Exploratory Analysis of House Prices in King County, Washington

A DATA ANALYTICS CAPSTONE PROJECT
BY GROUP 7 WITH TECHCRUSH

Presentation Outline

- Project Brief
- Understanding the Real Estate Sector
- Who are the stakeholders In the Real Estate Industry?
- House Pricing in Real Estate
- House Pricing Problems
- The Data set: Sourcing, Cleaning, Analysis, Visualization
- Actionable Insights
- Conclusion

Project Brief

Project Title: Exploratory Analysis of House Prices in King County, Washington

Problem Statement:

Real estate pricing in King County is influenced by multiple factors (location, size, condition, etc.), but stakeholders often lack data-driven insights to guide decisions. Inaccurate pricing leads to suboptimal outcomes for buyers, sellers, and investors. This project analyzes historical sales data to uncover pricing trends, identify high-value areas, and quantify the impact of key features like renovations and seasonality.

Project Objectives:

1. To identify key factors that influence house prices in King County.
2. To analyze trends in house pricing across different zip codes, sizes, and house features.
3. To create an interactive dashboard for realtors or buyers to understand market behavior.
4. To derive actionable insights that can support property valuation and investment decisions.

Data Source:

- Dataset: King County House Sales (Kaggle)

ROCCC Analysis:

1. **Reliable**: Publicly available, curated dataset with 21,000+ records.
2. **Original**: Sourced from King County government records.
3. **Comprehensive**: Includes price, location, size, grade, renovation status, and sale dates.
4. **Current**: Covers sales from 2014–2015 (relevant for baseline trends).
5. **Cited**: Properly attributed to Kaggle/license terms.

Tools Used:

- Excel: Data cleaning, preliminary analysis.
- SQL: Querying and aggregating data (e.g., average prices, average prices by zip code).
- Power BI: Interactive dashboards for visualizations (price distributions, Impact of house features on pricing.

Our Audience: This project addresses;

1. **Real Estate Executives & Portfolio Managers**
2. **Investors & Developers** considering new markets
3. **Internal Strategy Teams** (marketing, product, operations)
4. **Sales & Customer Engagement Leads** looking to better position listings

Understanding the Real Estate Sector

What Is Real Estate?.

According to Investopia.com, Real estate is defined as the land and any permanent structures, like a home, or improvements attached to the land, whether natural or artificial.

Real estate is considered real property that includes land and anything permanently attached to it or built on it, whether natural or artificial. There are five main categories of real estate, which include **residential**, **commercial**, **industrial**, **raw land**, and **special use**.

Investing in real estate includes purchasing a home, a rental property, or land.
Indirect investment in real estate can be made via REITs or through pooled real estate investment.

Real Estate: Definition, Types, How to Invest in It

How to Invest in Real Estate

Some of the most common ways to invest in real estate include homeownership, investment or rental properties, and house flipping which is a real estate strategy that involves buying, renovating, and selling homes quickly for a profit.

The earnings from investing in real estate are generated from rent or leases, as well as an appreciation of the real estate's value. Real estate is dramatically affected by its location, and factors such as employment rates, the local economy, crime rates, transportation facilities, school quality, municipal services, and property taxes can affect the value of the real estate.

Stakeholders in real estate include:

1. **Homeowners**: Individuals or families who own and occupy properties.
2. **Buyers**: Individuals or entities purchasing properties.
3. **Sellers**: Individuals or entities selling properties.
4. **Real Estate Agents/Brokers**: Professionals facilitating transactions.
5. **Developers**: Companies or individuals building new properties.
6. **Investors**: Individuals or entities investing in properties.
7. **Lenders**: Banks, mortgage companies, or private lenders providing financing.
8. **Government Agencies**: Local, state, or federal agencies regulating real estate.

Why is Real Estate Important?

Real estate is crucial to society for several reasons:

1. Economic Benefits: Job creation, Economic growth, Tax revenue
2. Social Benefits: Shelter and housing, Community development, Investment opportunities
3. Infrastructure and Development: Urban planning, Community facilities, etc.
4. Personal and Financial Security: Wealth accumulation, Stability and security, etc.

Overall, Real estate plays a very important role in national development, personal investment, and economic stability. Whether in King County, USA, Nigeria or any state, understanding trends in the housing market helps governments, investors or even homeowners make better decisions.

House Pricing in Real Estate

House pricing in real estate refers to the value or cost of a property, influenced by factors like:

Location:
1. Neighborhood quality
2. Proximity to schools, public transportation, and amenities
3. Zoning and land use

Property Characteristics:
1. Size and layout
2. Age, condition, and construction quality
3. Number of bedrooms and bathrooms
4. Amenities (pool, garden, garage)

Market Conditions:
1. Supply and demand
2. Local market trends
3. Economic conditions (interest rates, employment)

House Pricing Problems

House pricing problems in real estate can arise from various factors, including:

