

Final Project Report Template

1. Introduction

This project focuses on analyzing **housing market trends** using **Tableau** to transform raw housing data into meaningful visual insights. The main objective is to understand how factors such as **sale price, renovation status, house age, and property features** influence the housing market.

1.1. Project overviews

The **Visualization of Housing Market Trends** project aims to provide a comprehensive overview of housing data through interactive and user-friendly dashboards created using **Tableau**. The project analyzes key attributes such as **sale prices, renovation impact, house age distribution, and property features** including bedrooms, bathrooms, and floors.

The dataset is first cleaned and transformed to ensure accuracy and consistency. Then, meaningful visualizations are designed to highlight important trends and comparisons across different housing characteristics. These visual insights are finally integrated into a web-based interface to enhance accessibility and user interaction.

This project helps stakeholders quickly understand market behavior, identify pricing patterns, and make informed decisions based on data-driven insights.

1.2. Objectives

- Identify **key factors** influencing house prices.
- Analyze the **effect of renovations** on property value.
- Explore the **distribution of house sales** across different price ranges.
- Create **interactive Tableau dashboards** to present findings effectively.

2. Project Initialization and Planning Phase

2.1. Define Problem Statement

Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	A data analyst / real-estate analyst	Analyze the effect of renovation on house sale prices	It is hard to compare renovated vs non-renovated houses	The data is large and not visually structured	Slow and less confident in my analysis
PS-2	A business stakeholder	Understand housing market trends across different house age groups	The data is not grouped clearly by house age	Trends are shown in raw format without insights	Confused while identifying market patterns
PS-3	A student learning data analytics	Understand how house features (bedrooms, bathrooms, floors) affect prices	Feature comparisons are not easy to interpret	There are no clear visual comparisons	Frustrated during learning and analysis
PS-4	A decision-maker / investor	Make quick and confident property decisions	I cannot see all insights in one place	There is no single interactive dashboard	Uncertain and inefficient in decision-making

2.2. Project Proposal (Proposed Solution)

Project Overview	
Objective	The primary objective of this project is to analyze and visualize housing market trends by examining sale prices and key property features using Tableau. The project aims to provide insights into market patterns, price fluctuations, and factors influencing home values, helping stakeholders make data-driven decisions.
Scope	<ul style="list-style-type: none">Geographical Coverage: Focus on a specific city, region, or nationwide data.Data Sources: Utilize publicly available datasets (e.g., Zillow, Kaggle, government housing data).Analysis Areas: Price trends over time, location-based variations, property features affecting prices.Visualization Methods: Interactive dashboards, Story, Pie chart, Bar charts, and Histogram.
Problem Statement	
Description	The housing market is influenced by multiple factors, including location, property size, number of bedrooms, and economic trends. However, without effective visualization, identifying key trends and insights can be challenging. This project aims to bridge that gap by providing a clear, interactive representation of housing market data.
Impact	<ul style="list-style-type: none">Helps potential buyers and sellers make informed decisions.Assists real estate professionals in understanding pricing trends.Provides policymakers with insights into housing affordability and market fluctuations.
Proposed Solution	
Approach	<ul style="list-style-type: none">Collect and preprocess historical housing market data.Use Tableau to create interactive visualizations.Apply data analysis techniques to identify patterns and correlations.Develop dashboards for easy interpretation of housing trends.
Key Features	<ul style="list-style-type: none">Dynamic Visualizations: Interactive dashboards showing trends over time.Geospatial Analysis: Heat maps highlighting price variations by location.Comparative Analysis: Charts comparing features like square footage vs. price.Predictive Insights: Basic forecasting of price trends using historical data.

