

An internship in
Data Analytics with Tableau
by
SmartInternz

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

Project Id: LTVIP2025TMID47431

Project Mentor: M. Ganesh

Team Members:

- | | |
|-------------------|---------------------|
| 1. N Saanvi Reddy | (Reg.No.22BFA32097) |
| 2. S Haritha | (Reg.No.22BFA32115) |
| 3. Thota Indu | (Reg.No.22BFA32128) |
| 4. M Keerthi | (RegNo.22BFA32087) |

DEPARTMENT OF CSE (Data Science),

Sri Venkateshwara College of Engineering, Tirupati
Mangalam Road,dist. Tirupati, Andhra Pradesh.

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Sri Venkateshwara College of Engineering,Tirupati** for their kind encouragement in doing this project and for completing it successfully. I convey my thanks to **Dr. N. Sudhakar Reddy**, principal of Sri Venkateshwara College of Engineering, Tirupati and **Dr. R. Suresh Ramayanam**, Head of the Department of CSD (Data Science) 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. M. Ganesh**, for his valuable guidance, suggestions and constant encouragement pavedway 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(Data Science)** 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. These dashboards were then embedded into a Flask web application, ensuring easy accessibility and deployment. 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 Map Canvas

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 data-driven decisions for buyers, sellers, and real estate professionals.

1.2. Objectives

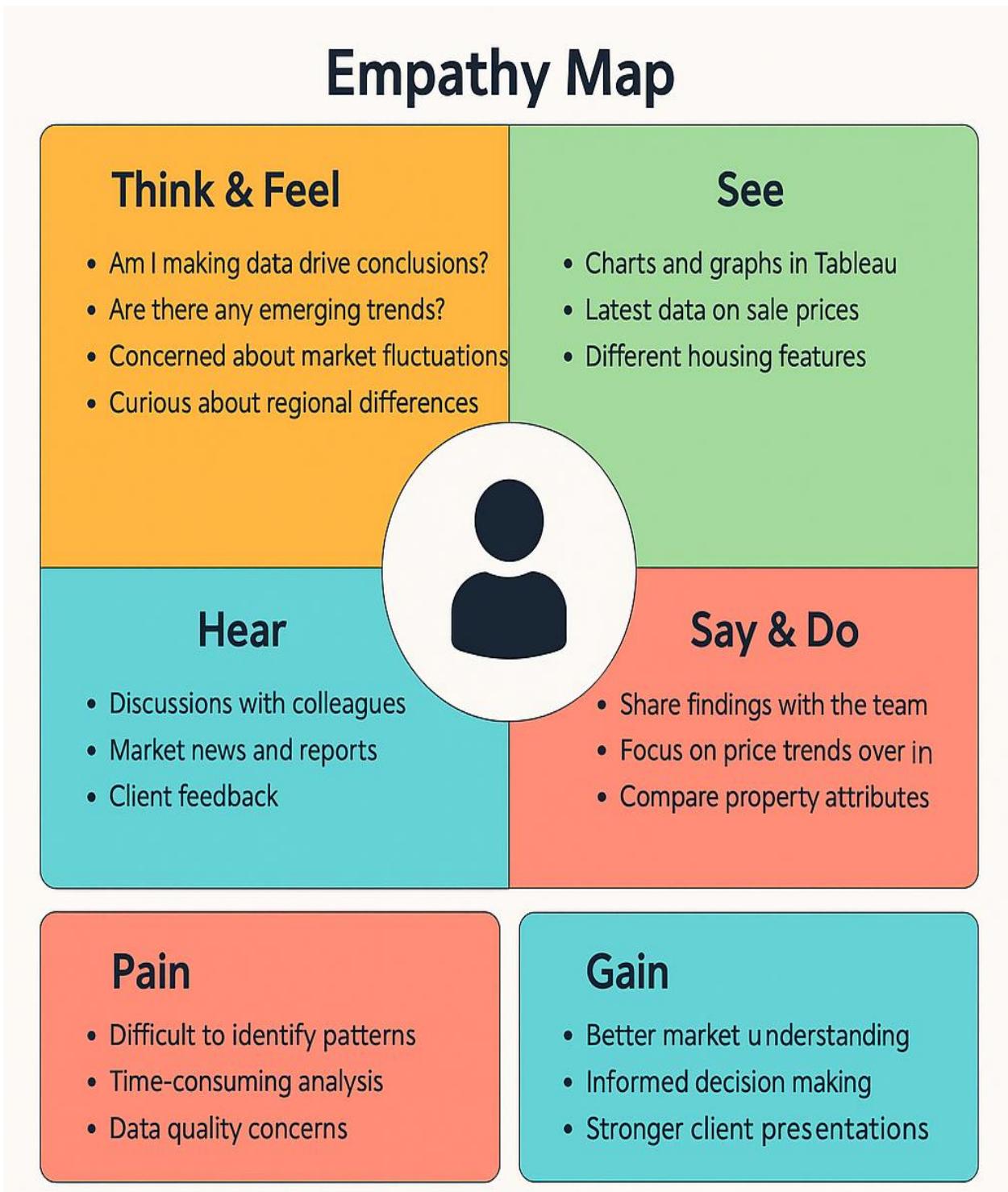
- Identify key factors influencing house prices.
- Analyze the effect of renovations on property value.
- Explore the distribution of house sales across different price ranges.
- Create interactive Tableau dashboards to present findings effectively.

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 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 analyze 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 Leader: N Saanvi Reddy
- Team Member: S Haritha
- Team Member: Thota Indu
- Team Member: M Keerthi

Step 2: Brainstorming, Idea Listing and Grouping

S.No	Idea Description	Category
1	Visualize average sale price by SalePriceBin	Pricing Insights
2	Analyze 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
5	Group insights by Zipcode Clusters	Geographical Comparison
6	Analyze 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

S.No	Idea Description	Category
11	Embed Dashboard in Web Application using Flask	Deployment & Integration
12	Add filters for Bedrooms, Condition, Renovation in Dashboard	Interactive Exploration

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 TotalArea vs Price (scatter plot)	High	Easy	High
4	Compare prices for renovated vs. non-renovated homes	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	Use Tableau dashboard to combine insights	High	Easy	High
10	Build a Story in Tableau	High	Medium	High
11	Embed Dashboard in Web Application	High	Hard	Medium
12	Add filters for Bedrooms, Condition, Renovation	Medium	Easy	Medium

3. Requirement analysis

3.1 Customer Journey map

Customer Journey Map: Housing Market Trends Dashboard

Stage	Actions & Touchpoints	Experience & Emotions	Pain Points	Opportunities	User Goals
Awareness	<ul style="list-style-type: none"> - 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	<ul style="list-style-type: none"> - 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	<ul style="list-style-type: none"> - 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 speed	Discover valuable insights
Decision	<ul style="list-style-type: none"> - 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	<ul style="list-style-type: none"> - Subscribes for updates - Revisits for new data - Leaves feedback 	Loyal, Empowered	No update notifications, feedback unacknowledged	Enable email updates, actively respond to feedback	Stay informed and engaged

