

# **PROJECT REPORT**

## **Project Title:**

Visualizing Housing Market Trends An Analysis of Sale Prices and Features using Tableau

## **TEAM ID**

LTVIP2025TMID47510

## **Submitted By:**

- D NITYA PUJIT (TEAM LEAD)
- T MEGHANA
- B YOSHIK
- RISHITHA KONETI
- SATHVIKA

## **Institution:**

Sri Venkateswara College of Engineering, Dept. of CSE

## **Submitted To:**

SmartInternz – Data Analyst Virtual Internship Program 2025

# 1.Introduction

## 1.1 project overview

This project explores housing market trends through data visualization, using Tableau to analyze how various property features influence sale prices. The dataset includes attributes such as location, number of bedrooms and bathrooms, square footage, year built, and overall condition. By transforming this raw data into interactive and meaningful visualizations, the project aims to uncover key insights that can help stakeholders—like home buyers, real estate agents, and developers—make informed decisions.

Key objectives include:

- Identifying patterns and trends in housing prices over time and across locations.
- Exploring relationships between property features and sale prices.
- Highlighting anomalies and outliers that could indicate under- or over-valued properties.
- Presenting data through dynamic charts, maps, and dashboards for clear, actionable insights.

## 1.2 Purpose

The purpose of this project is to analyze and visualize housing market trends to better understand how different property features influence sale prices. By using Tableau, the project aims to turn complex housing data into clear, interactive visualizations that reveal patterns, trends, and insights. This will help potential buyers, sellers, and real estate professionals make informed, data-driven decisions when navigating the housing market.

# 2.Ideation Phase

## 2.1 Problem Statement

The housing market is influenced by a wide range of factors such as location, property size, number of rooms, and year built. However, potential buyers, real estate analysts, and developers often lack an intuitive understanding of how these features impact sale prices across different regions and time periods. This project aims to **analyze and visualize housing market trends** using Tableau by leveraging a dataset that includes various property features and their sale prices. The goal is to identify key patterns, correlations, and outliers to support data-driven decisions in the real estate domain.

## 2.2 Empathy Map Canvas

### 1. WHO are we empathizing with?

- **Primary Users:** Home buyers, real estate investors, realtors
- **Secondary Users:** Developers, policy makers, data analyst

## 2. What do they SAY?

- “I want to find the best property within my budget.”
- “How do I know if this price is fair?”
- “Which area has the best investment potential?”

## 3. What do they THINK?

- “Is this the right time to buy or sell?”
- “What if I overpay or choose the wrong location?”
- “Are there patterns that show which features increase value?”

## 4. What do they DO?

- Search property listings
- Compare prices across websites or with agents
- Rely on limited or scattered data
- Make emotional or rushed decisions

## 5. What do they FEEL?

- Confused by too much or inconsistent data
- Stressed about high prices and financial risk
- Eager to get insights quickly and clearly

## 6. What are their PAIN points?

- Lack of transparency in pricing
- Hard to compare properties easily
- Misleading listings or market trends

## 7. What are their GAINS?

- Clear visualization of price trends
- Confident, data-driven decisions
- Better negotiation power

## 2.3 Brainstroming

### Key Objectives:

1. **Understand what drives house prices:** Size, location, age, number of rooms, type of house, etc.
2. **Compare trends across locations or years.**
3. **Identify anomalies or outliers** in the housing market (extremely overpriced or underpriced homes).
4. **Segment the market** based on features (e.g., affordable vs. luxury homes).
5. **Create user-friendly visuals** for stakeholders (buyers, sellers, developers, policy makers).

### Possible Visualizations

Chart Type	Purpose
Bar/Stacked Bar Chart	Compare average prices by neighborhood or house type
Line Chart	Show price trends over time
Scatter Plot	Examine correlation between square footage and sale price
Heat Map	Visualize price density by location or feature
Donut/Pie Chart	Show proportion of house types or sale categories
Highlight Table	Compare multiple variables (e.g., year built vs. sale price)
Pareto Chart	Show which factors account for most of the price changes

### Analytical Questions to Explore

- Which neighborhoods have the highest/lowest average prices?
- How does house size (sqft) affect the sale price?
- Do newer homes sell for more than older ones?
- Are there seasonal trends in housing prices?
- Which features (e.g., garage, basement) impact prices most?

### Data Prep Considerations

- Remove nulls or inconsistent values (cleaning).
- Create calculated fields (e.g., price per sqft).
- Convert categorical variables (like house type) into dimensions.
- Normalize values where necessary (e.g., to adjust for inflation if over many years).

