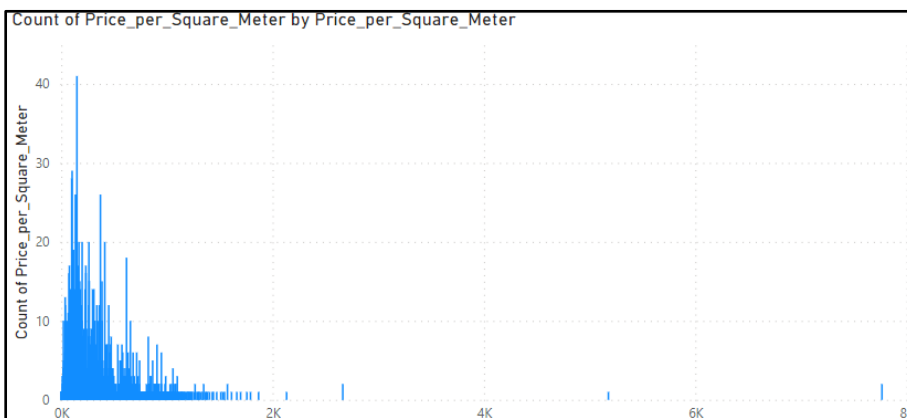


Nashville Housing Analysis

Introduction

This analysis focuses on cleaning and visualizing the dataset collected from [GitHub](#), which contains roughly 56,000 rows of housing data in Nashville between 1/2/2013 and 12/13/2019. When looking at the dashboards created to be intuitive and informative, I chose to use median over average because the price in this dataset is largely skewed to the left, which made the median a more accurate representation of the central tendency with less influence from the outliers.



◀ Figure 1

The distribution of price per square meter.

Click [here](#) for detail.

Technical Approach

Starting with the data cleaning using **MSSQL**, I standardized the sale dates by converting them into date values and separated the address and city from the property address column.

```
ALTER TABLE Portfolio_Project_NashvilleHousing..Nashville_Housing
ADD Sale_Date_Modified DATE
```

```
UPDATE Portfolio_Project_NashvilleHousing..Nashville_Housing
SET Sale_Date_Modified = CONVERT(DATE, SaleDate)
```

◀ Figure 2

One of the SQL queries that shows the process of datetime value conversion.

Click [here](#) for detail.

```
ALTER TABLE Portfolio_Project_NashvilleHousing..Nashville_Housing
ADD Property_address NVARCHAR(255),
    Property_city NVARCHAR(255)

UPDATE Portfolio_Project_NashvilleHousing..Nashville_Housing
SET Property_address = SUBSTRING(PropertyAddress, 1, CHARINDEX(',', PropertyAddress) - 1),
    Property_city = SUBSTRING(PropertyAddress, CHARINDEX(',', PropertyAddress) + 2, LEN(PropertyAddress))
```

▲ Figure 3

One of the SQL queries that shows the process of slicing the property address. Click [here](#) for detail.

Handling the missing property addresses by joining the table with itself and replace null values with the property address that shares identical parcel ID.

```
UPDATE t1
SET t1.PropertyAddress = ISNULL(t1.PropertyAddress, t2.PropertyAddress)
FROM Portfolio_Project_NashvilleHousing..Nashville_Housing AS t1
JOIN Portfolio_Project_NashvilleHousing..Nashville_Housing AS t2
ON t1.ParcelID = t2.ParcelID AND t1.[UniqueID ] != t2.[UniqueID ]
WHERE t1.PropertyAddress IS NULL
```

Figure 2

Substituting the missing addresses with addresses that have same parcel ID using inner join.

Click [here](#) for detail.

Imported the data queried from SQL database and created the following dashboards.

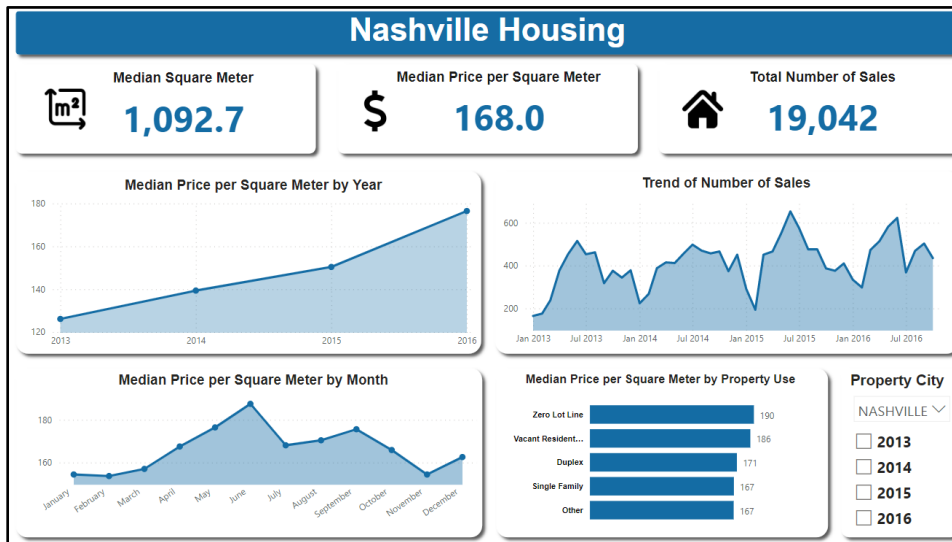


Figure 3

A brief market overview that shows key housing information, such as price, number of sales, trends, etc.

Click [here](#) for detail.

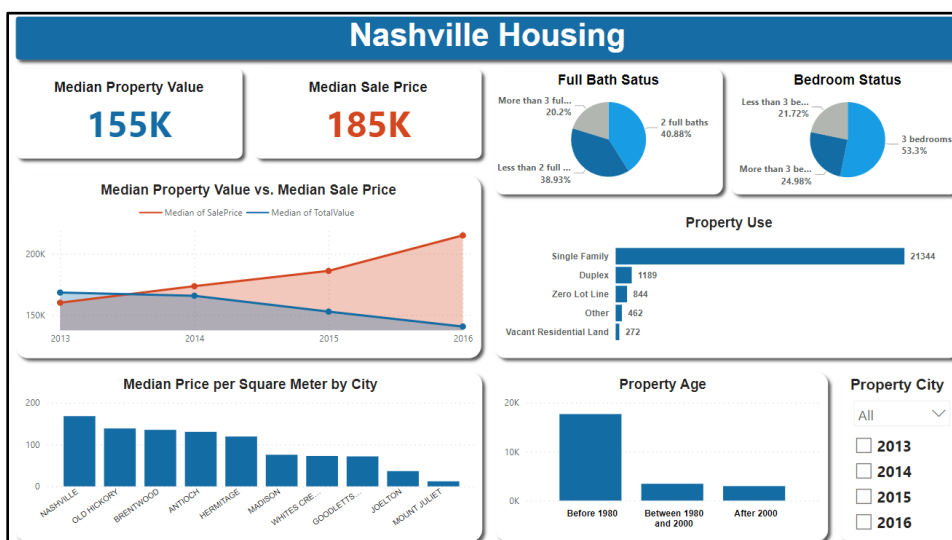


Figure 4

This dashboard includes more property details, such as property use and age, property value vs. sale price, etc.

Click [here](#) for detail.