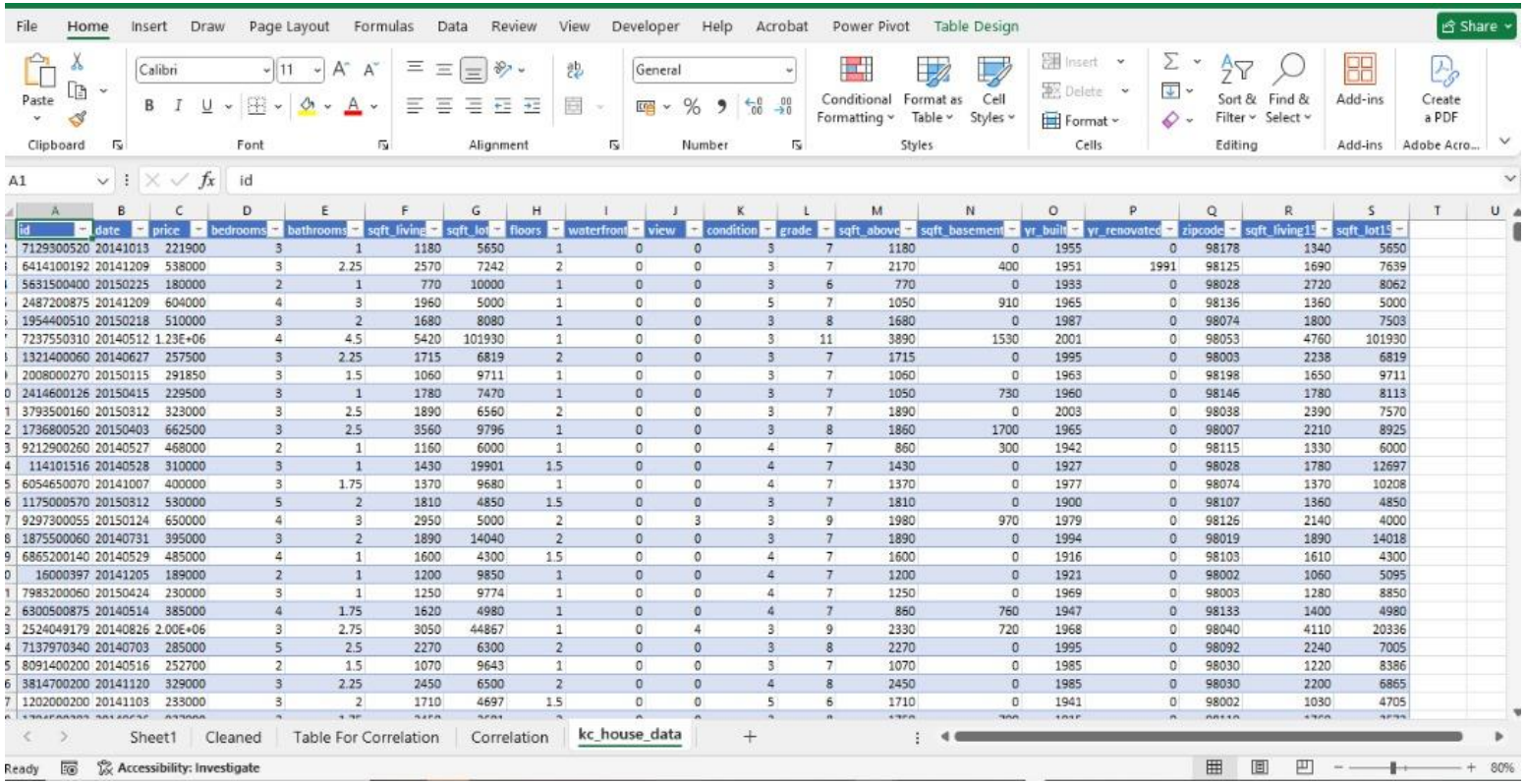
1. **Overpricing**: Setting an unrealistic price can deter potential buyers.
2. **Underpricing**: Undervaluing a property can result in lost revenue.
3. **Market fluctuations**: Changes in supply and demand, economic conditions, or interest rates can impact pricing.
4. **Comparable sales issues**: Difficulty finding comparable properties can make pricing challenging.
5. **Property condition**: Properties in poor condition may require price adjustments.
6. **Location**: Desirable or undesirable locations can significantly impact pricing.
7. **Lack of transparency**: Hidden defects or issues can lead to pricing disputes.
8. **Appraisal issues**: Discrepancies between market value and appraised value can affect pricing.

Sourcing the Dataset

As stated in our brief, we sourced our dataset from Kaggle where among some datasets, the King County real estate data made sense because of its completeness and relevance to Nigeria's urban real estate dynamics.

Exploring the Raw Data

The dataset contained over 21,000 records of property sales with attributes such as price, location (as zipcodes), number of bedrooms, bathrooms, square footage, renovation status, and some others.



