Understanding Real Estate Market Dynamics:

A Multi-Factor Analysis of Property Sales Patterns

Abstract

This study investigates the complex dynamics of the real estate market by analyzing the relationships between time on market, sales prices, and regional variations. Using a dataset of 38,057 real estate transactions across different regions of the United States, we examine how various factors influence property sales outcomes. Our findings indicate that time on market has a significant negative correlation with properties selling above list price, and that regional and property-type variations play crucial roles in determining sales outcomes. The analysis provides insights valuable for both real estate professionals and market participants.

1. Introduction

The real estate market represents one of the most significant sectors of the U.S. economy, with housing transactions affecting everything from individual wealth to broader economic indicators. Understanding the factors that influence property sales outcomes is crucial for market participants, policymakers, and researchers alike.

1.1 Research Context

Previous research has established that real estate markets are influenced by a variety of factors, including regional economic conditions, property characteristics, market timing, seasonal variations, and local market dynamics. However, the interplay between these factors and their relative importance in determining sales outcomes remains a subject of ongoing investigation.

1.2 Research Questions

This study focuses on three primary research questions:

Is there a significant correlation between time on market and the percentage of homes sold above list price?

How do seasonal adjustments impact housing inventory and prices across different regions?

What is the relationship between price drops and time on market across different property types?

1.3 Dataset Overview

Our analysis utilizes a comprehensive dataset containing 38,057 real estate transactions across multiple property types, including Single Family Residential, Condo/Co-op, Multi-Family, and Townhouse. The dataset covers regional variations across the Northeast, South, and West regions and includes temporal data that captures seasonal variations along with detailed metrics on price and market duration.

2. Exploratory Data Analysis

2.1 Time on Market and Sales Outcomes

Our initial exploration revealed that the relationship between time on market and sales outcomes varies significantly by region and property type. Key findings indicate that properties spending less time on the market are more likely to sell above list price, with distinct regional patterns influencing these outcomes. Additionally, different property types exhibit varying sensitivities to market duration.

2.2 Regional Market Characteristics

The analysis highlighted distinct regional patterns:

Northeast Region: This region shows the highest proportion of homes selling above list price (7.4% higher than baseline), with more volatile seasonal patterns and stronger price sensitivity to market duration.

South Region: Here, there is a lower overall proportion of above-list sales (1.8% below baseline), with more stable seasonal patterns and moderate price sensitivity to market duration.

West Region: This region's patterns align closely with national averages, showing no statistically significant deviation from baseline and consistent seasonal effects.

2.3 Property Type Analysis

Different property types displayed unique market behaviors:

Single Family Residential: Exhibited a 0.6% higher probability of selling above list price with consistent patterns across regions and stronger seasonal effects.

Condos/Co-ops: Showed a 4.4% lower probability of selling above list price with more variable regional patterns and less sensitivity to seasonal changes.

Multi-Family Properties: Had a 1.5% lower probability of selling above list price with mixed regional patterns and moderate seasonal sensitivity.

Townhouses: Demonstrated a 0.8% higher probability of selling above list price with patterns similar to single-family homes but strong regional variations.

3. Inferential Analysis

3.1 Regression Model Specifications

We employed a multiple linear regression model to analyze the relationship between time on market and sales outcomes while controlling for property type and regional variations.

3.2 Key Findings

The regression analysis revealed several significant relationships:

Time on Market Effect: The coefficient is -0.0008 ($p < 0.001$), indicating that for each additional day on the market, the probability of selling above list price decreases by 0.08 percentage points; this relationship is consistent across regions and property types.

Property Type Effects: Condos/Co-ops show the largest negative effect at -4.4%, while Single Family Residential and Townhouses exhibit positive effects of 0.6% and 0.8%, respectively; all property type effects are statistically significant ($p < 0.001$).

Regional Effects: The Northeast Region shows the strongest positive effect at 7.4%, whereas the South Region has a negative effect of -1.8%. The West Region shows no significant difference from baseline.