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

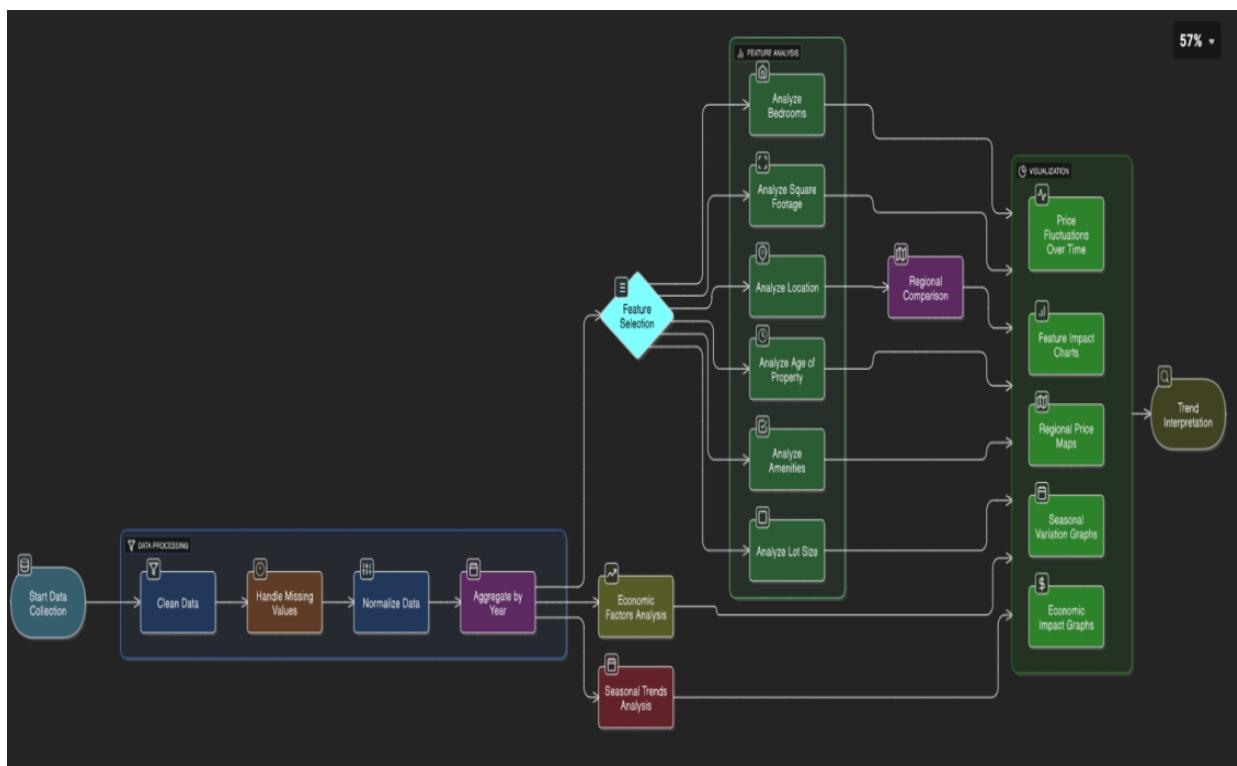
Non-Functional Requirements (NFRs)

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.

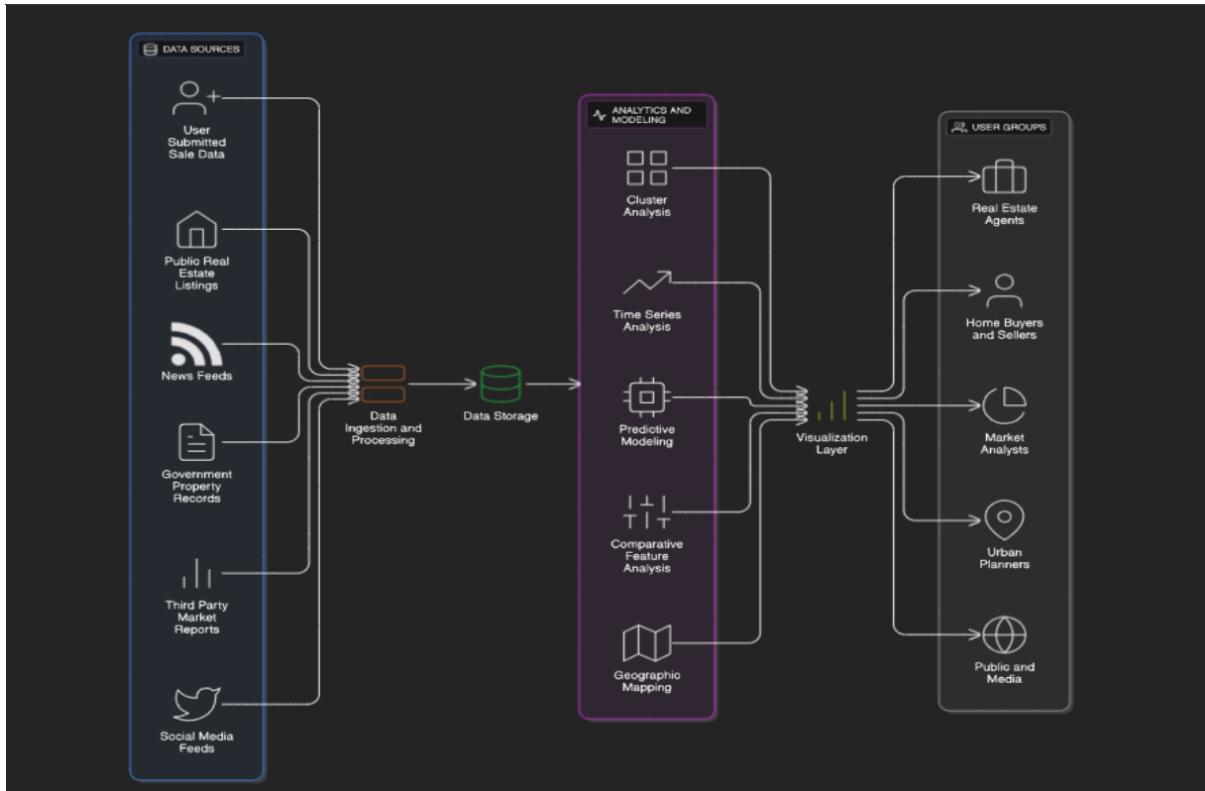
1. Data collected from POSOCO in CSV format.
2. Cleaned and transformed, with calculated fields like Year and Lockdown.
3. Visualizations built in Tableau using multiple worksheets.
4. Users review the dashboard and may request changes.
5. Final version archived after approval.



User Stories Table:

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance Criteria	Priority	Release
Analyst / Policymaker	View Electricity Trends	USN-1	As a user, I want to view electricity usage trends by region and year.	I can filter and view charts for specific regions and years.	High	Sprint-1
Analyst	Compare States	USN-2	As a user, I want to compare top and bottom electricity-consuming states.	I can view bar charts with top N and bottom N states.	Medium	Sprint-1
Policy Maker	Forecast Planning	USN-3	As a user, I want to analyze seasonal variations in usage.	I can view monthly and quarterly consumption trends.	High	Sprint-2
Energy Consultant	View Impact of Lockdown	USN-4	As a user, I want to compare electricity usage before and after lockdown.	A before-after lockdown chart is available for selection.	High	Sprint-2
Developer	Connect Data	USN-5	As a user, I want the dashboard to be connected to a live database.	Data refreshes automatically from MySQL to Tableau.	Medium	Sprint-2
Developer	Export Insights	USN-6	As a user, I want to export dashboard views for presentations.	I can download dashboards as images or PDFs.	Low	Sprint-3

