Group Coursework Submission Form

Specialist Masters Programme

Please list all names of group members: (Surname, first name) 1. Martin, Nathaniel 2. Aliieva, Niiara 3. Alshaikh, Budour	4. Kiefer, Marius 5. 6. 7. **GROUP NUMBER:** **14**
MSc in: Business Analytics	
Module Code: SMM635	
Module Title: Data Visualization	
Lecturer: Dr Simone Santoni	**Submission Date: 12/11/24**

Declaration:

By submitting this work, we declare that this work is entirely our own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work we acknowledge that we have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. We also acknowledge that this work will be subject to a variety of checks for academic misconduct.

We acknowledge that work submitted late without a granted extension will be subject to penalties, as outlined in the Programme Handbook. Penalties will be applied for a maximum of five days lateness, after which a mark of zero will be awarded.

Marker's Comments (if not being marked on-line):

Deduction for Late Submission: ____

Final Mark: ____ %

Real Estate Investment Analysis in King County

Map of King County, Washington

Business Case: We are an analytics group within a real estate investment firm who is looking to expand their holdings into Seattle and the surrounding area. Our dataset is from King County, which contains this area in and around the city.

Dataset: Our data tracks the apartment sales in King County. Important information being: Sale Date, Sale Month, Sale Price (USD), Quality Grade, Quartile Zone, Number of Bedrooms, Living Space m^2, and Renovation Status. We wanted to target the variable's effect on price, where the data uncovers information regarding the potential return value of the various independent variables we are investigating.

