29
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	2015-04-14
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	2013-12-12
NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	2015-07-13
629 GAYLEMORE DR, GOODLETTSVILLE, TN	1.07	GENERAL SERVICES DISTRICT	40000	203600	243600	2006	4	4	0	2013-06-07
633 GAYLEMORE DR, GOODLETTSVILLE, TN	1.03	GENERAL SERVICES DISTRICT	40000	115400	155400	1974	4	2	0	2014-06-30
637 GAYLEMORE DR, GOODLETTSVILLE, TN	1.03	GENERAL SERVICES DISTRICT	40000	180300	234000	1974	3	3	0	2014-06-30
2001 SADIE LN, GOODLETTSVILLE, TN	1.08	GENERAL SERVICES DISTRICT	40000	129200	169200	1974	4	2	0	2016-07-15
644 GAYLEMORE DR, GOODLETTSVILLE, TN	1.03	GENERAL SERVICES DISTRICT	40000	127400	167400	1974	3	2	0	2015-03-31
1921 NORMERLE DR, GOODLETTSVILLE, TN	1.03	GENERAL SERVICES DISTRICT	40000	286500	346300	1974	3	3	1	2013-10-11
1916 NORMERLE DR, GOODLETTSVILLE, TN	1	GENERAL SERVICES DISTRICT	40000	130400	170400	1986	3	2	1	2013-06-28
2050 GRACELAND DR, GOODLETTSVILLE, TN	1.28	GENERAL SERVICES DISTRICT	40000	238900	278900	1985	3	3	0	2015-02-10
2034 GRACELAND DR, GOODLETTSVILLE, TN	2.61	GENERAL SERVICES DISTRICT	52900	330700	383600	1997	5	4	1	2016-04-14
811 BENTON CT, GOODLETTSVILLE, TN	1.1	GENERAL SERVICES DISTRICT	40000	158000	198000	1989	3	3	0	2013-07-08
331 VIEW RIDGE DR, GOODLETTSVILLE, TN	1.69	GENERAL SERVICES DISTRICT	35100	116600	151700	1988	3	2	0	2015-08-31

-- Populate Property Address data (using self join)

```sql
SELECT *
FROM [Portfolio SQL].dbo.NashvilleHousing
-- WHERE PropertyAddress is NULL
ORDER BY ParcelID;




SELECT a.ParcelID, a.PropertyAddress, b.ParcelID, b.PropertyAddress,
ISNULL(a.PropertyAddress, b.PropertyAddress)
FROM [Portfolio SQL].dbo.NashvilleHousing a
JOIN [Portfolio SQL].dbo.NashvilleHousing b
        ON a.ParcelID = b.ParcelID
        AND a.[UniqueID ] <> b.[UniqueID ]
WHERE a.PropertyAddress IS NULL;




UPDATE a  --Updating Original Table using its alias a
SET PropertyAddress = ISNULL(a.PropertyAddress, b.PropertyAddress)
FROM [Portfolio SQL].dbo.NashvilleHousing a
JOIN [Portfolio SQL].dbo.NashvilleHousing b
        ON a.ParcelID = b.ParcelID
        AND a.[UniqueID ] <> b.[UniqueID ]
WHERE a.PropertyAddress IS NULL;
```