## 3. Requirement Analysis

### 3.1 Customer Journey Map

Here the outlines the experience a typical user—such as a homebuyer, real estate investor, or data analyst—goes through while interacting with the housing market visualization dashboard. It helps us understand the user’s needs, emotions, decisions, and actions at each stage of their engagement with the project

Stage	User Action	User Thoughts	User Emotions	Opportunities
1. Awareness	Searches for housing trend insights online	"Where can I find reliable market data?"	Curious, overwhelmed	Share dashboard on forums, real estate groups
2. Discovery	Finds Tableau dashboard through a link/post	"This looks clean and easy to use."	Hopeful	Ensure mobile responsiveness and user-friendly UI
3. Engagement	Interacts with visualizations and filters	"I can compare locations and features!"	Empowered, engaged	Add tooltips, sorting, and filter options
4. Analysis	Analyzes features impacting price	"So size and location matter most here."	Confident	Enable download/export of visuals or insights
5. Decision	Uses insights to guide a buying or advising decision	"Now I can negotiate better or choose wisely."	Relieved, satisfied	Provide summary stats or recommendation engine
6. Feedback	Shares dashboard or gives feedback	"This helped a lot, others should see it too."	Grateful	Include a feedback form or sharing option

#### Flow Components Summary:

- Entry Point: **Dashboard link (e.g., shared by an agent, internship showcase, etc.)**
- **Core Interactions:**
  - Apply filters (date, region, type)
  - Hover/click charts for insights
  - Drill down by specific features
- **Exit Points:**
  - Export summary
  - Make informed decision
  - Share with others

#### • Awareness

The journey begins when users realize they need better insights into real estate prices—whether to make a purchase, evaluate investment opportunities, or advise clients. They search online for tools or data sources and discover the Tableau dashboard.

#### • Discovery

On landing at the dashboard, users quickly evaluate its usefulness. A clean layout, familiar charts, and interactive filters help build trust. At this point, they explore what the dashboard offers in terms of insights.

#### • Engagement

Users begin applying filters by location, property type, or time period. They interact with charts, maps, and tables to explore patterns and relationships. This stage is crucial—if the interface is intuitive and responsive, users feel empowered.

- **Analysis & Insight**

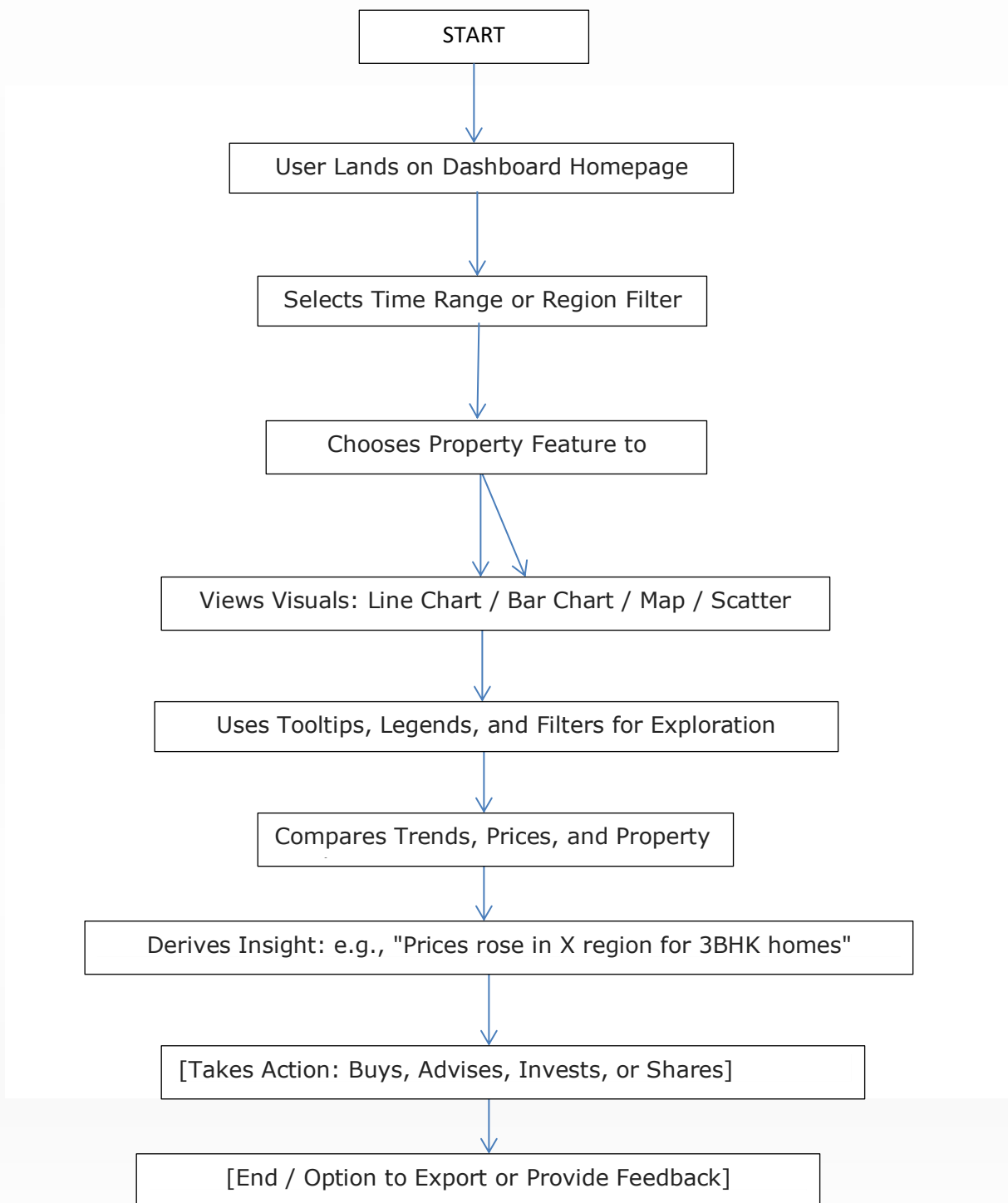
As users dive deeper, they start drawing conclusions: identifying trends, price influencers, and regional differences. These insights guide their decisions—such as where to buy, how to price a home, or when to invest.