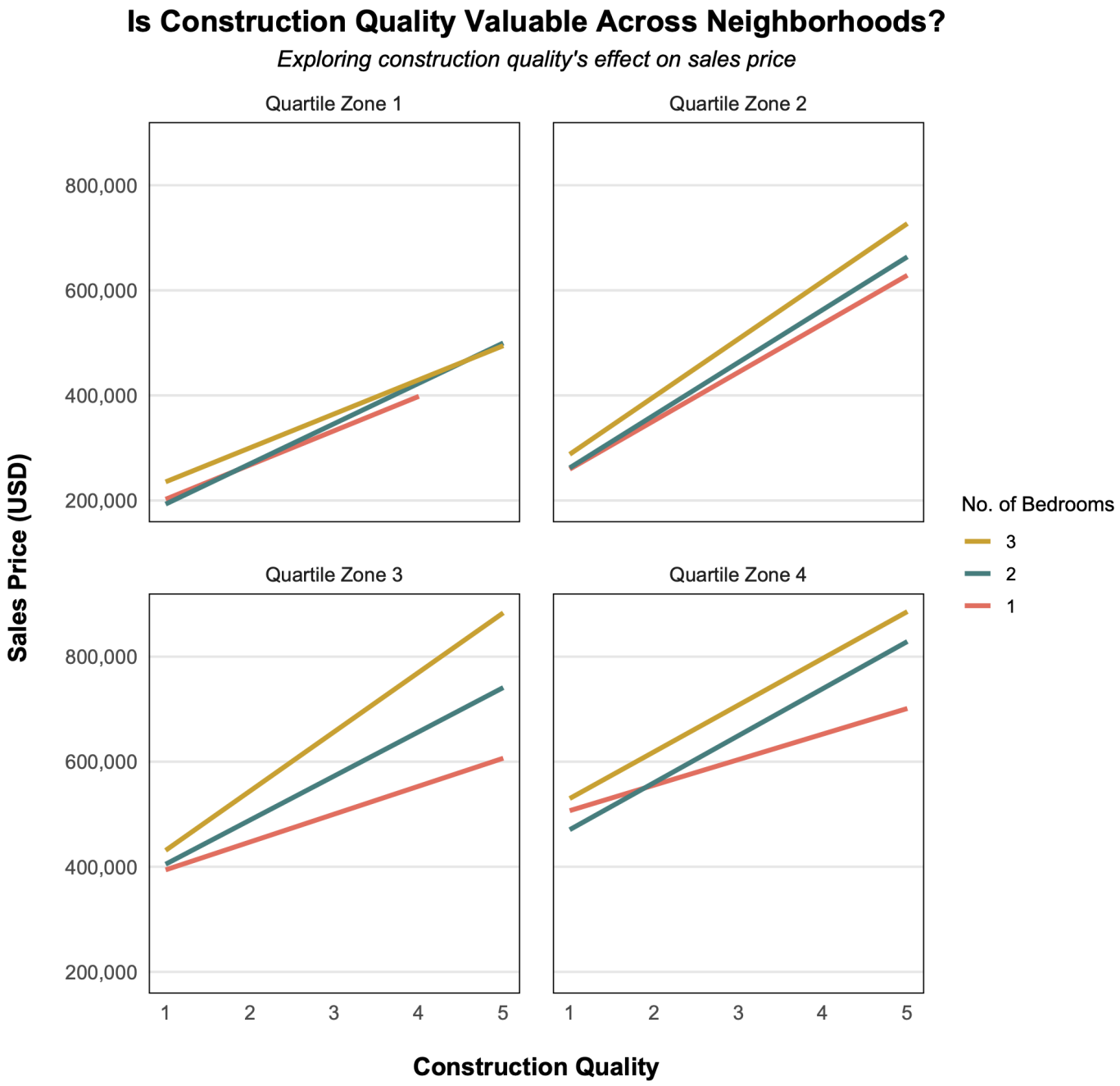

Figure 1: Linear regression visualizing the relationship of construction quality and sales price colored by number of bedrooms, faceted by quartile zone

We chose linear regression plots to show how construction quality directly influences home prices across different neighborhoods (Quartile Zones). By plotting quality scores against sale prices, we can instantly see the strength and direction of these relationships. The parallel display of four zones, from most affordable to premium areas, makes it easy to compare how quality improvements affect prices differently across neighborhoods.

The color coding for bedroom counts adds another layer, explaining how space and quality interact to drive prices in each zone. The straightforward linear relationships we observe are precious for real estate investment decisions, as they show consistent price increases with quality improvements. This helps us predict potential returns on construction investments across different market segments.

A positive correlation is demonstrated between construction quality and sale price across all neighborhoods. This trend holds for all numbers of bedrooms in apartments. However, the price varies depending on the neighborhood and bedroom count. The visualization illustrates that the neighborhood is sensitive to price and construction quality. Quartile Zone 4 has a sharper increase, meaning that quality improvements result in larger price increase compared to lower socioeconomic level in Quartile Zone 1. That indicates, customers targeting wealthier neighborhoods value construction quality more than lower income areas.

The bedroom count also influences prices, where apartments with more bedrooms command a higher price at every level of construction quality across all neighborhoods. In Quartile Zone 4, the jump in sale price between 1-bedroom and 2-bedroom is substantial, showing the willingness to purchase at a significantly higher price for larger living spaces and additional privacy features and flexibility that comes with higher bedroom count. In Zones 1 and 2, the increase in price with bedroom count is less steep. This suggests that in lower socioeconomic neighborhoods, buyers are unwilling to pay significantly more for additional bedrooms, prioritizing affordability over space.

These insights will allow us to make a case for business expansion in the area and help us to recommend targeting different classifications of apartments in the different Quartile Zone, to gain return on investment and meet demand based on market segments. Our recommendation is to focus on construction quality for higher socioeconomic neighborhoods; investing in high-quality materials, finishes, and craftsmanship for apartments in these areas will be beneficial. On the other hand, in lower quartiles, focus on developing smaller, more budget-friendly units with moderate quality. Another recommendation is, based on the potential client's needs; we could tailor our marketing strategy to Quartile Zones while emphasizing affordability and practicality in lower socioeconomic neighborhoods and luxury and construction quality in higher socioeconomic neighborhoods.

By aligning construction quality, pricing, and marketing strategies with customer preferences for each neighborhood, a business can better meet customer expectations and optimize its revenue.

