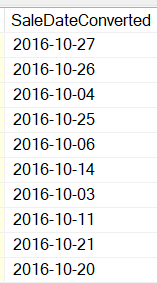
**Data Analyst**

**Aldiyar Baiserik Project**

Nashville Housing Query Document

SQL Data Cleaning

**Updating the Sale Date :**

Select SaleDate, Convert(Date, SaleDate) as NewDate 

From PortfolioProjectAldiyar..NashvilleHousing

ALTER TABLE NashvilleHousing

Add SaleDateConverted date ;

Update NashvilleHousing

Set SaleDateConverted = Convert(Date, SaleDate)

**Joining The Missing Addresses in with Same Parcel IDs :**

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

From PortfolioProjectAldiyar..NashvilleHousing a

join PortfolioProjectAldiyar..NashvilleHousing b

on a.ParcelID = b.ParcelID and a.[UniqueID ] <> b.[UniqueID ]

Where a.PropertyAddress is null

**A screenshot of a computer

Description automatically generated**

**Updating the Row with missing Data’s :**

Update a

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

From PortfolioProjectAldiyar..NashvilleHousing a

join PortfolioProjectAldiyar..NashvilleHousing b

on a.ParcelID = b.ParcelID and a.[UniqueID ] <> b.[UniqueID ]

Where a.PropertyAddress is null

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

Select SUBSTRING(PropertyAddress, 1, Charindex(',',PropertyAddress) -1) as Address, SUBSTRING(PropertyAddress, Charindex(',',PropertyAddress) +1, Len(PropertyAddress))

From NashvilleHousing

ALTER TABLE NashvilleHousing

Add PropertySplitAddress nvarchar(255) ;

Update NashvilleHousing

Set PropertySplitAddress = SUBSTRING(PropertyAddress, 1, Charindex(',',PropertyAddress) -1)

ALTER TABLE NashvilleHousing

Add PropertySplitCity nvarchar(255) ;

Update NashvilleHousing

Set PropertySplitCity = SUBSTRING(PropertyAddress, Charindex(',',PropertyAddress) +1, Len(PropertyAddress))

Select OwnerAddress

From NashvilleHousing

Select

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

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

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

From NashvilleHousing

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)

**A screenshot of a computer

Description automatically generated**

**A screenshot of a cell phone

Description automatically generatedChange Y and N to Yes and No in "Sold as Vacant" field:**

Update NashvilleHousing

Set SoldAsVacant =

Case When SoldAsVacant= 'Y' then 'Yes'

When SoldAsVacant= 'N' then 'No'

Else SoldAsVacant

End

**Deleting Duplicates :**

WITH RowNumCTE AS(

Select \*,

ROW\_NUMBER() OVER (

PARTITION BY ParcelID,

PropertyAddress,

SalePrice,

SaleDate,

LegalReference

ORDER BY

UniqueID

) row\_num

From PortfolioProjectAldiyar..NashvilleHousing

)

Delete

From RowNumCTE

Where row\_num > 1

**Delete Unused Columns:**

ALTER TABLE PortfolioProjectAldiyar.dbo.NashvilleHousing

DROP COLUMN OwnerAddress, TaxDistrict, PropertyAddress, SaleDate

A screenshot of a computer

Description automatically generated