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

MENTOR: K.RATHNA KUMARI MEDAM

TEAM ID: LTVIP2025TMID52598

TEAM DETAILS:

**TEAM LEADER**:BADAMSETTI INDRANI



**REG**: SBAP0052598 **EMAIL ID**: indhubadhamsetti12@gmail.

TEAM MEMBERS:

1.GANDI ANUSHA



**REG**: SBAP0052602 **EMAIL ID**: [anushagandi321@gmail.com](mailto:anushagandi321@gmail.com)

2.BOTTA KRUPA



**REG:** SBAP0052600 **EMAILID:** bottakrupa142@gmail.com

3. ALTHI NAGA VENKATA DEVIKA



**REG**: SBAP0052596 **EMAILID:** [nvdevika47@gmail.com](mailto:nvdevika47@gmail.com)

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

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

In this project for ABC Company, the goal is to address challenges in understanding the factors that influence house prices and sales trends. By analyzing comprehensive housing data, including the total sales by years since renovation, house age distribution by the number of bathrooms, bedrooms, and floors, and the impact of renovations on house age, the company aims to uncover key insights. Utilizing Tableau for this analysis, the objective is to visualize and interpret patterns in the housing market to inform strategic decisions, optimize pricing strategies, and enhance overall market competitiveness. Key stakeholders include real estate analysts, marketing teams, and company executives, all of whom will benefit from a deeper understanding of housing market dynamics.

**Scenario 1: Overall Data Overview**

This visualization presents a summary of the dataset, showing the count of transformed housing data records, the average sales price, and the total area of houses from the basement in square feet. This overview provides a quick snapshot of the dataset's scale and key metrics, offering stakeholders a foundational understanding of the data being analyzed.

**Scenario 2: Total Sales by Years Since Renovation**

This histogram illustrates the distribution of total sales based on the number of years since a house was renovated. The bars represent different sales price bins, highlighting how recently renovated houses correlate with varying price ranges. This scenario helps stakeholders understand the impact of renovations on house prices and identify trends in buyer preferences regarding renovated homes.

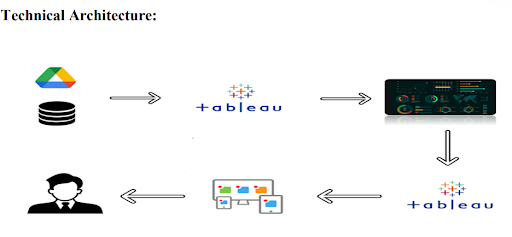
**Scenario 3: Distribution of House Age by Renovation Status**

This pie chart shows the distribution of houses based on their age and renovation status. Each segment of the pie represents a different age group, providing insight into how the age of houses is spread across the dataset and the proportion of houses that have been renovated versus those that have not. This visualization assists in assessing the age characteristics of the housing inventory and the prevalence of renovations.

**Scenario 4: House Age Distribution by Number of Bathrooms, Bedrooms, and Floors**

This grouped bar chart displays the distribution of house ages categorized by the number of bathrooms, bedrooms, and floors. It shows how houses of different ages are distributed according to these attributes, offering a detailed view of how house features vary with age. This scenario helps stakeholders identify patterns in housing characteristics and preferences related to house features over time.C

### Technical Architecture



### Project Flow

To accomplish this, we have to complete all the activities listed below,

? Data Collection & Extraction from Database

o Collect the dataset,

o Connect data with Tableau

? Data Preparation

o Prepare the Data for Visualization

? Data Visualizations

o No of Unique Visualizations

? Dashboard

o Responsive and Design of Dashboard

? Story

 o No of Scenes of Story

 ? Performance Testing

o Amount of Data Loaded

o Utilization of Data Filters

o No of Calculation Fields

o No of Visualizations/ Graphs

? Web Integration

o Dashboard and Story embed with UI With Flask

? Project Demonstration & Documentation

o Record explanation Video for project end to end solution

o Project Documentation-Step by step project development procedure

### Extraction from Database

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypotheses, evaluate outcomes and generate insights from the data.

### Downloading the dataset

Understand the data

Data contains all the meta information regarding the columns described in the CSV files

Column Description of the Dataset:

1. Overall Data Overview
2. Total Sales by Years Since Renovation
3. Distribution of House Age by Renovation Status
4. House Age Distribution by Number of Bathrooms,Bedrooms and Floors

### Data Preparation

Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete.

* **Explanation video links**

Links

**Data Visualization**

* + Data visualization is the process of creating graphical representations of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

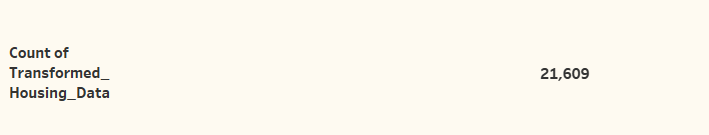
### No.of Unique Visualizations

Data visualization is the process of creating graphical representations of data to help people understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

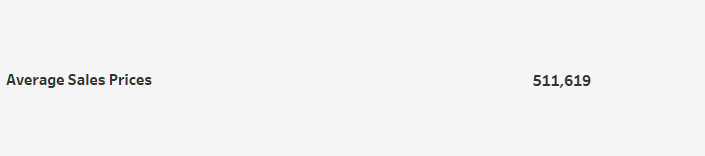
### Visualizations

Visualizations

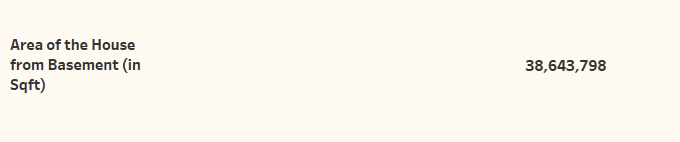
**Activity 1.1: Count of Transformed\_Housing\_Data**