3.3 Model Performance

The regression model demonstrates reasonable explanatory power with an R-squared value of 0.2099, an adjusted R-squared of 0.2097, and an F-statistic of 1261 on 8 and 37975 degrees of freedom ($p < 2.2e-16$). This indicates that while the model captures some variance in sales outcomes, there remains room for further exploration of additional influencing factors.

4. Discussion and Implications

4.1 Key Insights

Our analysis uncovers several significant patterns in real estate market dynamics. Firstly, there is a strong negative correlation between market duration and the likelihood of above-list sales, indicating that properties that remain on the market longer tend to sell for less than their listing price. This effect is consistent across various regions, highlighting the importance of an effective initial pricing strategy for sellers. Secondly, regional variations are pronounced, with Northeast markets exhibiting distinct behaviors that persist even after controlling for other factors, underscoring the significance of local market conditions. Lastly, different property types demonstrate consistent patterns, with single-family homes and townhouses performing similarly, while condos exhibit unique market behavior.

4.2 Practical Implications

The findings from our analysis have several practical implications for market participants. For sellers, establishing a strong initial pricing strategy is essential, as it can greatly influence the success of a sale. The type of property also affects the optimal strategy, and sellers must consider their regional location to enhance their chances of favorable outcomes. For buyers, understanding the duration a property has been on the market can provide valuable negotiating leverage, allowing them to make more informed offers. Additionally, regional variations necessitate tailored buying strategies to effectively navigate different market conditions, while property type influences price dynamics and should be factored into purchasing decisions.

Real estate professionals need to adopt different strategies depending on the property type they are dealing with. Local market knowledge is crucial to understand regional variations and effectively advise clients. Moreover, time on market emerges as a critical factor in price negotiations, emphasizing the need for professionals to stay informed about how long properties remain available in their markets.

5. Conclusions and Future Research

5.1 Summary of Findings

Our analysis reveals several key insights into the real estate market. Firstly, time on market has a significant influence on sales outcomes, indicating that the duration a property remains available can affect its likelihood of selling above list price. Secondly, regional variations are crucial, as different areas exhibit distinct market behaviors that impact sales performance. Additionally, property types demonstrate unique characteristics in their market behavior, further emphasizing the complexity of the real estate landscape. Overall, market dynamics vary systematically across regions, highlighting the importance of localized analysis in understanding real estate trends.

5.2 Limitations

While our findings provide valuable insights, several limitations must be acknowledged. The model accounts for only 21% of the variance in sales outcomes, suggesting that other important factors influencing the market may not have been captured. Furthermore, potential interaction effects between variables were not explored in this study, which could provide deeper insights into market dynamics. Seasonal effects were also not fully accounted for, potentially overlooking fluctuations in market behavior throughout the year. Lastly, the geographic coverage of our dataset may be limited, which could affect the generalizability of our findings.

5.3 Future Research Directions

Future research could enhance our understanding of real estate dynamics by exploring several avenues. Investigating the effects of economic indicators on market dynamics could provide insights into broader economic trends influencing property sales. Additionally, examining local market competition and its impact on sales outcomes would be beneficial. Understanding the interaction between property characteristics and market outcomes could also yield valuable information for buyers and sellers alike. Finally, analyzing temporal changes in market patterns would help identify evolving trends and shifts in consumer behavior over time.

5.4 Additional Data Needs

To improve future analyses, several additional data sources would be beneficial. Incorporating local economic indicators can provide context for market conditions and trends. Ratings for school districts may influence buyer decisions and property values, making them an important factor to consider. Assessments of property condition can offer insights into how physical attributes affect sales outcomes. Furthermore, neighborhood demographics can help contextualize buyer preferences and trends within specific areas. Lastly, metrics on market competition would be valuable for understanding how competitive dynamics influence pricing and sales strategies in various regions.