2.3. Initial Project Planning

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Preparation	USN-1	As a user, I want the housing dataset to be cleaned and prepared for analysis.	2	High	Jaya Surya
Sprint-1	KPI Dashboard	USN-2	As a user, I want to view KPI overview (total houses, average sale price).	1	High	Sri Hari
Sprint-2	Renovation Analysis	USN-3	As a user, I want to analyze the impact of renovation on house sale prices.	2	High	Sri Hari
Sprint-2	House Age Analysis	USN-4	As a user, I want to view house age distribution by renovation status.	2	Medium	Jaya Surya
Sprint-3	Feature Analysis	USN-5	As a user, I want to compare bedrooms, bathrooms, and floors across house age groups.	1	Medium	Sri Hari
Sprint-3	Dashboard Integration	USN-6	As a user, I want all scenarios combined into a single interactive dashboard.	2	High	Sri Hari
Sprint-4	Website Integration	USN-7	As a user, I want to view the dashboard through a web application.	2	High	Padmaleela
Sprint-4	Documentation & Testing	USN-8	As a reviewer, I want clear documentation and a tested dashboard.	1	Medium	Sowmya

3. Data Collection and Preprocessing Phase

3.1. Data Collection Plan and Raw Data Sources Identified

Section	Description
Project Overview	<p>Visualizing Housing Market Trends: -</p> <p>This project, "Visualizing Housing Market Trends: An Analysis of Sale Prices and Features using Tableau," aims to explore and analyze housing market trends using the Transformed Housing Data 2 dataset from Kaggle. The objective is to identify key factors influencing house prices, such as location, size, number of bedrooms, bathrooms, floors and basement area.</p> <p>By leveraging Tableau, the project will create interactive dashboards, story, bar chart, histogram, summary dashboard to visualize patterns, compare regional price variations, and gain insights into how different features impact house sale prices. The analysis will help in making data-driven decisions for buyers, sellers, and real estate professionals.</p> <p>Objectives:</p> <ol style="list-style-type: none">1. Understand Sale Price Trends – Identify price distribution and variations based on property features.2. Feature Analysis – Examine factors influencing house prices, such as area, bedrooms, and location.3. Market Insights – Discover trends and patterns to assist buyers, sellers, and real estate professionals.4. Data-Driven Decision Making – Use visualizations to interpret market conditions effectively.

Data Collection Plan	<p>The Data is collected from the “KAGGLE”</p> <p>Visualizing Housing Market Trends: An Analysis of Sale Prices and Features using Tableau</p> <hr/>
Raw Data Sources Identified	<p>Transformed Housing Data 2 from KAGGLE.</p> <p>This dataset contains 21,609 house sale records with various attributes such as sale price, number of bedrooms, square footage, location, and additional features. The data has been pre-processed and transformed for analytical purposes. There are a total of 31 columns, out of which Sale Price can be supposedly taken as a dependent variable. The other variables are different features, locations and date, etc. regarding the houses.</p>

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Transformed Housing Data 2	This dataset contains 21,609 house sale records with details such as price, area, number of bedrooms, and other property features.	Kaggle Dataset Link	CSV	10 MB	Public

3.2. Data Quality Report

Data Source	Data Quality Issue	Severity	Resolution Plan
Transformed Housing Data 2	<ul style="list-style-type: none"> Missing values in key attributes. Duplicate records. Outliers affecting analysis. Incorrect data types or formats. Inconsistent data entry. 	Moderate	<ul style="list-style-type: none"> Filling missing values using mean/median/mode. Removing or correcting duplicates. Normalizing or transforming data types. Handling outliers using statistical techniques.
Transformed Housing Data 2	Duplicate records found (11 instances).	Moderate	Remove duplicate records using data cleaning techniques.
Transformed Housing Data 2	'No of Times Visited' column contains only zero values, which may indicate missing or incorrect data.	High	Verify data source or remove column if it lacks meaningful information.
Transformed Housing Data 2	Zipcode representation is in multiple binary columns instead of a single categorical variable.	Moderate	Convert multiple binary columns into a single categorical 'Zipcode Group' column for better efficiency.
Transformed Housing Data 2	Potential outliers in 'Sale Price' and 'Lot Area' columns.	High	Use statistical methods to identify and handle outliers (e.g., IQR method).

