Findings & Recommendations – Real Estate Analysis

1.1 Analysis 1: EDA – Seasonal Transaction Trends

1.1.1 Setup for the analysis

The selected data includes historical records from 2007 to 2017. During preprocessing, the data was grouped by year and month to calculate transaction counts, with missing values appropriately filled. The analysis logic involved using a heatmap to visualize transaction trends across months and years.

1.1.2 Analytics results

The exploratory data analysis revealed sales peaks in 2007, a decline starting in 2008, and recovery with stability from 2013 to 2016. Monthly trends showed lower transactions in early months, improving mid-year, reflecting seasonal variations.

1.1.3 Conclusion

This analysis demonstrated a clear trend of decline, recovery, and stability in sales over the years. Seasonal low sales in January and February presented opportunities for targeted strategies. These findings highlighted areas for further analysis and strategic decision-making.

1.2 Analysis 2: EDA - Old vs. New Property Price and Transaction Trends

1.2.1 Setup for the analysis

The selected data includes historical records for old and new properties from 2007 to 2017. During preprocessing, the data was grouped by property age (old or new) and year to calculate average prices and transaction counts. The analysis used bar charts to compare the average price and transaction count of old and new properties over time.

1.2.2 Analytics results

From 2007 to 2017, new properties initially had lower average prices than old properties, matched in 2014, and surpassed them from 2015 onward, while old properties consistently had higher transaction volumes, reflecting their stronger market presence.

1.2.3 Conclusion

The analysis revealed old properties' dominance in transaction volumes, while rising prices of new properties after 2014 highlighted shifting buyer preferences, offering insights for targeted pricing strategies and market growth opportunities.

1.3 Analysis 3: EDA - Property Type Price and Transaction Trends

1.3.1 Setup for the analysis

The selected data includes historical records from 2007 to 2017 for property types categorized as Detached, Semi-Detached, Terraced, Flats, and Other. During preprocessing, property type codes were mapped to descriptive names, and data was grouped by property type and year to calculate average prices and transaction counts. Line charts were used to analyze price and transaction trends over time.

1.3.2 Analytics results

From 2007 to 2017, "Other" properties saw a price spike from 2011 before declining, while "Terraced" properties consistently had the lowest prices but dominated transaction volumes, reflecting their market popularity.

1.3.3 Conclusion
The analysis revealed that "Other" property types had unique price dynamics but limited market activity, reflecting their niche appeal. In contrast, "Terraced" properties emerged as the most traded type, driven by their affordability and wider market appeal, evidenced by lower average prices and higher transaction volumes.

1.4 Analysis 4: EDA - Seasonal Price Trends

1.4.1 Setup for the analysis
The selected data includes historical records from 2015 to 2017. During preprocessing, data was grouped by month to calculate average prices and by year and month to calculate transaction counts, with missing values appropriately filled. Combined bar and line plots were used to visualize monthly price and transaction trends.

1.4.2 Analytics results
Property prices dropped in February and May but peaked in September and December, while transaction volumes remained stable and did not align with price fluctuations.

1.4.3 Conclusion
The analysis showed seasonal price fluctuations, with low prices in February and May and peaks in September and December, suggesting external factors influence pricing trends, while transaction volumes remained stable. These findings highlighted areas for further analysis.

1.5 Analysis 5: Prediction – Price and transaction trend by Property Type

1.5.1 Setup for the analysis
The data (1995–2015) was filtered by year and property type to calculate average property prices and transaction counts. Using Linear Regression, forecasts for property prices and transaction counts were made for 2017–2022. Cross-validation (KFold) was applied to measure model accuracy using Mean Absolute Error (MAE). Predictions for property prices and transaction counts were visualized in line plots, showing trends for each property type.

1.5.2 Analytics results
Property prices are forecasted to grow steadily, with 'Other' properties valued highest by 2022, while detached properties remain the most transacted despite declining volumes across all type.

Validation of the model has shown moderate accuracy with room for improvement,
Property Price Cross-Validation MAE: 257840.15
Transaction Count Cross-Validation MAE: 134435.65

1.5.3 Conclusion
Invest in Detached, Semi-Detached, Terraced, and Flats for steady growth. Consider 'Other' properties cautiously for long-term gains.

1.6 Analysis 6: Prediction – Seasonal Price and Transaction Trend

1.6.1 Setup for the analysis
The data (1995–2015) was filtered by year and grouped by month to calculate transaction counts and average prices. Using SARIMAX, forecasts for monthly prices and transactions were generated for 2016–

2022. Cross-validation with TimeSeriesSplit measured model accuracy using RMSE. Heatmaps visualized forecasted transactions and prices (2017–2022) by year and month.

1.6.2 Analytics results

Transactions peak in mid-year (June to August) and drop in January and February. Average prices follow a similar pattern, with the lowest values early in the year and a steady rise, peaking in 2022.

Validation of the model has shown acceptable errors with,

Average RMSE for Price: 7305.66, RMSE for each fold (Price): [1548.35, 9053.06, 2513.72, 17299.57, 6113.60]

Average RMSE for Transactions: 16362.43, RMSE for each fold (Transactions): [5091.83, 10082.84, 27589.583, 32348.42, 6699.50]

1.6.3 Conclusion

Invest during off-peak months (January and February) for better deals and long-term gains. Complete renovations before peak months to attract high-paying tenants or buyers and maximize ROI.

1.7 Analysis 7: Predictive analysis – Regional Transaction growth

1.7.1 Setup for the analysis

The data (1995–2015) was filtered by year and month to calculate transaction counts. Using SARIMAX, forecasts for transactions were made for 2016–2022. Backtesting with historical data measured forecast accuracy using MAE and RMSE. Year-over-Year (YoY) growth was calculated for 2018–2022, and a bar chart visualized the top 15 cities by average YoY growth.

1.7.2 Analytics results

Bristol and London lead with the highest average YoY growth rates (around 14%), making them top-performing cities. Sheffield, Northampton, and Manchester also show significant growth. Cities in England are expected to have more growth in transactions.

Validation of the model has shown acceptable errors with,

[London: MAE = 1045.26, RMSE = 1126.18] [Manchester: MAE = 191.96, RMSE = 219.13] [Bristol: MAE = 96.21, RMSE = 114.69] [Birmingham: MAE = 54.06, RMSE = 69.88] [Southampton: MAE = 46.38, RMSE = 56.66] [Norwich: MAE = 53.76, RMSE = 63.92] [Reading: MAE = 35.83, RMSE = 44.24] [Newcastle Upon Tyne: MAE = 47.26, RMSE = 58.28]

1.7.3 Conclusion

Prioritize property investments in cities with high transaction growth than lower ones. Focus on top performers from both England and Wales.