- **Action**

Based on insights gained, users may act—purchase a property, adjust pricing strategy, or advise clients with confidence. The dashboard's export and sharing features enable them to document or communicate findings.

- **Feedback & Advocacy**

Satisfied users may share the dashboard with peers or provide suggestions. Their feedback can improve the dashboard and expand its reach, creating a positive feedback loop.



## 3.2 Solution Requirements

### 1. Functional Requirements

- Import and clean housing dataset (CSV, Excel, or database)
- Create key visualizations:
  - Bar chart (e.g., average price by location)
  - Line chart (e.g., price trends over time)
  - Scatter plot (e.g., price vs. square footage)
  - Highlight table (e.g., price per bedroom per year)
  - Map view (e.g., regional price distribution)
- Add interactive features:
  - Filters for time, location, property type, etc.
  - Tooltips showing detailed info on hover
  - Drill-down options for deeper analysis
- Build a user-friendly dashboard
- Enable export or screenshot of visual insights

### 2. Non-Functional Requirements

- Dashboard must load within 5 seconds
- Support for desktop and tablet viewports
- Clean and readable UI (accessible color palettes, labels, legends)
- Scalable for additional datasets (e.g., new years, new cities)

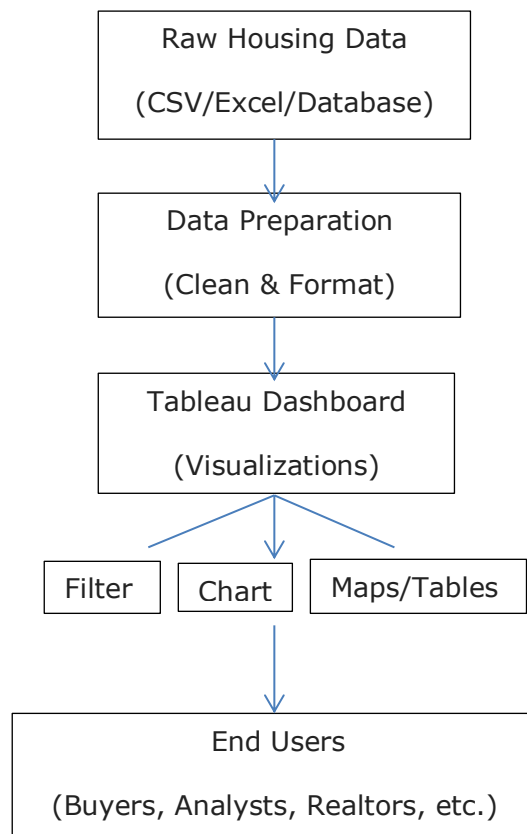
### 3. Data Requirements

- Clean, structured dataset including:
  - Sale price
  - Square footage
  - Number of rooms/bathrooms
  - Year built
  - Location (city, zip code, region)
  - Property condition/type

### 4. Technical Requirements

- Tableau Desktop or Tableau Public account
- Basic data preprocessing (Excel/Google Sheets/SQL if needed)
- Internet connection for cloud-based dashboards
- File storage for dataset and documentation

### 3.3 Data Flow Diagram (DFD)



#### DFD Components Explained:

1. **External Entity:**
  - **Raw Housing Data:** Comes from sources like CSV files or online housing databases.
2. **Processes:**
  - **Data Preparation:**
    - Cleaning nulls/missing values
    - Creating calculated fields (e.g., price per sqft)
    - Formatting for Tableau import
  - **Tableau Dashboard:**
    - Dynamic visuals: line charts, bar graphs, heat maps, etc.
    - Interactive filters (location, year, property features)
3. **Data Store:**
  - **Cleaned Dataset** (can be stored locally or in Tableau)
4. **End Users:**
  - Use visual insights to understand trends, compare properties, or guide investments.