3.3. Data Exploration and Preprocessing

Section	Description
Data Overview	The dataset contains Transformed housing data and 21,609 house sale records, including Property features such as Sales price, area, bedrooms, bathrooms, floors and location. There are a total of 31 columns, out of which Sale Price can be supposedly taken as a dependent variable. The other variables are different features, locations and date, etc. regarding the houses.
Data Cleaning	Missing values were handled, duplicates were removed, and data inconsistencies were corrected. Removed 11 duplicate records to ensure data integrity. Verified that no missing values exist.
Data Transformation	Applied filtering, sorting, pivoting, and created calculated fields such as Sales price (Average), price per square foot (SUM) and Age of houses (Average) and filters like Sales price (Top 25), Ages of houses (Top 10) and Measure Names. Converted multiple binary zip code group columns into a single categorical variable. Removed or adjusted potential outliers in ‘Sale Price’ and ‘Lot Area’ using statistical techniques.
Data Type Conversion	Corrected data types, ensuring numerical fields (e.g., price, area) are properly formatted. Ensured that numerical fields (e.g., Sale Price, Lot Area) are in the correct format. Converted categorical variables like Zipcode Groups into appropriate data types.
Column Splitting and Merging	Split combined address columns and merged relevant features for better analysis. Merged zip code binary columns into a single categorical field for better data structure.
Data Modeling	Established relationships between features and variables such as house size, number of floors, bathrooms, bedrooms, sale price and Age of houses (in years). Identified key relationships between features such as Sale Price, Lot Area, and House Features for further analysis in Tableau.

4. Data Visualization

4.1. Framing Business Questions

1. What is the average sale price of houses?
 2. Which basement size shows the highest sale price?
 3. Which price range has the lowest number of sales?
 4. What does the color coding represent?
 5. How many different houses age categories are shown?
 6. What age groups of houses are included in the chart?
 7. What is the count of houses in the dataset for each house age?
 8. Which price bins have the least number of houses?

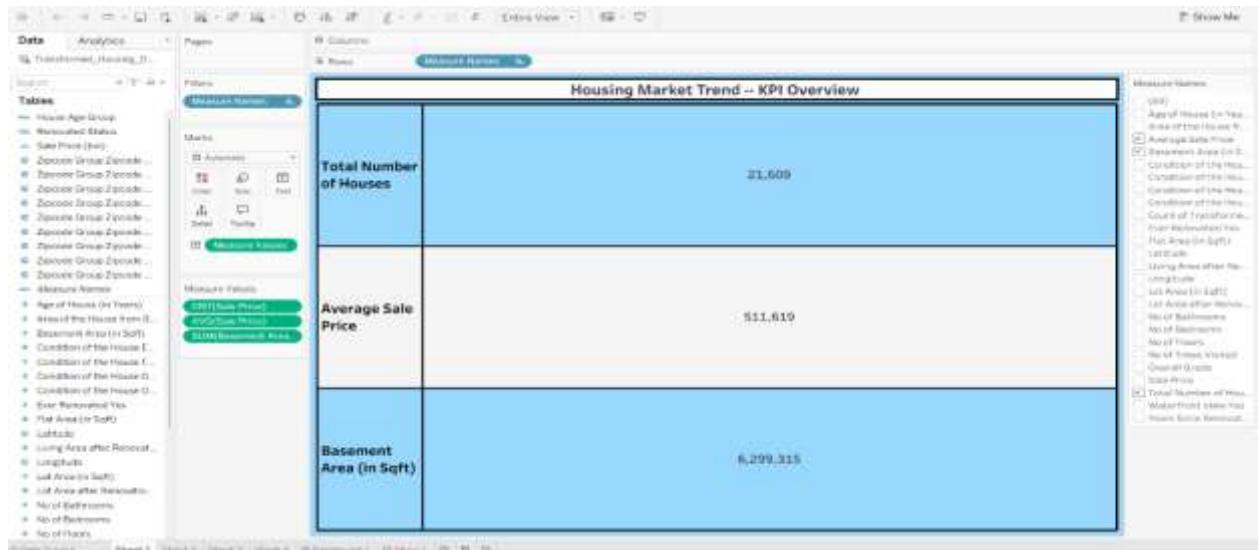
4.2. Developing Visualizations

Visualizations play a crucial role in analyzing and interpreting housing market trends. They help in uncovering patterns, relationships, and insights from the dataset. Here's how you can approach developing visualizations for your project:

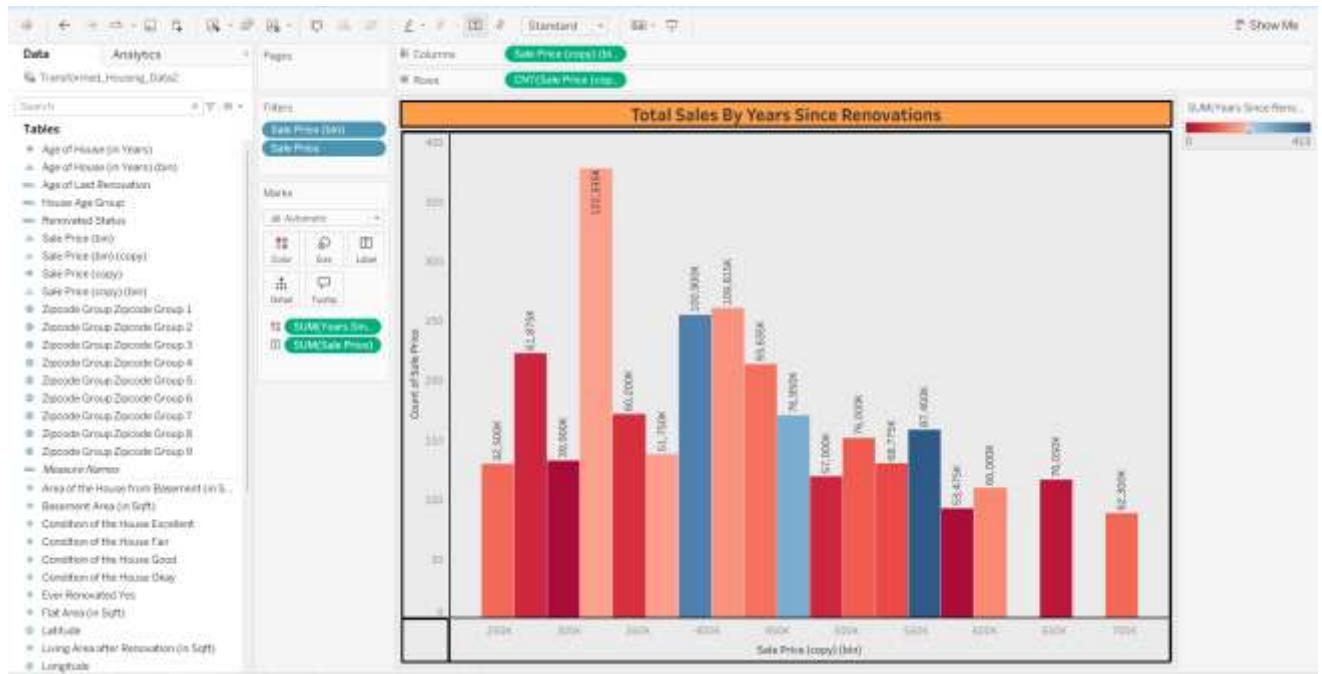
- **Bar Charts** – Compare categorical variables like the number of houses by neighbourhood.
 - **Scatter Plots** – Show relationships between sale prices and features like square footage or number of bedrooms.
 - **Histograms** – Display distribution of sale prices to understand market trends.
 - **Heatmaps** – Represent correlations between different features.
 - **Box Plots** – Identify outliers in sale prices and feature distributions.
 - **Geospatial Maps** – Show housing price trends across different locations.

