

An internship in
Data Analytics with Tableau
by
SmartInternz

Project Name: Visualizing housing market trends: an analysis of sale prices and features

Project Id: LTVIP2025TMID47597

Project Mentor: Anji Babu

Team Members:

- | | |
|---------------------------|---------------------|
| 1.Sudeshna Pendyala | (Reg.No.22P31A4250) |
| 2.Venkateswara Rao Allaka | (Reg.No.22P31A4266) |
| 3. Prakash Kodi | (Reg.No.23P31A0427) |

DEPARTMENT OF CSE-Artificial Intelligence & Machine Learning,
Aditya College Of Engineering and Technology, Surampalem, Andhra Pradesh.

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Institute of Aditya College of Engineering and Technology** for their kind encouragement in doing this project and for completing it successfully.

I convey my thanks to **Dr. A. Ramesh MTech., Ph.D.**, principal of Aditya College of Engineering And technology and **Dr. K. Naga Bhargavi, MTech., Ph.D.**, Head of the Department of CSE (Artificial Intelligence and Machine Learning) for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project mentor **Mr. Anji Babu**, for his valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **CSE(AI&ML)** who were helpful in many ways for the completion of the project.

ABSTRACT

This project, titled “Visualizing Housing Market Trends: An Analysis of Sale Prices and Features using Tableau,” focuses on transforming complex real estate data into clear, actionable insights through interactive data visualization. By cleaning and preparing a dataset containing various housing attributes—such as sale price, area, number of bedrooms, renovation status, and location—key trends were uncovered using Tableau’s powerful visual analytics. The project involved the creation of calculated fields (e.g., TotalAreaSqft, SalePriceBin), the use of filters (e.g., condition, renovation status, zipcode group), and the development of dashboards and stories that narrate insights across multiple dimensions. The resulting solution empowers users—including buyers, real estate agents, and policy makers—to make data-driven decisions. With its scalability and modular structure, the project lays a foundation for further enhancements like live data integration, predictive analytics, and expanded geographic coverage.

Key Words:

- Tableau Dashboard
- Housing Market Analysis
- Data Visualization
- Sale Price Prediction
- Property Features
- Renovation Insights

Project Report Format

1. INTRODUCTION

- 1.1 Project Overview
- 1.2 Purpose

2. IDEATION PHASE

- 2.1 Problem Statement
- 2.2 Empathy MapCanvas
- 2.3 Brainstorming

3. REQUIREMENT ANALYSIS

- 3.1 Customer Journey map
- 3.2 Solution Requirement
- 3.3 Dataflow Diagram
- 3.4 Technology Stack

4. PROJECT DESIGN

- 4.1 Problem Solution Fit
- 4.2 Proposed Solution
- 4.3 Solution Architecture

5. PROJECT PLANNING & SCHEDULING

- 5.1 Project Planning

6. FUNCTIONAL AND PERFORMANCE TESTING

- 6.1 Performance Testing

7. RESULTS

- 7.1 Output Screenshots

8. ADVANTAGES & DISADVANTAGES

9. CONCLUSION

10. FUTURE SCOPE

11. APPENDIX

Source Code(if any)

Dataset Link

GitHub & Project Demo Link

1. Introduction

The real estate market is influenced by various factors such as house age, renovation status, number of bedrooms and bathrooms, and overall size. This project aims to analyze housing market trends and visualize key insights using Tableau to better understand how different features impact sale prices.

1.1. Project overviews

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. 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.

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 datadriven decisions for buyers, sellers, and real estate professionals.

1.2. Objectives

- Identify key factors influencing house prices.
- Analyse 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 first-time homebuyer who wants to make an informed decision	Find a home within my budget that meets my needs	The available market data is difficult to interpret and scattered across multiple sources	There is no centralized, easy-to-use tool that visualizes housing trends based on historical sales data	Confused and overwhelmed, making me hesitant to proceed
PS-2	A real estate investor looking for high-return properties	Identify profitable properties based on price trends and key influencing factors	Existing datasets require extensive manual analysis and lack clear insights	No interactive visualization tool allows me to compare property appreciation trends effectively	Frustrated and uncertain about making investment decisions
PS-3	A real estate agent aiming to assist clients efficiently	Provide accurate and insightful recommendations based on market data	The data is time-consuming to analyse and spread across various reports	There is no comprehensive tool to aggregate and visualize pricing trends for quick insights	Less efficient, unable to provide quick, data-backed advice to clients

2.2 Empathy Map Canvas



2.3 Brain Storming

Step 1: Team Gathering, Collaboration and Problem Statement

Our team collaborated to identify pressing challenges in the real estate market, particularly in understanding how various property features influence housing sale prices. After exploring themes like housing affordability, real estate investment planning, urban development, and smart property insights, we narrowed down our focus to uncover actionable insights hidden in housing data. The objective was to visually explore trends using Tableau that would help buyers, sellers, investors, and policy makers understand patterns of sale prices based on features like area, bedrooms, renovation status, condition, location (zipcode groups), and more.

Problem Statement:

How can housing sale price trends and property characteristics be visualized and analyzed using Tableau to identify patterns, improve buyer/seller decision-making, and uncover insights that support strategic real estate planning?

Team Members:

- Team Member: Sudeshna Pendyala
- Team Member: Venkateswara Rao Allaka
- Team Leader: Prakash Kodi

Step 2: Brainstorming, Idea Listing and Grouping

S.No	Idea Description	Category
1	Visualize average sale price by SalePriceBin	Pricing Insights
2	Analyse impact of number of bedrooms on sale price	Property Features
3	Explore relationship between Total Area and Price (scatter plot)	Size-Based Pricing
4	Compare prices for renovated vs. non-renovated homes	Renovation Analysis

S.No	Idea Description	Category
5	Group insights by Zip code Clusters	Geographical Comparison
6	Analyse house condition vs. price using dummy variables	Quality-Based Pricing
7	Add calculated field: TotalAreaSqft	Data Preparation
8	Create SalePriceBin with 100k intervals	Binning / Categorization
9	Use Tableau dashboard to combine insights	Dashboard Design
10	Build a Story in Tableau for narrative	Storytelling & Reporting

Step 3: Idea Prioritization Table

S.No	Idea Description	Impact	Feasibility	Priority
1	Visualize average sale price by SalePriceBin	High	Easy	High
2	Analyze impact of number of bedrooms on sale price	High	Easy	High
3	Explore Total Area vs Price (scatter plot)	High	Easy	High
4	Compare renovated vs. non-renovated home prices	High	Medium	High
5	Group insights by Zipcode Clusters	Medium	Medium	Medium
6	Analyze house condition vs. price	High	Medium	High
7	Add calculated field: TotalAreaSqft	Medium	Easy	High
8	Create SalePriceBin with ₹100k intervals	Medium	Easy	High
9	Combine insights using Tableau dashboard	High	Easy	High
10	Build a Tableau Story for business narrative	High	Medium	High