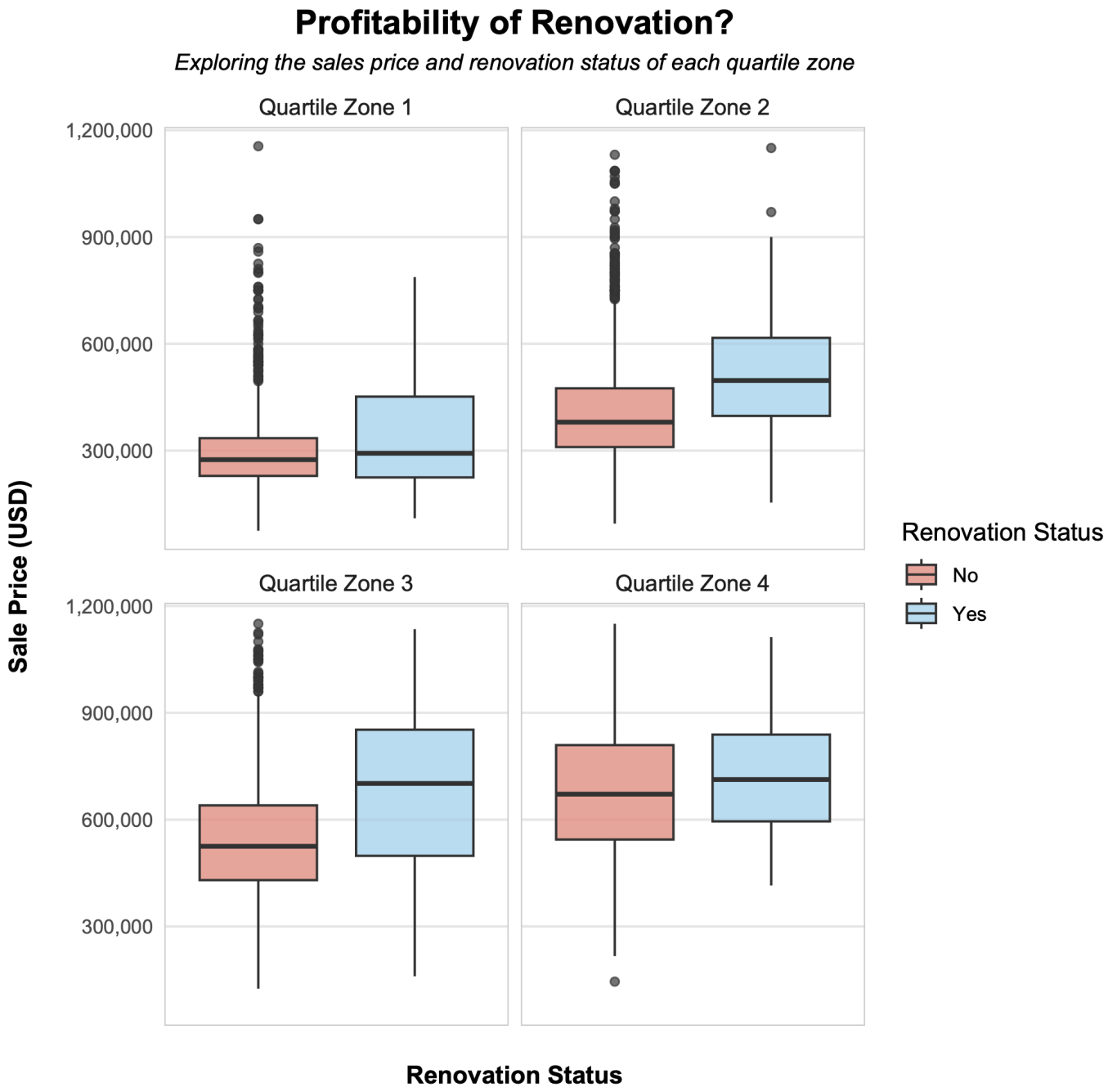

Figure 2: Boxplot visualizing the effect of renovations on sales price, faceted by quartile zone

Figure 2 compares house prices in King County by renovation status across quartile zones. We'll begin by explaining the value of a boxplot for this type of data. A boxplot is particularly useful for examining how house prices are distributed within each Quartile Zone, divided from less to more expensive areas (Quartile Zones). It provides a clear summary of each zone's median price, range, and spread, highlighting trends and differences between renovated and non-renovated homes.

Each box represents the interquartile range (IQR) - the middle 50% of values - allowing us to see where most prices cluster for each category. The line inside the box marks the median price, giving a quick sense of the typical sale price in each group. By comparing medians, we can easily assess if renovated homes tend to sell for more than non-renovated ones in each split. The "whiskers" extend to capture the range of prices, and individual outlier dots beyond the whiskers mark homes sold at notably higher or lower prices.

This format not only shows central trends but also reveals variability and outliers, giving a nuanced view of renovation effects across King County's market. For instance, the plot lets us observe that Quartile Zones 2 and 3 show the widest range of prices for renovated houses, while Quartile Zone 4 exhibits little difference in median prices between renovated and non-renovated homes. This visual tool is thus ideal for understanding how renovation impacts sale prices across varying quartile zones, highlighting areas where renovations add the most value and where they may be less impactful. Considering this information we are able to assess whether investing into renovations in each area is profitable or if we should already buy renovated property.