Visualization:

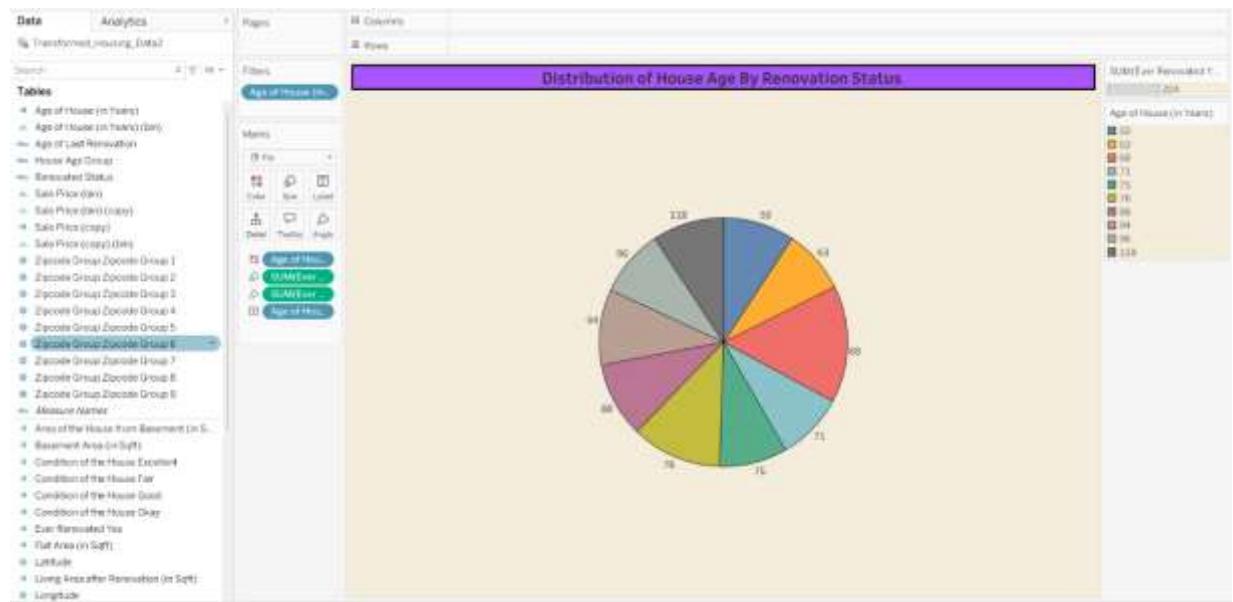
1.KPI Overview :



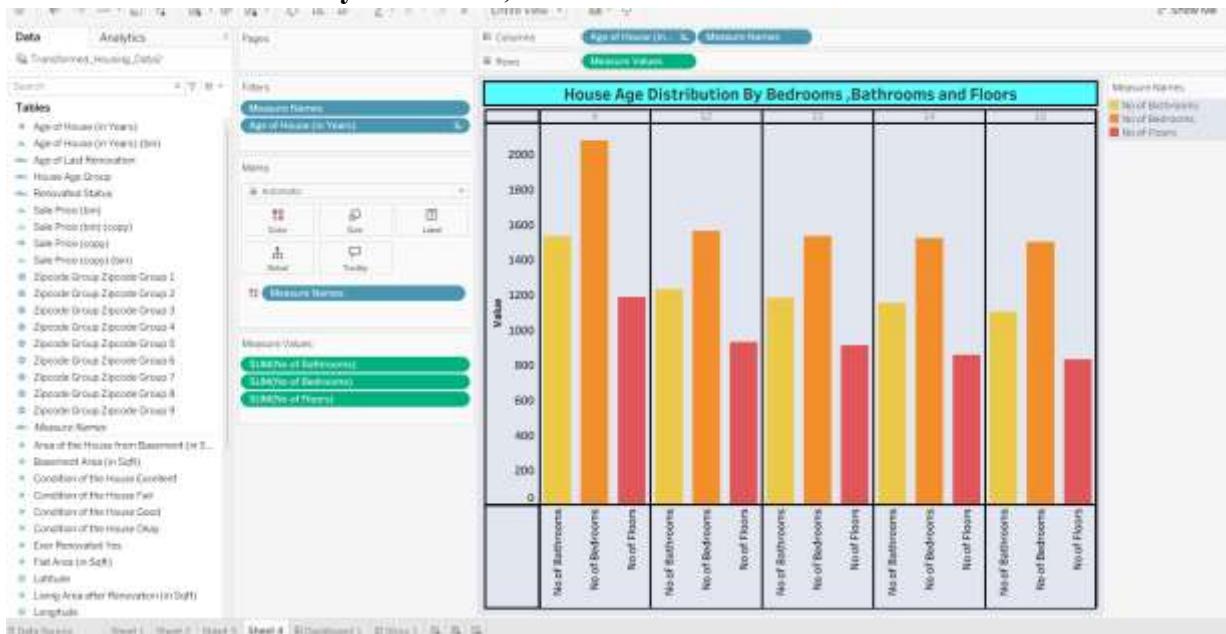
2. impact of Renovation on Sales Prices:



3.Distribution of House Age By Renovation Status:

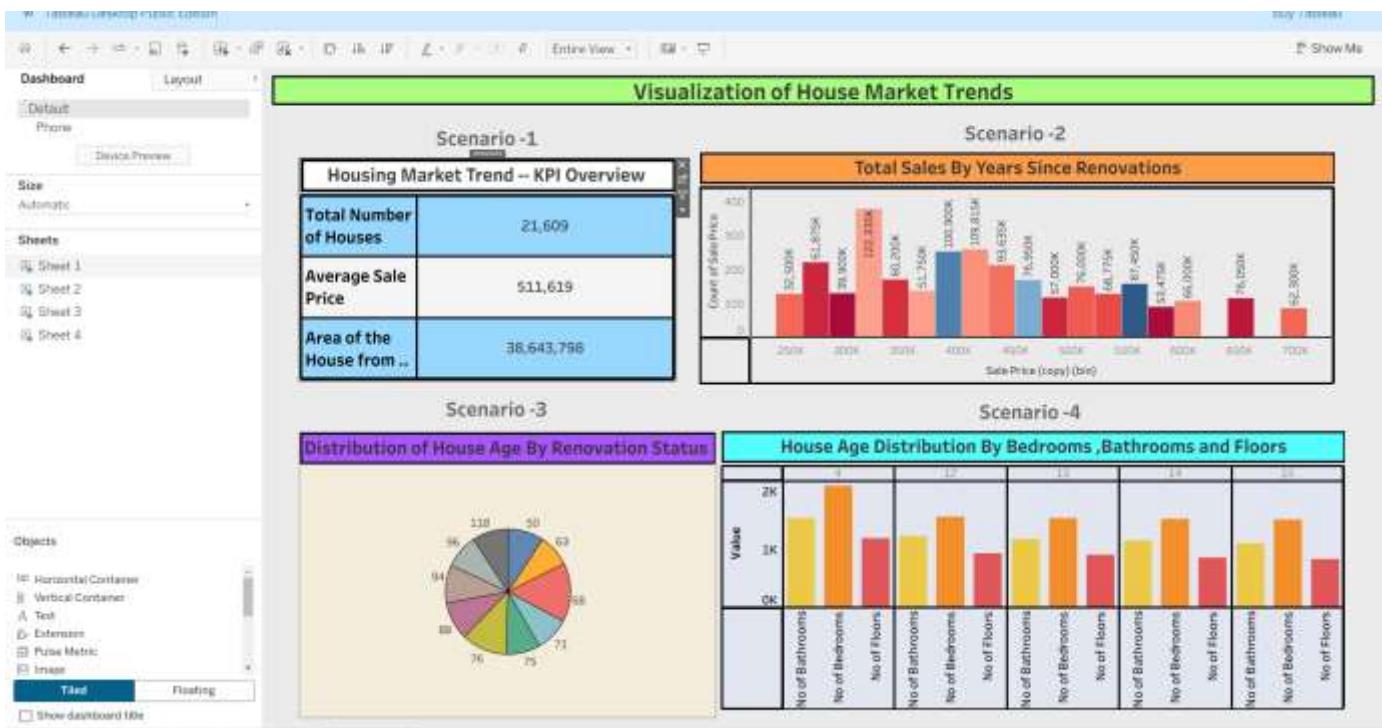


4. House Distribution By Bedrooms, Bathrooms and Floors:



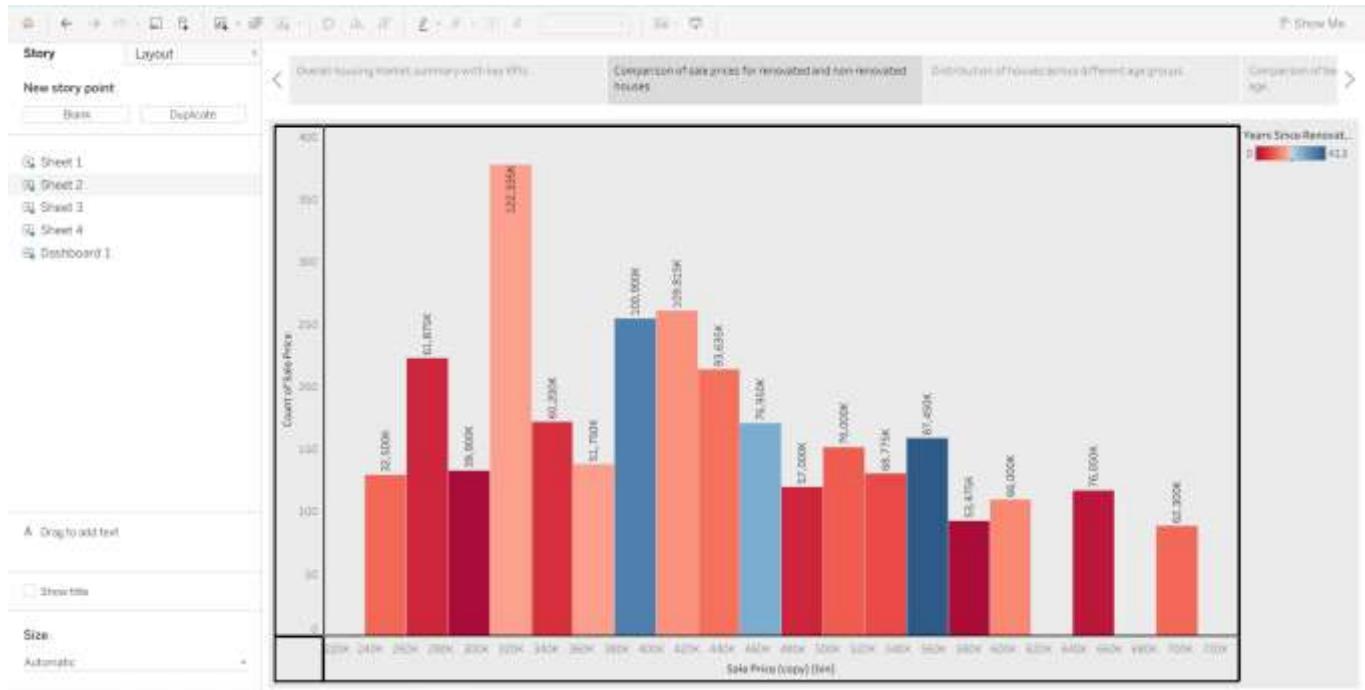
5. Dashboard

5.1. Dashboard Design File



6. Report

6.1. Story Design File



7. Performance Testing

7.1 Utilization of Data filters

Selected Top (10) Highest Sales Price

Selected Top (10) Age of House (in years)

Selected Top (10) No of Floors, Bathrooms, Bedrooms

7.2 No of Calculation Field

Total no of calculation fields are :3

7.2.1 Renovation Status

- Logical calculation:
 - Renovated → Renovation Year > 0
 - Not Renovated → Renovation Year = 0
- Used to compare renovated vs non-renovated houses.

7.2.2 House Age

- **Calculation:**
House Age = Current Year – Year Built
- **Used for user:**
Helps the user understand how old the properties are and compare prices based on age.