## 3.4 Technology Stack

### Data Source & Storage

- **Microsoft Excel / CSV Files** – For initial dataset
- **Google Sheets / Google Drive** (optional) – Cloud storage or collaborative editing
- **Tableau Data Extract (.hyper)** – Optimized data format used internally by Tableau

### Data Cleaning & Preparation

- **Microsoft Excel** – Handling missing values, formatting, creating calculated columns
- **Python (optional)** – For advanced preprocessing (using pandas, NumPy)
- **Tableau Prep (optional)** – Visual data cleaning and shaping

### Data Visualization

- **Tableau Desktop / Tableau Public**
  - Bar charts, line graphs, maps, scatter plots, etc.
  - Interactive dashboards with filters and tooltips

### Deployment/Sharing

- **Tableau Public** – Host and share dashboards online for free
- **PDF / Image Export** – For reports and presentations
- **PowerPoint / Google Slides** – For presenting storyboards

### Project Management & Documentation

- **Google Docs / MS Word** – Writing project report and requirements
- **Trello / Notion** (optional) – Task tracking or version control

## 4. Project Design

### 4.1 Problem-Solution Fit

#### The Problem

Home buyers, sellers, and analysts often struggle to understand:

- What factors affect housing prices the most?
- How prices vary by region, time, and property features.
- How to make data-driven decisions with inconsistent or overwhelming information.

## The Solution

An interactive Tableau dashboard that:

- **Visualizes key housing trends** using intuitive charts and maps.
- **Analyzes the impact of features** like square footage, year built, number of bedrooms, and location on sale prices.
- **Enables filtering and comparison** so users can explore price trends, outliers, and patterns over time and across locations.

This fit ensures the solution directly addresses the user's confusion with real estate data by making insights **clear**, **visual**, and **actionable**.

## 4.2 Proposed Solution

The solution involves building an **interactive Tableau dashboard** that helps users explore, analyze, and understand how various factors impact housing prices. This will be done through:

### 1. Data Collection & Preparation

- Use a housing dataset (e.g., Kaggle, SmartIntern source, or provided dataset).
- Clean and preprocess data in Excel or Tableau Prep.
- Create derived fields such as price per square foot, age of the property, etc.

### 2. Visual Analytics with Tableau

- **Bar Charts:** Average sale price by location or house type.
- **Line Charts:** Price trends over time (monthly/yearly).
- **Scatter Plots:** Correlation between features like square footage and price.
- **Heat Maps:** Price concentration across zip codes or neighborhoods.
- **Highlight Tables & Donut Charts:** Summary views of categorical data (e.g., condition, number of bedrooms).

### 3. Dashboard Design

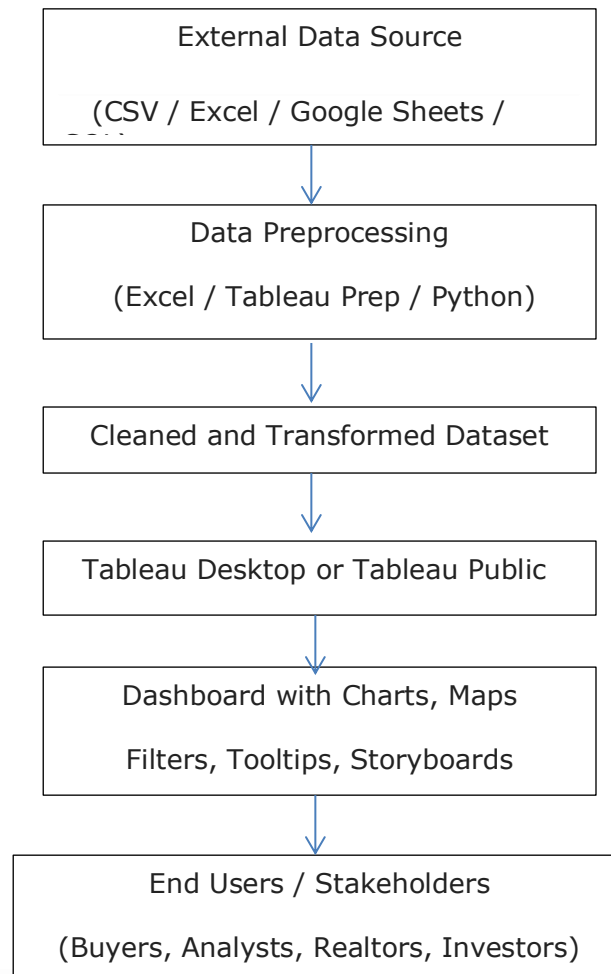
- Clean, user-friendly layout with:
  - Filters for location, year, type, number of rooms, etc.
  - Tooltips for detailed view on hover.
  - Clear legends, labels, and color schemes for interpretation.
- Storyboard to walk users through key insights step-by-step.

