Final Project Report: House Sales Trend Analysis Using Tableau

1.INTRODUCTION

1.1 Project OverviewThis project involves analyzing house sales data using Tableau to provide insights on market trends, renovation impacts, area-wise pricing, and seasonal demand patterns. It helps real estate professionals make informed decisions.

1.2 PurposeTo visualize and interpret key metrics from house sales data across various cities to support strategic planning and investment decisions in the housing market.

Ideation Phase
Brainstorm & Idea Prioritization Template

Brainstorm & Idea Prioritization Template:

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

Reference: https://www.mural.co/templates/brainstorm-and-idea-prioritization

Step-1: Team Gathering, Collaboration and Select the Problem Statement

This project aims to analyze a comprehensive housing dataset using Tableau to identify patterns in sale prices across different renovation timelines, house ages, and feature distributions. The goal is to provide a data-driven overview that supports strategic decision-making in areas such as pricing, marketing, and investment planning.

To analyze and visualize factors influencing housing sale prices and property characteristics using Tableau. The goal is to understand how features like renovations, house age, and property attributes affect sales patterns for better decision-making in the housing market.

Step-2: Brainstorm, Idea Listing and Grouping

| Idea No. | Idea Description | Group/Theme |

```
--------------------------------------------------------------
|  1    | Show total basement area as a KPI            | High-level Metrics    |
|  2    | Count the number of houses in the dataset    | High-level Metrics    |
|  3    | Show average sale price as a KPI             | High-level Metrics    |
|  4    | Bar chart: Sales by years since renovation   | Renovation Impact   |
|  5    | Pie chart: House age by renovation status    | House Age Analysis   |
|  6    | Grouped bars: Age by bedrooms/bath/floors    | Feature Comparison  |
|  7    | Use color for clarity (age, renovation)      | Visual Enhancement   |
--------------------------------------------------------------
```

Step-3: Idea Prioritization

Idea Description	Priority	Reason for Priority
Total Basement Area KPI	High	Key metric; shows property size
Total Count of Records KPI	High	Indicates dataset volume
Average Sale Price KPI	High	Crucial for price insight
Sales by Years Since Renovation (Bar Chart)	High	Reveals renovation trends
House Age by Renovation (Pie Chart)	Medium	Helps segment property types
Age by Features (Grouped Bar Chart)	High	Shows trends across features
Use of Colors and Legends	Medium	Improves readability

