1. **Introduction**:
   * Provide a brief overview of the project, including its objectives and scope.
   * Introduce the dataset used for the analysis.

It times to move further with more analysis part to understand and predict our model to have better insight for **Real Estate Pricing** such as

Feature Engineering and Size Impact

* Discuss the process of feature engineering, including the creation of new variables or transformations of existing ones.
* Explore the impact of property size (e.g., number of bedrooms/bathrooms) on pricing.
* Use visualizations and statistical analysis to demonstrate relationships between property size and pricing.

Market Trends and Historical Pricing

* Analysis historical pricing trends over time using time series analysis techniques.
* Identify factors influencing market trends, such as economic conditions, interest rates, or housing supply and demand dynamics.
* Visualize historical pricing data to highlight trends, seasonal variations, and long-term patterns.

Customer Preferences and Amenities

* Investigate customer preferences and their influence on property pricing.
* Analysis the impact of amenities such as location, proximity to parks and transportation hubs.

1. **Exploratory Analysis Summary**:
   * Summarize the key findings, trends, and relationships identified during the exploratory analysis.
   * Highlight important insights gained from data exploration, such as distribution of variables and patterns observed in the data.

# Feature Engineering and Size Impact

# Insight from total room and total square footage with comparison with house price with the help for scatterplot

# 

# 

# Feature Engineering: Two new features have been engineered - ‘TotalRooms’ and ‘TotalSquareFootage’. ‘TotalRooms’ is the sum of total rooms above grade, full bathrooms above grade, and bedrooms above grade. ‘TotalSquareFootage’ is the sum of the total square feet of basement area, first floor square feet, and second-floor square feet.

# Relationship Exploration: The scatterplot shows the relationship between ‘TotalSquareFootage’ and ‘SalePrice’. It indicates a positive correlation, meaning as the total square footage of a house increases, its sale price tends to increase as well.

# The red line in the scatterplot represents the relationship between the total number of rooms (‘TotalRooms’) and the sale price (‘SalePrice’). However, it’s a vertical line that intersects the x-axis at approximately 2000 rooms. This suggests that there is a house with around 2000 total rooms in the dataset.

# Impact on House Prices: The title of the plot suggests that the goal is to understand the impact of features and size on house prices. From the available data, it can be inferred that the size of the house (in terms of total square footage) has a significant impact on the sale price.

# Relationship and insight from above grade living area and sale price

# 

# The scatterplot shows a positive correlation between the ‘Above Grade Living Area’ and ‘Sale Price’. This means as the living area of a house increases, its sale price also tends to increase. Most of the data points are clustered in the area where the living area is between approximately 1000 and 2000 sqft, and the sale price is between approximately $100,000 and $300,000. There are fewer homes with larger living areas and higher prices in this dataset.

# Relationship and insight from over all quality of material and finished of the house and sale price

# 

# The image is a boxplot graph titled “Overall Quality vs. Sale Price”. Here are some insights:

# Overall Quality: The x-axis represents the overall quality of items, rated from 0 to 9.

# Sale Price: The y-axis represents the sale price of items, ranging from 0 to 700,000.

# Positive Correlation: The median sale price increases as the overall quality rating increases, indicating a positive correlation between these two variables. This means items with higher quality tend to have higher sale prices.

# Outliers: There are outliers present in several quality ratings, particularly noticeable at ratings 4 through 8. These are items that have a sale price significantly different from others with the same quality rating.

# Relationship and insight from total square feet of basement area and sale price

# 

# Scatter plot graph titled “Total Basement Area vs. Sale Price”. Here are some insights:

# Total Basement Area: The x-axis represents the total basement area of properties, ranging from 0 to 6000 sqft.

# Sale Price: The y-axis represents the sale price of properties, ranging from 0 to 800,000.

# Positive Correlation: The scatter plot and the red trend line indicate a positive correlation between the total basement area and the sale price. This means properties with larger basements tend to have higher sale prices.

# Confidence Intervals: The shaded red area around the trend line likely represents confidence intervals or prediction error, providing a range of values that are likely to contain the true regression line.

**Market Trends and Historical Pricing**

**Historic sale of houses as per months with house sale prices**

# 

# The graph title is “Historical Sales Trends Over Time.”

# The y-axis represents “Sale Price” and ranges from 175,000 to 195,000.

# The x-axis represents “Month Sold (MM)” and spans from February to January of the following year.

# The yellow line depicts the “Average Sale Price (Month)”, showing fluctuations over time.

# Notable points:

# Peaks occur in December and July, indicating higher sale prices during these months.

# Troughs or low points are observed in February, May, and November.

# In summary, the graph illustrates the variations in sale prices throughout the year, with distinct patterns emerging during specific months.

**Historic sale of houses as per year with house sale prices**

# 

# Average Sale Price:

# The graph shows the average sale price per year.

# From 2005 to 2007, there was a significant drop in sale price.

# In 2008, there was a sharp increase in sale price.

# Afterward, the sale price remained steady until 2010.

# Yearly Variation:

# The fluctuations in sale price indicate market dynamics.

# The peak in 2008 might be influenced by external factors.

# Overall Trend:

# Despite short-term variations, the overall trend suggests a gradual decline in sale price over the years.

# Customer Preferences and Amenities

**Insight of size of garage in car capacity with house sale prices**

# 

# The box plot graph illustrates the relationship between house prices and the size of the garage. Here are the key takeaways:

# Garage Size and House Prices:

# As the garage size increases, there is a general trend of higher house prices.

# Houses with garage size 3 exhibit the widest range of prices, indicating greater variability.

# Median prices follow this order: Size 3 > Size 2 > Size 1 > Size 4 > Size 0.

# Interpretation:

# A larger garage is often considered a desirable feature, leading to increased property value.

# However, the wide range for size 3 suggests that other factors also influence house prices.

**Recommendations based on the analysis and insights that could inform decision-making**

**Insights from Univariate and Multivariate Analysis**:

* Univariate and multivariate analysis techniques have revealed significant correlations between various property features and sale prices.
* Stakeholders should leverage these insights to prioritize property features and amenities that drive value and appeal to potential buyers

**Market Trends and Historical Pricing**:

* + Analysis of historical pricing trends reveals seasonal variations and long-term patterns in the real estate market.
  + Stakeholders should consider the timing of property listings and sales to optimize pricing strategies and maximize returns.

**Customer Preferences and Amenities**:

* + Understanding customer preferences and the influence of amenities on property pricing is essential for attracting buyers and maximizing property value.

Development and evaluation of predictive models can provide stakeholders with valuable tools for forecasting property prices and making informed investment decisions.

The insights gained from the exploratory analysis of the real estate pricing dataset provide valuable information for stakeholders in the real estate industry. By leveraging these insights and recommendations, stakeholders can optimize pricing strategies, identify investment opportunities, and make informed decisions to maximize returns and achieve their business objectives.