Renovation costs vary greatly, ranging from $15 to $60 per square foot, depending on the scope (e.g., kitchen, bathroom, and bedroom renovations) and materials used (Togal AI, 2024). This wide range makes it difficult to assess profitability consistently across all quartiles, as the price of materials and extent of renovations can significantly influence the final investment. The boxplot reveals that Quartile Zones 2 and 3 show the greatest variability in sale prices for renovated homes . This might indicate higher demand for updated properties in mid-tier neighborhoods, where buyers appear willing to pay more for renovations. In these quartiles, the median prices for renovated homes are substantially higher than for non-renovated ones, suggesting that renovation adds more value in these areas compared to Quartiles 1 and 4.

Based on renovation ROI (Return on investment) trends in the U.S. (typically averaging about 70%) (RenoFi, 2024), Quartile Zones 1 through 3 could be profitable if we as investors purchase non-renovated homes, renovate them, and then sell at a higher price. This aligns with the plot, which shows increased sale prices for renovated houses in Quartiles 1-3. However, Quartile Zone 4, representing the most expensive areas, shows little difference in price between renovated and non-renovated homes, suggesting renovations may not yield substantial returns in high-end areas. Buyers may prioritize property size and location over recent upgrades, thus lowering the potential profit from renovations.

Relying on 2014 data to estimate renovation profitability has limitations, as market conditions, renovation costs, and consumer preferences have changed significantly over the past decade. Housing markets have evolved, and comparing 2014 house prices with 2024 renovation costs highlights potential discrepancies due to inflation and rising material prices. According to

Togal AI (2024), renovation costs per square foot have been heavily influenced by inflation, supply chain disruptions, and labor shortages in recent years. For example, the cost of materials such as lumber, steel, and concrete has surged, leading to higher renovation expenses. Additionally, the increased demand for renovations post-pandemic has further driven up prices, making it challenging to rely on older data for accurate profitability estimates. More recent data would provide a better picture of current renovation trends, taking into account the heightened costs and shifting market dynamics.

The observed trends suggest that we may benefit most from buying non-renovated homes in Quartile Zones 1-3, renovating, and reselling. For Quartile Zone 4, where renovated and non-renovated homes have similar prices, renovations may not justify the expense, as the resale price doesn't increase significantly.