3. Requirement analysis

3.1 Customer Journey map

Stage	Actions & Touchpoints	Experience & Emotions	Pain Points	Opportunities	User Goals
Awareness	- Sees dashboard via social media, newsletter, Tableau Public- Reads title/summary	Curious, Interested	Unclear if dashboard is relevant	Use benefit-driven titles, visual thumbnails	Attract interest and clarify purpose
Consideration	- Clicks dashboard link- Reads introduction, explores layout	Engaged, Cautious	Overwhelmed by layout, unsure where to start	Add guided walkthrough, simplify navigation	Understand the dashboard and its features
Exploration	- Uses filters for location, price, features- Views charts (bar, scatter, pie, etc.)	Excited, Inquisitive	Filters not intuitive, charts slow to load	Add example queries, improve performance/speed	Discover valuable insights
Decision	- Exports visuals- Shares dashboard- Bookmarks or downloads insights	Satisfied, Confident	Limited export options or unclear formats	Enable easy download/share, offer export guides	Preserve and share findings
Retention	- Subscribes for updates- Revisits for new data-	Loyal, Empowered	No update notifications, feedback unacknowledged	Enable email updates, actively respond to feedback	Stay informed and engaged

Stage	Actions & Touchpoints	Experience & Emotions	Pain Points	Opportunities	User Goals
	Leaves feedback				

3.2 Solution Requirement

Functional Requirements (FRs) FR

No. Functional Requirement (Epic) Sub Requirement (Story / Sub-Task)

FR-1	Data Import	- Import data from CSV - Enable live database integration (MySQL)
FR-2	Data Cleaning & Transformation	- Handle missing values - Add calculated fields like Year, Lockdown
FR-3	Data Visualization	- Create Tableau worksheets - Build multiple dashboards
FR-4	User Interaction	- Enable filtering by region, year - View comparative bar charts - Analyze pre/post-lockdown trends
FR-5	User Access	- Role-based views for Analyst, Policy Maker, Developer - Download/export options
FR-6	Feedback Loop	- Allow stakeholder feedback and change requests - Implement revision cycles

Non-Functional Requirements (NFRs)

NFR No.	Non-Functional Requirement	Description
NFR-1	Usability	Dashboard must be intuitive with clear filters, legends, and guided walkthroughs
NFR-2	Security	Implement role-based access and secure backend/database connectivity
NFR-3	Reliability	System must handle unexpected data formats and maintain high accuracy