Appendix A: Technical Details

[Code appendix to be added separately]

```
# Load required libraries

library(tidyverse)

library(lubridate)

library(scales)

library(stats)
```

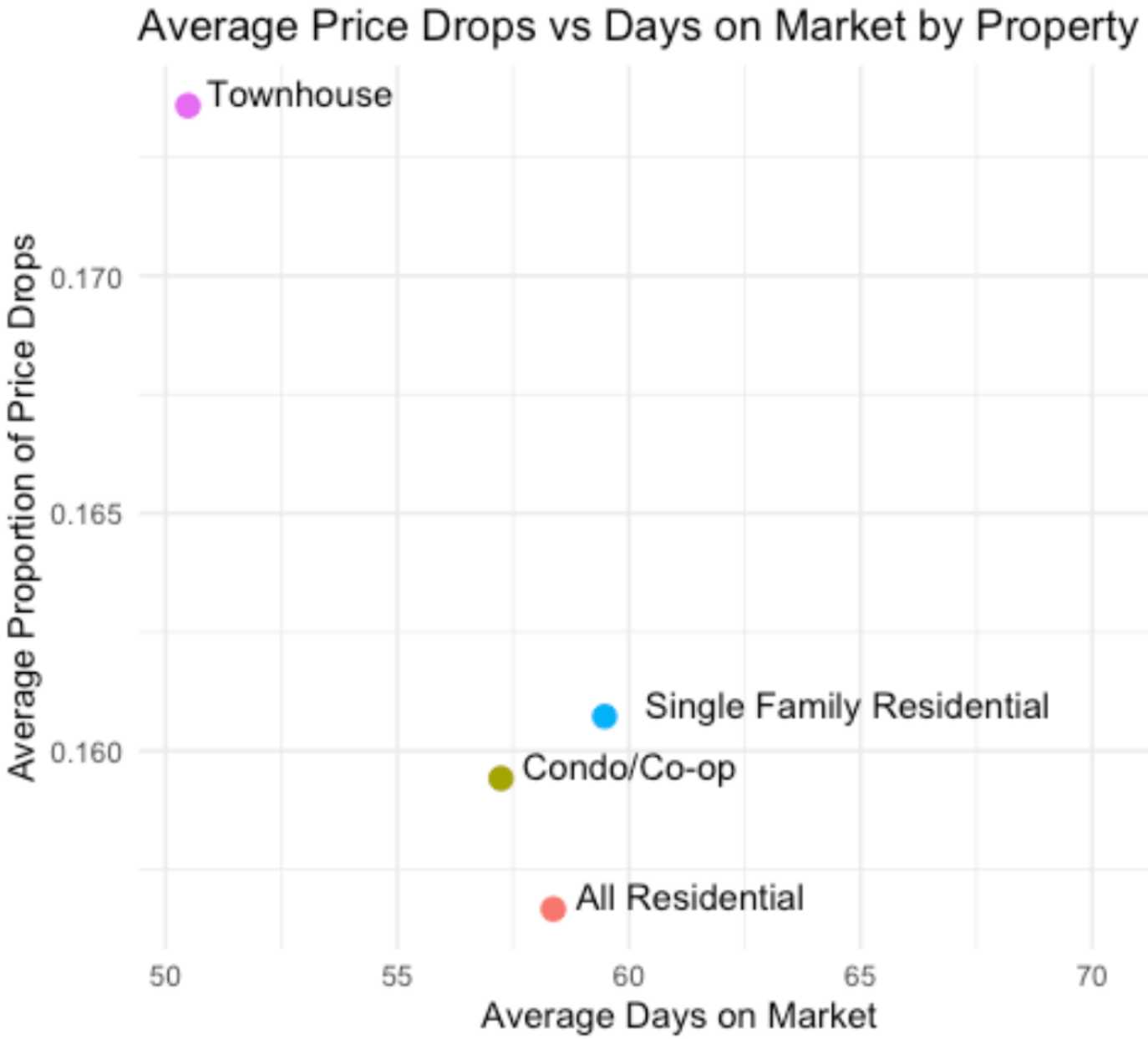
```
# Read and clean data
real_estate_data <- read.csv("your_data.csv") %>%
 # Convert dates
 mutate(
  period_begin = as.Date(period_begin),
  period_end = as.Date(period_end)
 ) %>%
 # Add season column
 mutate(
  season = case_when(
   month(period_begin) %in% c(12, 1, 2) ~ "Winter",
   month(period_begin) %in% c(3, 4, 5) ~ "Spring",
   month(period_begin) %in% c(6, 7, 8) ~ "Summer",
   month(period_begin) %in% c(9, 10, 11) ~ "Fall"
  )
 )

# Analysis for Question 2: Time on Market vs Sold Above List
market_time_analysis <- real_estate_data %>%
 filter(!is.na(median_dom) & !is.na(sold_above_list)) %>%
 select(period_begin, region, property_type, median_dom,
        sold_above_list, parent_metro_region)

# Question 2 visualization
ggplot(market_time_analysis,
       aes(x = median_dom, y = sold_above_list)) +
 geom_point(aes(color = parent_metro_region), alpha = 0.5) +
```

