-- Change Date Format

SELECT SaleDateConverted, CONVERT(Date, SaleDate)

FROM Housing

ALTER TABLE Housing

Add SaleDateConverted Date;

Update Housing

SET SaleDateConverted = CONVERT(Date,SaleDate)

--Populate Property Address Data

SELECT \*

FROM Housing

ORDER BY ParcelID

SELECT a.ParcelID, a.PropertyAddress, b.ParcelID, b.PropertyAddress, ISNULL(a.PropertyAddress, b.PropertyAddress)

FROM Housing AS a

Join Housing AS b

ON a.ParcelID= b.ParcelID

AND a.[UniqueID]<> b.[UniqueID]

WHERE a.PropertyAddress IS NULL

UPDATE a

SET PropertyAddress = ISNULL (a.PropertyAddress, b.PropertyAddress)

FROM Housing AS a

Join Housing AS b

ON a.ParcelID= b.ParcelID

AND a.[UniqueID]<> b.[UniqueID]

WHERE a.PropertyAddress IS NULL

--Separating Address into individual columns

SELECT

SUBSTRING(PropertyAddress, 1,CHARINDEX(',',PropertyAddress)-1) AS Address,

SUBSTRING(PropertyAddress, CHARINDEX(',',PropertyAddress)+1,LEN(PropertyAddress)) AS Address

FROM Housing

ALTER TABLE Housing

ADD PropertySplitAddress Nvarchar(255)

UPDATE Housing

SET PropertySplitAddress = SUBSTRING(PropertyAddress, 1,CHARINDEX(',',PropertyAddress)-1)

ALTER TABLE Housing

ADD PropertySplitCity Nvarchar(255)

UPDATE Housing

SET PropertySplitCity = SUBSTRING(PropertyAddress, CHARINDEX(',',PropertyAddress)+1,LEN(PropertyAddress))

--Separate Owner Address (Best Of The Two)

SELECT

PARSENAME (REPLACE(OwnerAddress,',','.'),3),

PARSENAME (REPLACE(OwnerAddress,',','.'),2),

PARSENAME (REPLACE(OwnerAddress,',','.'),1)

FROM Housing

ALTER TABLE Housing

ADD OwnerSplitAddress Nvarchar(255)

UPDATE Housing

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

ALTER TABLE Housing

ADD OwnerSplitCity Nvarchar(255)

UPDATE Housing

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

ALTER TABLE Housing

ADD OwnerSplitState Nvarchar(255)

UPDATE Housing

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

--Change Y and N to Yes and No from Column "Sold as vacant"

SELECT Distinct(SoldAsVacant), Count(SoldAsVacant)

FROM Housing

GROUP BY SoldAsVacant

ORDER BY 2

SELECT SoldAsVacant

, CASE WHEN SoldAsVacant = 'Y' THEN 'Yes'

WHEN SoldAsVacant = 'N' THEN 'No'

ELSE SoldAsVacant

END

FROM Housing

UPDATE Housing

SET SoldAsVacant = CASE WHEN SoldAsVacant = 'Y' THEN 'Yes'

WHEN SoldAsVacant = 'N' THEN 'No'

ELSE SoldAsVacant

END

--Delete Duplicates

WITH RowNumCTE AS(

SELECT \*,

ROW\_NUMBER() OVER(

PARTITION BY ParcelID,

PropertyAddress,

SalePrice,

SaleDate,

LegalReference

ORDER BY

UniqueID

)row\_num

FROM Housing

)

SELECT\*

FROM RowNumCTE

WHERE row\_num >1

--Delete Unused Columns

ALTER TABLE Housing

DROP COLUMN OwnerAddress, TaxDistrict, PropertyAddress

ALTER TABLE Housing

DROP COLUMN SaleDate

SELECT \*

FROM Housing