3.4 Technology Stack



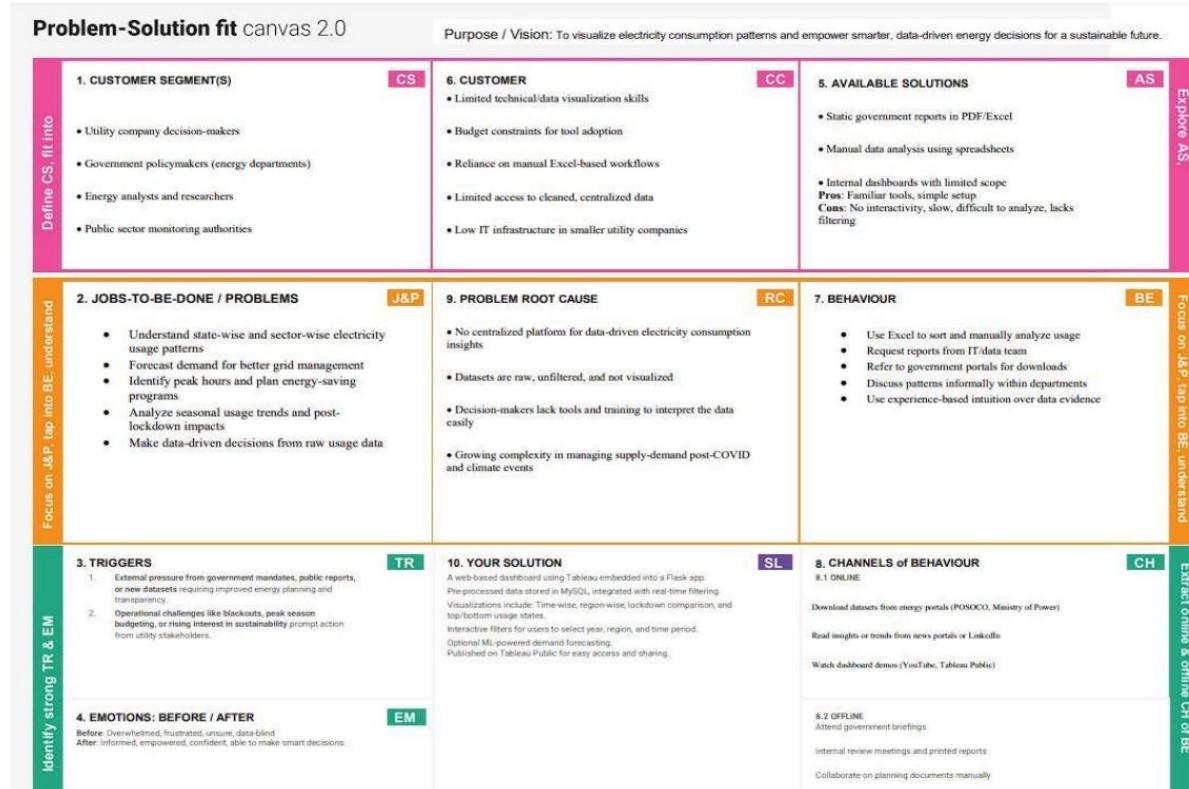
4. Project design

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 behavior.
- 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

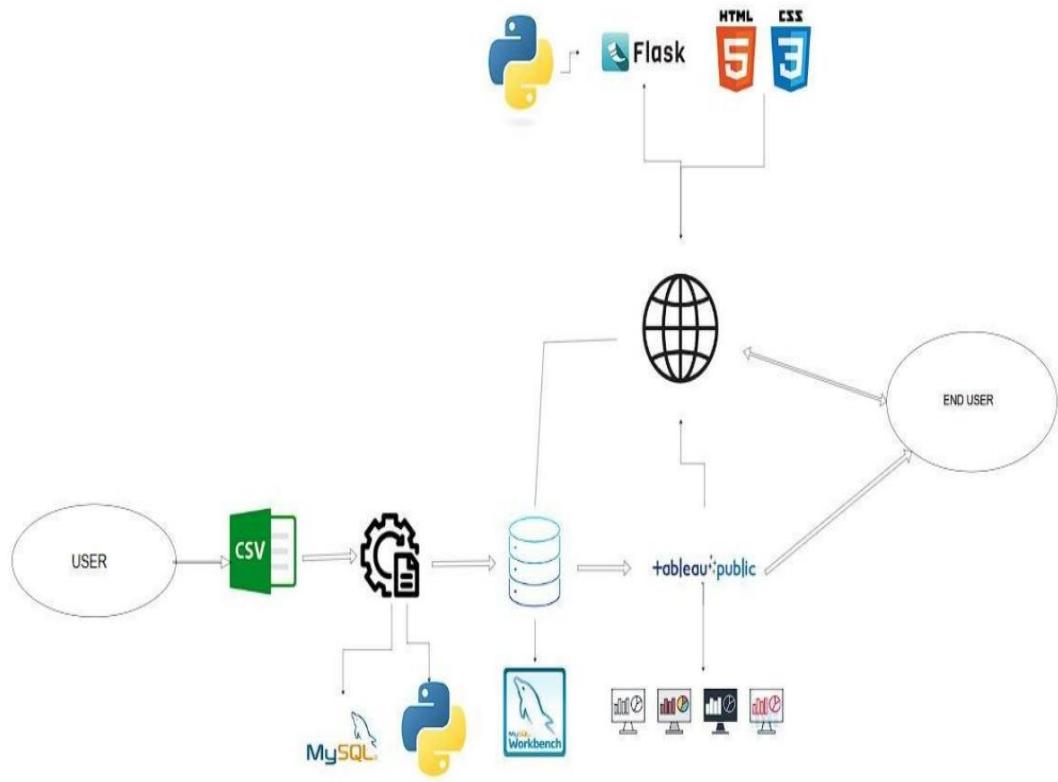
Proposed Solution Template

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), 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	Point Priority
Sprint 1 Data Setup	USN-1	As a user, I can upload housing data in CSV format	3 High
Sprint 1 Data Cleaning	USN-2	As a developer, I can clean and preprocess housing data in Tableau	4 High
Sprint 1 Field Creation	USN-3	As a user, I can create calculated fields like TotalAreaSqft	2 Medium
Sprint 1 Price Binning	USN-4	As a user, I can create SalePriceBin for grouping houses	2 Medium
Sprint 2 Data Visualization	USN-5	As a user, I can create sheets with charts: price vs features	5 High
Sprint 2 Dashboard Creation	USN-6	As a user, I can build an interactive Tableau Dashboard with filters	3 High
Sprint 2 Dashboard Styling	USN-7	As a user, I can style the dashboard for better readability and navigation	2 Medium
Sprint 3 Storytelling	USN-8	As a user, I can create a Tableau Story showing insights step by step	2 Medium
Sprint 3 Flask Integration	USN-9	As a developer, I can embed Tableau dashboard into a Flask web app	4 High
Sprint 3 Embed Testing	USN-10	As a user, I can test and review the embedded dashboard UI	2 Medium
Sprint 4 Documentation	USN-11	As a team, we can prepare final project documentation	3 High
Sprint 4 Demo Preparation	USN-12	As a team, we can prepare and rehearse a full demo walkthrough	2 Medium

Sprint Epic	User Story No.	User Story / Task	Point Priority
-------------	----------------	-------------------	----------------

Sprint 4	Bug Fixing / Q4 USN-13	As a team, we can test the full system and fix visual/logic bugs	2	Medium
----------	------------------------	--	---	--------

Project Tracker, Velocity & Burndown Chart

Sprint	Total Story Points	Duration	Start Date	End Date	Points Completed	Release Date
Sprint 11	4 Days	11 June 2021	14 June 2021		11	14 June 2021
Sprint 10	4 Days	15 June 2021	18 June 2021		10	18 June 2021
Sprint 7	4 Days	19 June 2021	22 June 2021		7	22 June 2021
Sprint 7	4 Days	23 June 2021	26 June 2021		7	26 June 2021