NFR-4 Performance

Ensure fast loading and responsive interaction across all dashboard elements

NFR-5 Availability

Dashboard should be accessible across browsers/devices with minimal downtime

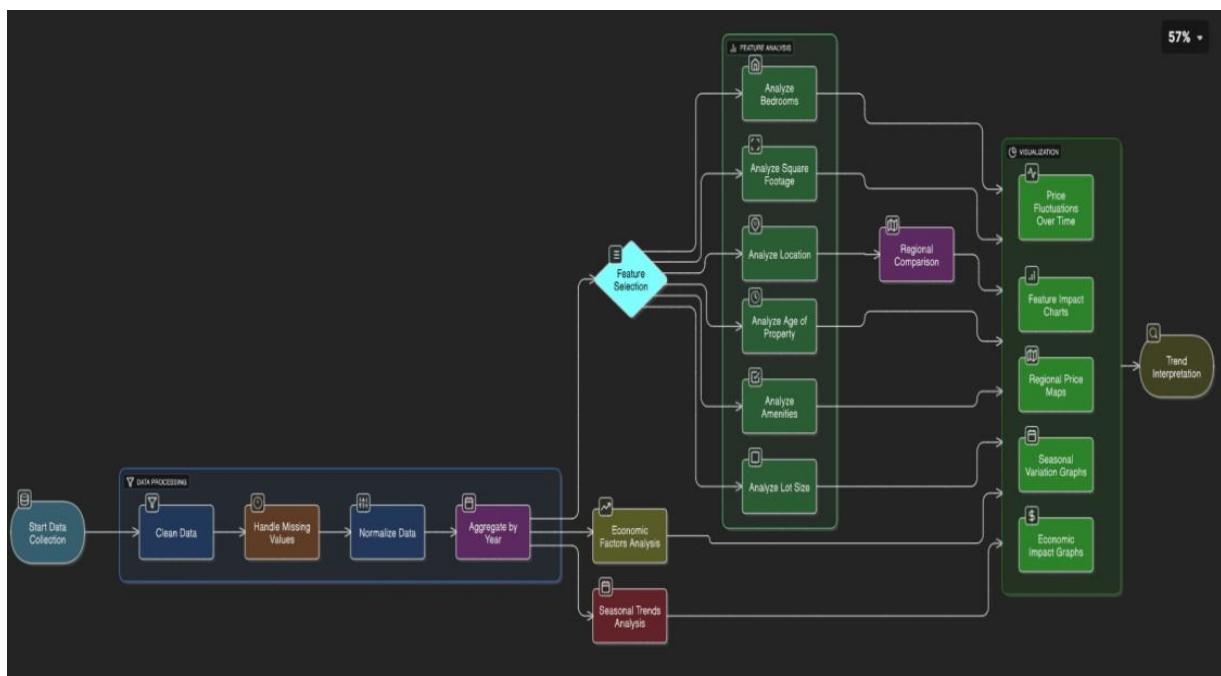
NFR-6 Scalability

Should scale for large datasets and support additional features/modules

3.3 Data Flow Diagram

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

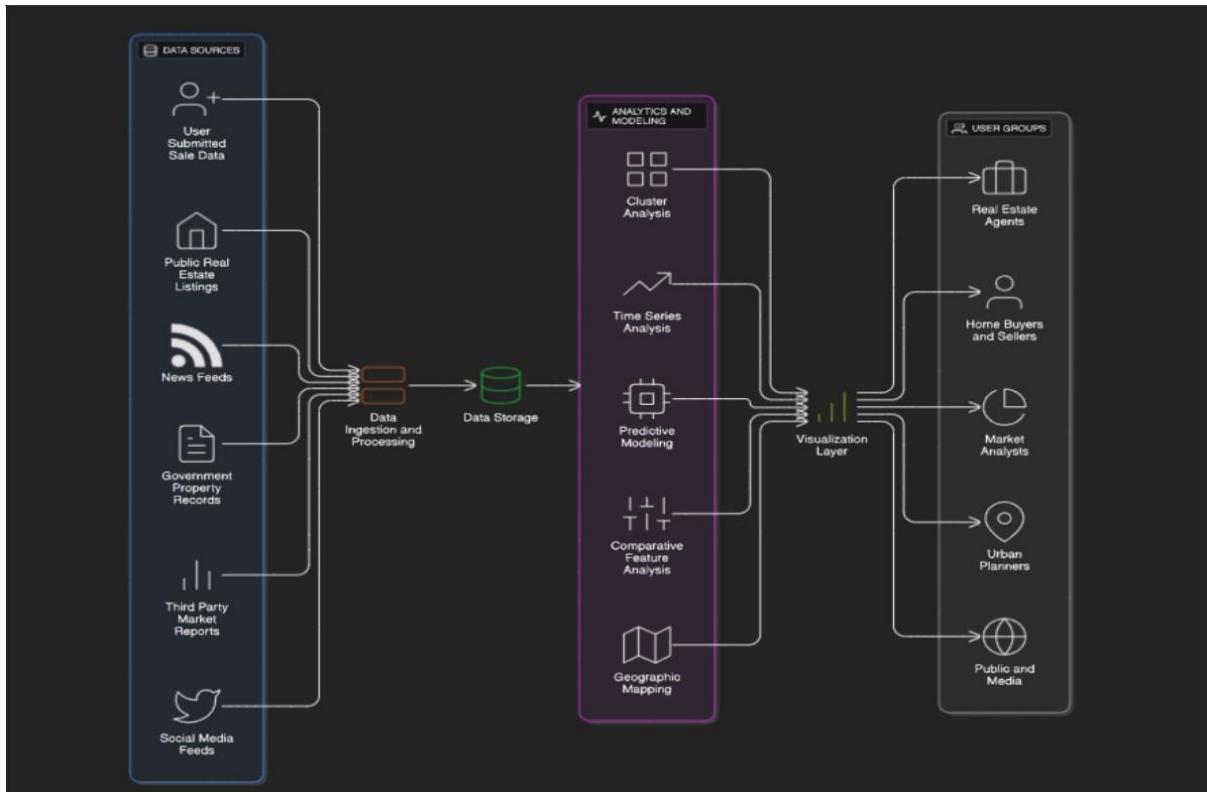
1. Data collected from POSOCO in CSV format.
2. Cleaned and transformed, with calculated fields like Year and Lockdown.
3. Visualizations built in Tableau using multiple worksheets.
4. Users review the dashboard and may request changes.
5. Final version archived after approval.



User Stories Table:

User Type	Requirement	Story Number	User Story / Task	Priority	Release (Epic)	Acceptance Criteria
Analyst / Buyer	View Housing Trends	USN-1	As a user, I want to filter and view average sale prices by region and year.	High	Sprint-1	I can use filters to view charts for specific locations and timeframes.
Analyst	Compare Locations	USN-2	As a user, I want to compare top and bottom performing zip codes based on price.	Medium	Sprint-1	I can view bar charts with top N and bottom N zip codes by sale price.
Policy Maker	Analyze Seasonal Trends	USN-3	As a user, I want to identify property price trends over time (monthly/quarterly).	High	Sprint-2	I can view time-series charts to analyze seasonal pricing trends.
Real Estate Agent	View Renovation Impact	USN-4	As a user, I want to compare prices of renovated vs. non-renovated properties.	High	Sprint-2	A before-and-after renovation price chart is available for comparison.
Developer	Connect Housing Data	USN-5	As a user, I want the dashboard to be connected to a live housing database (MySQL).	Medium	Sprint-2	Data refreshes automatically from MySQL to Tableau.
Developer	Export Dashboard Insights	USN-6	As a user, I want to export dashboard views for presentations.	Low	Sprint-3	I can download visualizations as images or PDFs.

3.4 Technology Stack



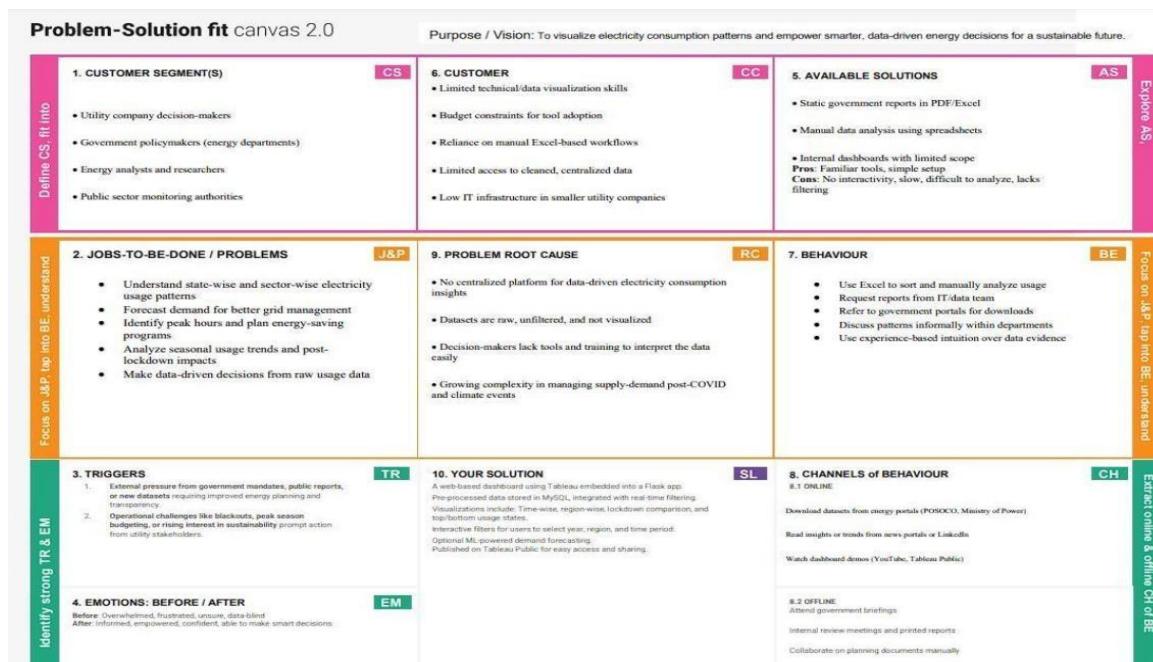
4. Project design

4.1 Problem Solution Fit

The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem. It helps entrepreneurs, marketers and corporate innovators identify behavioral patterns and recognize what would work and why.

Purpose:

- Solve complex problems in a way that fits the state of your customers.
- Succeed faster and increase your solution adoption by tapping into existing mediums and channels of behaviour.
- Sharpen your communication and marketing strategy with the right triggers and messaging.
- Increase touch-points with your company by finding the right problem-behavior fit and building trust by solving frequent annoyances, or urgent or costly problems.
- Understand the existing situation in order to improve it for your target group.