id	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	grade	sqft_above	sqft_basement	yr_built	yr_renovated	zipcode	sqft_living15	sqft_lot15
7129300520	20141013	221900	3	1	1180	5650	1	0	0	3	7	1180	0	1955	0	98178	1340	5650
6414100192	20141209	538000	3	2.25	2570	7242	2	0	0	3	7	2170	400	1951	1991	98125	1690	7639
5631500400	20150225	180000	2	1	770	10000	1	0	0	3	6	770	0	1933	0	98028	2720	8062
2487200875	20141209	604000	4	3	1960	5000	1	0	0	5	7	1050	910	1965	0	98136	1360	5000
1954400510	20150218	510000	3	2	1680	8080	1	0	0	3	8	1680	0	1987	0	98074	1800	7503
7237550310	20140512	1.23E+06	4	4.5	5420	101930	1	0	0	3	11	3890	1530	2001	0	98053	4760	101930
1321400060	20140627	257500	3	2.25	1715	6819	2	0	0	3	7	1715	0	1995	0	98003	2238	6819
2008000270	20150115	291850	3	1.5	1060	9711	1	0	0	3	7	1060	0	1963	0	98198	1650	9711
2414600126	20150415	229500	3	1	1780	7470	1	0	0	3	7	1050	730	1960	0	98146	1780	8113
3793500160	20150312	323000	3	2.5	1890	6560	2	0	0	3	7	1890	0	2003	0	98038	2390	7570
1736800520	20150403	662500	3	2.5	3560	9796	1	0	0	3	8	1860	1700	1965	0	98007	2210	8925
9212900260	20140527	468000	2	1	1160	6000	1	0	0	4	7	860	300	1942	0	98115	1330	6000
114101516	20140528	310000	3	1	1430	19901	1.5	0	0	4	7	1430	0	1927	0	98028	1780	12697
6054650070	20141007	400000	3	1.75	1370	9680	1	0	0	4	7	1370	0	1977	0	98074	1370	10208
1175000570	20150312	530000	5	2	1810	4850	1.5	0	0	3	7	1810	0	1900	0	98107	1360	4850
9297300055	20150124	650000	4	3	2950	5000	2	0	3	3	9	1980	970	1979	0	98126	2140	4000
1875500060	20140731	395000	3	2	1890	14040	2	0	0	3	7	1890	0	1994	0	98019	1890	14018
6865200140	20140529	485000	4	1	1600	4300	1.5	0	0	4	7	1600	0	1916	0	98103	1610	4300
16000397	20141205	189000	2	1	1200	9850	1	0	0	4	7	1200	0	1921	0	98002	1060	5095
7983200060	20150424	230000	3	1	1250	9774	1	0	0	4	7	1250	0	1969	0	98003	1280	8850
6300500875	20140514	385000	4	1.75	1620	4980	1	0	0	4	7	860	760	1947	0	98133	1400	4980
2524049179	20140826	2.00E+06	3	2.75	3050	44867	1	0	4	3	9	2330	720	1968	0	98040	4110	20336
7137970340	20140703	285000	5	2.5	2270	6300	2	0	0	3	8	2270	0	1995	0	98092	2240	7005
8091400200	20140516	252700	2	1.5	1070	9643	1	0	0	3	7	1070	0	1985	0	98030	1220	8386
3814700200	20141120	329000	3	2.25	2450	6500	2	0	0	4	8	2450	0	1985	0	98030	2200	6865
1202000200	20141103	233000	3	2	1710	4697	1.5	0	0	5	6	1710	0	1941	0	98002	1030	4705

Data Cleaning and Standardization

The dataset from Kaggle was already clean and well-organized—free from missing values, duplicate entries, or obvious inconsistencies.
However, to make it more suitable for analysis, especially in Excel and Power BI, we needed to standardize certain columns.
We focused mainly on formatting where We converted the date column from a general data type to a proper Date format, so time-based trends could be easily analyzed.
We also changed the price column from general to Currency format. This little standardization is to ensure that all the tools (Excel, SQL Server, and Power BI) would be able to interpret the data accurately.
We also grouped numeric data into categories. For example, we binned square footage into ranges and converted some variables into readable formats.