Velocity Calculation

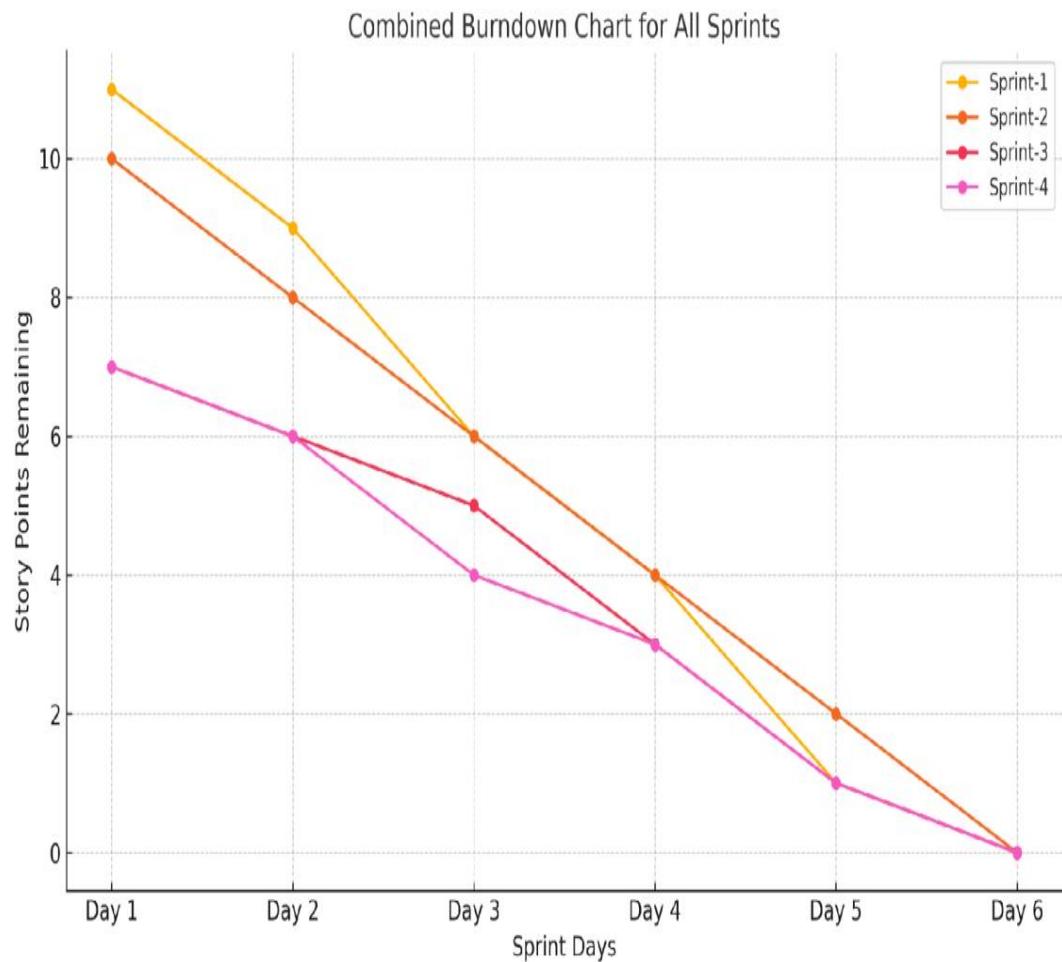
Total Points Completed: $11 + 10 + 7 + 7 = 35$