4.2 Proposed Solution

Proposed Solution Template

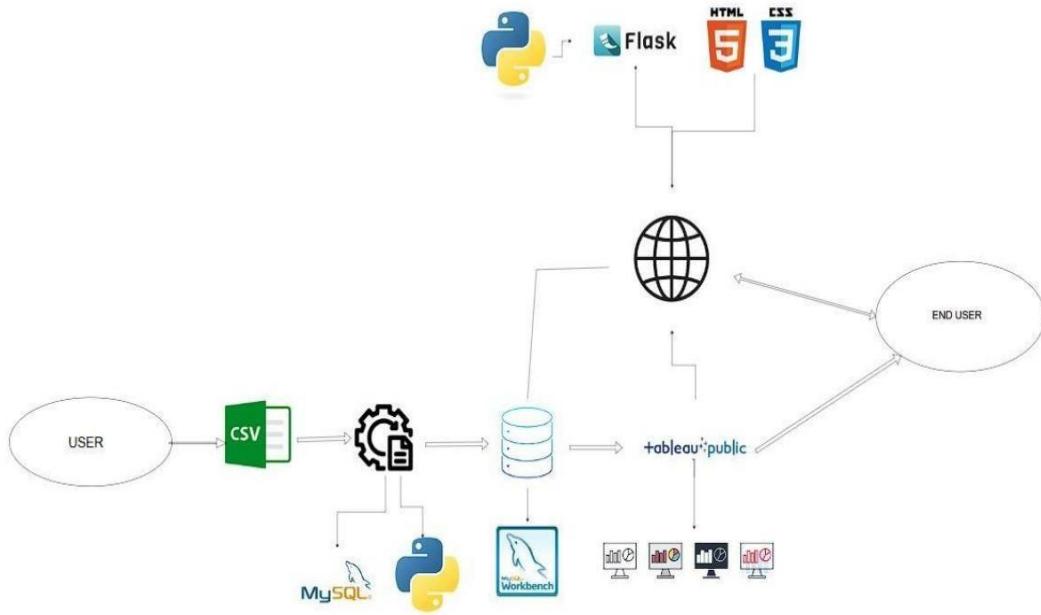
S.No.	Parameter	Description
1	Problem Statement	The real estate market involves vast and complex datasets on housing features and sale prices. These datasets are often underutilized due to lack of effective visualization, making it difficult for buyers, sellers, and analysts to draw insights or forecast trends.
2	Idea / Solution Description	Our solution transforms static housing datasets into interactive, insightful visualizations using Tableau. The project involves cleaning and transforming the data, creating calculated fields and KPIs, and developing a dashboard that highlights key trends, comparisons, and location-based analyses. The solution is deployed via a Flask web app.
3	Novelty / Uniqueness	This project leverages Tableau's powerful visual capabilities to go beyond basic data analytics. By combining calculated fields, condition segmentation, and geographic mapping, the dashboard offers a dynamic exploration of how features like bedrooms, area, renovation, and location influence housing prices.
4	Social Impact / Customer Satisfaction	This solution enables real estate buyers, sellers, agents, and market researchers to make informed decisions. It improves housing transparency, supports better urban planning, and enhances user engagement with clear visuals and actionable insights.
5	Business Model (Revenue Model)	This dashboard can be scaled and offered as a subscription-based SaaS tool to real estate companies, market research firms, or housing consultancies. Advanced forecasting modules, API integrations, and custom dashboards can be monetized as premium features.
6	Scalability of the Solution	The system is designed to be scalable and adaptable. It can incorporate new datasets (like rental trends or economic indicators),

S.No.	Parameter	Description
1	Market Expansion	extend to new regions or cities, and integrate with ML models for price predictions, thereby offering long-term growth potential.

4.3 Solution Architecture

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- The architecture separates data preprocessing, storage, visualization, and UI layers—making it easy to maintain, scale, and enhance.
- Cleaned data from MySQL is visualized using Tableau dashboards, offering region-wise, year-wise, and seasonal insights with filtering capabilities.
- Dashboards are embedded into a Flask-based web interface, allowing end users to interact with visual data through a user-friendly portal.
- The solution supports future extensions like forecasting models and can be deployed locally or on cloud platforms like Heroku or AWS.



5. Project planning & scheduling

5.1 Project Planning

Sprint	Epic	User Story No.	User Story / Task	Points	Priority	Assigned To
Sprint-1	Data Setup	USN-1	As a user, I can upload housing data in CSV format	3	High	Sudeshna
Sprint-1	Data Cleaning	USN-2	As a developer, I can clean and preprocess housing data in Tableau	4	High	Sudeshna
Sprint-1	Field Creation	USN-3	As a user, I can create calculated fields like TotalAreaSqft	2	Medium	Sudeshna
Sprint-1	Price Binning	USN-4	As a user, I can create SalePriceBin for grouping houses	2	Medium	Sudeshna
Sprint-2	Data Visualization	USN-5	As a user, I can create sheets with charts: price vs features	5	High	Sudeshna
Sprint-2	Dashboard Creation	USN-6	As a user, I can build an interactive Tableau Dashboard with filters	3	High	Sudeshna
Sprint-2	Dashboard Styling	USN-7	As a user, I can style the dashboard for better readability and navigation	2	Medium	Sudeshna
Sprint-3	Storytelling	USN-8	As a user, I can create a Tableau Story showing insights step by step	2	Medium	Sudeshna
Sprint-4	Documentation	USN-11	As a team, we can prepare final project documentation	3	High	Sudeshna

Sprint	Epic	User Story No.	User Story / Task	Points	Priority	Assigned To
Sprint-4	Demo Preparation	USN-12	As a team, we can prepare and rehearse a full demo walkthrough	2	Medium	Sudeshna
Sprint-4	Bug Fixing / Final QA	USN-13	As a team, we can test the full system and fix visual/logic bugs	2	Medium	Sudeshna

Project Tracker, Velocity & Burndown Chart

Sprint	Total Story Points	Duration	Start Date	End Date	Points Complete	Release Date
Sprint-11	1	4 Days	11 June 2021	14 June 2021	1	14 June 2021
Sprint-10	0	4 Days	15 June 2021	18 June 2021	0	18 June 2021
Sprint-7	7	4 Days	19 June 2022	22 June 2022	7	22 June 2022
Sprint-7B	7	4 Days	23 June 2022	26 June 2022	7	26 June 2022