Ideation Phase
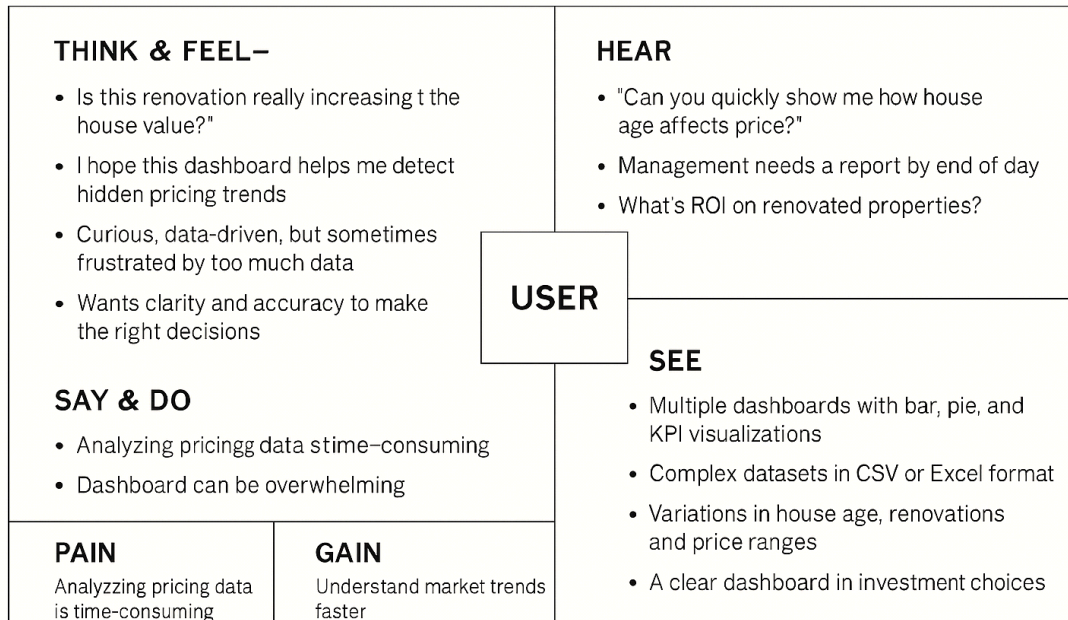## Empathize & Discover

Empathy Map Canvas:

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviours and attitudes.

It is a useful tool to helps teams better understand their users.
Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.

EMPATHY MAP CANVAS

THINK & FEEL–

- Is this renovation really increasing t the house value?"
- I hope this dashboard helps me detect hidden pricing trends
- Curious, data-driven, but sometimes frustrated by too much data
- Wants clarity and accuracy to make the right decisions

USER

HEAR

- "Can you quickly show me how house age affects price?"
- Management needs a report by end of day
- What's ROI on renovated properties?

SEE

- Multiple dashboards with bar, pie, and KPI visualizations
- Complex datasets in CSV or Excel format
- Variations in house age, renovations and price ranges
- A clear dashboard in investment choices

SAY & DO

- Analyzing pricingg data stime–consuming
- Dashboard can be overwhelming

PAIN	GAIN
Analyzzing pricing data is time-consuming	Understand market trends faster

Ideation Phase
Define the Problem Statements

Customer Problem Statement Template:

Create a problem statement to understand your customer's point of view. The Customer Problem Statement template helps you focus on what matters to create experiences people will love.

A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

Problem Statement (PS)	I am (Customer)	I'm trying to	But	Because	Which makes me feel
PS-1	a real estate analyst who works with large housing datasets.	identify how renovation and property features (like bathrooms, floors, bedrooms) affect house sale prices.	it's hard to spot trends in raw CSV or Excel files.	the data is too complex to interpret without visual tools or summaries.	frustrated, time-constrained, and unsure whether I'm making the right insights.
PS-2	a decision-maker at a real estate firm focused on pricing strategies.	compare houses by age, renovation status, and feature count to make informed pricing decisions.	I don't have an easy way to visualize these comparisons quickly.	traditional data formats lack interactive views and filtering options.	uncertain about pricing and hesitant to act without clear insight

Customer Journey Map – Housing Market Trends

SCENARIO:

Using a Tableau dashboard to analyze housing sale prices based on age, renovation history, and features.

ENTICE

How does someone become aware of the process?

The user receives a large dataset and a request from management to generate insights on market trends.

Goal: Help me understand what the data can reveal about pricing and renovation value.

Positive Moment: Opportunity to work with visual tools instead of Excel.

Negative Moment: Feels unsure where to begin with so many fields and filters.

Opportunity: Provide a sample dashboard or story outline to follow.

ENTER

What happens when they begin the analysis process?

They open Tableau and load the dataset. They scan the available fields: sale price, renovation year, basement area, etc.

Interactions: Tableau UI, data source connection.

People: Working alone or reporting to a manager.

Places: Workplace or remote workstation.

Goal: Help me load and prep the data without confusion.

Positive Moment: Seeing data load successfully and field names appear.

Negative Moment: Overwhelm from too many columns or unclear field names.

Opportunity: Provide a "data field dictionary" with short descriptions.

ENGAGE

What happens in the core interaction moments?

User creates visualizations like KPI cards, bar charts, and pie charts to find trends.

Interactions: Filtering by "Years Since Renovation", using SUM(Sale_Price), dragging fields into columns and rows.

Goal: Help me extract the right patterns and understand relationships visually.