7.2.3 Years Since Renovation

- **Calculation:**
Years Since Renovation = Current Year – Renovation Year
- **Used for user:**
Helps the user analyze how recent renovations affect house sale prices.

7.3 No of Visualization

- Bar Chart
- Pie Chart
- Bubble Chart
- Donut Chart
- Text Table
- Word Cloud
- Funnel Chart
- Horizontal Chart

8. Conclusion/Observation

Conclusion:

Developing visualizations for housing market trends provides deep insights into the relationships between different property features and sale prices. By leveraging Tableau, interactive and visually appealing dashboards can be created, making it easier to analyze market dynamics. The integration of these visualizations into a Flask-based web application allows users to explore trends effectively. Properly selected charts, such as scatter plots, heatmaps, and geospatial maps, help in identifying patterns, correlations, and anomalies within the dataset.

Observations:

1. **Price Distribution** – Sale prices are not evenly distributed; they tend to cluster around certain price ranges, indicating popular price segments.
2. **Impact of Features** – Factors like square footage, lot size, and the number of bedrooms significantly influence sale prices, while some features have minimal impact.
3. **Neighbourhood Trends** – Certain neighbourhoods consistently show higher property values, possibly due to better infrastructure, schools, or amenities.
4. **Seasonal Variation** – Housing prices fluctuate across different months or years, indicating seasonal demand shifts.
5. **Outliers in Data** – Some properties are priced significantly higher or lower than average, which could be due to unique characteristics or anomalies in data collection.

9. Future Scope

The project on "**Visualizing Housing Market Trends: An Analysis of Sale Prices and Features using Tableau**" can be expanded in various ways to enhance its impact and usability. Some potential future directions include:

1. Advanced Predictive Analytics

- Implement **Machine Learning models** (e.g., regression, decision trees, neural networks) to predict future house prices based on historical trends.
- Use **Time Series Forecasting** (e.g., ARIMA, LSTM) to anticipate market fluctuations.

2. Enhanced Interactive Visualizations

- Develop **real-time dashboards** in Tableau by integrating live data sources such as real estate APIs.
- Introduce **user-driven filtering** to allow users to explore data based on their preferences (e.g., price range, neighborhood).

3. Integration with External Data Sources

- Combine **economic indicators** (e.g., interest rates, inflation, job market data) to analyze their impact on housing prices.
- Use **Geospatial Data & GIS Mapping** to show property trends based on location, infrastructure, and amenities.

4. Web Application Enhancements

- Deploy the **Flask-based website** on a cloud platform (e.g., AWS, Heroku) for wider accessibility.
- Add **User Authentication & Personalization**, allowing users to save preferences and compare properties.

10. Appendix

10.1. Source Code :

1. Tableau Workbook

- **File type:** .twb or .twbx
- **Purpose:**
 - Contains all dashboards, stories, calculated fields, filters, and visualizations
- **Example:**

https://public.tableau.com/views/SmartBridgeProjectdashboard/Sheet4?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link.

2 Dataset File

- **File type:** .csv
- **Purpose:**
Transformed_Housing_Data.csv

3.Calculated Fields (Inside Tableau)

- Created using Tableau's calculation editor
- Examples:
 - House Age
 - Years Since Renovation
 - Average Sale Price

4. Web Integration Code (Optional / If Used)

If you integrated Tableau into a website:

a) Backend

- **Language:** Python
- **Framework:** Flask
- **File:**
app.py

b) Frontend

- **Files:**
 - index.html

5. Embedding Code

- Tableau embed iframe code
- Used to display dashboards on the website

10.2. GitHub & Project Demo Link

GitHub Link :

<https://github.com/Jayasurya-2324/Visualizing-Housing-Market-Trends-An-Analysis-of-Sale-Prices-and-Features-using-Tableau>

Project Demo Link :

<https://drive.google.com/file/d/1trymW6ADKyZkOzGMWJIFO8OE4XBzb0dkG/view?usp=sharing>