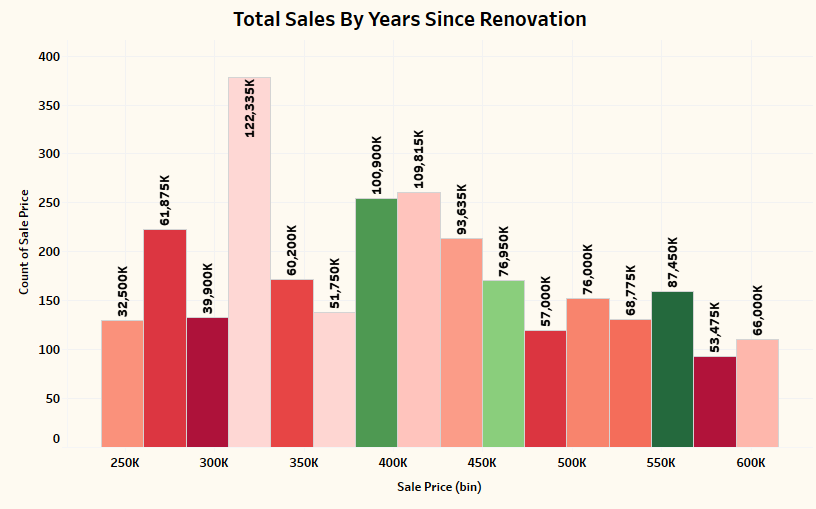
**Activity 1.2:  Average Sale Prices**



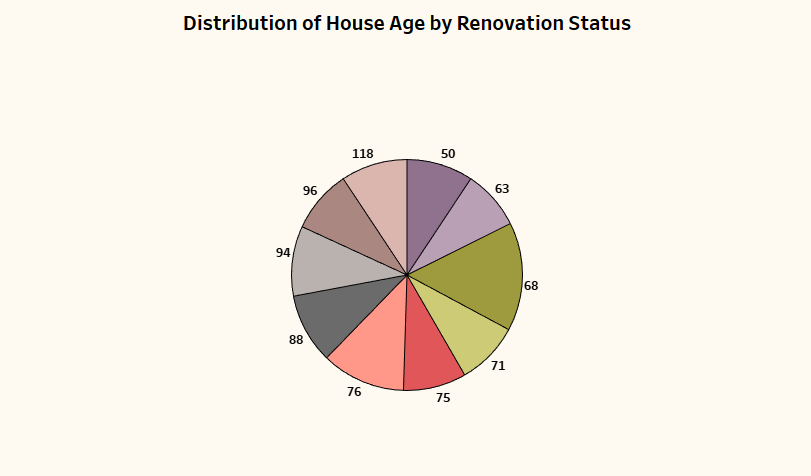
 **Activity 1.3: Area of House from Basement(in Sqft)**



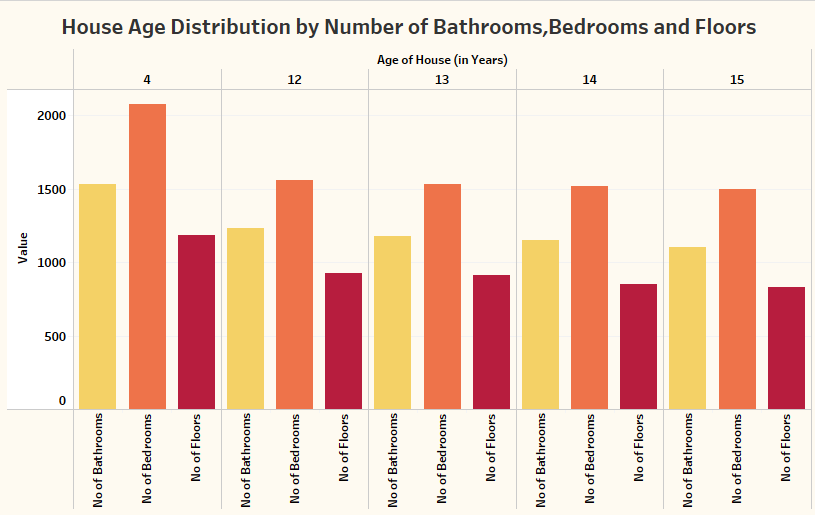
**Activity 1.4 Total Sales by Years Since Renovation**



**Activity 1.5 Distribution of House Age by Renovation Status**

****

**Activity 1.6 House Age Distribution by Number of Bathrooms, Bedrooms, and Floors**

****

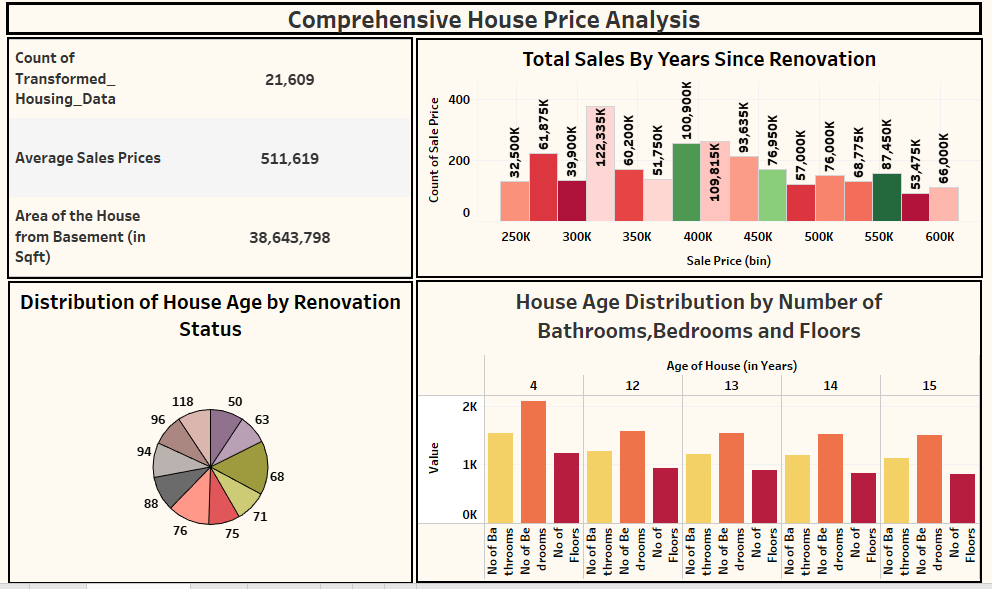
**VISUALIZATIONS VIDEO:-**[**Link**](https://drive.google.com/file/d/1coDrSmwggP2JjtnUN3BivFr9Ljzss74Y/view?usp=sharing)