Positive Moments: Seeing how recent renovations lead to higher sale prices.

Creating a pie chart that cleanly separates house age groups.

Negative Moments:Trouble grouping years logically (e.g., renovation bins).

Difficulty comparing across multiple filters.

Opportunity:Pre-built templates for renovation charts.

Use grouped calculations for bins (e.g., 0–5 years, 6–10).

EXIT

What happens when the dashboard is complete?

User exports visuals, compiles them into a Word or PDF report.

Goal: Help me clearly communicate my findings to others.

Positive Moment: Final visuals look polished and insightful.

Negative Moment: Copy-pasting charts to WordPad or PowerPoint can be clunky.

Opportunity: Offer export options or dashboard story layouts with auto-caption.

EXTEND

What happens after the experience is over?

User receives feedback or requests follow-ups. The dashboard may be reused or updated.

Goal: Help me refine or build on this dashboard in future cases.

Positive Moment: Dashboard becomes a repeatable tool for future analysis.

Negative Moment: Data gets outdated or structure changes, requiring rework.

Opportunity: Allow users to plug in updated data with same dashboard layout.

Project Design Phase-II
Data Flow Diagram & User Stories

Data Flow Diagrams:

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.
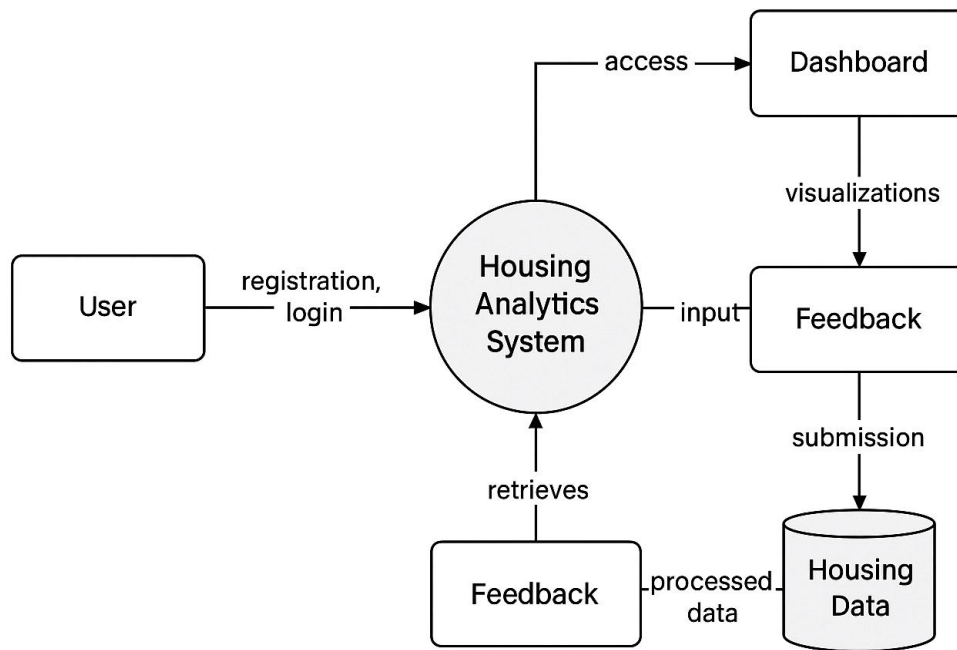
User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1

		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
Customer (Web user)	Registration	USN-1	As a user, I can register by entering my email, password, and confirming password	I can access my dashboard after signup	High	Sprint-1
		USN-2			High	Sprint-1
		USN-3	As a user, I will receive a confirmation email after registration	I receive an email and can confirm my account	Low	Sprint-2
			As a user, I can register using LinkedIn I log in successfully via LinkedIn	I log in successfully via LinkedIn		
Administrator	User Access Control	USN-1	As an admin, I can assign roles (Viewer, Analyst, Admin)	Users have access based on their roles	High	Sprint-2
	Report Management	USN-2	As an admin, I can view and export usage reports	Admin can download PDF of dashboard insight	Medium	Sprint-3