	UniqueID	ParcelID	LandUse	PropertyAddress	SaleDate	SalePrice	LegalReference	SoldAsVacant	OwnerName
223	22721	026 01 0 069.00	VACANT RESIDENTI...	141 TWO MILE PIKE, GOODLETTSVILLE	2014-10-27 00:00:00.000	59800	20141031-0100471	Yes	SHACKLEFORD, MICHAEL C., JR.
224	39432	026 01 0 069.00	VACANT RESIDENTI...	141 TWO MILE PIKE, GOODLETTSVILLE	2015-10-23 00:00:00.000	153000	20151028-0109602	No	SHACKLEFORD, MICHAEL C., JR.
225	31082	026 01 0 109.00	SINGLE FAMILY	130 TWO MILE PIKE, GOODLETTSVILLE	2015-05-27 00:00:00.000	43000	20150610-0054642	No	KAVANAGH, SAMANTHA & SOSA, MARI
226	33028	026 01 0 112.00	SINGLE FAMILY	136 TWO MILE PIKE, GOODLETTSVILLE	2015-06-01 00:00:00.000	133000	20150604-0052410	No	GOOLSBY, BERRY & CARLENE
227	22722	026 01 0 153.00	SINGLE FAMILY	143 TWO MILE PIKE, GOODLETTSVILLE	2014-10-27 00:00:00.000	59800	20141031-0100471	Yes	MEADOR, ERIC M.
228	29471	026 01 0 153.00	SINGLE FAMILY	143 TWO MILE PIKE, GOODLETTSVILLE	2015-04-15 00:00:00.000	153000	20150417-0034215	No	MEADOR, ERIC M.
229	13378	026 02 0 005.00	SINGLE FAMILY	313 WITHAM CT, GOODLETTSVILLE	2014-03-14 00:00:00.000	177900	20140321-0023642	No	PATTERSON, GEORGE, III & PATRICIA &
230	46902	026 02 0 006.00	SINGLE FAMILY	121 LIDGATE TER, GOODLETTSVILLE	2016-04-29 00:00:00.000	169000	20160503-0043379	No	CLAXTON, JONATHAN & HOLLEE
231	28178	026 02 0 008.00	SINGLE FAMILY	409 NEWBERRY CT, GOODLETTSVILLE	2015-03-13 00:00:00.000	165000	20150319-0023524	No	DARNELL, KAY LEA
232	10782	026 02 0 010.00	SINGLE FAMILY	405 NEWBERRY CT, GOODLETTSVILLE	2013-12-31 00:00:00.000	172000	20140106-0001151	No	JACKSON, JOHNNY M. & REDONNICA
233	54610	026 02 0 010.00	SINGLE FAMILY	405 NEWBERRY CT, GOODLETTSVILLE	2016-09-07 00:00:00.000	197000	20160909-0094999	No	JACKSON, JOHNNY M. & REDONNICA
234	26231	026 02 0 015.00	SINGLE FAMILY	111 LIDGATE TER, GOODLETTSVILLE	2015-01-30 00:00:00.000	174000	20150206-0011210	No	OSBORNE, KIMBERLY
235	45312	026 02 0 016.00	SINGLE FAMILY	515 HANOVER CT, GOODLETTSVILLE	2016-03-04 00:00:00.000	194500	20160307-0021349	No	CHRISTION, KEVIN
236	7017	026 02 0 017.00	SINGLE FAMILY	513 HANOVER CT, GOODLETTSVILLE	2013-08-15 00:00:00.000	160000	20130823-0089241	No	ROGERS, CARROLL & CHRISTY
237	2081	026 02 0 018.00	SINGLE FAMILY	511 HANOVER CT, GOODLETTSVILLE	2013-04-22 00:00:00.000	164900	20130424-0040730	No	MARLAR, JOESPH ADAM & LORI GAIL
238	2082	026 02 0 022.00	SINGLE FAMILY	501 HANOVER CT, GOODLETTSVILLE	2013-04-15 00:00:00.000	186900	20130417-0038218	No	SNIDER, CHARLOTTE E.
239	16942	026 02 0 023.00	SINGLE FAMILY	201 ASHTEAD CT, GOODLETTSVILLE	2014-06-27 00:00:00.000	164900	20140703-0058681	No	SMITH, LEE ANN
240	8004	026 03 0 006.00	SINGLE FAMILY	331 WITHAM CT, GOODLETTSVILLE	2013-09-18 00:00:00.000	189000	20130923-0099641	No	CHAGANI, HUSSAIN
241	673	026 03 0 007.00	SINGLE FAMILY	329 WITHAM CT, GOODLETTSVILLE	2013-02-25 00:00:00.000	145000	20130228-0019937	No	WHITE, LINDSAY E.
242	12410	026 03 0 008.00	SINGLE FAMILY	327 WITHAM CT, GOODLETTSVILLE	2014-02-28 00:00:00.000	154000	20140306-0018734	No	FRIZZELL, VELMA C.
243	18387	026 03 0 009.00	SINGLE FAMILY	325 WITHAM CT, GOODLETTSVILLE	2014-07-29 00:00:00.000	154000	20140731-0068358	No	JAMES, ROBERT & ROSSI, SABRINA
244	12411	026 03 0 016.00	SINGLE FAMILY	311 CHISWICK CT, GOODLETTSVILLE	2014-02-18 00:00:00.000	194900	20140225-0015908	No	WALKER, DEBRA K.
245	39433	026 05 0 013.00	SINGLE FAMILY	207 B EAST AVE, GOODLETTSVILLE	2015-10-28 00:00:00.000	95000	20151029-0110063	No	SMITH, BENJAMIN BLAKE INHERITANCE
246	4521	026 05 0 017.00	SINGLE FAMILY	208 EAST AVE, GOODLETTSVILLE	2013-06-24 00:00:00.000	129900	20130626-0065757	No	TRIPP, MARVIN S. & DEBORAH YOUNG