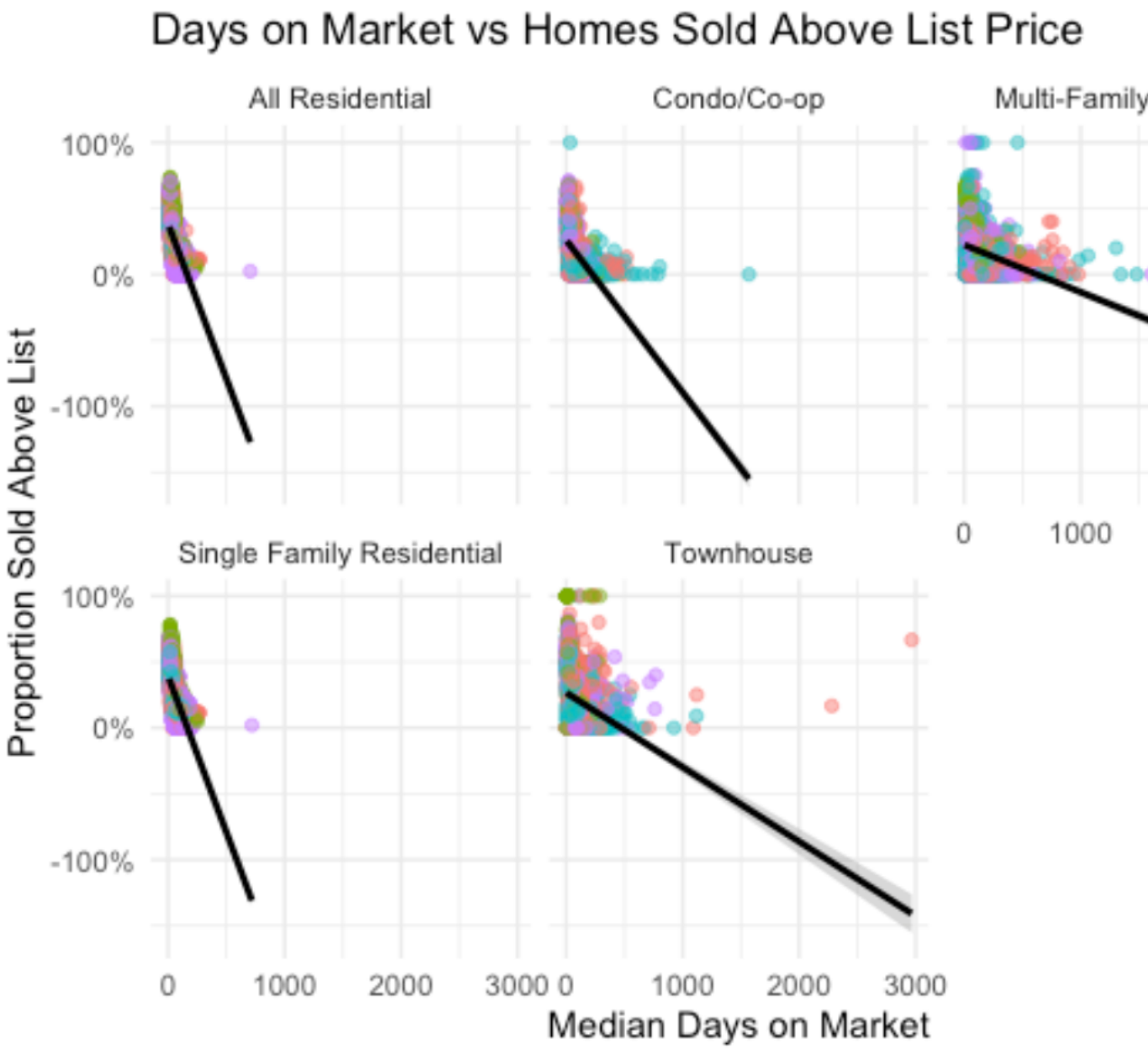
```
 geom_smooth(method = "lm", color = "black") +
 facet_wrap(~property_type) +
 scale_y_continuous(labels = scales::percent) +
 theme_minimal() +
 labs(title = "Days on Market vs Homes Sold Above List Price",
     x = "Median Days on Market",
     y = "Proportion Sold Above List",
     color = "Region")

# Analysis for Question 3: Seasonal Effects
# Updated to handle grouping explicitly
seasonal_analysis <- real_estate_data %>%
 filter(!is.na(inventory) & !is.na(median_sale_price)) %>%
 group_by(season, parent_metro_region) %>%
 summarize(
  avg_inventory = mean(inventory, na.rm = TRUE),
  avg_price = mean(median_sale_price, na.rm = TRUE),
  n = n(),
  .groups = "drop"  # Explicitly drop grouping
 )

# Question 3 visualization
ggplot(seasonal_analysis,
     aes(x = season, y = avg_price)) +
 geom_bar(stat = "identity", aes(fill = parent_metro_region),
        position = "dodge") +
 theme_minimal() +
 theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
```