**NOTE: Video Explanations for the above Visualizations are in Dashboard and Report sections**.

### Dashboard

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

* **Responsive and Design of Dashboard**

Dashboard: 

**Story**

A storyboard is a visual representation of a sequence of events, typically used in multimedia projects such as films, animations, advertisements, or presentations. It consists of a series of drawings or images arranged in a sequence, often accompanied by annotations or descriptions, to outline the flow of the story or concept.

### No of Scenes of Story

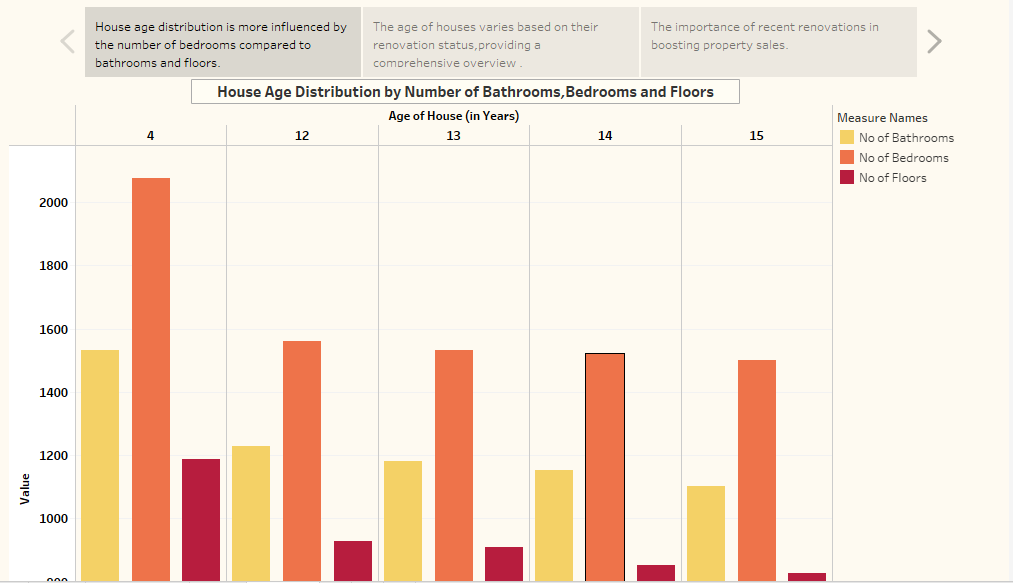
Design of Storyboard

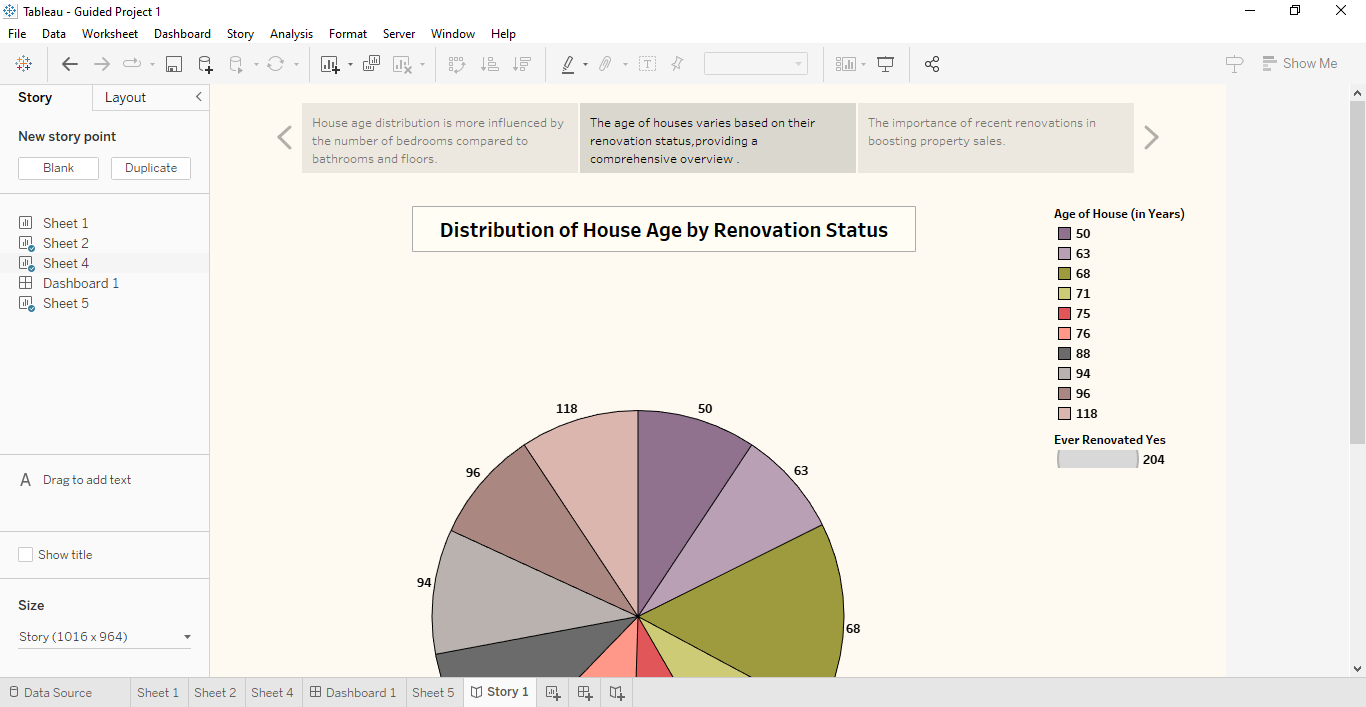
To design a storyboard in Tableau, start by defining the key elements and flow of your narrative. Utilize Tableau Analytics' capabilities to create a series of interactive visualizations, such as charts, graphs, and KPI cards, that represent each scene or stage of your story. Arrange these visualizations sequentially on a dashboard, ensuring a logical progression from beginning to end. Incorporate text boxes, annotations, and navigation buttons to provide context, explanation, and interactivity within the storyboard. Customize the design and layout to enhance clarity and engagement, leveraging Tableau' formatting options and templates.