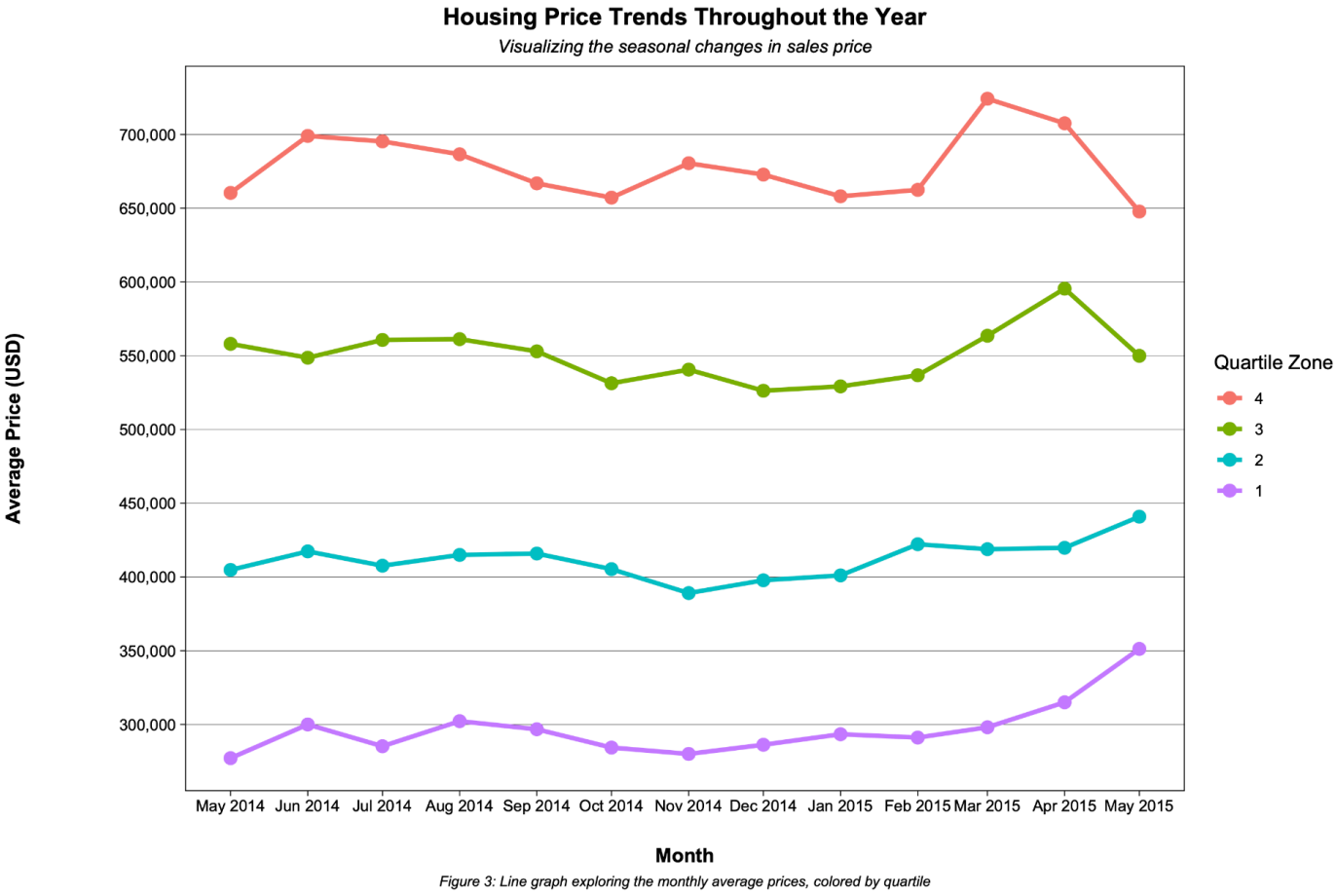

Figure 3: Line graph exploring the monthly average prices, colored by quartile

We chose a multiple-line graph to visualize Housing Price Trends as it effectively reveals price patterns across different market segments in King County. Each line, represented by a distinct color, tracks one of the four neighborhood types (Quartile Zones), allowing us to compare price movements simultaneously. The clean design helps us quickly identify how prices

move over the year, showing us both the big-picture trends and sudden market shifts from May 2014 to May 2015. By connecting the monthly prices with smooth lines, we can easily follow price movements in each neighborhood type and spot seasonal changes. This approach gives us a clear view of how different market segments relate to each other, making it practical for making smart investment choices.

Looking at the overall market behavior in King County, we see all four Quartile Zones showing a clear upward trend from January 2015 onwards, with the sharpest increases occurring between February and April 2015. Beyond these shared patterns, each market segment tells its own story. The luxury market (Zone 4) shows the most dramatic movements, starting at $660,000 in May 2014, reaching a peak of $720,000 in February 2015, before dropping sharply to $650,000 by May 2015. The upper-middle market (Zone 3) mirrors these swings but with less volatility, ranging from $530,000 to $600,000. Meanwhile, the more affordable segments show steadier growth. Zone 2 maintains a consistent climb from $400,000 to $440,000. In contrast, the entry-level market (Zone 1) demonstrates the most stable yet positive trajectory, rising steadily from $280,000 to $350,000 during the study period, showing particularly strong momentum in the final months.

Based on the year-long analysis of King County's housing market, we recommend focusing our investment strategy on Zone 1 properties for several compelling reasons. First, we are observing a remarkable 25% price appreciation in this segment from late 2014 to May 2015, with strong momentum continuing upward. This growth outperforms other zones in percentage terms and requires lower initial capital, making it an accessible entry point for our expansion into the Seattle market. Our data suggests optimal timing would be purchasing properties in October-November when prices traditionally dip, then holding through the substantial spring growth period in early 2015. The steady price growth in Zone 1 makes it more reliable than other areas, providing us with a solid foundation as we enter the Seattle market.

We see strategic opportunities in Zone 2 properties for a diversified approach, especially for value-add investments. The consistent price gaps between zones, approximately $100,000 to $150,000, create clear targets for property improvements. By purchasing Zone 1 properties and strategically renovating them, we could capture both natural market appreciation, which refers to the increase in property value over time due to market factors, and renovation value-add. However, we advise against heavy investment in Zone 4 properties at this time. While the luxury market shows the largest price movements - with swings up to $70,000 between March and May 2015 - these fluctuations represent significant risk, especially for a new market entry. The higher capital requirements and increased volatility of Zone 4 properties make them less suitable for our initial expansion strategy in the King County area.

Reference:

Togal AI. 2024. *Average Renovation Costs Per Square Foot.* Available at: https://www.togal.ai/blog/average-renovation-costs-per-square-foot [Accessed 12 Nov. 2024].

RenoFi. 2024. *Home Renovations with the Best Return on Investment.* Available at: https://www.renofi.com/home-renovations/renovation-best-return-on-investment/ [Accessed 12 Nov. 2024].