-- *Breaking out Address into Individual Columns (Address, City, State)*

```sql
SELECT PropertyAddress
FROM [Portfolio SQL].dbo.NashvilleHousing
--Where PropertyAddress is null
--order by ParcelID

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

```sql
FROM [Portfolio SQL].dbo.NashvilleHousing;

ALTER TABLE [Portfolio SQL].dbo.NashvilleHousing
ADD PropertySplitAddress Nvarchar(255);

UPDATE [Portfolio SQL].dbo.NashvilleHousing
SET PropertySplitAddress = SUBSTRING(PropertyAddress, 1, CHARINDEX(',', PropertyAddress)
-1);

ALTER TABLE [Portfolio SQL].dbo.NashvilleHousing
ADD PropertySplitCity Nvarchar(255);

UPDATE [Portfolio SQL].dbo.NashvilleHousing
SET PropertySplitCity = SUBSTRING(PropertyAddress, CHARINDEX(',', PropertyAddress) +1,
LEN(PropertyAddress));




SELECT * FROM [Portfolio SQL].dbo.NashvilleHousing;




SELECT OwnerAddress FROM [Portfolio SQL].dbo.NashvilleHousing;




SELECT
PARSENAME(REPLACE(OwnerAddress, ',', '.'), 3),
PARSENAME(REPLACE(OwnerAddress, ',', '.'), 2),
PARSENAME(REPLACE(OwnerAddress, ',', '.'), 1)
FROM [Portfolio SQL].dbo.NashvilleHousing;




ALTER TABLE [Portfolio SQL].dbo.NashvilleHousing
ADD OwnerSplitAddress Nvarchar(255);

UPDATE [Portfolio SQL].dbo.NashvilleHousing
SET OwnerSplitAddress = PARSENAME(REPLACE(OwnerAddress, ',', '.'), 3);




ALTER TABLE [Portfolio SQL].dbo.NashvilleHousing
ADD OwnerSplitCity Nvarchar(255);

UPDATE [Portfolio SQL].dbo.NashvilleHousing
SET OwnerSplitCity = PARSENAME(REPLACE(OwnerAddress, ',', '.'), 2);




ALTER TABLE [Portfolio SQL].dbo.NashvilleHousing
ADD OwnerSplitState Nvarchar(255);
```

```sql
UPDATE [Portfolio SQL].dbo.NashvilleHousing
SET OwnerSplitState = PARSENAME(REPLACE(OwnerAddress, ',', '.'), 1);
```

