**Final Report**

**Subject: Real Estate Sales Records of State of Connecticut**

**Data From**:

<https://www.kaggle.com/datasets/anoopjohny/real-estate-sales-2001-2020-state-of-connecticut>

**For the Source Code, please refer to GitHub Link**:

<https://github.com/CzXiong1024/INFSCI_2415_Final_Project/tree/main>

**Data and method text describing the data and method used in this process:**

* Data sourced from **Real Estate Sales Records of State of Connecticut** spanning 2001-2020.
* Data Consists of: Serial Number, List Year, Date Recorded, Town, Address, Assessed Value, Sale Amount, Sales Ratio, Property Type, Residential Type, Non-Use Code, Assessor Remarks, OPM remarks, Location.
* Utilized Python with libraries: Pandas for data manipulation, Matplotlib and Seaborn for visualization, and NumPy for numerical operations.

**Significance statement on why the presented figures are important:**

* Provides a comprehensive historical perspective on real estate trends over two decades, serving as a foundation for predictive analytics.
* Enables stakeholders in the real estate industry to understand town-specific market dynamics and shifts.
* Assists industry professionals in anticipating future demand by revealing long-term preferences in residential types.
* Empowers investors and developers with data-driven insights to make informed decisions for future projects and investments.

Figure 1: Total Real Estate Sales Transactions by Top 20 Towns (2001-2020) (Bar Chart)

A graph showing a number of transactions

Description automatically generated

* **Real Estate Sales Transactions by Top 20 Towns**:
  + Bridgeport has the highest number of real estate sales transactions, close to 35000.
  + Middletown has the lowest among the top 20 towns.
  + The majority of towns have transactions between 10,000 and 20,000.

Figure 2: Average Real Estate Sales Price Trend (2001-2020) (Line Plot)

A graph showing the average sales of a home

Description automatically generated with medium confidence

* **Legend description:**
  + **Bridgeport (Blue Line):** Average sales price trend for Bridgeport from 2001-2020**.**
  + **Stamford (Orange Line):** Average sales price trend for Stamford from 2001-2020.
  + **Waterbury (Green Line):** A**v**erage sales price trend for Waterbury from 2001-2020.
  + **Norwalk (Red Line):** Averagesales price trend for Norwalk from 2001-2020.
  + **New Haven (Purple Line):** Average sales price trend for New Haven from 2001-2020.
  + **Danbury (Brown Line):** Average sales price trend for Danbury from 2001-2020.
  + **West Hartford (Light Purple Line):** Average sales price trend for West Hartford from 2001-2020.
  + **Hartford (Grey Line):** Average sales price trend for Hartford from 2001-2020.
  + **Milford (Light Green Line):** Average sales price trend for Milford from 2001-2020.
* **Average Real Estate Sales Price Trend (2001-2020)**:
  + Stamford has the highest average sale price for most of the years during 2001 to 2020.
  + Waterbury consistently has the lowest average sale price among the highlighted towns.
  + There's a noticeable dip in average sales prices around 2009 for all towns, likely indicative of the **global financial crisis.**
  + Sale prices generally started to recover post-2011, with a notable increase in 2016-2018.
  + There's a decline in average sale prices for all highlighted towns in 2020.

Figure 3: Distribution of Residential Type (Stem Plot)

A graph of different types of residential types

Description automatically generated

* **Legend description:**
  + **Condominium:** Color-coded with light blue and shaped in an upward-pointing triangle.
  + **Four Family:** Color-coded with orange and shaped in a rectangle.
  + **Single Family:** Color-coded with green and shaped in a pentagon.
  + **Three Family:** Color-coded with red and shaped in a diamond.
  + **Two Family:** Color-coded with purple and shaped in an oval.
* **Distribution of Residential Type**:
  + **Condominiums (Condo)** are the most popular residential type across all the selected towns, with the highest frequency observed in towns like Stamford, Norwalk, and Milford.
  + The town of **Waterbury** has the highest frequency of Condos, reaching over 10,000.
  + **Four Family** homes, represented by the orange square, also have a significant presence in many towns but are generally less frequent than Condos.
  + **Three Family** and **Two Family** homes have varying distributions across towns but are typically less frequent than Condos and Single Family homes.
  + The town of **New Heaven** has a noticeable diversity in its residential types, with a balanced distribution between Four Family, Single Family, Three Family.
  + **Milford** shows a distinct spike in Condos, surpassing all other residential types by a significant margin.