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



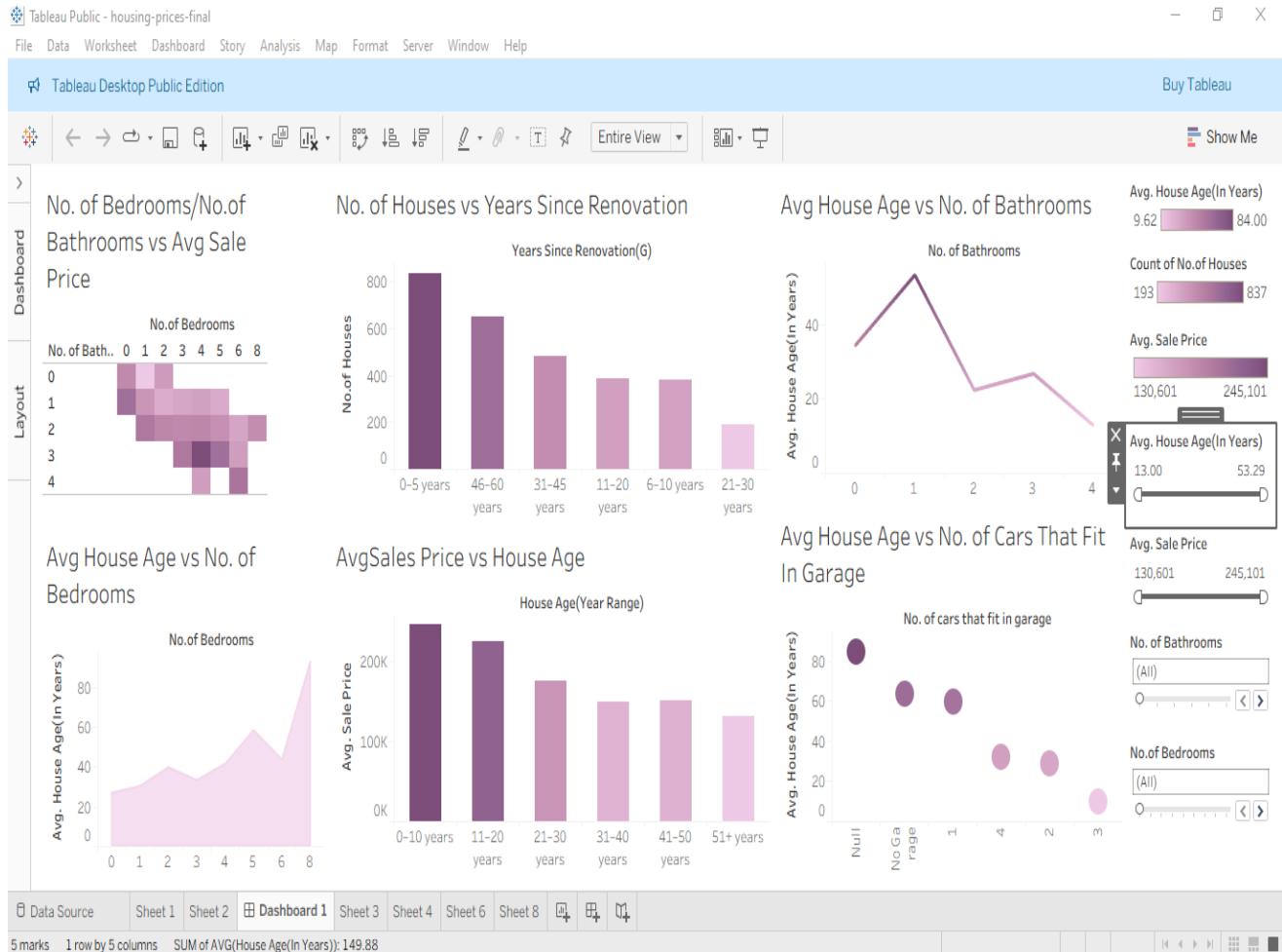
6. Functional and performance testing

6.1 Performance Testing

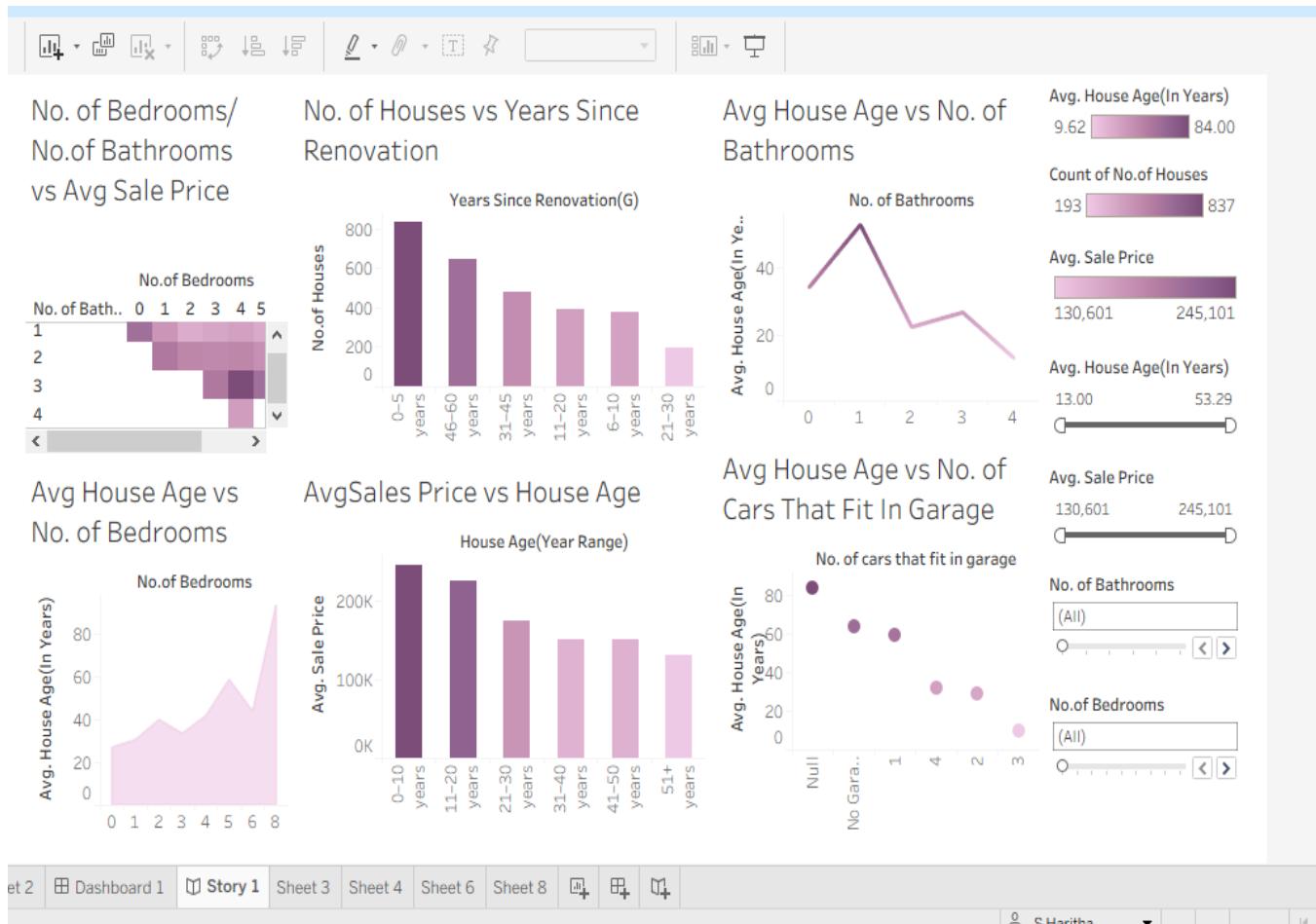
S.No	Parameter	Screenshot / Values
1.	Data Rendered	<p>The dataset used contains housing sales data with fields such as Sale Price, Number of Bedrooms, Bathrooms, Flat Area, Lot Area, Basement Area, House Age, Condition, Renovation Status, Zipcode Group, and others. The data was provided in .csv format and includes derived and transformed columns suitable for advanced analytics and visualizations in Tableau.</p>
2.	Data Preprocessing	<p>Before importing the data into Tableau, preprocessing was done using Python (Pandas). The following steps were performed:</p> <ul style="list-style-type: none">• Removed null or missing values.• Renamed columns for clarity (e.g., “No of Bedrooms” → “Bedrooms”).• Created calculated fields like “TotalAreaSqft” (sum of flat, lot and basement areas).• Generated dummy variables for house conditions and renovation status.• Transformed categorical fields to improve Tableau usability. <p>The final cleaned dataset was stored and imported into Tableau for visualization.</p>

3.	Utilization of Filters	<p>Multiple filters were implemented in Tableau to improve interactivity and user exploration. These include:</p> <ul style="list-style-type: none"> • Number of Bedrooms • Number of Bathrooms • House Condition • Renovation Status (Yes/No) • Zipcode Group • Sale Price Bins <p>These filters allow users to drill down and compare trends across different property types and regions.</p>
4	Calculated Fields Used	<p>Several calculated fields were created in Tableau to enhance analysis and interactivity:</p> <ul style="list-style-type: none"> • TotalAreaSqft → [FlatAreaSqft] + [LotAreaSqft] + [BasementAreaSqft] • SalePriceBin → Binning Sale Price into ₹100,000 intervals • Condition_Excellent, Condition_Good, etc. → Dummy fields (0/1) • Ever_Renovated_Yes → Dummy field to identify renovated home • AvgPrice → AVG([SalePrice]) for grouped insights • HouseAge → Difference between year built and sale date if available <p>(or derived field if pre-calculated)</p> <p>These fields enable comparisons across pricing, condition, and space utilization.</p>