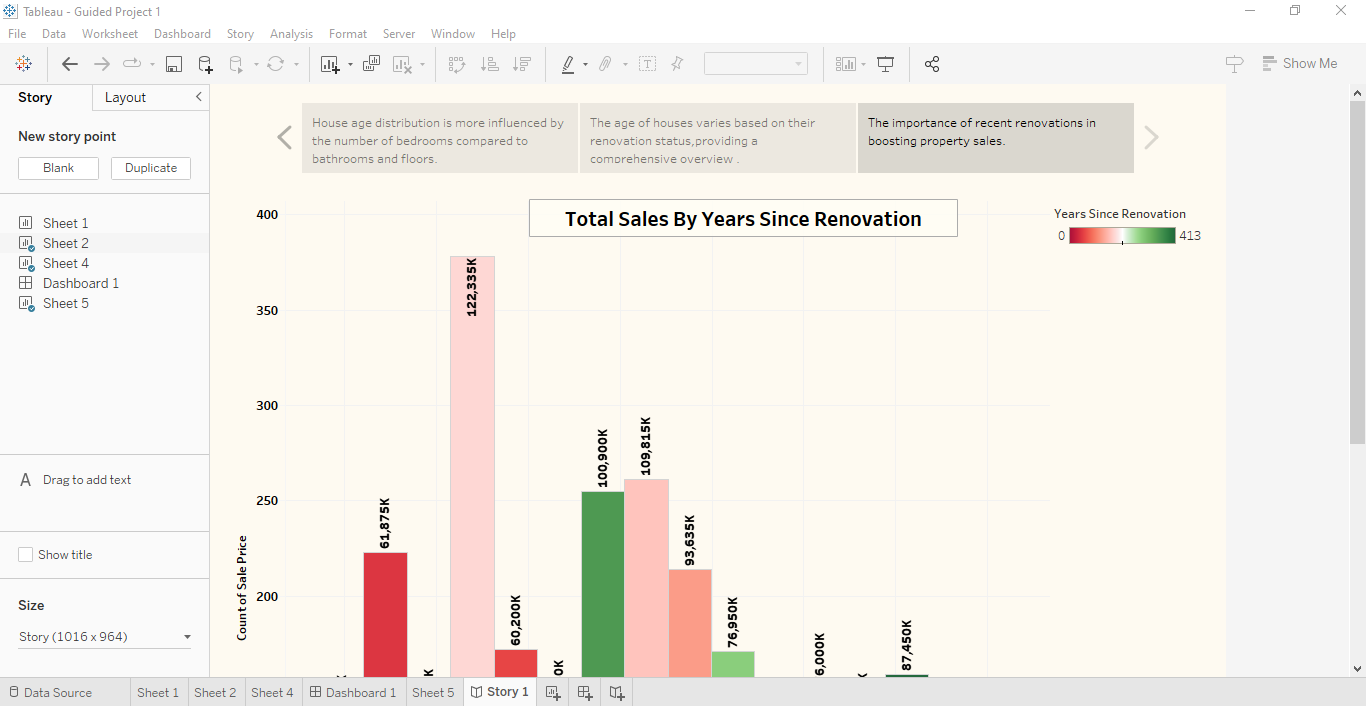
Explanation video link:

<https://drive.google.com/file/d/1nIyItMopZbD3nFYmGvRiSgElvhytTDs6/view?usp=sharing>

Storyboard:



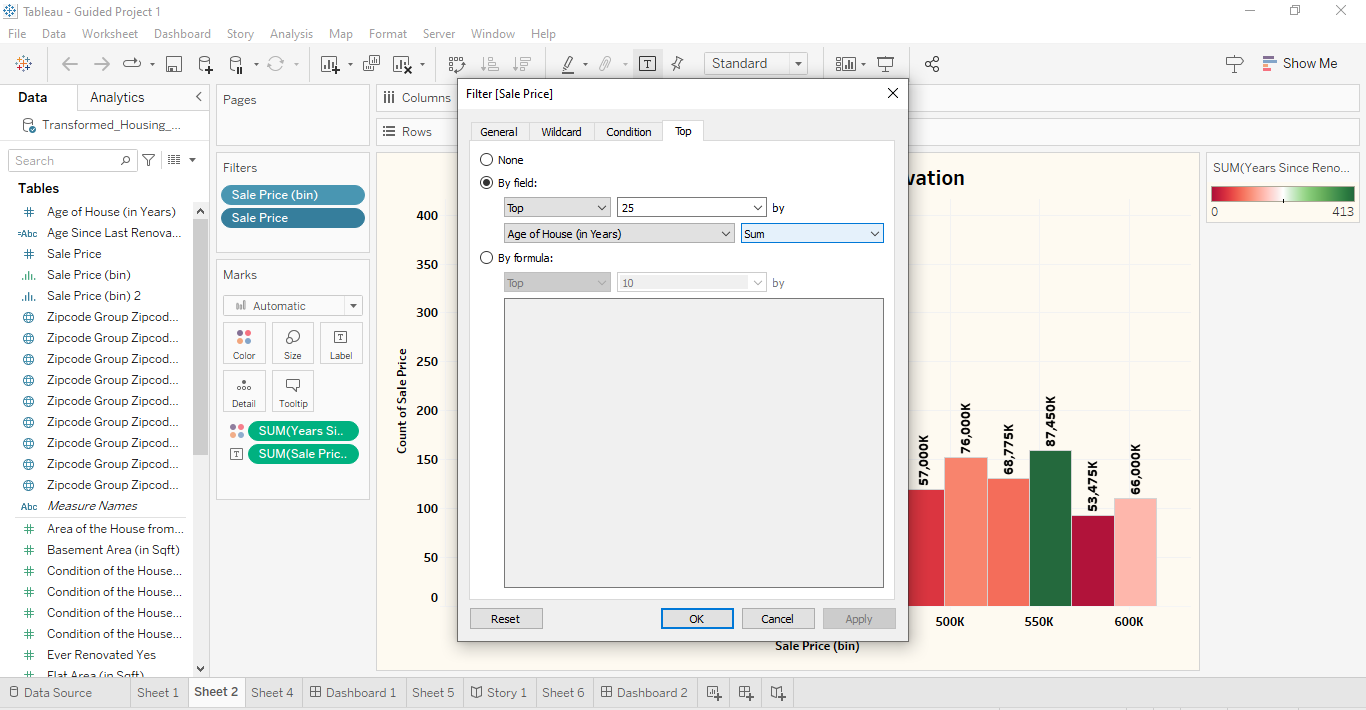




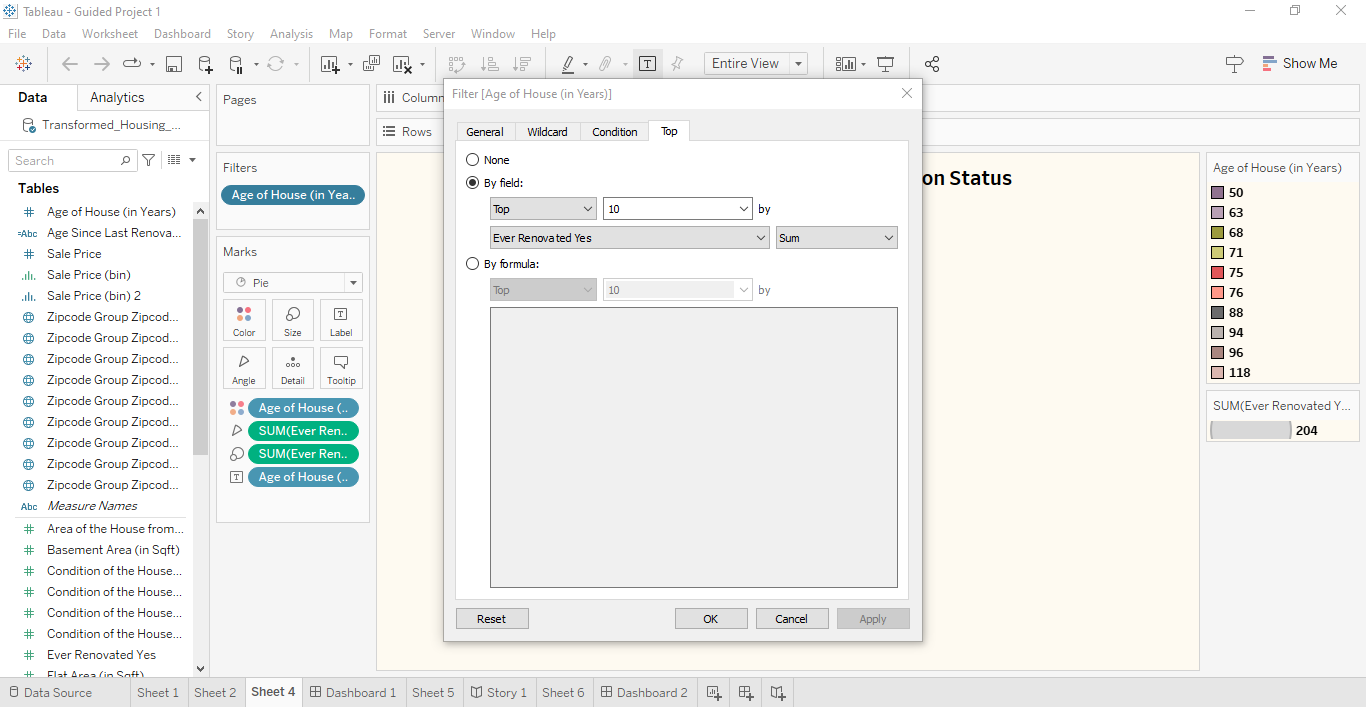
* **Utilization of Filters**

"Utilization of Filters" refers to the application or use of filters within a system, software application, or data processing pipeline to selectively extract, manipulate, or analyze data based on specified criteria or conditions.