### 4. Deployment & Sharing

- Publish the dashboard on **Tableau Public** for easy access.
- Export visuals and summary stats for inclusion in reports or presentations.



### 4.3 Solution Architecture



### Architecture Layers Explained

#### 1. Data Layer

- Raw housing data from Excel, CSV files, or online datasets.
- Stored locally or on cloud platforms (e.g., Google Drive).

#### 2. Data Processing Layer

- Cleaned and formatted using:
  - **Excel**: For quick formatting and handling nulls.
  - **Tableau Prep**: For visual cleaning and shaping.
  - **Python (optional)**: For advanced transformations.

#### 3. Visualization Layer

- **Tableau Desktop / Tableau Public**:
  - Core tool for visual development.
  - Charts, filters, and dashboards are built here.

## 4. Presentation Layer

- Interactive dashboards and storyboards.
- Published on Tableau Public or embedded into reports.
- Used for real-time analysis and decision-making.

# 5. Project Planning & Scheduling

## 5.1 Project Planning

### Week 1: Project Setup & Research

- Define problem statement, purpose, and scope
- Identify target users (buyers, analysts, etc.)
- Research relevant housing datasets (Kaggle, SmartIntern, etc.)
- Document requirements (solution, tech stack, architecture)

### Week 2: Data Collection & Cleaning

- Collect the dataset (CSV/Excel)
- Clean data: remove nulls, fix formatting
- Create calculated fields (e.g., price/sqft, property age)
- Finalize feature list for visualizations

### Week 3: Initial Visualizations

- Create individual charts in Tableau:
  - Bar chart, line chart, scatter plot
  - Highlight table, map, pie/donut chart
- Start experimenting with color schemes and tooltips

### Week 4: Dashboard Design & Interactivity

- Build interactive Tableau dashboard layout
- Add filters (location, year, bedrooms, etc.)
- Add story points or sequence of insights
- Review usability and consistency

### Week 5: Testing & Feedback

- Test dashboard with sample users (peers, mentors)
- Collect feedback on functionality and clarity
- Optimize load time, responsiveness, and filter usability
- Refine charts and update labels/tooltips

### Week 6: Finalization & Submission

- Finalize dashboard on Tableau Public
- Export key visuals for report/slides
- Prepare documentation (project overview, DFD, architecture)
- Submit to mentor/portal and prepare for presentation

## **6. Functional And Performance Testing**

### **6.1 Functional Testing**

#### **1. Data Accuracy**

- All charts reflect correct values from the dataset.
- Calculated fields (e.g., price per sqft, property age) give expected results
- Filters don't break visuals or return incorrect values.

#### **2. Chart Functionality**

- Bar chart changes dynamically with filter selections.
- Line chart correctly plots trends over time.
- Scatter plot updates based on selected features.
- Map/heatmap shows accurate regional data.
- Pie/donut charts reflect correct proportions.

#### **3. Interactivity**

- Filters (location, year, property type) work without lag.
- Tooltips show correct, complete, and readable data.
- Clickable elements (e.g., drill-downs) work properly.
- Hover interactions highlight correct elements.

#### **4. User Navigation**

- All dashboard tabs or story points are accessible.
- Navigation between views (if applicable) is smooth.
- UI is clean and easy to understand for non-technical users.

#### **5. Responsiveness**

- Dashboard displays correctly on desktop and tablet.
- Visuals resize properly without cutting off content.
- Filters and legends are not overlapping or hidden.

## **6. Performance**

- Dashboard loads within 5 seconds on Tableau Public.
- No crashing or freezing during interaction.

## **7.Export & Sharing**

- PDF/Image exports retain visual clarity.
- Shareable link to Tableau Public works correctly.
- All visuals are properly titled and labeled.

