**Background Information:**

The dataset of NYC Property Sales contains information about real estate transactions in New York City for the financial year 2016-2017. Government agencies, such as the NYC Department of Finance or similar organizations, typically collect and maintain this dataset. It includes a wide range of information about property sales, such as property addresses, sale prices, sale dates, property types, and more. This dataset contains the location, address, type, sale price, and sale date of building units sold.

**NYC Department of Finance:** This is a government agency responsible for collecting property-related taxes and maintaining property records in New York City. They collect data on property sales as part of their tax assessment and collection process.

**Real Estate Industry:** The real estate industry in New York City is a critical part of the economy. It includes real estate developers, brokers, property management companies, and investors. Access to accurate and comprehensive property sales data is crucial for various stakeholders in this industry to make informed decisions.

**Project Methods and Improvements:**

Leveraging advanced analytics and machine learning on the NYC Property Sales dataset can provide valuable insights and improvements in various aspects of the real estate industry and government operations. These methods can lead to more informed decision-making, reduced risks, and increased efficiency in business operations.

Methods like Clustering methods like k-means can group neighborhoods or properties with similar characteristics, assisting in market segmentation and targeted marketing strategies. BOROUGH and NEIGHBORHOOD columns can provide insights into the geographic distribution of properties and sales. We can analyze which boroughs or neighborhoods have the highest or lowest sales activity.

We can explore correlations between variables to identify relationships, such as how property size (square footage) relates to sales price. Linear regression and more advanced regression techniques can be used to model property prices based on various features such as property size, location, and more. Time series models can predict property price trends over time, helping investors and developers make informed decisions about when to buy or sell properties.

**Objectives:**

The main objective of the project is to present a Sale Price column and time-based analysis using Sale Dates, and how property prices have evolved over time. Additionally, by examining the Borough and Neighborhood columns, I can assess the spatial distribution of real estate activity, identifying which areas are experiencing heightened demand.

BUILDING CLASS CATEGORY and BUILDING CLASS AT PRESENT/TIME OF SALE columns allow for categorization and analysis of different property types. We can identify which types of properties are in high demand or have higher sales prices. Columns like RESIDENTIAL UNITS, COMMERCIAL UNITS, LAND SQUARE FEET, GROSS SQUARE FEET, and YEAR BUILT offer insights into property characteristics. We can analyze the size, type (residential or commercial), and age of properties sold.

**Data Description:**

|  |  |  |
| --- | --- | --- |
| **Name of the Attribute** | **Data type** | **Descriptions of the Attribute** |
| BOROUGH |  | The name of the borough in which the property is located. Manhattan (1), Bronx (2), Brooklyn (3), Queens (4), and Staten Island (5). |
| NEIGHBORHOOD | object | Department of Finance assessors determine the neighborhood name while valuing properties |
| BUILDING CLASS CATEGORY | object | This is a field that we are including so that users of the Rolling Sales Files can easily identify similar properties by broad usage (e.g., One Family Homes) without looking up individual Building Classes. |
| TAX CLASS AT PRESENT | object | Every property is assigned to one of four tax classes (Classes 1, 2, 3, and 4) |
| BLOCK | int64 | sub-division of the borough on which real properties are located |
| LOT | int64 | A Lot is a subdivision of a Tax Block and represents the property unique location |
| EASE-MENT | object | An easement is a right, such as a right of way, which allows an entity to make limited use of another’s real property |
| BUILDING CLASS AT PRESENT | object | The first position of the Building Class is a letter that is used to describe a general class of properties that signifies office buildings. The second position, a number, adds more specific information about the property’s use or construction style. |
| ADDRESS | object | The street address of the property as listed on the Sales File |
| APARTMENT NUMBER | object | Apartment Number allocated |
| ZIP CODE | int64 | The property’s postal code |
| RESIDENTIAL UNITS | int64 | The number of residential units at the listed property |
| COMMERCIAL UNITS | int64 | The number of commercial units at the listed property. |
| TOTAL UNITS | int64 | The total number of units at the listed property. |
| LAND SQUARE FEET | object | The land area of the property is listed in square feet. |
| GROSS SQUARE FEET | object | The total area of all the floors of a building as measured from the exterior surfaces of the outside walls of the building, |
| YEAR BUILT | int64 | Year the structure on the property was built. |
| TAX CLASS AT THE TIME OF SALE | int64 | tax classes at the time of sales |
| BUILDING CLASS AT THE TIME OF SALE | object | Building classes at the time of sales |
| SALE PRICE | object | Price paid for the property. |
| SALE DATE | object | Date the property sold. |