Dashboard



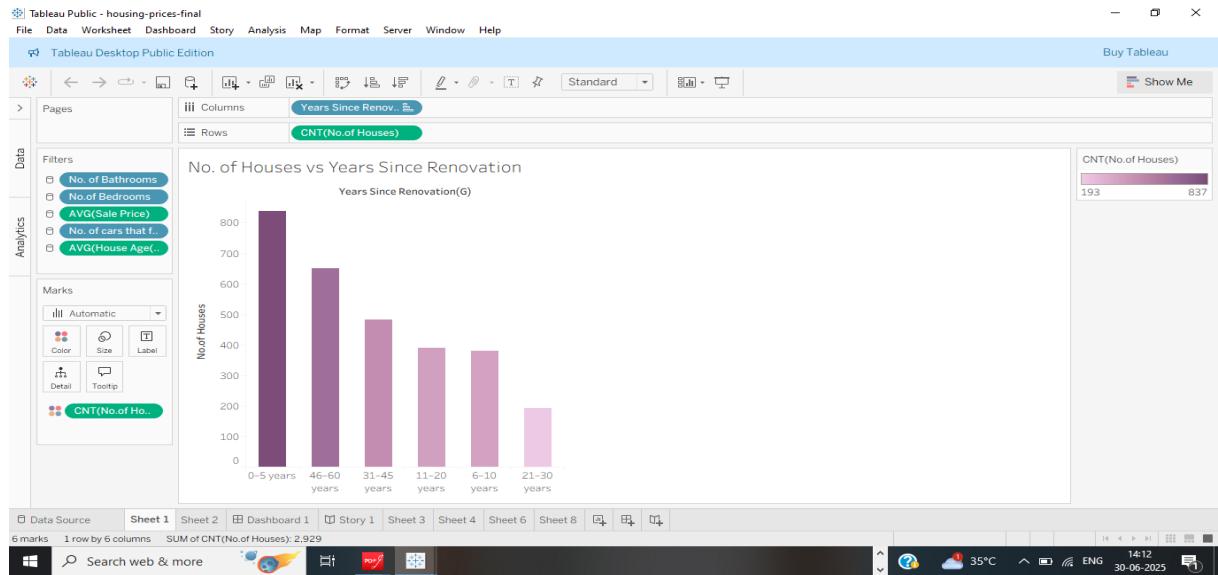
STORY



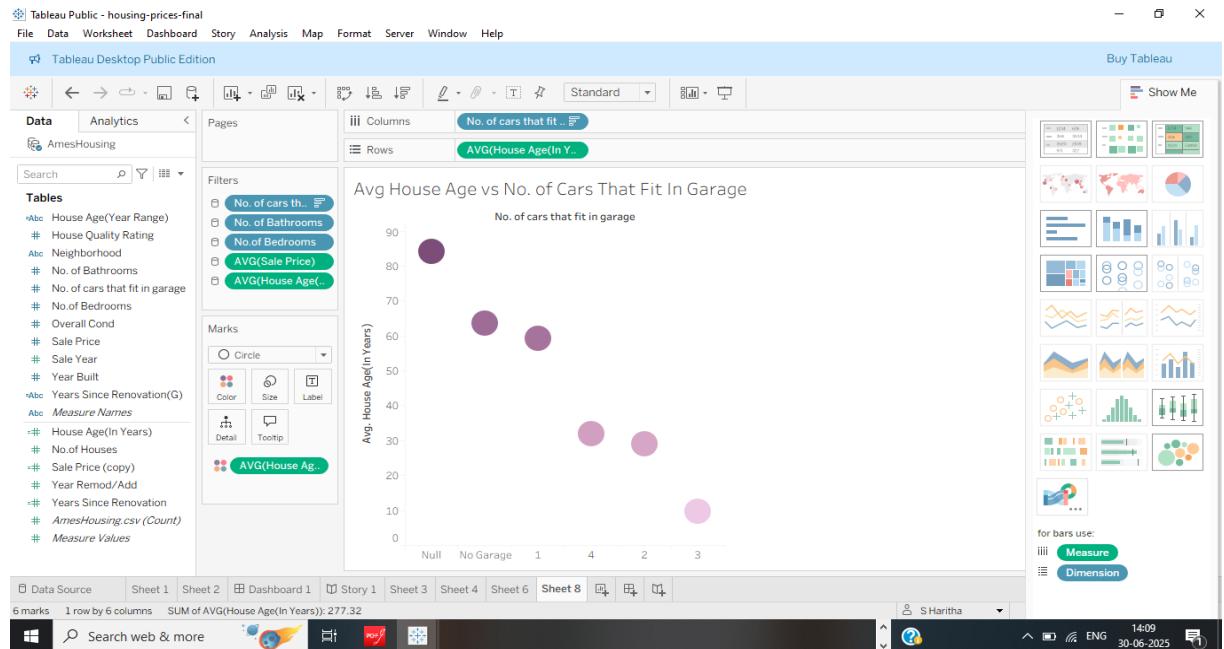
7. Results

7.1 Output Screenshots

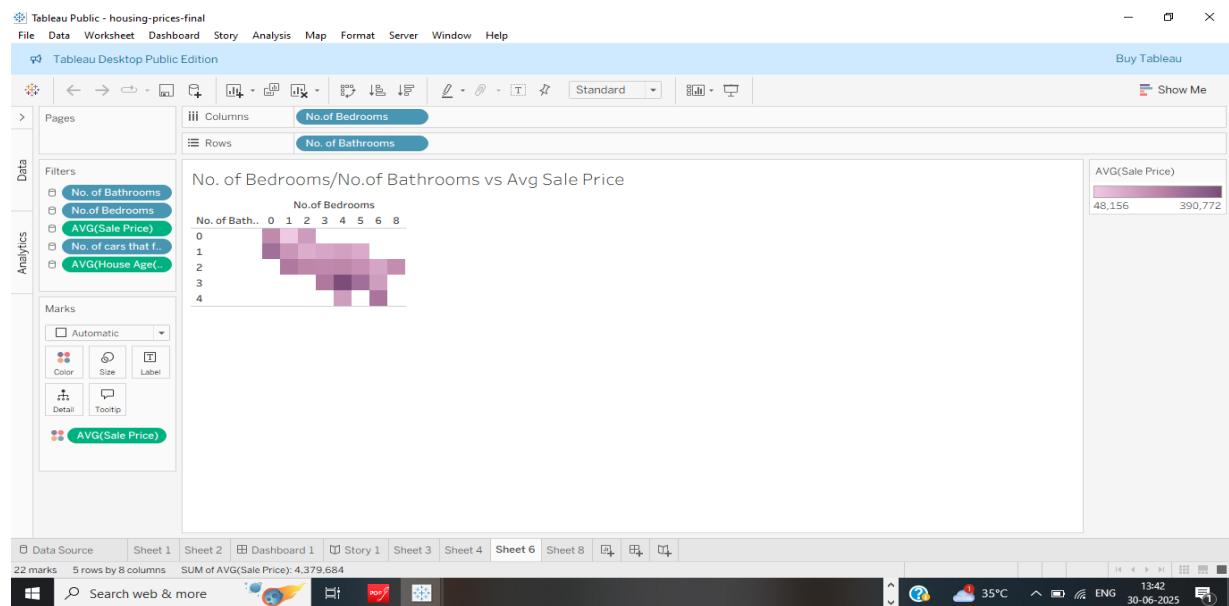
Output of Sheet 1



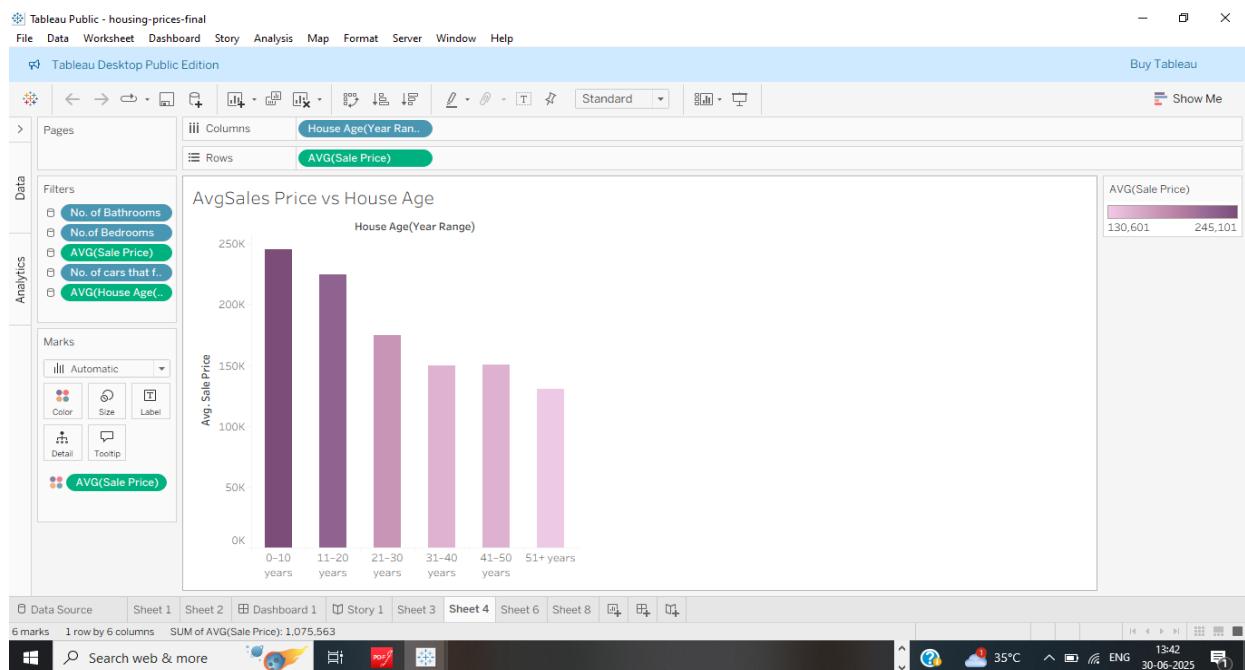
Output of Sheet 2



