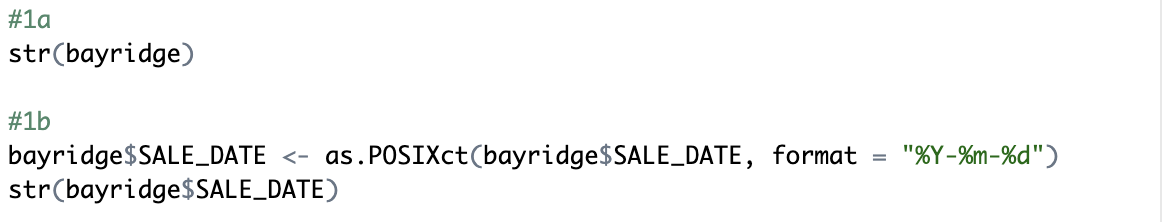
Meruyert Balgabek

MET AD 571 Assignement 5

Time Series Analytics

***Data Preparation***

**

*A screenshot of a computer code

Description automatically generated*

1. The dataset was loaded, and its structure was explored using the str() function. The data consists of 5,001 observations with 18 variables
2. The SALE\_DATE column was initially recognized as a POSIXct -a date format. The column was reprocessed to ensure it’s in the correct datetime format using the as.POSIXct() function with format "%Y-%m-%d" (for example, 2003-01-15)