Velocity Calculation

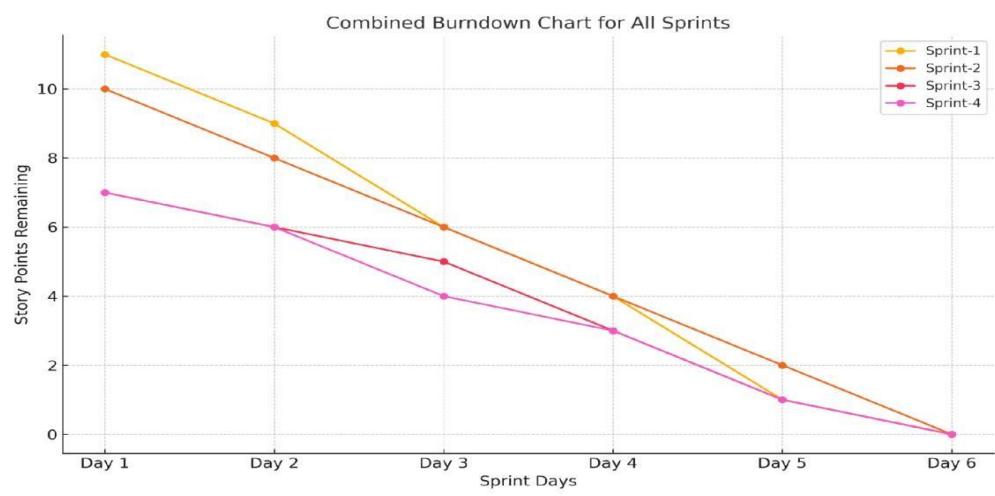
Total Points Completed: $11 + 10 + 7 + 7 = 35$

Total Duration: $4 + 4 + 4 + 4 = 16$ days

Average Velocity = Total Points Completed / Total Days = $35 / 16 = 2.19$ points/day

Burndown Chart Insight

- Initial Total Story Points: 35
- Sprint-wise burn (Remaining Points):
 - After Sprint-1: 24
 - After Sprint-2: 14
 - After Sprint-3: 7
 - After Sprint-4: 0



6. Functional and performance testing

6.1 Performance Testing

S.No	Parameter	Screenshot / Values
1	Data Rendered	<p>The dataset used contains housing sales data with fields such as: - Sale Price, Number of Bedrooms, Bathrooms, Flat Area, Lot Area, Basement Area, House Age, Condition, Renovation Status, Zipcode Group, and others.</p> <p>The data was provided in .csv format and includes derived and transformed columns for analytics and Tableau visualizations.</p>
2	Data Preprocessing	<p>Preprocessing was done using Python (Pandas) before importing into Tableau. Steps:</p> <p>Removed null/missing values</p> <ul style="list-style-type: none">Renamed columns for clarity (e.g., "No of Bedrooms" → "Bedrooms")Created calculated fields like TotalAreaSqftGenerated dummy variables for house conditions and renovation statusTransformed categorical fields for better Tableau usability.
3	Utilization of Filters	<p>Multiple filters were implemented in Tableau to improve interactivity:</p> <ul style="list-style-type: none">Number of BedroomsNumber of BathroomsHouse ConditionRenovation Status (Yes/No)Zipcode GroupSale Price Bins <p>These filters allow users to drill down and compare trends across different property types and locations.</p>
4	Calculated Fields Used	<p>Calculated fields created in Tableau include:</p> <ul style="list-style-type: none">TotalAreaSqft → [FlatAreaSqft] + [LotAreaSqft] + [BasementAreaSqft]SalePriceBin → Binning sale price into ₹100,000 intervals

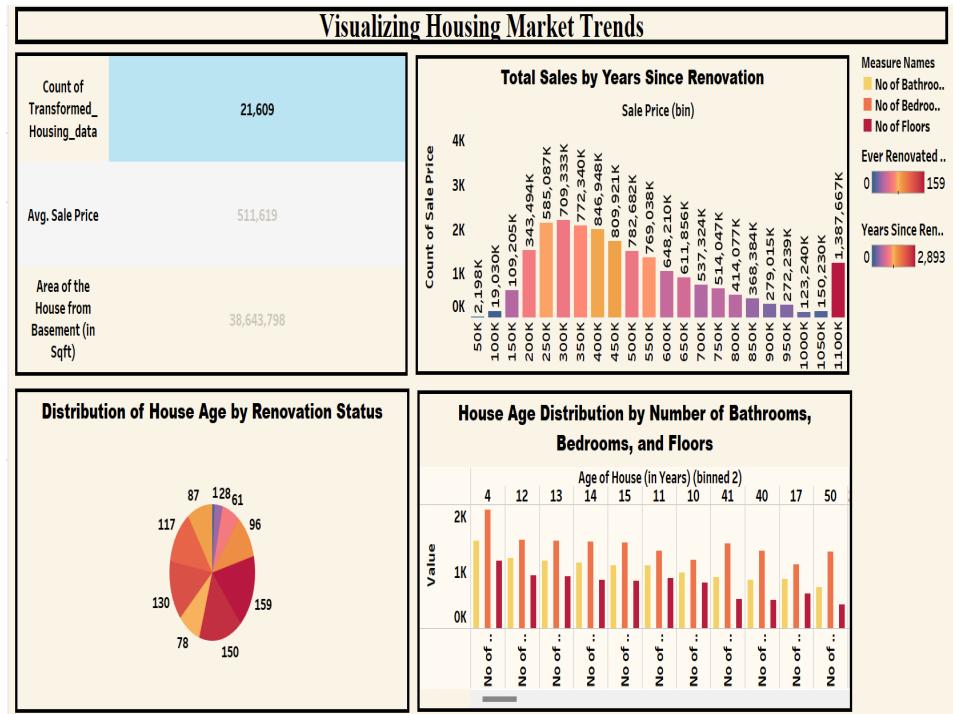
S.No Parameter

Screenshot / Values

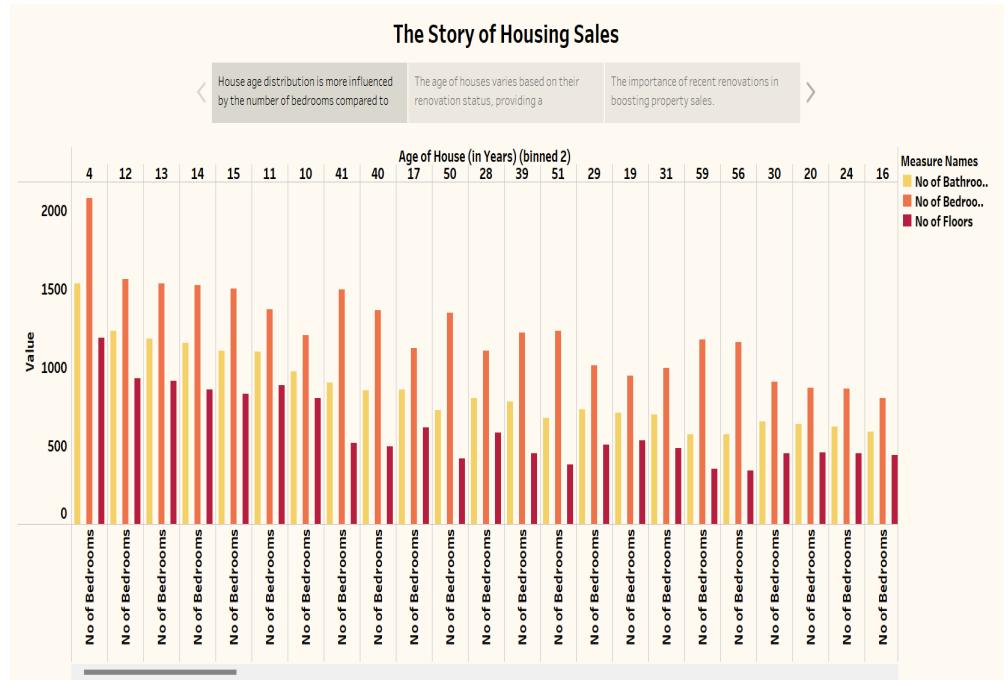
- Condition_Excellent, Condition_Good, etc. → Dummy fields
- Ever_Renovated_Yes → Dummy field for renovated homes
- AvgPrice → Average sale price for group insights
- HouseAge → Difference from year built and sale year