```sql
SELECT * FROM [Portfolio SQL].dbo.NashvilleHousing;
```

ndValue	BuildingValue	TotalValue	YearBuilt	Bedrooms	FullBath	HalfBath	SaleDateConverted	PropertySplitAddress	PropertySplitCity	OwnerSplitAddress	OwnerSplitCity	OwnerSplitState
000	168200	235700	1986	3	3	0	2013-04-09	1808 FOX CHASE DR	GOODLETTSVILLE	1808 FOX CHASE DR	GOODLETTSVILLE	TN
000	264100	319000	1998	3	3	2	2014-06-10	1832 FOX CHASE DR	GOODLETTSVILLE	1832 FOX CHASE DR	GOODLETTSVILLE	TN
000	216200	298000	1987	4	3	0	2016-09-26	1864 FOX CHASE DR	GOODLETTSVILLE	1864 FOX CHASE DR	GOODLETTSVILLE	TN
000	147300	197300	1985	3	3	0	2016-01-29	1853 FOX CHASE DR	GOODLETTSVILLE	1853 FOX CHASE DR	GOODLETTSVILLE	TN
000	152300	202300	1984	4	3	0	2014-10-10	1829 FOX CHASE DR	GOODLETTSVILLE	1829 FOX CHASE DR	GOODLETTSVILLE	TN
000	190400	259800	1980	3	3	0	2014-07-16	1821 FOX CHASE DR	GOODLETTSVILLE	1821 FOX CHASE DR	GOODLETTSVILLE	TN
000	137900	177900	1976	3	2	0	2014-08-28	2005 SADIE LN	GOODLETTSVILLE	2005 SADIE LN	GOODLETTSVILLE	TN
000	157900	197900	1978	3	2	0	2016-09-27	1917 GRACELAND DR	GOODLETTSVILLE	1917 GRACELAND DR	GOODLETTSVILLE	TN
400	176900	222300	2000	3	2	1	2015-08-14	1428 SPRINGFIELD HWY	GOODLETTSVILLE	1428 SPRINGFIELD HWY	GOODLETTSVILLE	TN
000	179600	219600	1995	5	3	0	2014-08-29	1420 SPRINGFIELD HWY	GOODLETTSVILLE	1420 SPRINGFIELD HWY	GOODLETTSVILLE	TN
JLL	NULL	NULL	NULL	NULL	NULL	NULL	2015-04-14	2209 KAYLA DR	GOODLETTSVILLE	NULL	NULL	NULL
JLL	NULL	NULL	NULL	NULL	NULL	NULL	2013-12-12	109 BAILEY VIEW CT	GOODLETTSVILLE	NULL	NULL	NULL
JLL	NULL	NULL	NULL	NULL	NULL	NULL	2015-07-13	1900 TINNIN RD	GOODLETTSVILLE	NULL	NULL	NULL
000	203600	243600	2006	4	4	0	2013-06-07	629 GAYLEMORE DR	GOODLETTSVILLE	629 GAYLEMORE DR	GOODLETTSVILLE	TN
000	115400	155400	1974	4	2	0	2014-06-30	633 GAYLEMORE DR	GOODLETTSVILLE	633 GAYLEMORE DR	GOODLETTSVILLE	TN
000	180300	234000	1974	3	3	0	2014-06-30	637 GAYLEMORE DR	GOODLETTSVILLE	637 GAYLEMORE DR	GOODLETTSVILLE	TN
000	129200	169200	1974	4	2	0	2016-07-15	1976 SADIE LN	GOODLETTSVILLE	2001 SADIE LN	GOODLETTSVILLE	TN
000	127400	167400	1974	3	2	0	2015-03-31	644 GAYLEMORE DR	GOODLETTSVILLE	644 GAYLEMORE DR	GOODLETTSVILLE	TN
000	286500	346300	1974	3	3	1	2013-10-11	1921 NORMERLE DR	GOODLETTSVILLE	1921 NORMERLE DR	GOODLETTSVILLE	TN
000	130400	170400	1986	3	2	1	2013-06-28	1916 NORMERLE DR	GOODLETTSVILLE	1916 NORMERLE DR	GOODLETTSVILLE	TN
000	238900	278900	1985	3	3	0	2015-02-10	2050 GRACELAND DR	GOODLETTSVILLE	2050 GRACELAND DR	GOODLETTSVILLE	TN
900	330700	383600	1997	5	4	1	2016-04-14	2034 GRACELAND DR	GOODLETTSVILLE	2034 GRACELAND DR	GOODLETTSVILLE	TN
000	158000	198000	1989	3	3	0	2013-07-08	811 BENTON CT	GOODLETTSVILLE	811 BENTON CT	GOODLETTSVILLE	TN
100	116600	151700	1988	3	2	0	2015-08-31	331 VIEW RIDGE DR	GOODLETTSVILLE	331 VIEW RIDGE DR	GOODLETTSVILLE	TN
300	112800	152600	1982	3	2	0	2013-10-21	361 VIEW RIDGE DR	GOODLETTSVILLE	361 VIEW RIDGE DR	GOODLETTSVILLE	TN

-- Change Y and N to Yes and No in "Sold as Vacant" field

```sql
SELECT DISTINCT(SoldAsVacant), COUNT(SoldAsVacant)
FROM [Portfolio SQL].dbo.NashvilleHousing
GROUP BY SoldAsVacant
ORDER BY 2;
```

```sql
SELECT SoldAsVacant,
CASE WHEN SoldAsVacant = 'Y' THEN 'Yes'
    WHEN SoldAsVacant = 'N' THEN 'No'
        ELSE SoldAsVacant
        END
FROM [Portfolio SQL].dbo.NashvilleHousing;
```

```sql
UPDATE [Portfolio SQL].dbo.NashvilleHousing
SET SoldAsVacant = CASE WHEN SoldAsVacant = 'Y' THEN 'Yes'
    WHEN SoldAsVacant = 'N' THEN 'No'
        ELSE SoldAsVacant
        END
```

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

→

	SoldAsVacant	(No column name)
1	Yes	4675
2	No	51802

-- *Remove Duplicates*

```sql
WITH RowNumCTE AS(
SELECT *,
        ROW_NUMBER() OVER(
        PARTITION BY ParcelID,
                        PropertyAddress,
                        SalePrice,
                        SaleDate,
                        LegalReference
                        ORDER BY
                                UniqueID
                                ) row_num