Project Design Phase-II
Solution Requirements (Functional & Non-functional)

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIN
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Dashboard Access & Visualization	Access user-specific dashboard View KPIs, charts, and summaries Filter data by renovation year, price range, features
FR-4	Report Export & Sharing	Export dashboard views as PDF or image Share reports via email/download link
FR-5	User Access & Role Management	Assign roles (Admin, Analyst, Viewer) Set access permissions for features and data visibility Login/logout/session control
FR-6	Feedback Collection	Allow users to submit feedback on dashboard usability Include rating system (stars or thumbs up/down) Capture user suggestions for future features

Dashboard Access & Visualization

Non-functional Requirements:

Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	Intuitive Tableau dahboard with clear filter and legends
NFR-2	**Security**	Role-based access and secure database connection
NFR-3	**Reliability**	System handles unepected data formats
NFR-4	**Performance**	Fast load times and responsiveness of visualizations
NFR-5	**Availability**	Dashboard is accessible across devices ith minimial downtime
NFR-6	**Scalability**	supports increasing data volume and new features

Project Design Phase-II
Technology Stack (Architecture & Stack)

Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table
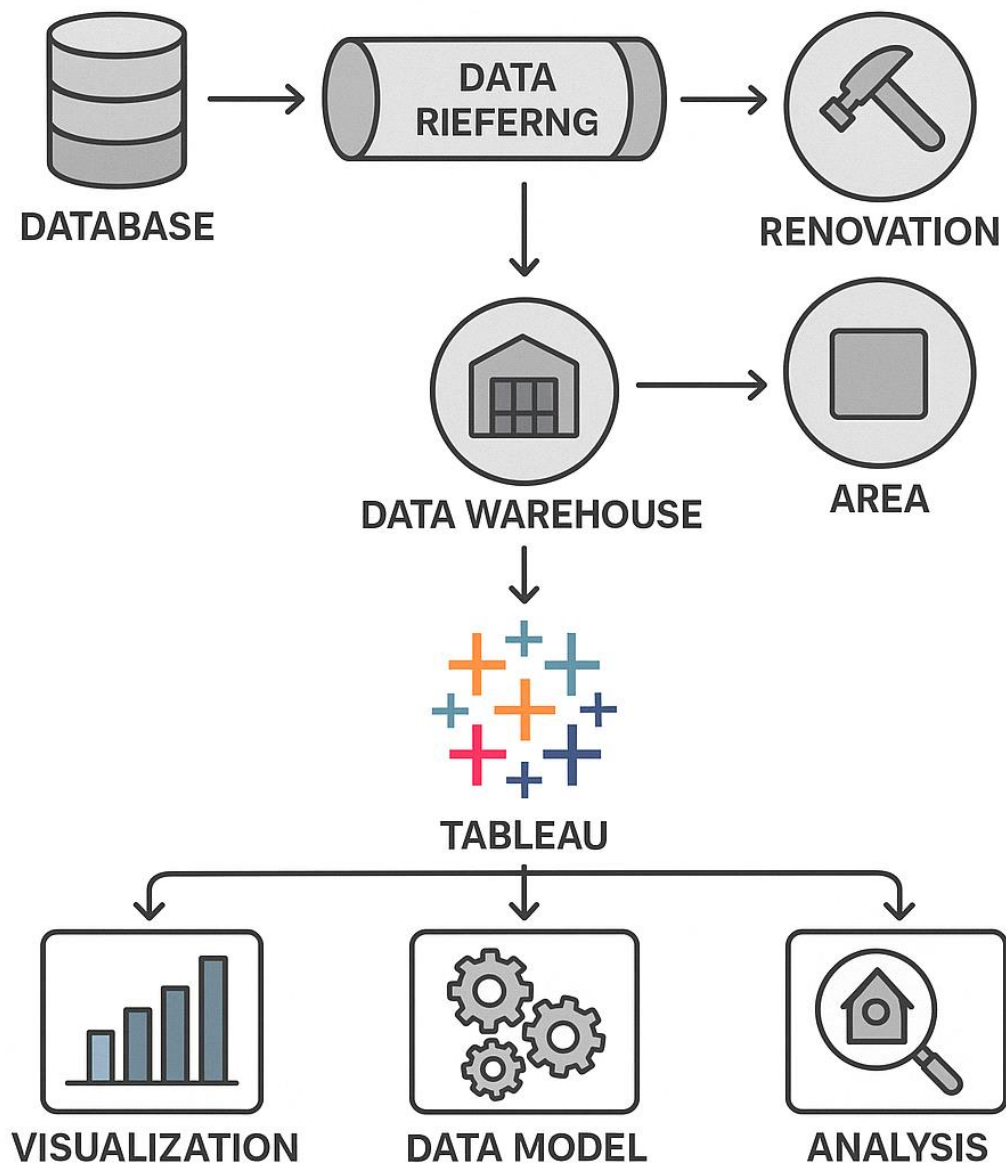
Table-1 : Components & Technologies:

S.No	Component	Description	Technology
•	User Interface	How users interact with the dashboard or web app	How users interact with the dashboard or web app

•	Application Logic-1	Logic for data transformation	Python (e.g., using pandas, numpy)
•	Application Logic-2	Data pipeline and preprocessing logic	ETL tools like Apache NiFi / Talend / IBM DataStage
•	Application Logic-3	Business rules implementation (e.g., filters, KPIs)	Tableau Calculated Fields / Parameters
•	Database	Raw data storage from multiple sources	MySQL / PostgreSQL / MongoDB
•	Cloud Database	Cloud-based storage for processed datasets	IBM Cloudant / IBM DB2
•	File Storage	Historical and large dataset storage	IBM Block Storage / AWS S3 / Local Filesystem
•	External API-1	Real-time external weather data	IBM Weather API
•	External API-2	Government registry verification or housing IDs	Aadhar API / RERA API
•	Machine Learning Model	Estimate house price or predict sales trends	Regression Model, Time Series Forecasting (Prophet)
•	Infrastructure (Server / Cloud)	Deployment platform for application and Tableau Server	Local, IBM Cloud Foundry, Kubernetes.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
•		Frameworks used in development and analysis	Python, Scikit-learn, ReactJS, Flask
•	Security Implementations	Security controls for database and API	SHA-256, HTTPS, OAuth2, IAM, JWT, Firewall Rules
•	Scalable Architecture	Modular architecture to support future features	Microservices, 3-Tier Architecture, Docker
•	Availability	Ensuring 24/7 availability of data services and dashboards	Load Balancers, Multi-Region Deployment, Redundancy
•	Performance	Handles multiple users and large data efficiently	Caching, Indexing, CDN, Tableau Extracts, Redis

Project Design Phase

Problem – Solution Fit Templat

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.**

Template:

Real estate analysts and stakeholders lack clear, visual insight into how features like renovation, house age, and structure impact sale prices.Project Design Phase Proposed Solution Template

Proposed Solution Template:

Project team shall fill the following information in the proposed solution template.

S.No.	Parameter	Description
•	Problem Statement (Problem to be solved)	Real estate analysts and stakeholders lack clear, visual insight into how features like renovation, house age, and structure impact sale prices.
•	Idea / Solution description	Build an interactive Tableau dashboard that visualizes house pricing trends based on key attributes (renovation year, basement size, no. of bedrooms, etc.)
•	Novelty / Uniqueness	Combines renovation impact, feature distribution, and price analytics into a single visualization suite. Uses KPIs, pie charts, and grouped bar charts.
•	Social Impact / Customer Satisfaction	Helps users (buyers, sellers, analysts) make data-informed decisions. Improves transparency in housing market trends.
•	Business Model (Revenue Model)	Can be offered as a subscription-based analytics platform for real estate companies, or embedded in CRM systems as a premium dashboard feature.
•	Scalability of the Solution	Can easily scale to different cities, more house attributes, or integrate APIs with live property data. Suitable for enterprise-wide rollout.