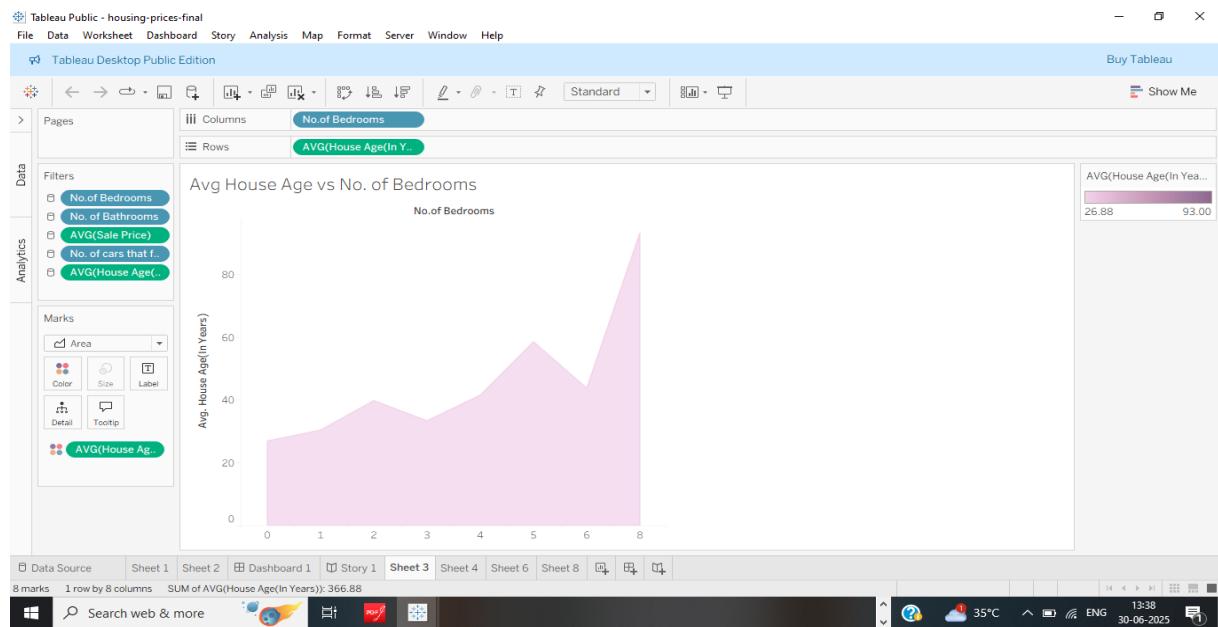
Output of Sheet 3



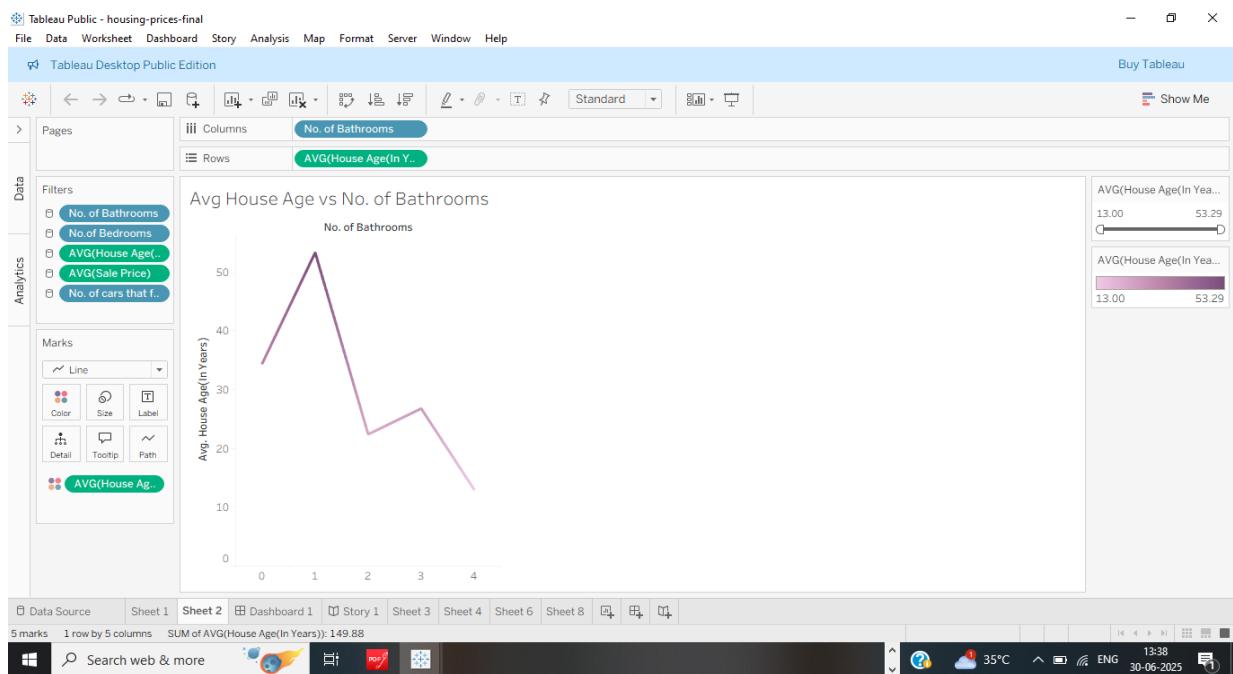
Output of Sheet 4



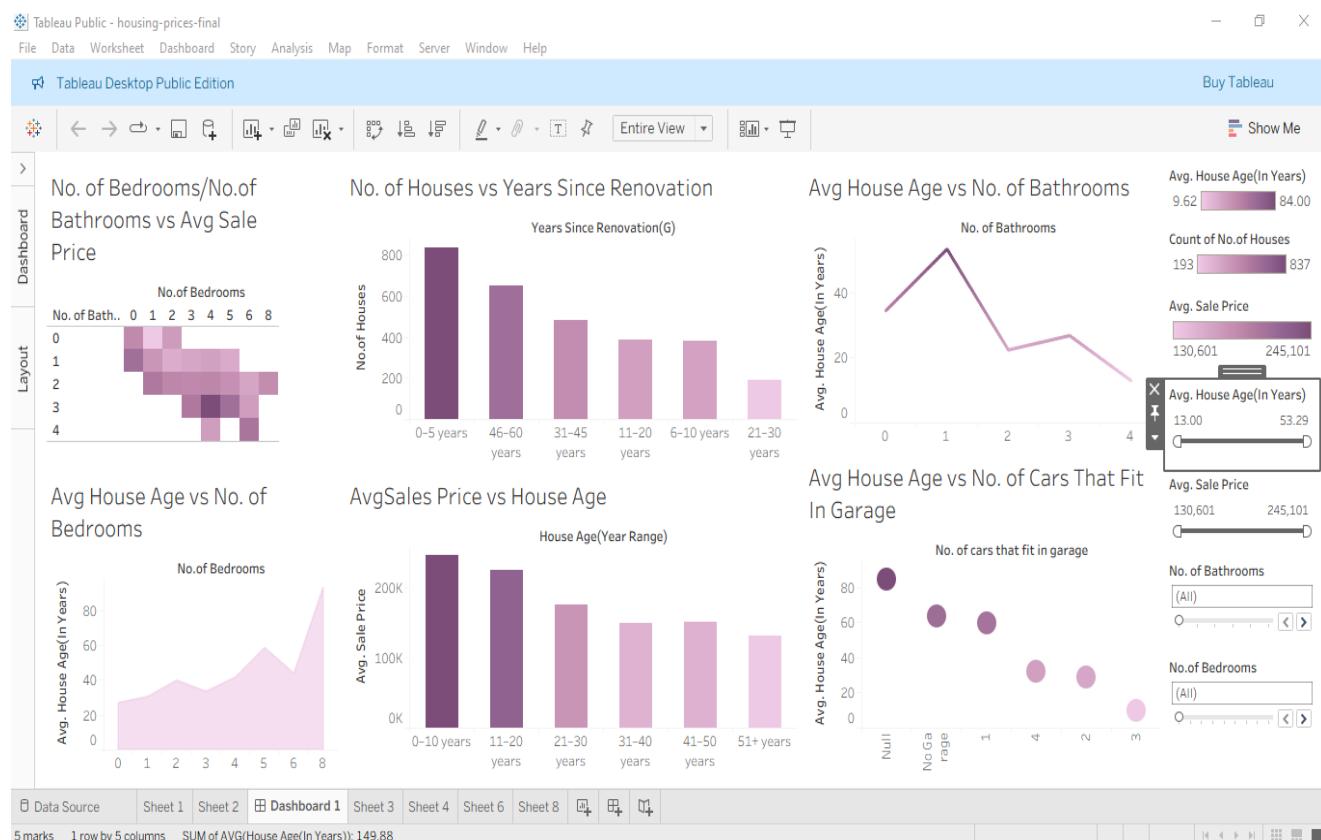
Output of Sheet 5



Output of Sheet 6



Output Dashboard



Output Story

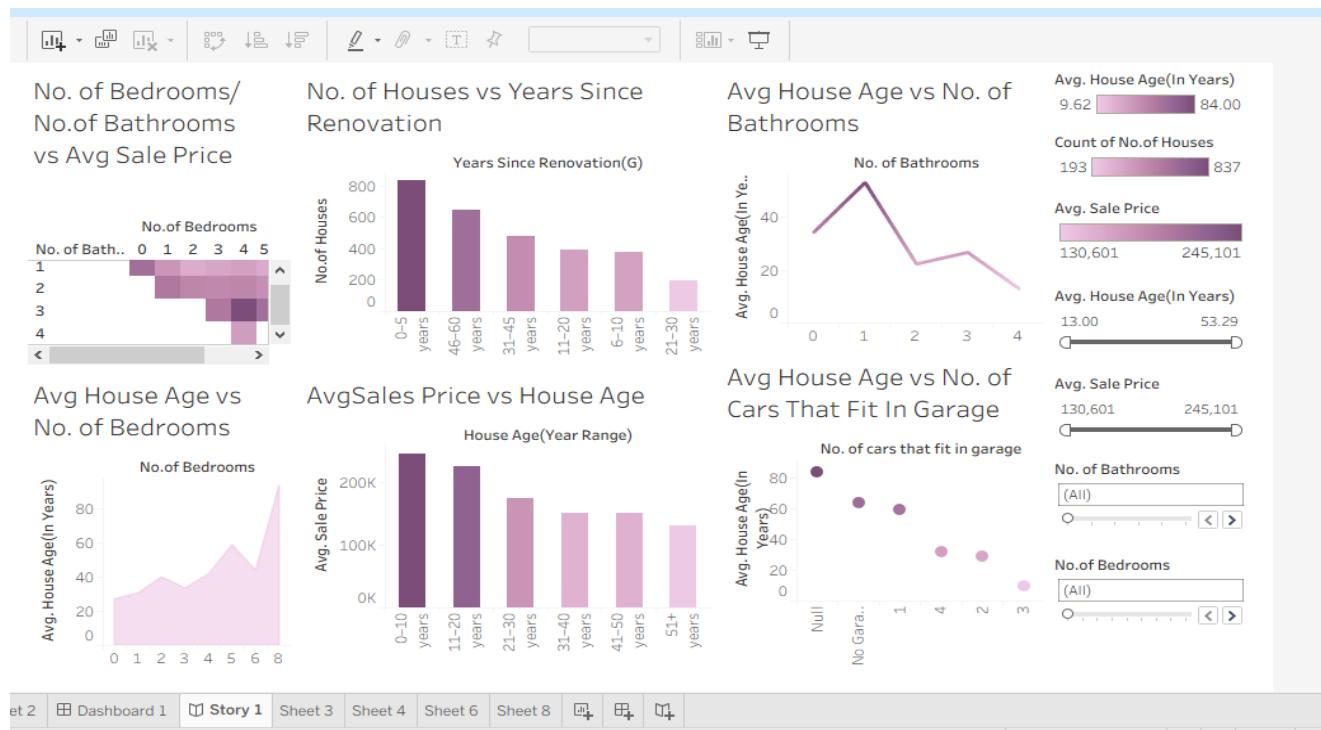
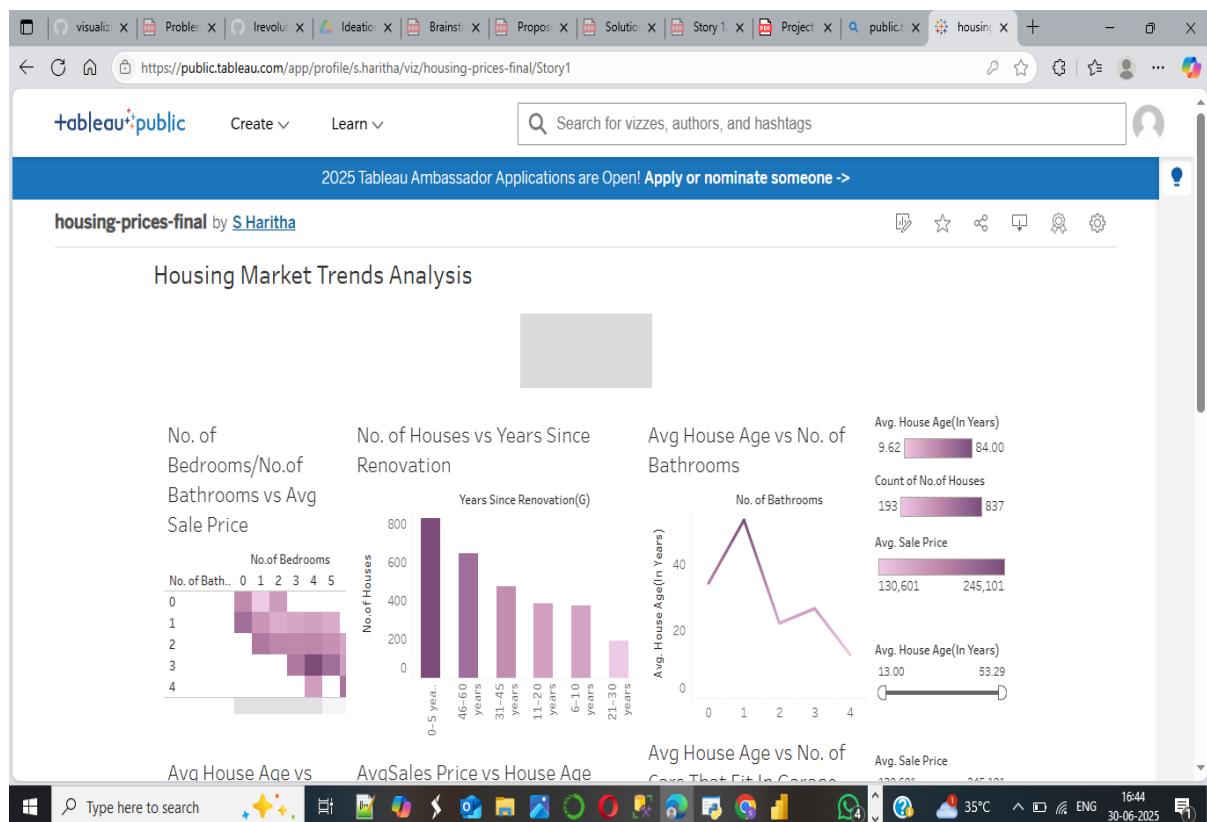


Tableau Public Link

https://public.tableau.com/views/housing-prices-final/Story1?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link



Appendix

Dataset:

https://docs.google.com/spreadsheets/d/1m1CBnGE2Wpe_m2g32ZArw09YRw8SbHJCKuBK_XCohBQ/edit?usp=sharing