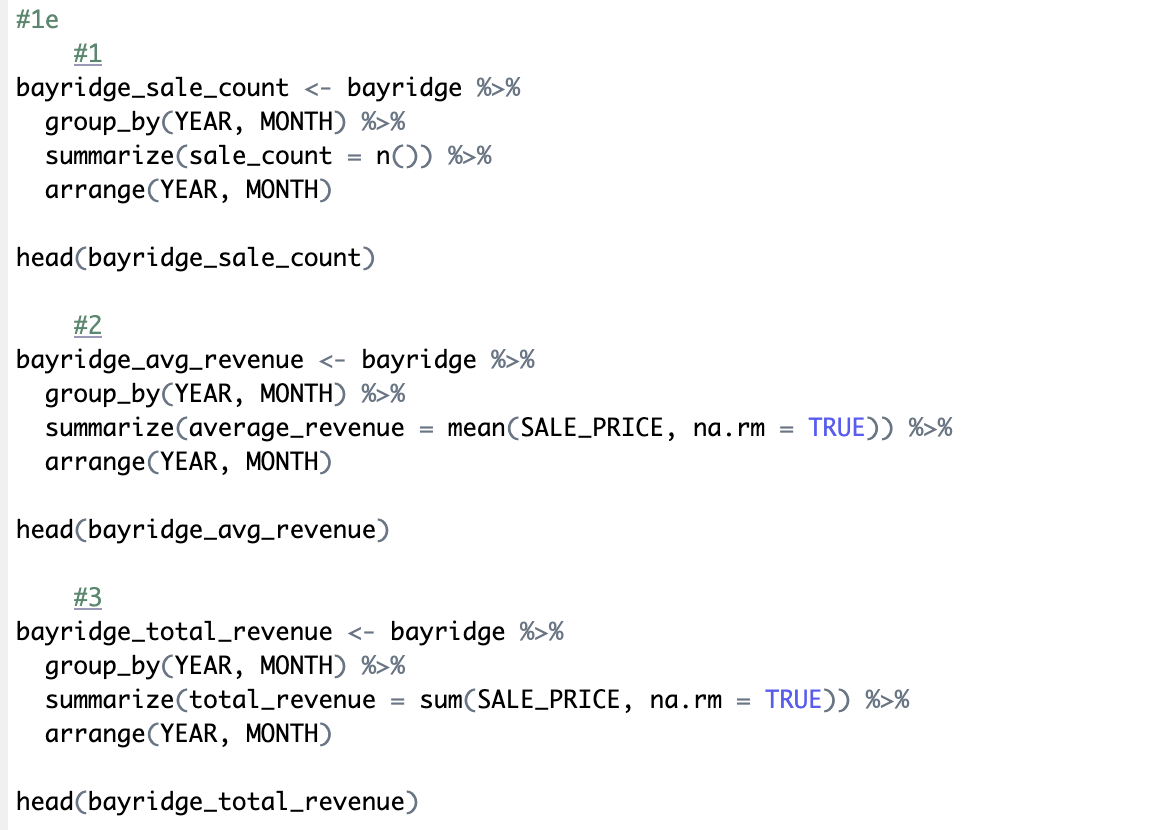
*A close-up of text

Description automatically generated*

*A screenshot of a computer

Description automatically generated*

1. Sum(is.na()) function confirmed there are no missing values in SALE\_PRICE
2. The YEAR and MONTH columns were derived from the SALE\_DATE column using the format() function



A white background with blue text

Description automatically generated

A white background with black dots

Description automatically generated

A blue and white background

Description automatically generated with medium confidence

1. Three new data frames were created:

Sale Count by Month, Average Revenue by Month, and Total Revenue by Month by grouping the data by YEAR and MONTH.

A screenshot of a computer code

Description automatically generated

A close-up of numbers

Description automatically generated

1. Each of the grouped data frames – sale count, average revenue, and total revenue –was converted into a time series using the ts() function.

Since all three data frames were derived from the same dataset, they share identical start and end dates. The start parameter was set to the first year and month present in the dataset. The frequency was set to 12 to reflect monthly data.

After the conversion, the structure of these time series was verified using the str() function to ensure they were properly formatted as time series objects.

***Exploratory Analysis***

A screenshot of a computer

Description automatically generated

*Plot 1: Monthly Sale Count in Bay Ridge, 2003-2021*

*A graph showing a number of years

Description automatically generated*

The chart shows the monthly sale count of properties in Bay Ridge over a period from 2003 to 2021. The sale counts fluctuate significantly, with peaks exceeding 40 sales in some months, particularly in the mid-2000s or 2014s, and lows below 10 sales around 2010 and 2020. The data exhibits a cyclical trend with varying degrees of volatility, indicating changes in the real estate market activity over time, possibly influenced by economic factors such as the 2008 financial crisis and the 2020 pandemic.

*Plot 2: Monthly Average Revenue in Bay Ridge, 2003-2021*

*A graph showing the average revenue

Description automatically generated*

Over time, the average revenue shows a general upward trend, with values rising from around 500,000 to peaks over 2,500,000. Noticeable spikes appear around 2015 and 2018, indicating months with exceptionally high revenues. Despite some fluctuations, there is an overall increase in average revenue, suggesting a long-term appreciation in property values or larger, higher-value sales in recent years.

*Plot 3: Monthly Total Revenue in Bay Ridge, 2003-2021*

A graph showing a number of revenue

Description automatically generated

The total revenue shows an increasing trend, with values fluctuating between 10 million and 70 million. Peaks in total revenue occur around 2015, 2018 and 2021. Despite the fluctuations, there is a noticeable upward trend, suggesting growth in the overall revenue generated from property sales in Bay Ridge over time.

*A screenshot of a computer code

Description automatically generated*

*Plot 4: Decomposition of Monthly Sale Count in Bay Ridge, 2003-2021*

*A graph showing the number of sales

Description automatically generated*

The trend shows some moderate shifts, including a decline in the mid-2000s till 2010 and a gradual recovery starting around 2012. But it didn’t reach the high counts as it reached prior 2005. The seasonal component reveals predictable, recurring cycles in sales throughout the year. The remainder captures irregular variations, suggesting periodic fluctuations due to external factors. Overall, the market shows steady activity with strong seasonality and some irregular variability.