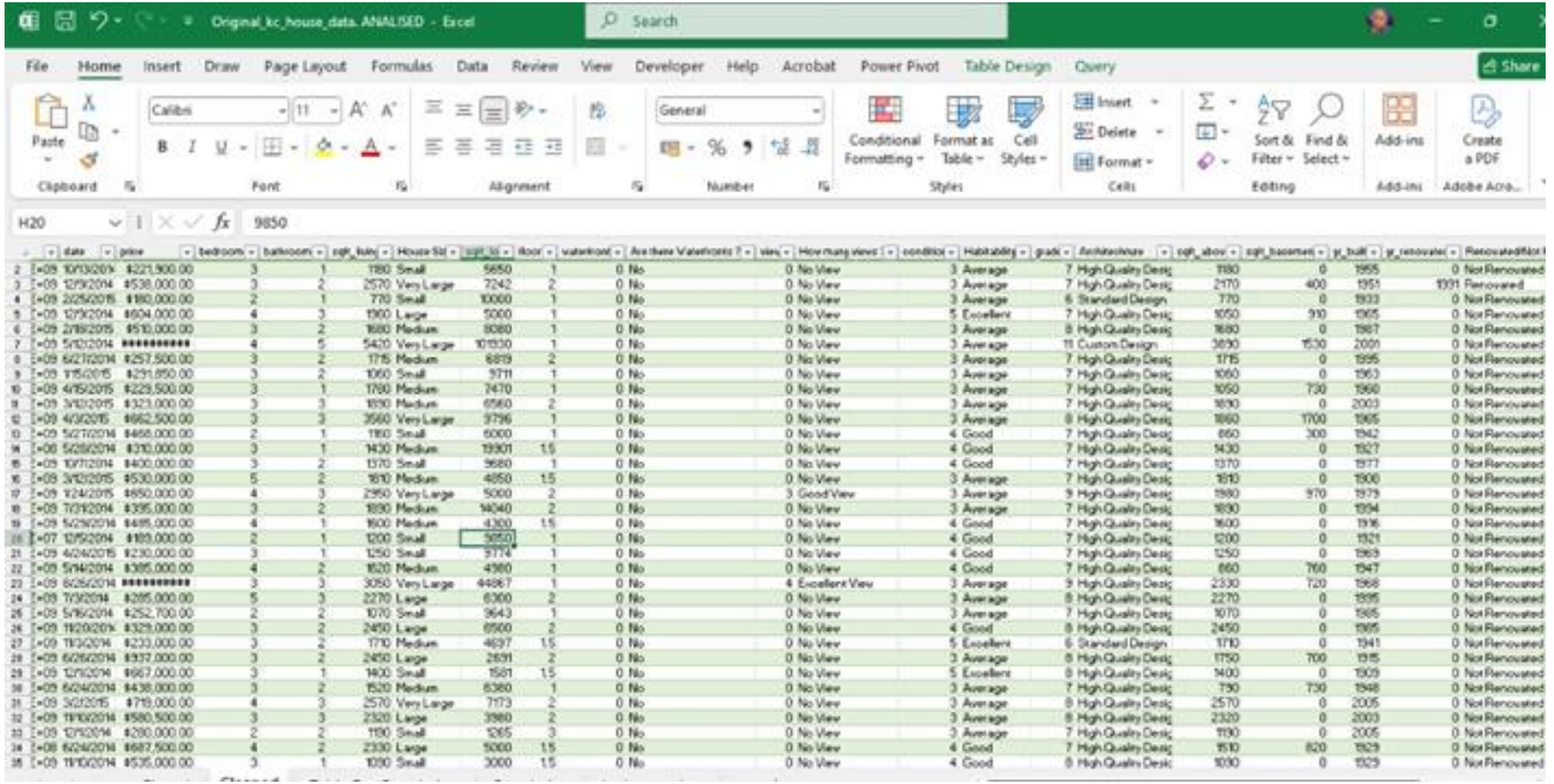
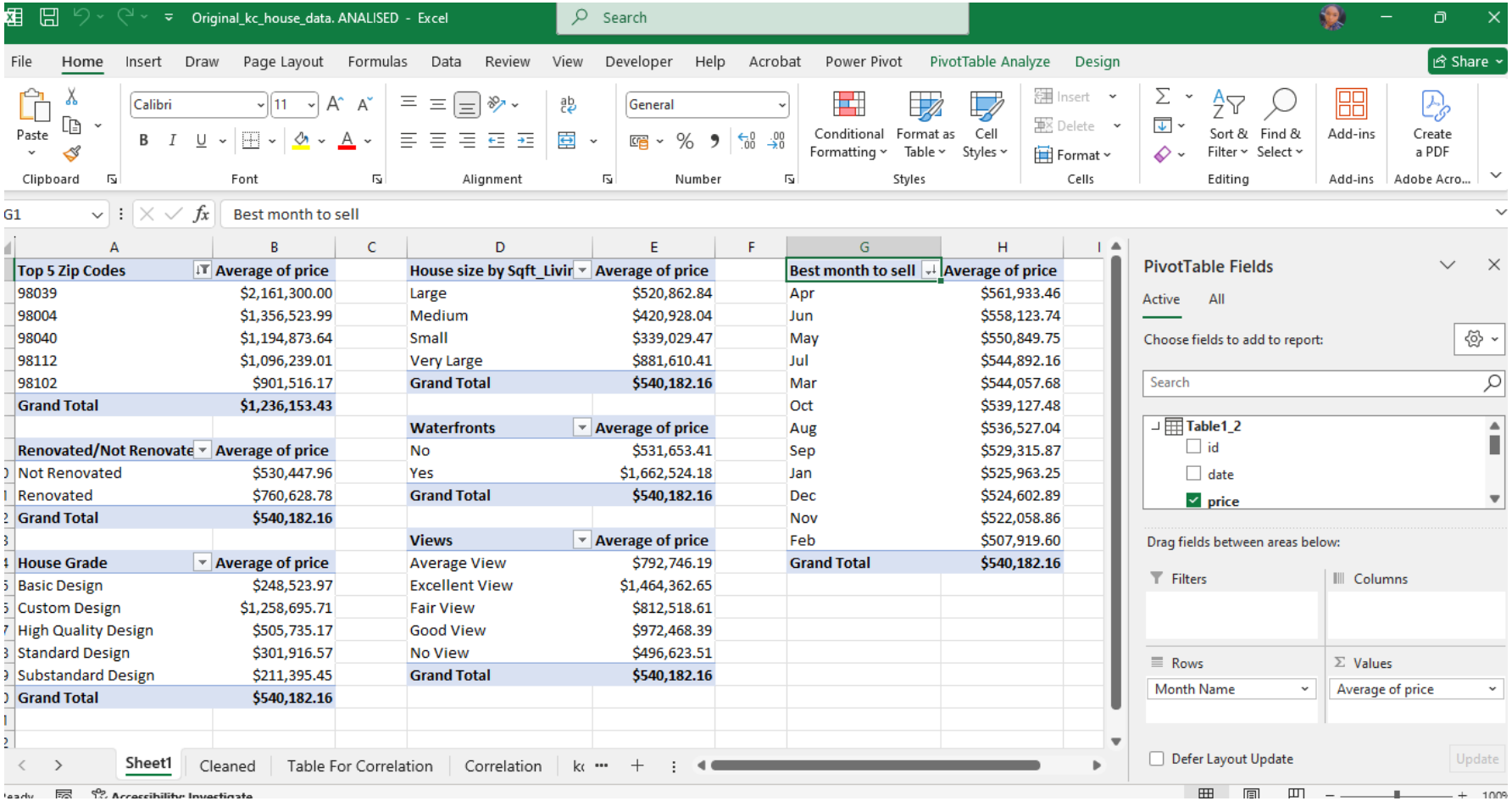


Image showing transformed data with standardization processes such as group numbered rows into fewer relatable grading systems using the IF and other excel formatting functions.