Project Design Phase

Solution Architecture:

Users

Frontend

Backend Services

Data Ingestion Service

Trend Analysis Engine

Reporting & Visualization Enggine

Authentication & Access Control

Storage

Object Storage

Relational Database

Social/News Sentiment

Processed Analytics Data

EXTERNAL DATA SOURCES

MLS / Zillow/ Redfin APIs

Government Datasets

Social Media News Feeds

IoT Sensors

:
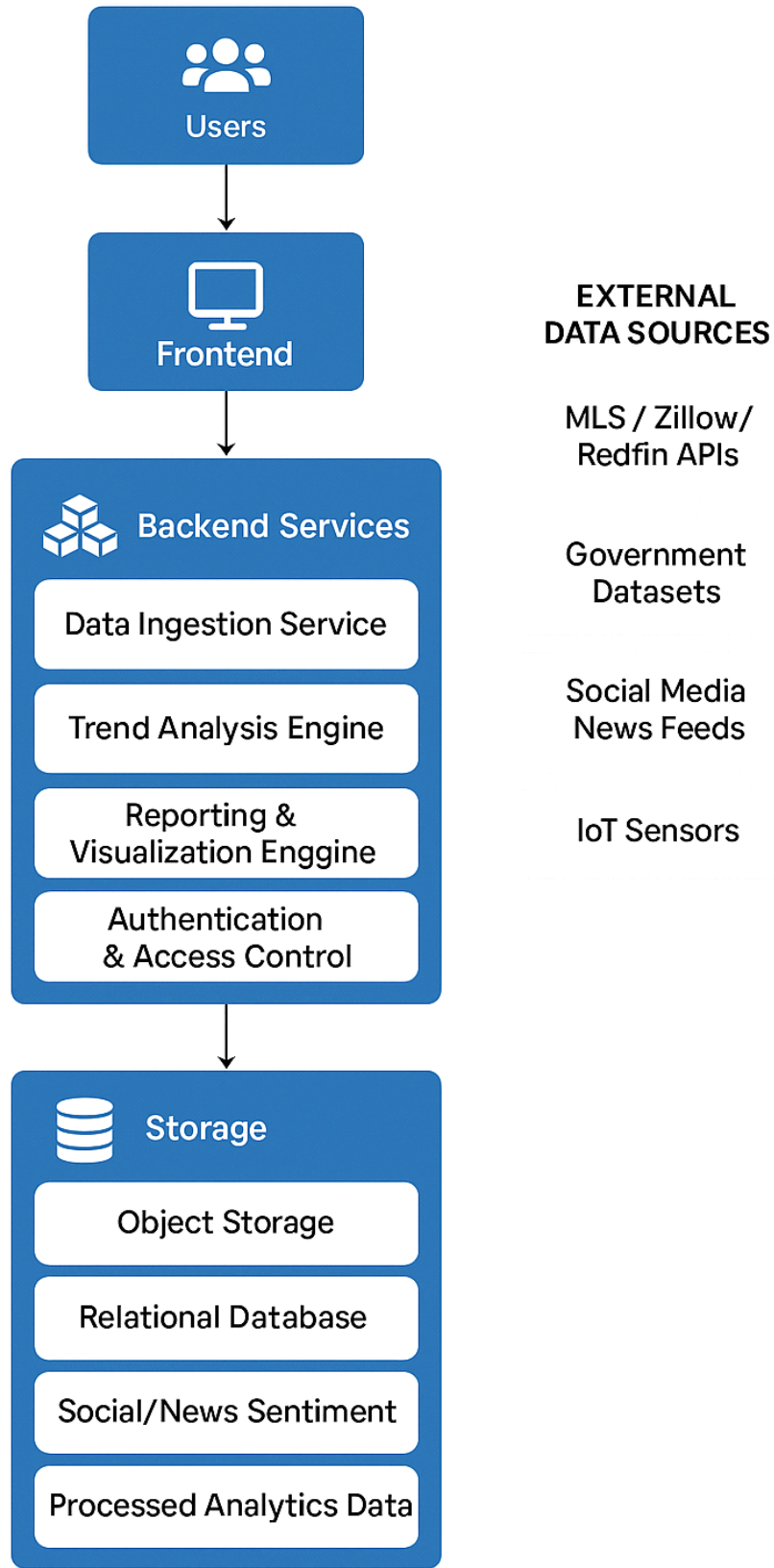
🏠 1. Users

Who: Analysts, real estate buyers, agents, and researchers.

What they do: Interact with the system via a dashboard or web app to view house trend analytics.

📋 2. Frontend

Technology: Web dashboard built using tools like React, Tableau, or Power BI.

Purpose: Provides interactive charts, filters, maps, and reports to visualize house trends, price forecasts, location insights, etc.

⬜ 3. Backend Services

These core services process, analyze, and serve data:

Data Ingestion Service: Collects housing data from APIs (like Zillow), CSV files, or government sources.

Trend Analysis Engine: Applies statistical and ML models to identify pricing patterns, demand trends, and area growth.

Reporting & Visualization Engine: Generates graphs and summaries for frontend dashboards.

Authentication & Access Control: Secures user access to sensitive data and analytics features.

🗄 4. Storage Layer

Stores raw and processed data used for analysis:

Object Storage: Holds raw files like CSVs, API dumps, and model outputs (e.g., AWS S3).

Relational Database: Stores structured data like sales records, location info (e.g., PostgreSQL).

Sentiment Data: Optional storage for news/social media sentiment if included.

Processed Data: Cleansed and aggregated data for visualization and trend modeling.

🌐 5. External Data Sources

Real Estate APIs: Zillow, Redfin, MLS for listings and price data.