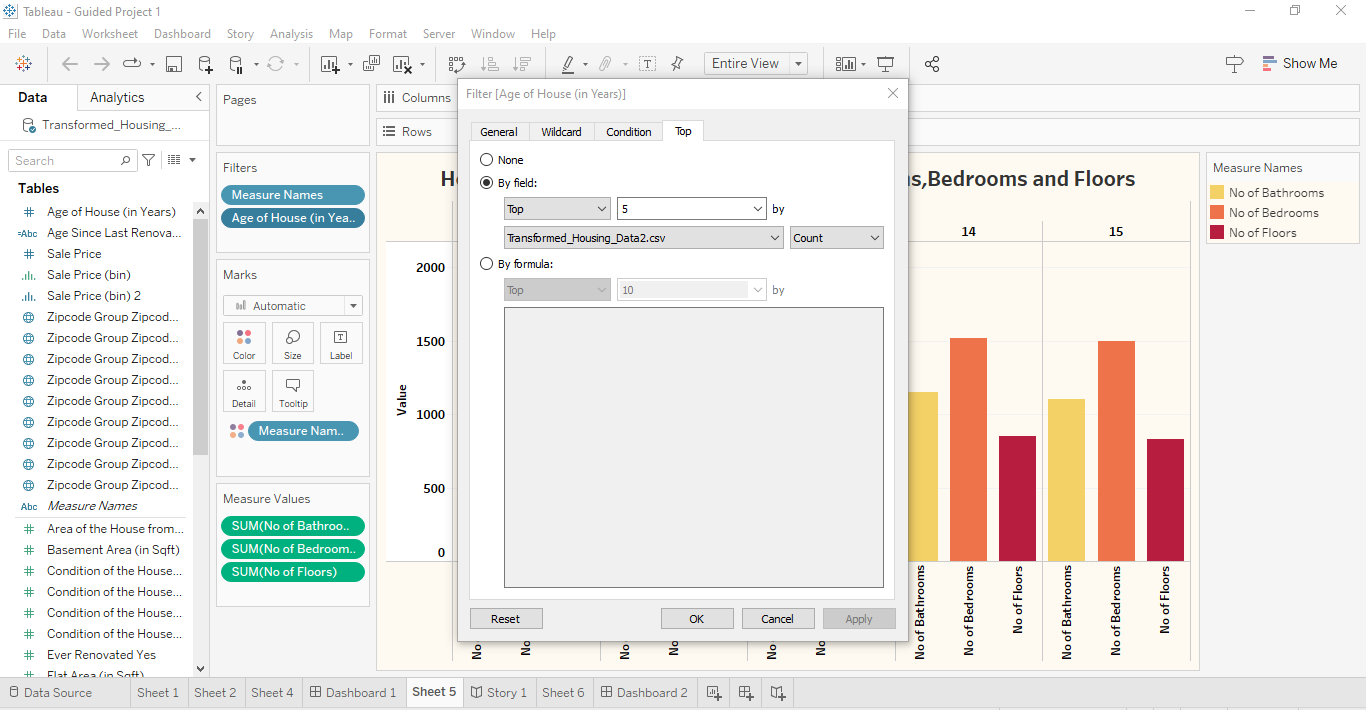
**Activity 2.1: Selected “Sale Price” as a Top N Filter**



**Activity 2.2: Selected “Age of House (in Years)” as a Top N Filter**

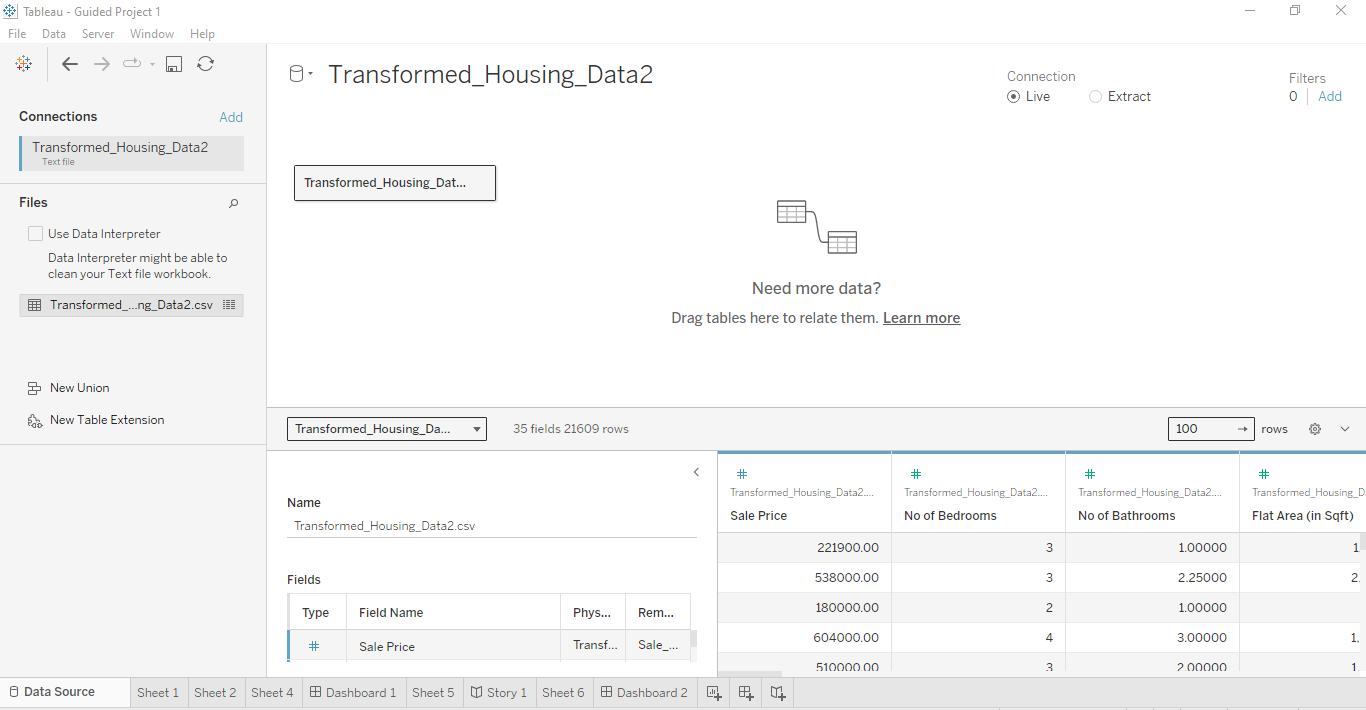


**Activity 2.1: Selected “Age of House (in Years)” as a Top N Filter**



**Performance Testing**

"Amount of Data Loaded" refers to the quantity or volume of data that has been imported, retrieved, or loaded into a system, software application, database, or any other data storage or processing environment. It's a measure of how much data has been successfully processed and made available for analysis, manipulation, or use within the system**.**