FROM [Portfolio SQL].dbo.NashvilleHousing
--ORDER BY ParcelID;
)
DELETE
FROM RowNumCTE
WHERE row_num > 1
--ORDER BY PropertyAddress
```

	BuildingValue	TotalValue	YearBuilt	Bedrooms	FullBath	HalfBath	SaleDateConverted	PropertySplitAddress	PropertySplitCity	OwnerSplitAddress	OwnerSplitCity	OwnerSplitState	row_num
81	177400	222400	2014	3	2	1	2015-02-05	5210 A MICHIGAN AVE	NASHVILLE	5210 A MICHIGAN AVE	NASHVILLE	TN	2
82	234500	279500	2016	4	3	1	2015-02-06	5222 KENTUCKY AVE	NASHVILLE	5222 KENTUCKY AVE	NASHVILLE	TN	2
83	219700	264700	2016	3	3	0	2015-02-06	5300 KENTUCKY AVE	NASHVILLE	5300 KENTUCKY AVE	NASHVILLE	TN	2
84	210300	259100	2014	3	2	1	2015-02-10	5301 A LOUISIANA AVE	NASHVILLE	5301 A LOUISIANA AVE	NASHVILLE	TN	2
85	210300	259100	2014	3	2	1	2015-02-10	5301 A LOUISIANA AVE	NASHVILLE	5301 A LOUISIANA AVE	NASHVILLE	TN	2
86	117900	167900	1920	2	1	0	2015-02-03	5312 TENNESSEE AVE	NASHVILLE	5312 TENNESSEE AVE	NASHVILLE	TN	2
87	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-25	5572 DORY DR	ANTIOCH	NULL	NULL	NULL	2
88	267400	312400	2014	3	2	1	2015-02-04	5610 KENTUCKY AVE	NASHVILLE	5610 KENTUCKY AVE	NASHVILLE	TN	2
89	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-06	5755 S NEW HOPE RD	HERMITAGE	NULL	NULL	NULL	2
90	NULL	NULL	NULL	NULL	NULL	NULL	2014-02-06	5963 MONROE XING	ANTIOCH	NULL	NULL	NULL	2
91	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-19	6203 NEW YORK AVE	NASHVILLE	NULL	NULL	NULL	2
92	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-20	6829 HICKORY RIM CT	ANTIOCH	NULL	NULL	NULL	2
93	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-17	713 44TH AVE N	NASHVILLE	NULL	NULL	NULL	2
94	95800	120800	1969	2	1	0	2015-02-13	743 CROLEY DR	NASHVILLE	743 CROLEY DR	NASHVILLE	TN	2
95	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-13	813 POINT BREAK CR S	ANTIOCH	NULL	NULL	NULL	2
96	49400	62400	1960	2	1	0	2015-02-20	815 31ST AVE N	NASHVILLE	815 31ST AVE N	NASHVILLE	TN	2
97	102000	129500	1987	3	2	0	2015-02-26	820 DRESDEN CT	ANTIOCH	820 DRESDEN CT	ANTIOCH	TN	2
98	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-13	8528 BEAUTIFUL VALLEY ...	NASHVILLE	NULL	NULL	NULL	2
99	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-19	8553 BEAUTIFUL VALLEY ...	NASHVILLE	NULL	NULL	NULL	2
100	NULL	NULL	NULL	NULL	NULL	NULL	2014-02-03	904 BIRCHMILL PT S	ANTIOCH	NULL	NULL	NULL	2
101	69400	95500	1954	2	1	0	2015-02-13	906 DELRAY DR	NASHVILLE	906 DELRAY DR	NASHVILLE	TN	2
102	NULL	NULL	NULL	NULL	NULL	NULL	2015-02-06	908 TOWNVIEW PL	NASHVILLE	NULL	NULL	NULL	2
103	NULL	NULL	NULL	NULL	NULL	NULL	2014-02-05	926 COARSEY DR	NASHVILLE	NULL	NULL	NULL	2
104	97300	111600	1958	2	1	0	2015-02-13	934 DRUMMOND DR	NASHVILLE	934 DRUMMOND DR	NASHVILLE	TN	2

-- *Delete Unused Columns*

```sql
SELECT *
FROM [Portfolio SQL].dbo.NashvilleHousing;

ALTER TABLE [Portfolio SQL].dbo.NashvilleHousing
DROP COLUMN OwnerAddress, TaxDistrict, PropertyAddress, SaleDate;
```