**Cleaning Data in SQL Queries**

In this project, I cleaned dirty excel file in SQL to

1. **STDARDIZE DATE FORMAT--** (To set the date in a more pleasant format).
2. **POPULATE PROPERTY ADDRESS DATA-- (**To put same address with same ParcelID under empty PropertyAddresses).
3. **BREAKING OUT ADDRESS INTO INDIVIDUAL COLUMNS--**(Address, City, State).
4. **CHANGE Y AND N TO YES AND NO IN.**
5. Delete unused Columns.
6. I decided not to remove NULL or duplicate values in SQL because the deleted columns might be useful in future, I will rather do that in Excel or Power BI.

**--To check the entire file before starting the cleaning process.**

SELECT \*

FROM NatshvilleHousing

A screenshot of a computer

Description automatically generated with medium confidence

**1--Remove unwanted space before or in between the names**

SELECT TRIM([OwnerName])

FROM NatshvilleHousing

UPDATE NatshvilleHousing

SET [OwnerName] = TRIM([OwnerName])

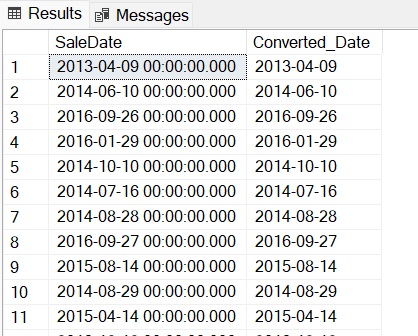
A screenshot of a computer

Description automatically generated with medium confidence

**2--STANDARDIZE DATE FORMAT-- (To set the date in a more pleasant format)**

SELECT SaleDate, CONVERT(DATE, SaleDate) AS Converted\_Date

FROM NatshvilleHousing



UPDATE NatshvilleHousing

SET SaleDate = CONVERT (DATE,SaleDate)

ALTER TABLE NatshvilleHousing

ADD SaleDateConverted DATE; --This will add another column known as SaleDateConverted

UPDATE NatshvilleHousing

SET SaleDateConverted = CONVERT (DATE,SaleDate) --This will set the converted dates into SaleDateConverted

A screenshot of a computer

Description automatically generated with medium confidence

**3--POPULATE PROPERTY ADDRESS DATA-- (To put same address with same ParcelID under empty PropertyAddresses)**

SELECT A.ParcelID, A.PropertyAddress, B.ParcelID, B.PropertyAddress, ISNULL(A.PropertyAddress,B.PropertyAddress) --Adds Address into the Empty PropertyAddress

FROM NatshvilleHousing A

JOIN NatshvilleHousing B

on A.ParcelID = B.ParcelID

AND A.[UniqueID ] <> B.[UniqueID ]

where A.PropertyAddress IS NULL

A screenshot of a computer

Description automatically generated with medium confidence

UPDATE A

SET PropertyAddress = ISNULL(A.PropertyAddress,B.PropertyAddress)

FROM NatshvilleHousing A

JOIN NatshvilleHousing B

on A.ParcelID = B.ParcelID

AND A.[UniqueID ] <> B.[UniqueID ]

where A.PropertyAddress IS NULL

A picture containing text, screenshot

Description automatically generated

**4--BREAKING OUT ADDRESS INTO INDIVIDUAL COLUMNS (ADDRESS, CITY, STATE)**

**--FOR PROPERTYADDRESS**

SELECT PropertyAddress

FROM NatshvilleHousing

A picture containing text, font, screenshot

Description automatically generated

SELECT

PARSENAME(REPLACE(PropertyAddress, ',','.'),2) Address,

PARSENAME(REPLACE(PropertyAddress, ',','.'),1) Address

FROM NatshvilleHousing

A screenshot of a computer

Description automatically generated with medium confidence

ALTER TABLE NatshvilleHousing

ADD PropertySplitAddress nvarchar(255);

UPDATE NatshvilleHousing

SET PropertySplitAddress = PARSENAME(REPLACE(PropertyAddress,',', '.'),2)

ALTER TABLE NatshvilleHousing

ADD PropertySplitCity nvarchar(255);

UPDATE NatshvilleHousing

SET PropertySplitCity = PARSENAME(REPLACE(PropertyAddress,',', '.'),1)

SELECT \*

FROM NatshvilleHousing

A screenshot of a computer

Description automatically generated with medium confidence

**5--CHANGE Y AND N TO YES AND NO IN [SOLD AS VACANT] FIELD**

SELECT SoldAsVacant,

CASE

WHEN SoldAsVacant = 'Y' then 'YES'

WHEN SoldAsVacant = 'N' then 'NO'

ELSE SoldAsVacant

END

FROM NatshvilleHousing

A screenshot of a computer

Description automatically generated with medium confidence

UPDATE NatshvilleHousing

SET SoldAsVacant = case

when SoldAsVacant = 'Y' then 'YES'

when SoldAsVacant = 'N' then 'NO'

when SoldAsVacant = 'Yes' then 'YES'

when SoldAsVacant = 'No' then 'NO'

ELSE SoldAsVacant

END

A picture containing text, screenshot, font, line

Description automatically generated

**6--Delete unused Columns**

select\*

from NatshvilleHousing

ALTER TABLE NatshvilleHousing

DROP COLUMN OwnerAddress, TaxDistrict, PropertyAddress

ALTER TABLE NatshvilleHousing

DROP COLUMN SaleDate

A screenshot of a computer

Description automatically generated with medium confidence