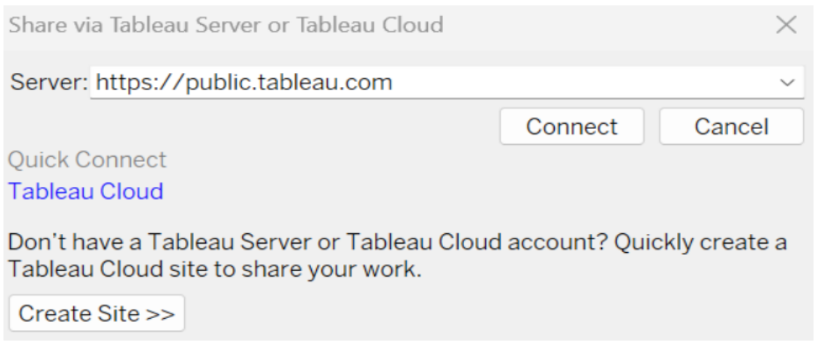


**Web integration**

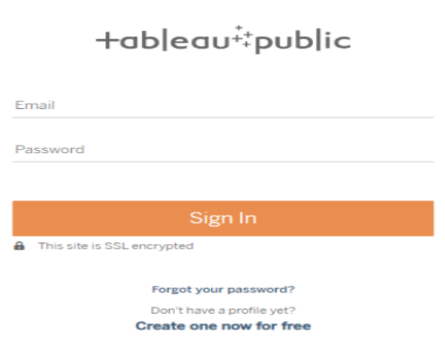
Publishing helps us to track and monitor key performance metrics, to communicate results and progress. help a publisher stay informed, make better decisions, and communicate their performance to others. Publishing dashboard and reports to tableau public

**Go to Dashboard/story, click on share button on the top ribbon**

**Give the server address of your tableau public account and click on connect.**

****

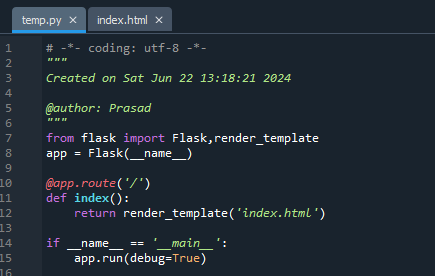
**Step 2: Once you click on connect it will ask you for tableau public user name and password**.



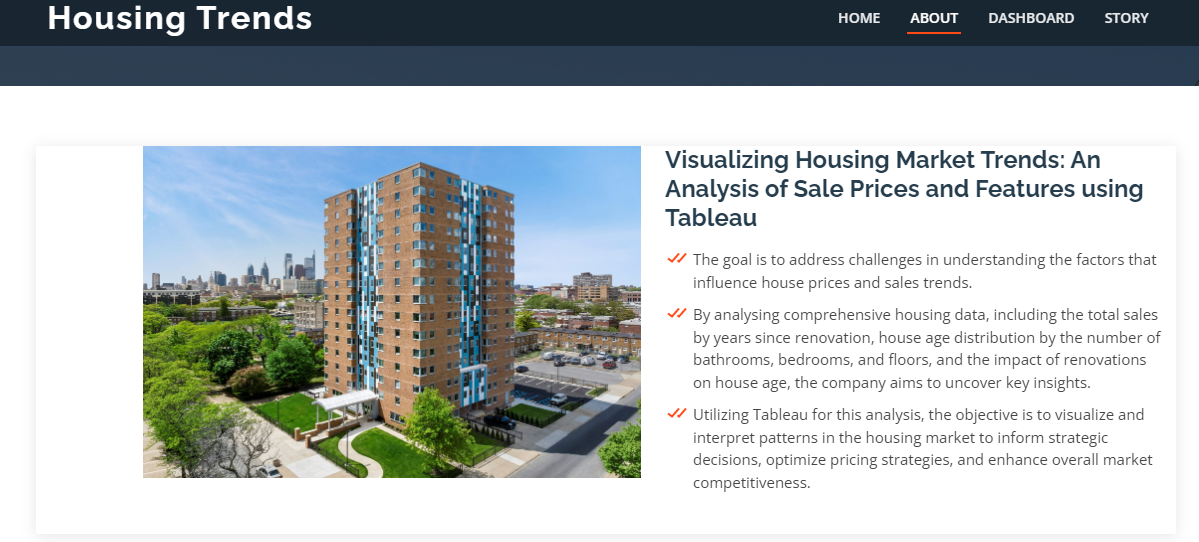
Once you login into your tableau public using the credentials, the particular visualization will be published into tableau public.