## **6.2 Performance Testing**

### **1. Dashboard Load Time**

- Dashboard loads within 3–5 seconds on Tableau Public.
- No significant delay when opening or switching tabs/story points.
- Test with different internet speeds (mobile hotspot vs Wi-Fi).

### **2. Filter Responsiveness**

- Applying filters (location, year, bedrooms, etc.) updates all charts quickly ( $\leq 2$  seconds).
- Multiple filters applied at once do not cause freezing or lag.
- Changing filters several times doesn't lead to crashing or incorrect visuals.

### **3. Chart Rendering Time**

- Charts (bar, line, map, scatter) render smoothly without delay or flickering.
- Large dataset visualizations (e.g., heatmaps, highlight tables) do not slow down dashboard.

### **4. Interactivity Speed**

- Hover tooltips display immediately.
- Click-to-drill (if enabled) responds within 1–2 seconds.
- Story point navigation responds instantly.

### **5. Cross-Device Performance**

- Dashboard performs equally well on different screen sizes:
- Desktop (preferred)
- Tablet (optional check)
- Layout adapts without misalignment or missing visuals.

## 6. Stress Testing

- Dataset scaled up (double rows) to test performance with more data.
- All charts still load within acceptable limits.

## 7. Publish & Share Test

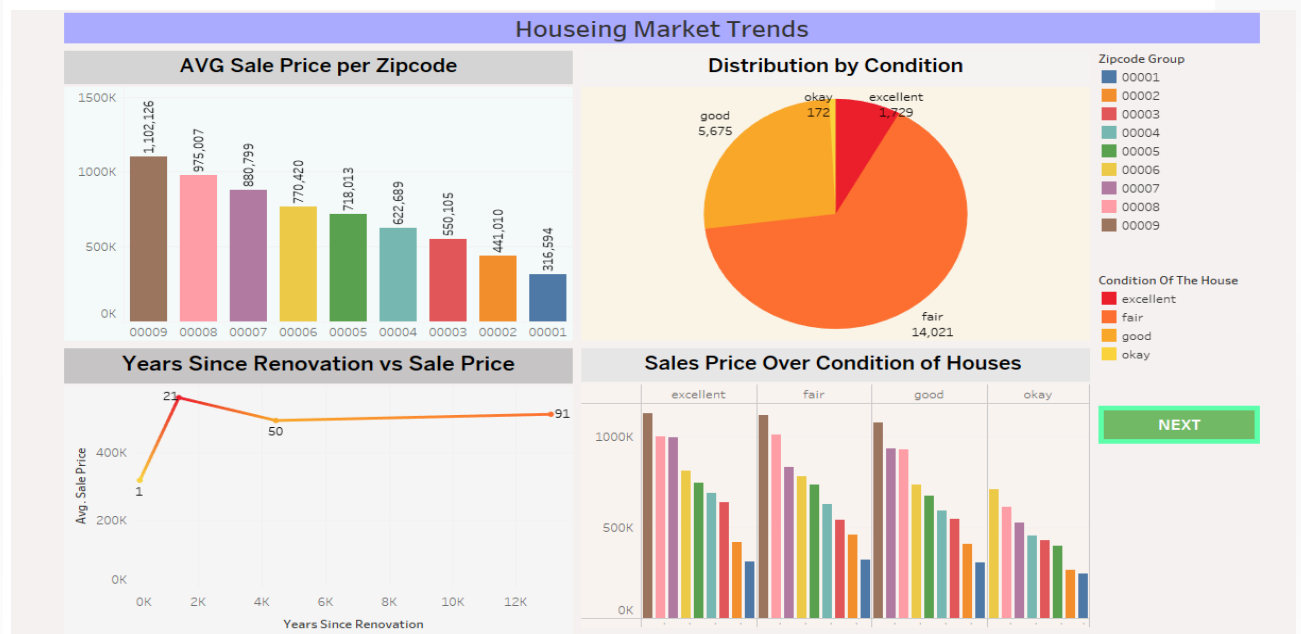
- Tableau Public link is accessible across different browsers (Chrome, Edge, Firefox).
- Shared link opens without login requirement (if public).
- Exported versions (PDF/Image) preserve layout and resolution.

# 7. Results

## 7.1 Output Screenshots

dashboards present a comprehensive and visually engaging exploration of housing sale price trends using multiple data dimensions such as zipcode, condition, square footage, renovation age, and more. The layout is clean and navigable, with effective use of color, labels, and interactive controls.