5

Dashboard



6 Story Design



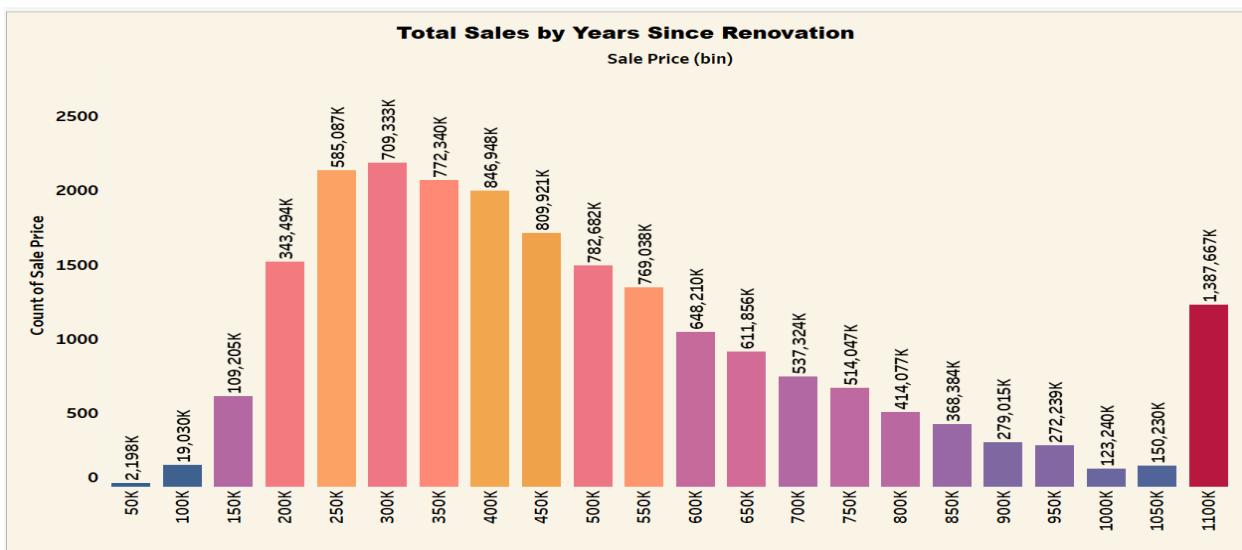
7. Results

7.1 Output Screenshots

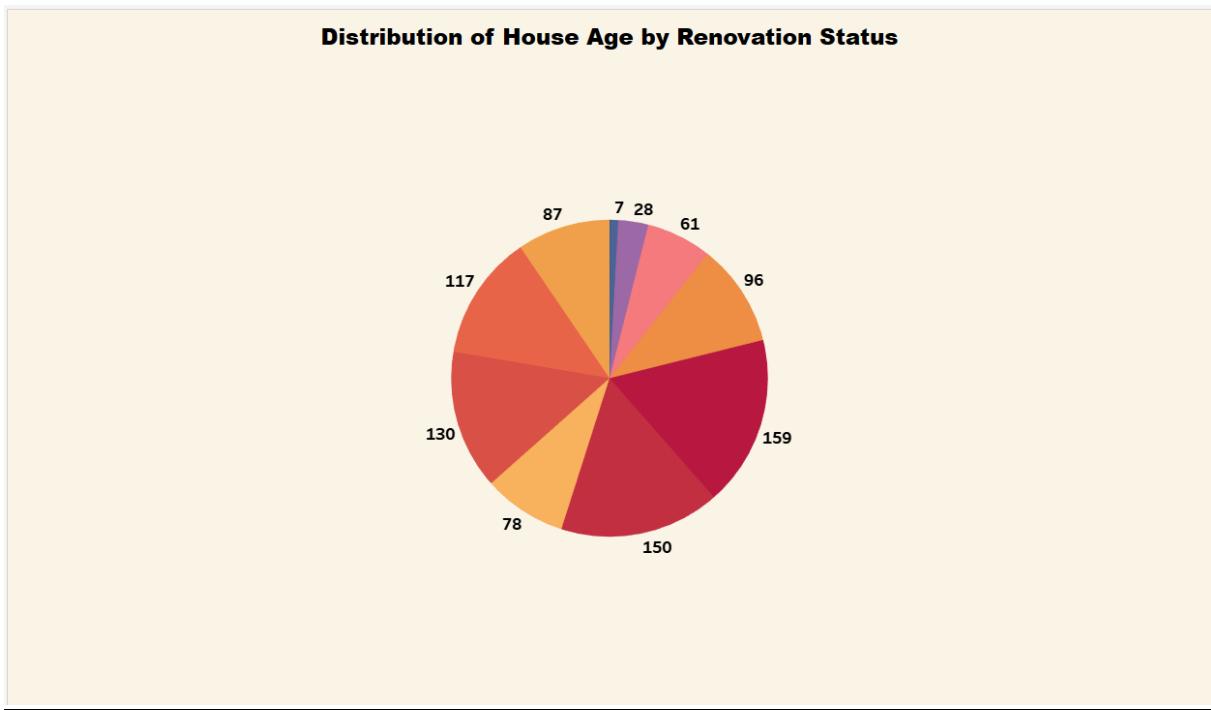
Output of Sheet 1:

Count of Transformed_ Housing_data	21,609
Avg. Sale Price	511,619
Area of the House from Basement (in Sqft)	38,643,798