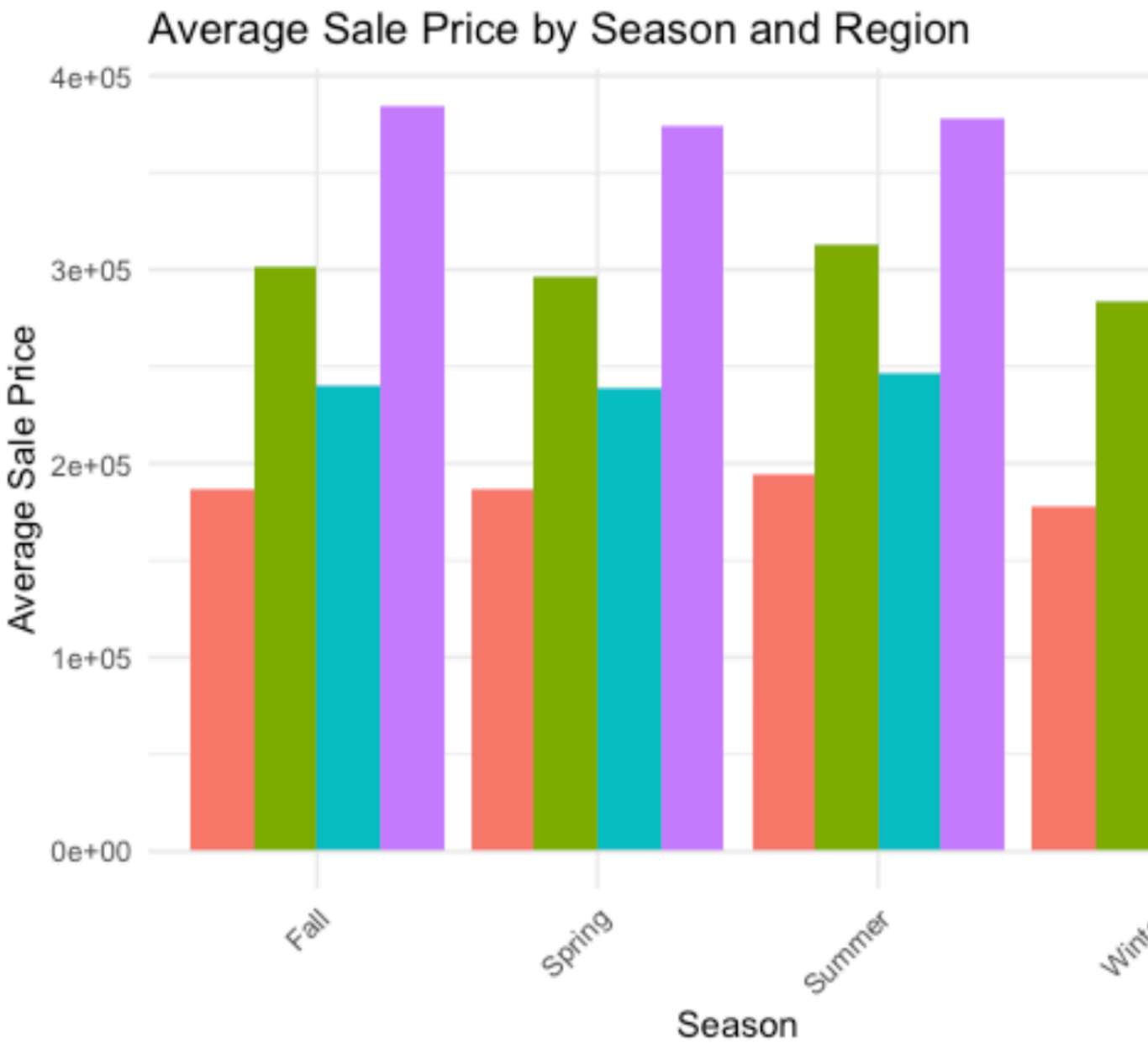
```
  labs(title = "Average Sale Price by Season and Region",
       x = "Season",
       y = "Average Sale Price",
       fill = "Region")

# Analysis for Question 4: Price Drops vs Time on Market
price_drop_analysis <- real_estate_data %>%
  filter(!is.na(price_drops) & !is.na(median_dom)) %>%
  group_by(property_type) %>%
  summarize(
    avg_price_drops = mean(price_drops, na.rm = TRUE),
    avg_dom = mean(median_dom, na.rm = TRUE),
    n = n(),
    .groups = "drop"  # Explicitly drop grouping
  )

# Question 4 visualization
ggplot(price_drop_analysis,
       aes(x = avg_dom, y = avg_price_drops)) +
  geom_point(aes(color = property_type), size = 3) +
  geom_text(aes(label = property_type), hjust = -0.1, vjust = 0) +
  theme_minimal() +
  labs(title = "Average Price Drops vs Days on Market by Property Type",
       x = "Average Days on Market",
       y = "Average Proportion of Price Drops")

# Inferential Analysis for Question 2
# Linear regression model
```

```
market_model <- lm(sold_above_list ~ median_dom + property_type +
                     parent_metro_region, data = market_time_analysis)

# Get model summary
summary(market_model)
```

Appendix B


Average Price Drops vs Days on Market by Property
Townhouse
Single Family Residential
Condo/Co-op
All Residential
Average Proportion of Price Drops
0.170
0.165
0.160
50
55
60
65
70
Average Days on Market


Days on Market vs Homes Sold Above List Price
All Residential
Condo/Co-op
Multi-Family
Single Family Residential
Townhouse
Proportion Sold Above List
100%
0%
-100%
0
1000
2000
3000
Median Days on Market


Average Sale Price by Season and Region
4e+05
3e+05
2e+05
1e+05
0e+00
Average Sale Price
Fall
Spring
Summer
Season