Government Datasets: Demographics, land use, crime stats, etc.

Social Media/News Feeds: For gauging market sentiment.

IoT Sensors: Optional, for smart home data like energy use, temperature trends.

✅ Purpose of Architecture

Enables scalable and automated collection and analysis of house trend data.

Supports real-time insights and historical comparisons.

Can be extended with machine learning models for price prediction and investment scoring.

Project Planning Phase

Project Planning Template (Product Backlog, Sprint Planning, Stories, Story points)

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	5	High	B. Sarala
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	4	High	B. Sarala
Sprint-2		USN-3	As a user, I can register for the application through Facebook	5	high	B. Sarala
Sprint-1		USN-4	As a user, I can register for the application through Gmail	5	Medium	B. Sarala
Sprint-1	Login	USN-5	As a user, I can log into the application by entering email & password	5	High	B. Sarala
Sprint-2	Dashboard	USN-6	As a user, I can see a dashboard after logging in	5	High	B. Sarala
Sprint-3	Analytics	USN-7	As a user, I can view recent trends and stats related to houses on the dashboard	4	Medium	B. Sarala
Sprint-3	Analytics	USN-8	As a user, I can see visual charts and graphs about housing trends	5	High	B. Sarala
Sprint-4	Prediction	USN-9	As a user, I can get predicted prices based on trends	5	High	B. Sarala
Sprint-4	Prediction	USN-10	As a user, I receive	3	Medium	B. Sarala

			email notifications for important changes			

Project Tracker, Velocity & Burndown Chart: (4 Marks)

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	19	5Days	11 june2025	15 june2025	19	15 june2025
Sprint-2	10	5 Days	16 june2025	20 june2025	10	20 june2025
Sprint-3	9	5 Days	21 june2025	25 june2025	9	25 june2025
Sprint-4	8	5 Days	26 june2025	30 june2025	8	30 june2025

Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

AV=Total story point/Total Days=46/20=2.3
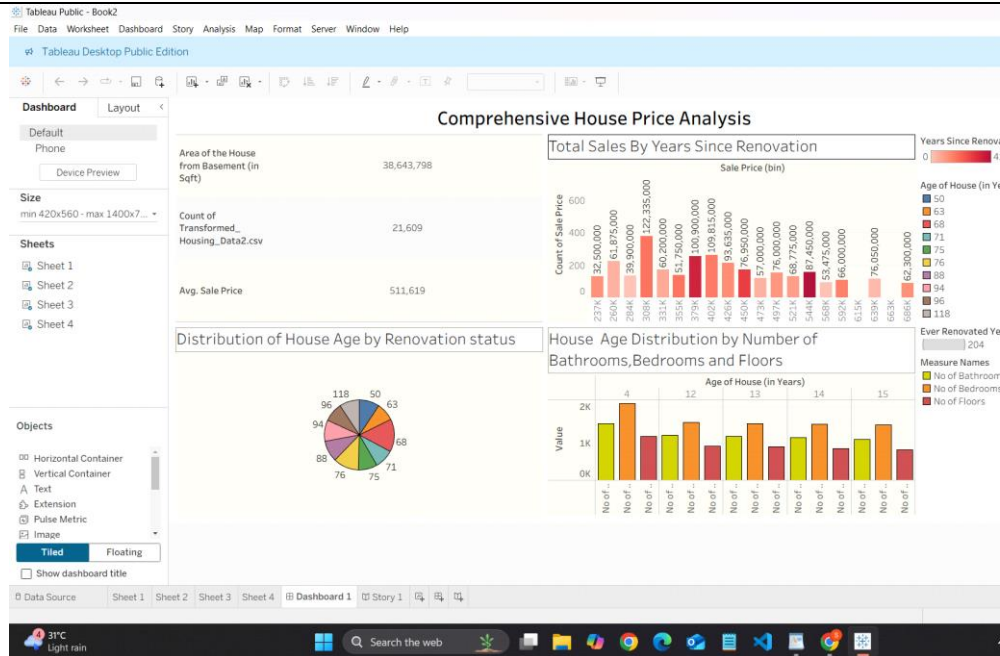
Project Development Phase
Model Performance Test

Date	10 February 2025
Team ID	LTVIP2025TMID50177
Project Name	Visualizing Housing Market Trends:An Analysis Of Sales prices And Features Using Tableau
Maximum Marks	

Model Performance Testing:

Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Screenshot / Values
•	Data	~10,000 house sales records across multiple cities

	Rendered	Fields: Sale Price, Renovation Year, Area (sqft), Bedrooms, Bathrooms, Location, Date
•	Data Preprocessing	- Removed nulls in critical fields (e.g. Sale Price) - Converted Renovation Year to categorical variable - Grouped Area into ranges - Extracted month/year from sale date
3	Utilization of Filters	- Branch Filter - House Type - Area Range - Year of Sale - Renovated (Yes/No)
4	Calculation fields Used	IF(renovated)=1 THEN "renovated" ELSE "Not renovated" END
5	Dashboard design	

6	Story Design	

7. RESULTS

8. ADVANTAGES & DISADVANTAGES

<u>Advantages:</u>

Easy to interpret visuals

Interactive filtering
Scalable design
<u>Disadvantages:</u>
Requires Tableau knowledge
Limited without internet/cloud

9. CONCLUSION

The dashboard successfully helps stakeholders gain deeper insights into house sales trends, improving their ability to make data-driven decisions.

10. FUTURE SCOPE

Add predictive modeling (ML)
Integrate live data feeds
Mobile app version
11. APPENDIX:
Dataset Link: Visualizing Housing Market Trends An Analysis of Sale Prices and Features using Tableau
https://www.kaggle.com/datasets/rituparnaghosh18/transformed-housing-data-2
GitHub & Demo Video:
https://drive.google.com/file/d/1KCLcYaAKiy-TuC4ZP5MiqAl8ETUhyUTL/view?usp=drivesdk