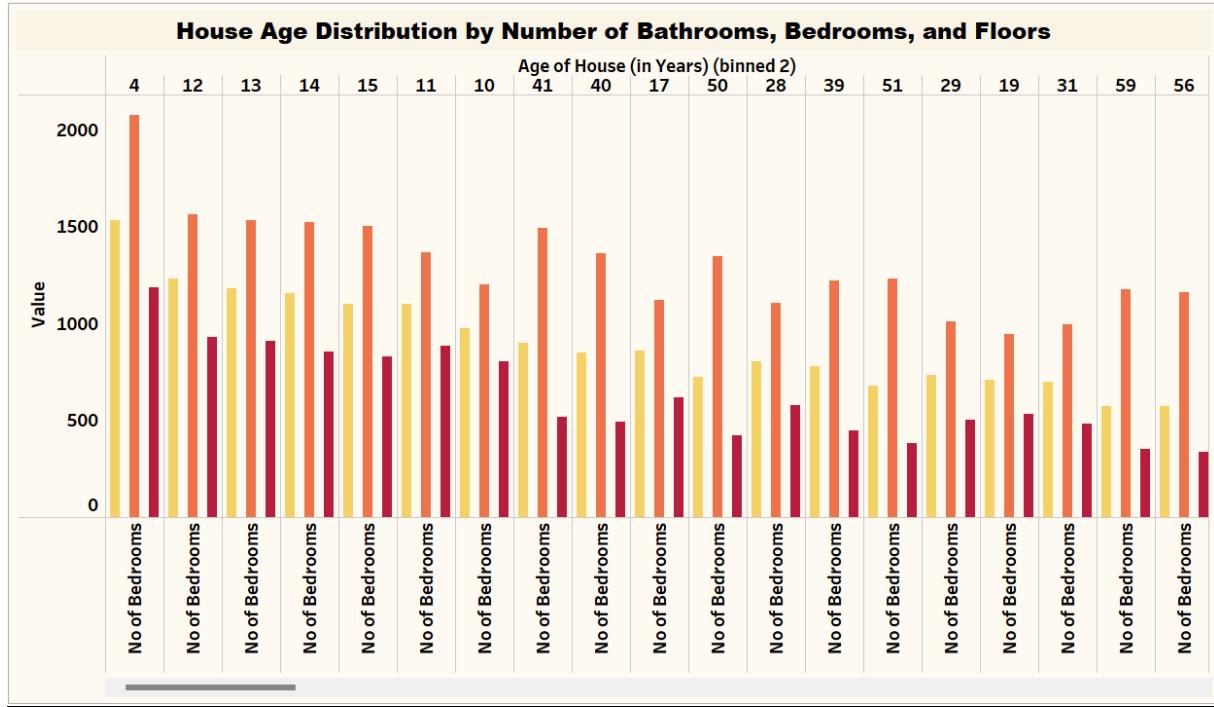
Output of Sheet 2:



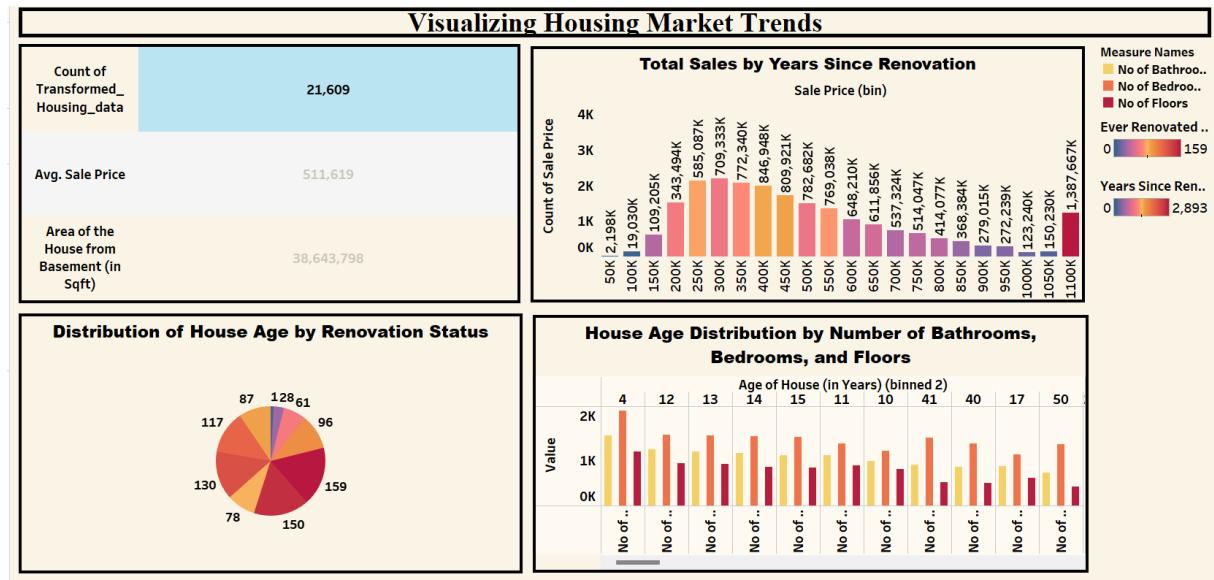
Output of Sheet 3:



Output of Sheet 4:



Output of Dashboard:



Output of Story:

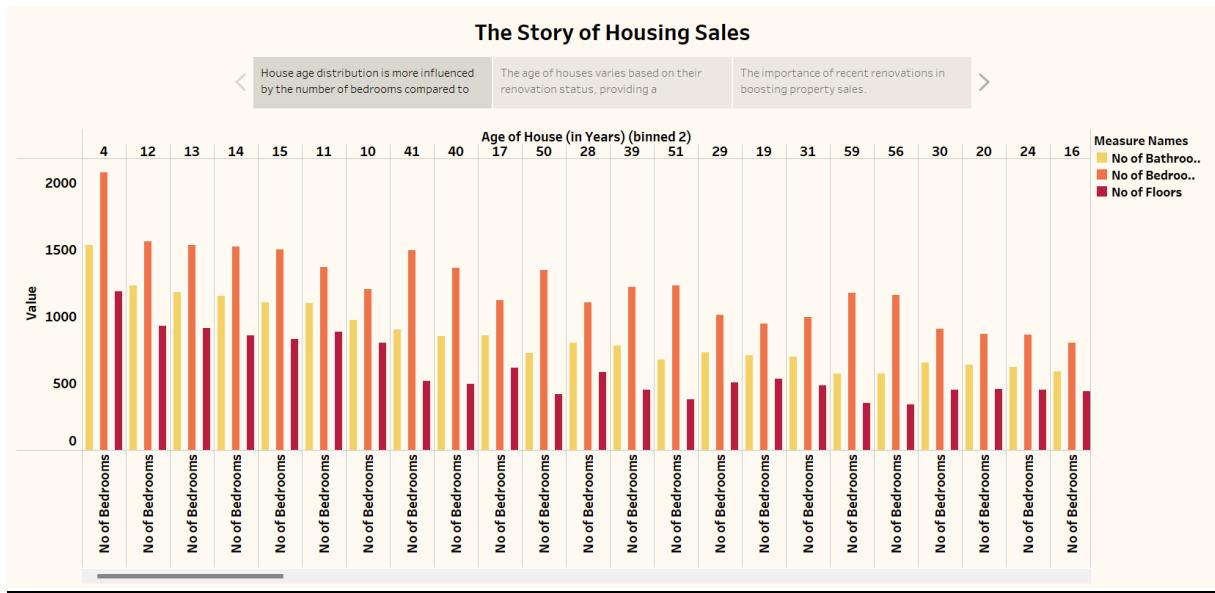
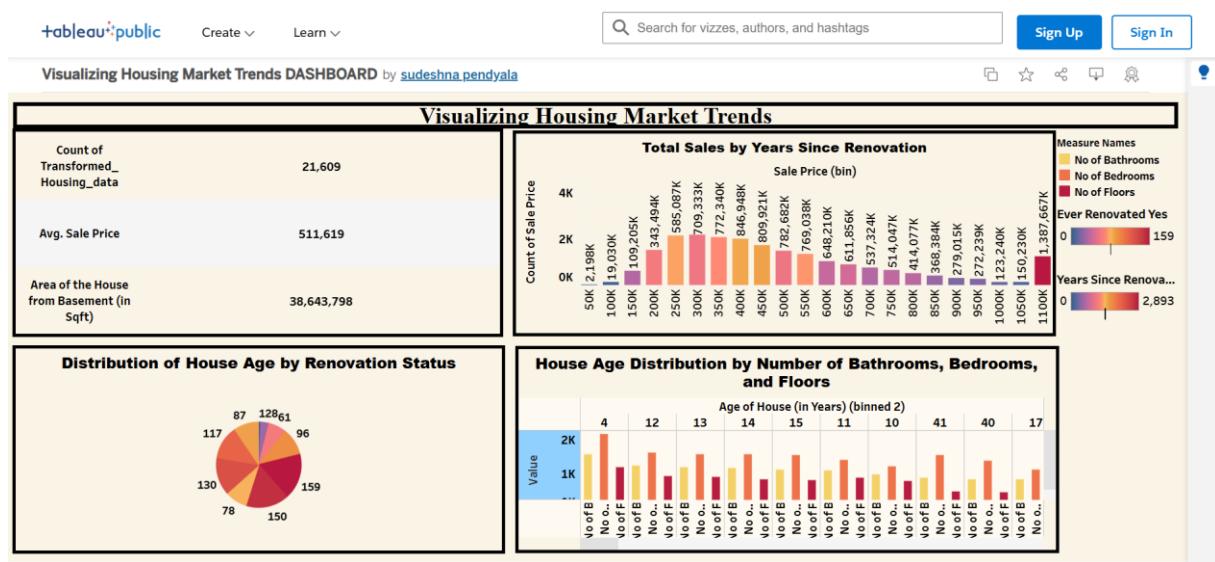


Tableau public link

<https://public.tableau.com/app/profile/sudeshna.pendyala/viz/VisualizingHousingMarketTrendsDASHBOARD/Dashboard1?publish=yes>

Output of Published Dashboard:



8. Advantages & disadvantages

Advantages:

1. Interactive Analysis:

The Tableau dashboard allows users to explore data with filters (e.g., bedrooms, renovation status, price bins), enhancing understanding through dynamic interactions.

2. Informed Decision-Making:

Buyers, sellers, agents, and investors can make data-driven decisions by identifying which features (e.g., area, renovations, number of floors) impact property value.

3. Data Storytelling:

The Tableau Story feature presents insights in a sequential, digestible narrative—great for business reports or stakeholder presentations.

4. Geographic Visualization:

Zipcode grouping allows regional comparison of price trends and property types, revealing market opportunities and local disparities.

5. Calculated Metrics & KPIs:

Metrics like Average Sale Price and Total Area improve business clarity and enable fast comparisons across categories.

6. Web Accessibility:

Embedding the dashboard into a Flask web app increases accessibility—users can view it from any browser without needing Tableau Desktop.

7. Modular & Scalable Design:

The project structure supports additional data (e.g., rental prices, future years), making it expandable to other regions or market conditions.

8. Minimal Coding Required:

Most of the visualizations are created using Tableau's drag-and-drop interface—making it ideal for analysts without deep programming expertise.

Disadvantages:

1. Static Dataset Limitation:

The analysis depends on a preloaded CSV file; it doesn't support real-time updates unless integrated with live databases or APIs.

2. Tool Dependency:

The system relies on Tableau Public, which has limitations like no row-level security and requires dashboards to be public.

3. Learning Curve for Tableau:

While Tableau is user-friendly, new users may need time to understand calculated fields, filters, and advanced charting options.

4. Limited Predictive Power:

This is a descriptive and visual analytics project—it does not use machine learning or predictive modeling to forecast housing prices.

5. Browser Compatibility:

Older browsers or low-resolution screens may not render complex dashboards optimally, especially if not designed responsively.

6. Manual Data Preprocessing:

Initial data cleaning, renaming, and transformation were done manually using Python or within Tableau, which might be error-prone at scale.

9. Conclusion:

The project "Visualizing Housing Market Trends: An Analysis of Sale Prices and Features using Tableau" successfully demonstrates how complex real estate data can be transformed into meaningful, interactive visual insights. By leveraging Tableau's powerful visualization capabilities, we have made it easier for buyers, sellers, investors, and analysts to understand the key factors influencing house prices. Our dashboard enables quick comparisons based on features like number of bedrooms, renovations, house age. Overall, this project bridges the gap between raw housing data and strategic real estate decision-making, allowing users to gain actionable insights with minimal technical expertise.

10. Future scope:

1. Live Data Integration:

Future versions can integrate live property listings or transaction data via APIs or real-time databases to provide up-to-date market insights.

2. Machine Learning Forecasting:

Incorporating regression models or time-series forecasting can help predict future housing prices based on historical trends and features.

3. Rental Market Visualization:

Extend the dashboard to include rental data analysis, enabling a broader comparison between buying vs. renting decisions.

4. Mobile Optimization:

Responsive design enhancements can be implemented to ensure the dashboard performs well across tablets and smartphones.

5. Advanced User Access Control:

By using Tableau Server or Tableau Online, dashboards can be secured with role-based access for different stakeholders.

6. Location Intelligence Enhancements:

Integration of geospatial data, satellite maps, or demographic overlays can improve locationbased insights (e.g., school zones, crime rates).

7. Recommendation Engine:

Develop a recommendation system to suggest optimal property types using user-input filters.

11. Appendix

Dataset Link:

- [Transformed Housing Data](#)

Project Demo Video Link:

- [Demo Video](#)

GitHub Repository Link:

- [Github Repository](#)