*Plot 5: Decomposition of Monthly Average Revenue in Bay Ridge, 2003-2021*

*A graph of a number of different numbers

Description automatically generated with medium confidence*

The trend shows steady growth in average revenue, with a notable rise after 2010, indicating a strengthening market. The seasonal component reveals predictable, recurring fluctuations throughout the year. The remainder captures irregular spikes and dips, suggesting occasional variability due to external factors. Overall, the market indicates a long-term increase in revenue with clear seasonality and some irregular variations.

*Plot 6: Decomposition of Monthly Average Revenue in Bay Ridge, 2003-2021*

*A graph of different numbers and lines

Description automatically generated with medium confidence*

The trend shows fluctuations, with a decline around the mid-2000s to 2010, followed by a significant recovery starting around 2012, reaching new highs more recently. The seasonal component demonstrates predictable, recurring cycles throughout the year. The remainder reveals irregular spikes and dips, suggesting occasional impacts from external factors. Overall, the market shows growth in total revenue, strong seasonality, and periodic variability.

***Forecasting***

***A screenshot of a computer program

Description automatically generated***

***A screenshot of a computer

Description automatically generated***

*Plot 7: 12-Month Forecast of Sale Count (ETS)*

***A graph of a graph showing a number of years

Description automatically generated with medium confidence***

The Forecasted Sale Count for the next 12 months, generated using Exponential Smoothing State Space Model (ETS) shows notable fluctuations, as high as 25.27 in August 2022 and as low as 18.73 in February 2022.

*Plot 8: 12-Month Forecast of Total Revenue (ETS)*

***A graph of a graph showing the amount of revenue

Description automatically generated with medium confidence***

Predicted monthly Total Revenue range significantly, with a high in September (44.58M) and a low in November (27.66M).

*Plot 9: 12-Month Forecast of Average Revenue (ETS)*

***A graph of a graph showing the average revenue

Description automatically generated***

Predicted Average Revenue remain steady at 1.46M per month.

***Reflection Questions***

1. Insights from Trend and Seasonality in Bay Ridge Property Sales

The trend analysis for Bay Ridge shows a general increase in property values and sales activity over time, punctuated by fluctuations likely tied to economic events (e.g., 2008 financial crisis, 2020 pandemic). The seasonality component reveals predictable cycles, with certain months consistently exhibiting higher or lower sales (like high sales in months August or September and lower in wintertime). These patterns suggest that both market cycles and external economic events significantly impact property sales in this area, with a clear long-term upward trend in average and total revenue.

1. Usefulness of Time Series Analysis for Real Estate Stakeholders

Time series analysis provides stakeholders—real estate investors, agents, and policymakers—with valuable foresight into market behavior. For agents and brokers, recognizing seasonal patterns can guide timing for listings and marketing strategies, while long-term trend analysis helps investors assess potential property value appreciation. Policymakers can leverage these insights not only to forecast demand and allocate resources for infrastructure but also to assess the housing market's sensitivity to broader economic cycles. This understanding can inform policies that help stabilize the economy through various business cycles.

Additionally, this analysis could prompt further studies on housing market funding and mortgage exposure. Given that the mortgage market is closely tied to economic health and liquidity, stakeholders could use these insights to evaluate risks and opportunities, making informed decisions to support sustainable market growth and stability.

1. Dataset Limitations and Impact on Analysis

Limitations in the dataset include the absence of detailed property characteristics, such as property condition, type, and renovation history, which can significantly impact sales and pricing. Additionally, broader economic factors like interest rates and inflation—key drivers in real estate markets—are not included, which limits the analysis's ability to account for external economic influences. Another limitation is that we are studying only the Bay Ridge area in New York and it won’t be representative of NY or whole US housing market.

Funding types, particularly the distinction between cash purchases and financed (mortgage-based) transactions, are also missing from the dataset. This gap is crucial, as mortgage trends have shown significant influence on housing markets, especially during periods of economic boom or downturn. Without this information, it’s challenging to fully assess how mortgage market changes might drive or dampen sales activity, which could reduce the precision of forecasts and insights derived from the time series analysis.