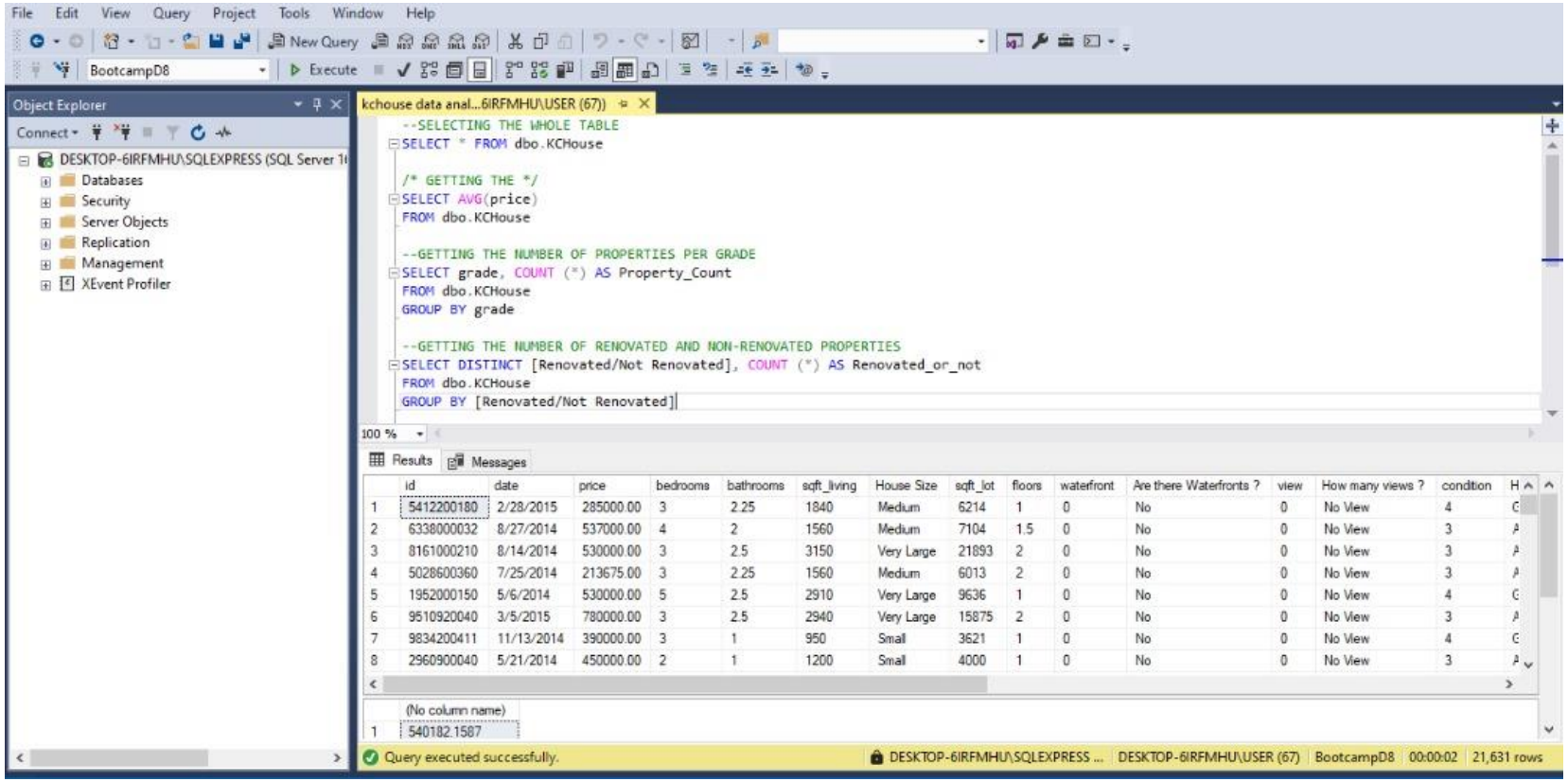
Analysis in Excel

Using Excel PivotTables, we answered key business questions about price, location, renovation, and seasonal trends.