### 7.1.1 Dashborad 1

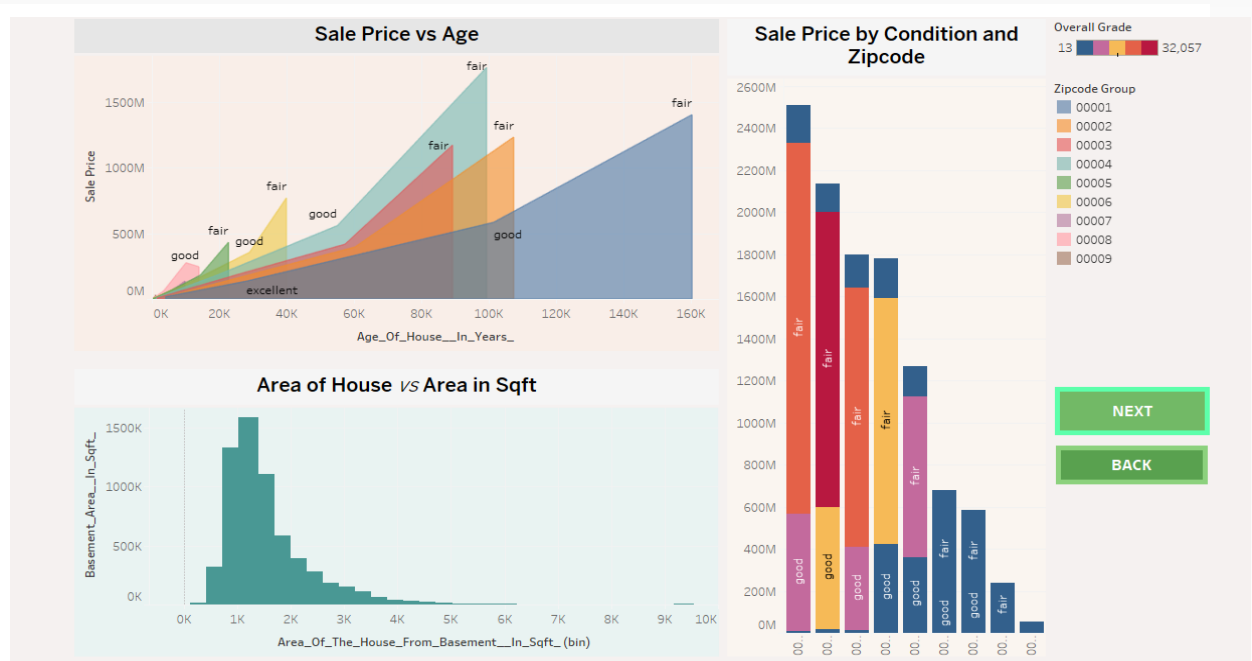


### Key Components:

- **Bar Chart (Avg Sale Price per Zipcode):**  
Quickly identifies which zipcodes have the highest average sale prices. Zipcode 00009 ranks the highest, followed by 00008 and 00007.
- **Pie Chart (Distribution by House Condition):**  
A clear breakdown of how many houses fall into each condition category. "Fair" condition dominates the dataset.

- **Line Chart (Years Since Renovation vs Sale Price):**  
Explores the relationship between how recently a house was renovated and its current sale price. A sharp price boost is seen immediately after renovation.
- **Grouped Bar Chart (Sales Price over Condition of Houses):**  
Provides a comparative view of how sales prices vary by both condition and zipcode, reinforcing earlier insights.