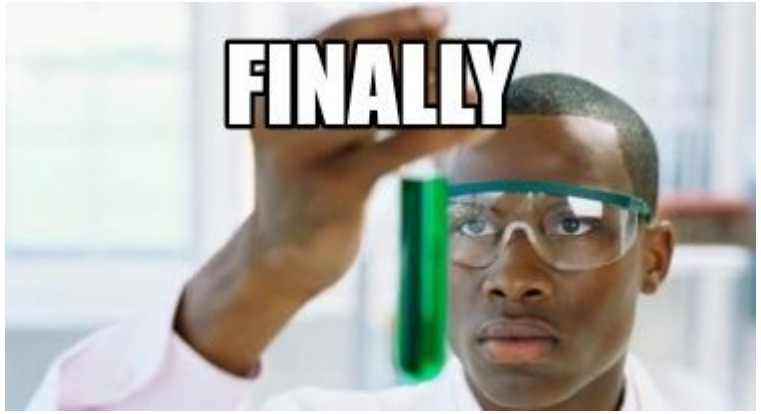
Moving to MS SQL Server

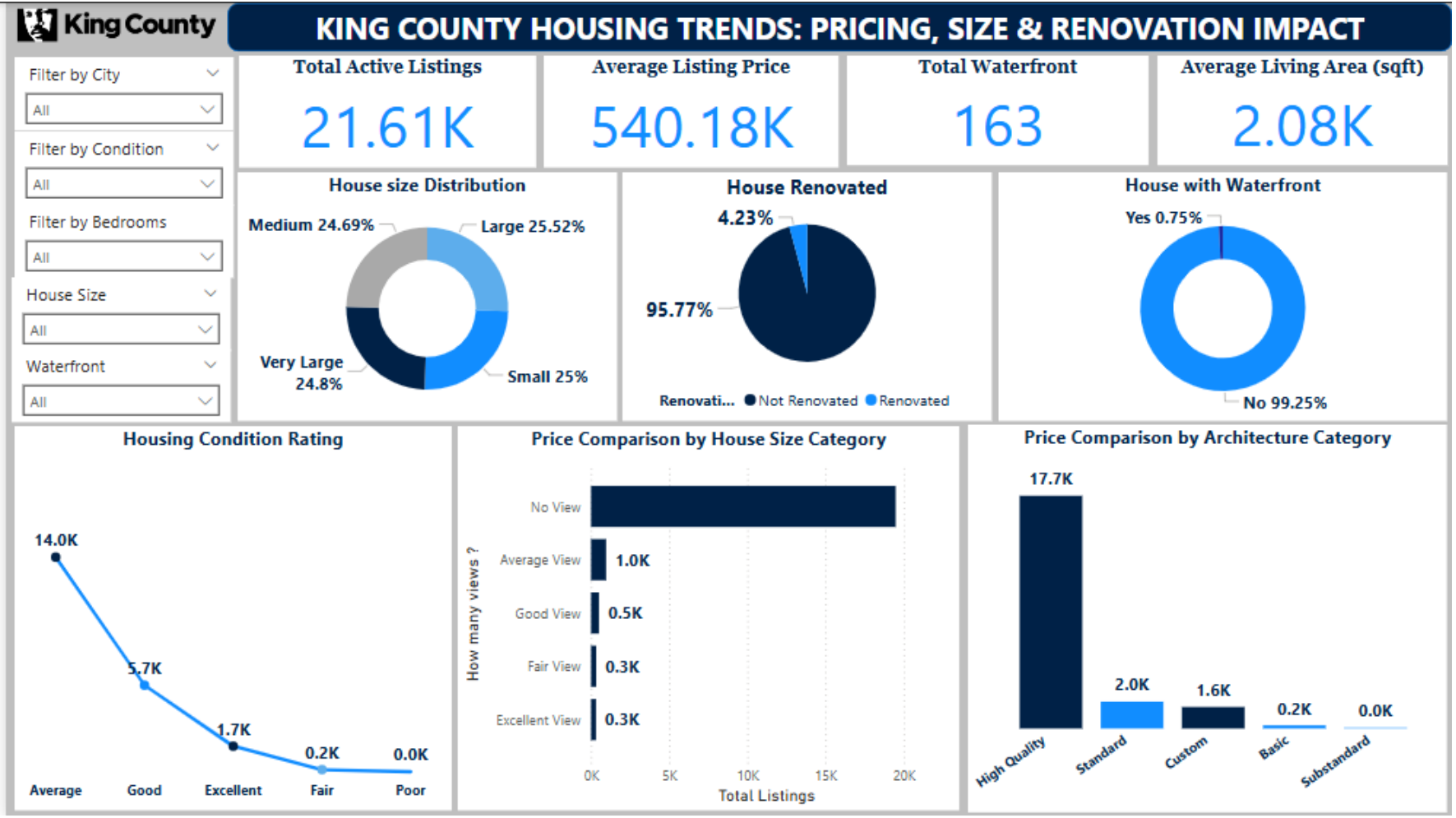
Using Excel PivotTables, we answered key business questions about price, location, renovation, and seasonal trends. We imported the cleaned data into MS SQL Server to calculate overall average and mode of house prices. This helped us understand the price distribution across the entire dataset.