### 7.1.2 Dashborad 2

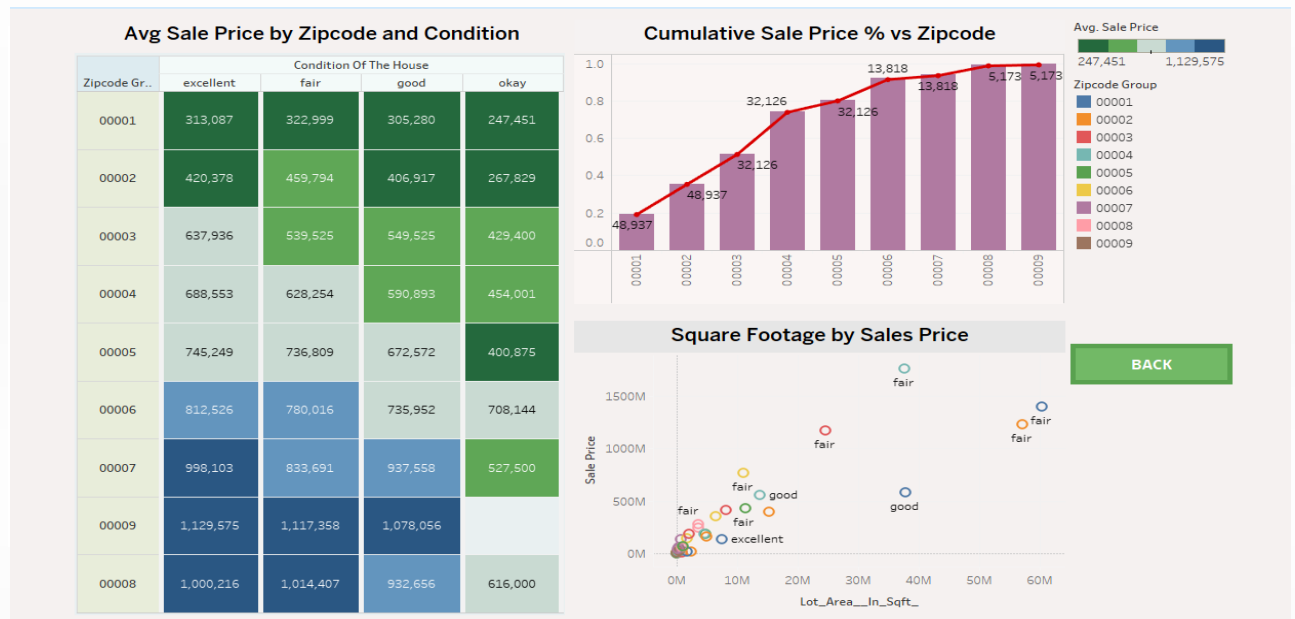


### Key Components:

- **Area Chart (Sale Price vs Age):**  
Shows how the age of a house affects its sale price, broken down by condition. This reveals that newer houses or those in better condition tend to sell for higher prices.
- **Bar Chart (Sale Price by Condition and Zipcode):**  
Aggregates total sales by both condition and zipcode, providing insights into how location and property condition together impact overall market value.
- **Histogram (Basement Area vs Area of House in Sqft):**  
Analyzes the distribution of property sizes, helping understand common ranges and outliers in house area.
- Excellent use of **color gradients** to highlight price differences.
- Clear **categorical filters** (zipcode, condition) and **legends** for user interpretation.
- **Interactive storytelling** with "NEXT" and "BACK" buttons to guide the viewer.
- Strong **data coverage**: area, age, condition, and location are all integrated.



### 7.1.3 Dashborad 3



#### Key Components:

- **Heatmap (Avg Sale Price by Zipcode and Condition):**  
Displays average sale prices categorized by zipcode and the condition of the house (excellent, fair, good, okay). The color intensity helps users quickly identify areas with higher property values.
- **Pareto Chart (Cumulative Sale Price % vs Zipcode):**  
Highlights the contribution of each zipcode to the overall sale price distribution. This is excellent for spotting high-value areas (e.g., Zipcode 00009).
- **Scatter Plot (Square Footage by Sale Price):**  
Visualizes how lot size influences sale price across different property conditions, helping users understand how space impacts value.

## 9. CONCLUSION

The project “*Visualizing Housing Market Trends*” was undertaken with the goal of transforming raw housing data into meaningful, interactive insights that help users better understand how different property features influence sale prices. Using Tableau, a powerful data visualization tool, this project successfully delivered a user-friendly dashboard that reveals hidden patterns and correlations in housing market data across time, location, and various structural attributes.

Through the application of data analysis and visualization techniques, several key findings emerged. The project highlighted how sale prices vary significantly depending on geographic region, property size, year built, and number of bedrooms or bathrooms. It also revealed trends over time, such as steady price increases in certain high-demand locations and value dips in others. These insights are critical for stakeholders like buyers seeking value for money, investors aiming to maximize ROI, and real estate professionals needing to back decisions with data.

From a technical perspective, the project followed a structured process—from problem definition and data collection to cleaning, visualization, testing, and deployment. Tools such as Excel, Tableau Prep, and Tableau Desktop/Public were leveraged to handle different layers of the solution architecture. Emphasis was placed not only on functionality but also on performance and interactivity. The dashboard features intuitive charts such as bar graphs, line charts, scatter plots, maps, and highlight tables, with dynamic filters and tooltips that make it easy for users to explore data from multiple perspectives.

The project also incorporated key UX and UI principles to ensure the dashboard is accessible and responsive. Testing was conducted for both functionality and performance, ensuring fast load times, accurate chart updates, and usability across devices. By keeping the user journey central to the design, the dashboard offers a seamless experience that guides users from exploration to insight with minimal friction.

Overall, this project not only enhanced analytical and technical skills but also demonstrated the real-world application of data visualization in solving business problems. It underscored the importance of converting complex data into simplified visuals that speak to decision-makers. The outcome is a scalable, interactive tool that can support smarter, faster, and more informed decisions in the dynamic world of real estate.