Final Dashboard Showcase

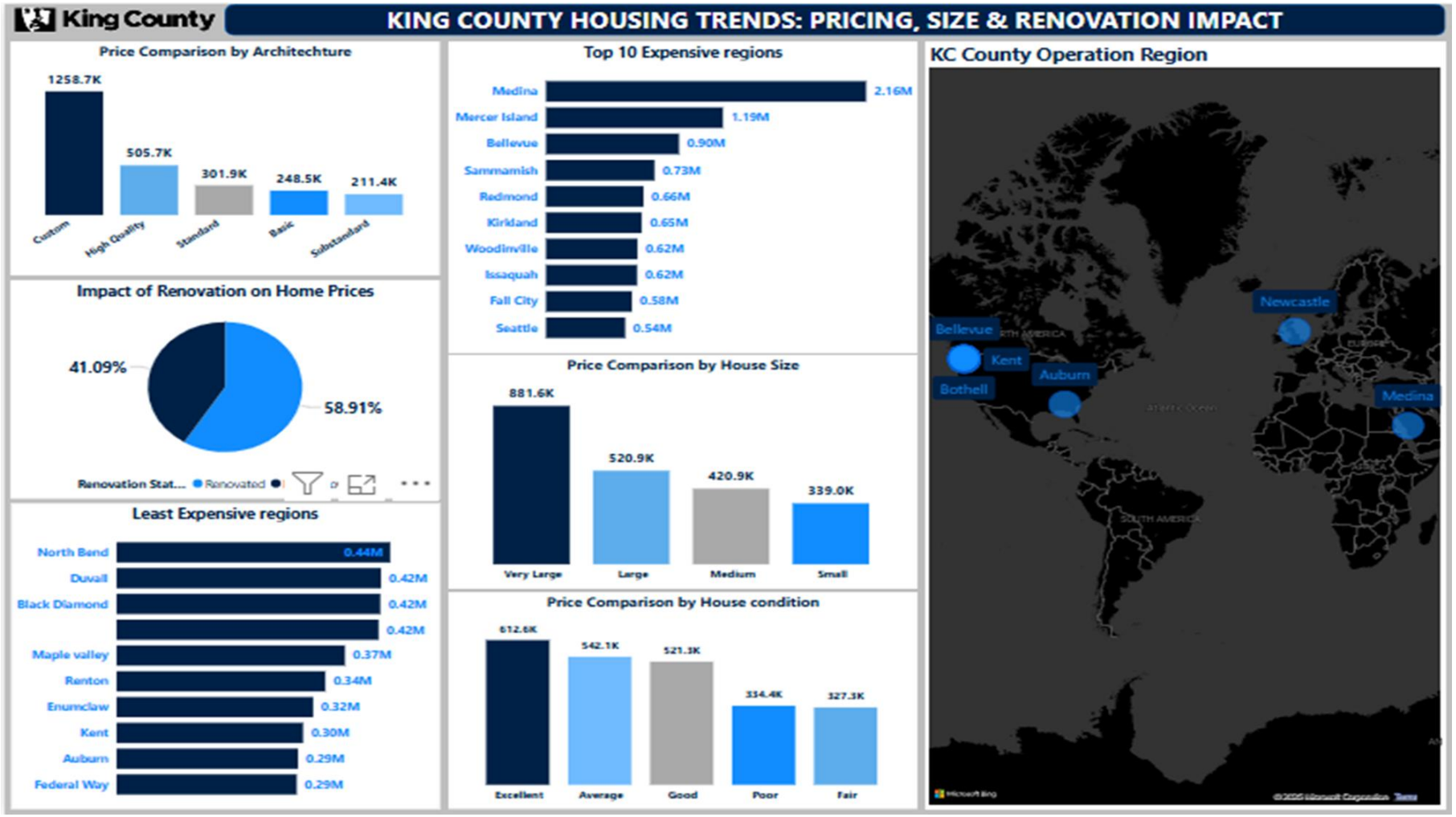
After the analysis on SQL, we imported the data to Power BI where we created interactive reports to summarizes everything—from top zip codes, average price by Architecture, to impacts of certain feature on the pricing of the house. This report can help Nigerian urban real estate stakeholders make data-informed decisions.





This dashboard reveals that pricing varies drastically by zip code, with waterfront and renovated properties showing premium pricing. Medium and large houses dominate urban centres like Seattle and Bellevue, while smaller homes cluster in more affordable zones like Kent and Auburn. This insight helps us position listings, tailor marketing, and optimize investments based on area-specific demand."

- This dashboard gives a broad overview of real estate listings in King County, with **21,613** active listings analyzed.
- Average Listing Price is approximately $**540,180**, and the average living area is around 2,080 sqft, showing that mid-sized homes dominate the market.
- The House Size Distribution shows a fairly even spread among Very Large (**24.8**%), Medium (**24.69**%), and Large (**25.52**%) homes, with Small homes taking the least portion.
- Only **4.23**% of the homes are renovated, while **95.77**% are not—this highlights a significant opportunity in property renovation, which also applies in Nigeria where older homes often have untapped potential.
- Waterfront homes are extremely rare (**0.75**%), and they command a premium. This is comparable to waterfront properties in Lagos, Lekki, or Port Harcourt, which are high-value due to location.
- Condition Rating shows that most homes are in "**Average**" or "**Good**" condition. In Nigeria, property buyers are also very sensitive to structural condition and visual appeal, which directly affects demand.
- In reference to Nigeria, especially in urban areas like Abuja, Lagos, or Ibadan, real estate stakeholders can learn from the King County Data that mid-sized and fairly priced homes have broader appeal. Also, renovation is a profitable opportunity in underperforming properties. Finally, Prime locations (like waterfronts or island properties) hold unique value.



This second dashboard shows the impact of Design, Renovation, location(regions), Size, Condition and the map shows the coverage of the location covered during data collection and also shows the pricing impact when expanded further.

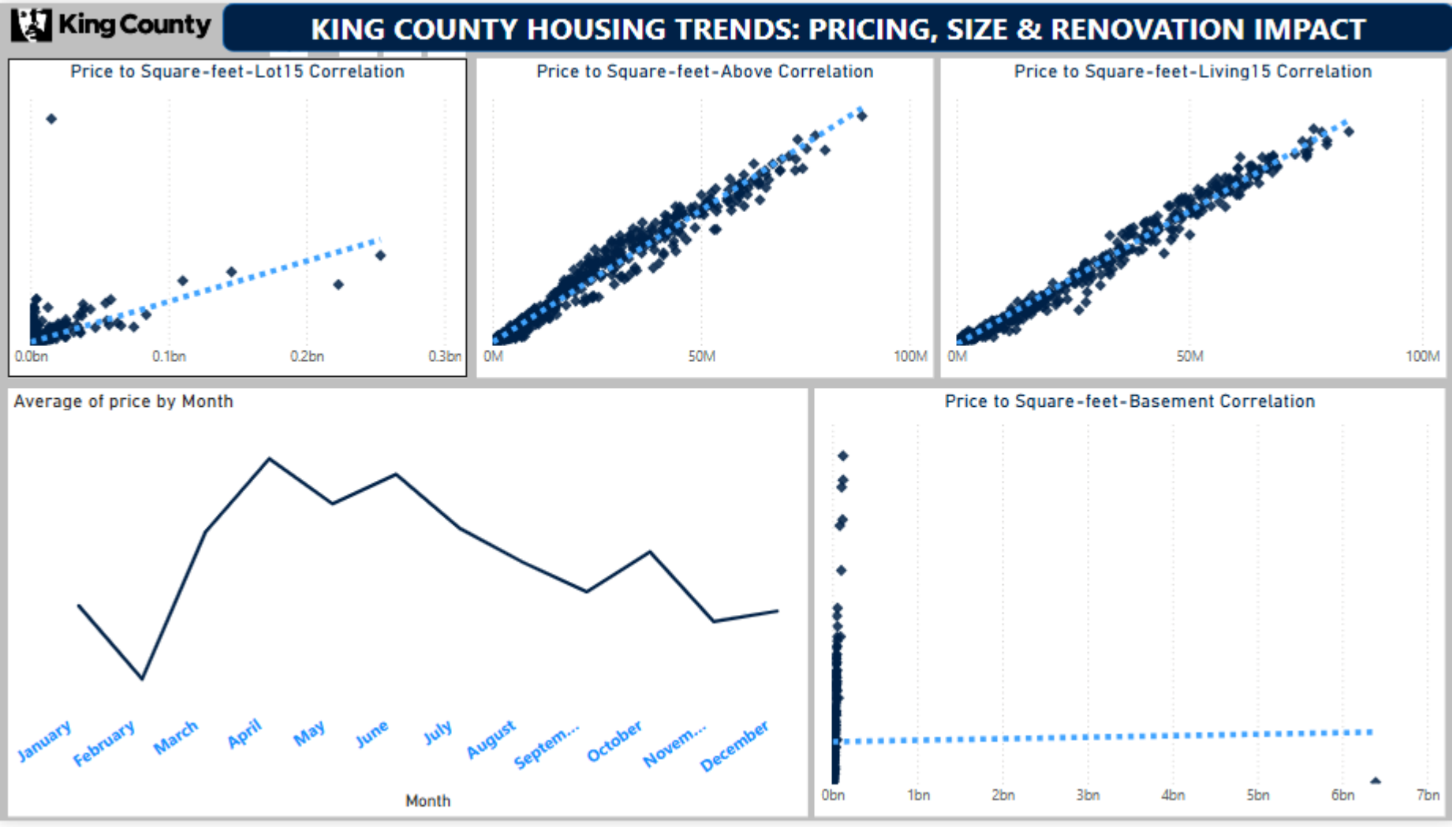
This dashboard explores what features impact house prices the most and regional price variation across King County.

Architecture Type Matters:
Custom homes lead in price, averaging $**1.25 Million**.
Standard and Substandard homes fall under $**250,000**.
This highlights the value buyers place on design and architectural uniqueness—a similar trend in Nigeria where duplexes or contemporary-designed buildings command higher prices.

Renovation Increases Value:
Homes that have been renovated are priced **58.91**% higher on average.
This supports a Nigerian trend where refurbished buildings in old areas like Surulere or Yaba gain significant resale value.

Location Makes a Difference:
Expensive areas: Medina, Mercer Island, and Bellevue have average prices well above $**1M**.
Least expensive regions include Federal Way, Kent, and Auburn.
On the map, expensive areas are clustered near bodies of water and commercial hubs, a trend relatable to Lekki, Victoria Island, and Gwarinpa in Nigeria.

Size and Condition Influence Pricing:
Very large homes fetch the highest average price.
Excellent-condition homes also command premium pricing.
This underlines the importance of space and maintenance, relevant to Nigerian buyers seeking value for long-term occupancy or rental income.



This image shows the correlation analysis between some of the house features with graphical representations of how strongly or weakly a house feature may impact house pricing.

This dashboard explores the relationship between house features and pricing using scatter plots, and also analyzes price trends over the months.

Correlation Insights:

- There’s a strong positive correlation between:

sqft_above, sqft_living15, and price – the more usable space, the higher the price.

- A weaker or unclear correlation appears between sqft_lot15 and price, meaning land size does not always guarantee a higher price.

This shows interior space and livability matter more than just the plot size, a lesson for Nigerian developers who focus on large plots without optimizing internal structure.

Monthly Price Trend:

- Prices peak around May to July, then drop from August to December.
- This implies the best time to sell is late spring to early summer, possibly due to school holidays, better weather, or relocation plans.
- For Nigeria, the equivalent period is often post-Easter to mid-year (April–June), when economic activity picks up before mid-year rains and budget delays.
- The best time to buy may be in November–December, when prices dip—possibly due to fewer transactions and end-of-year financial tightening.

Actionable Insights

1. Investors seeking premium short-term returns should focus on the top-tier cities. Those looking for longer-term appreciation or affordable housing projects may find better margins in the outer cities.
2. Waterfront listings should be positioned as luxury products. Consider highlighting these aggressively in marketing and premium renovation projects.
3. Renovations aren't just aesthetic — they are a strategic value driver. Flipping opportunities, upgrade packages, or post-purchase renovation services can be monetized.
4. Customize listings and campaigns based on the dominant house size profile in each area. Also consider modular designs or add-on services that help small homes grow with buyers.
5. This insight supports targeted investment in home staging, minor repairs, and listing enhancements especially in mid-tier markets.

“Knowledge becomes wisdom only after it has been put to good use”

- MARK TWAIN

THANK YOU