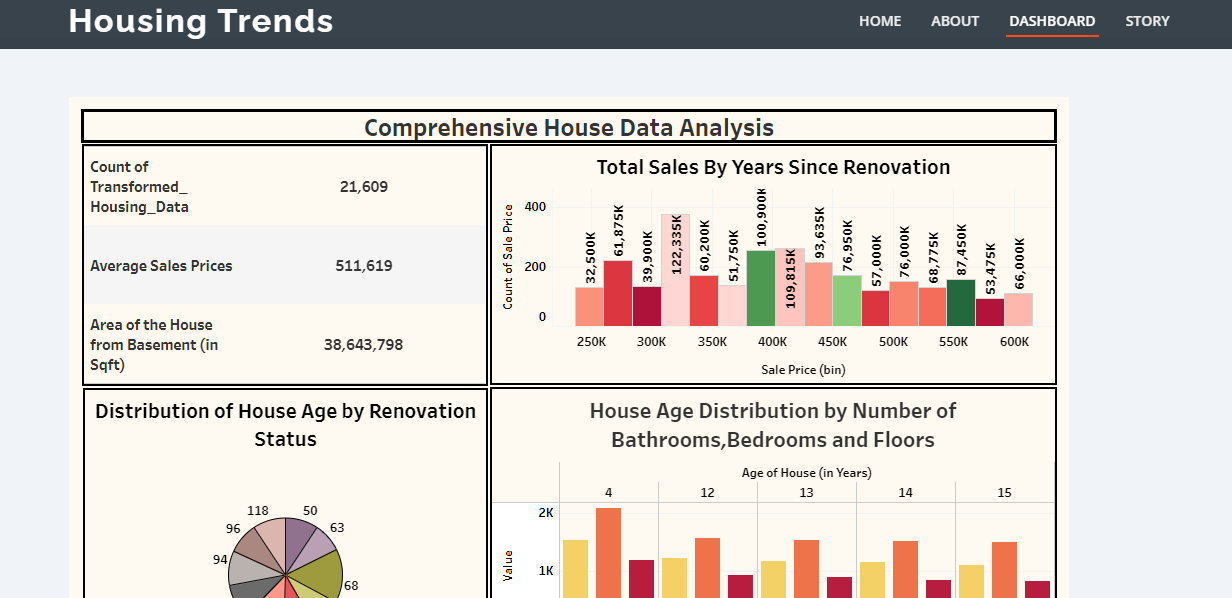
**Note:** While publishing the visualization to the public, the respective sheet will get published when you click on share option.

* **Dashboard and Story embed with UI With Flask**











**Project Demonstration & Documentation**

Below mentioned deliverables to be submitted along with other deliverables

Activity 1: - Record explanation Video for the project's end-to-end solution

Creating a record explanation video for a project's end-to-end solution is crucial for ensuring clarity and transparency in its implementation.

Activity 2: - Project Documentation-Step by step project development procedure

Create document as per the template provided

## Clean Data from Excel, CSV, PDF, and Google Sheets with Data Interpreter

*Applies to: Tableau Cloud, Tableau Desktop, Tableau Server*

When you track data in Excel spreadsheets, you create them with the human interface in mind. To make your spreadsheets easy to read, you might include things like titles, stacked headers, notes, maybe empty rows and columns to add white space, and you probably have multiple tabs of data too.

When you want to analyze this data in Tableau, these aesthetically pleasing attributes make it very difficult for Tableau to interpret your data. That’s where Data Interpreter can help.

**Tip:** Though Tableau's Excel add-in is no longer supported, Data Interpreter can help you reshape your data for analysis in Tableau.

## What does Data Interpreter do?

Data Interpreter can give you a head start when cleaning your data. It can detect things like titles, notes, footers, empty cells, and so on and bypass them to identify the actual fields and values in your data set.

It can even detect additional tables and sub-tables so that you can work with a subset of your data independently of the other data.

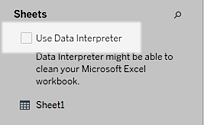
After Data Interpreter has done its magic, you can check its work to make sure it captured the data that you wanted and identified it correctly. Then, you can make any necessary adjustments.

After you select the data that you want to work with, you might also need to do some additional cleaning steps like pivoting your data, splitting fields, or adding filters to get the data in the shape you want before starting your analysis.

**Note**: If your data needs more cleaning than what Data Interpreter can help you with, try [Tableau Prep(Link opens in a new window)](https://www.tableau.com/products/prep).

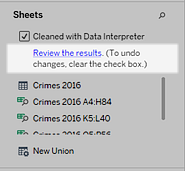
## Turn on Data Interpreter and review results

1. From the **Connect** pane, connect to an Excel spreadsheet or other connector that supports Data Interpreter such as Text (.csv) files, PDF files or Google sheets.
2. Drag a table to the canvas (if needed), then on the **Data Source** page, in the left pane, select the **Use Data Interpreter** check box to see if Data Interpreter can help clean up your data.

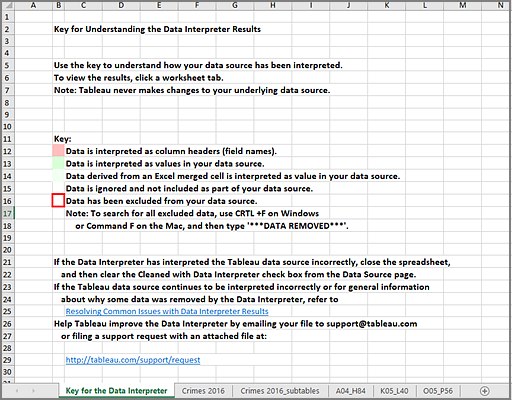


**Note:** When you clean your data with Data Interpreter, Data Interpreter cleans all the data associated with a connection in the data source. Data Interpreter does not change the underlying data.

1. In the Data pane, click the **Review the results** link to review the results of the Data Interpreter.



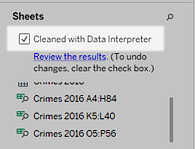
A copy of your data source opens in Excel on the **Key for the Data Interpreter** tab. Review the key to find out how to read the results.



1. Click each tab to review how Data Interpreter interpreted the data source.

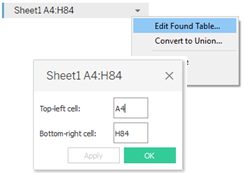
If Data Interpreter found additional tables, also called found tables or sub-tables, they are identified in the <sheet name>\_subtables tab by outlining their cell ranges. A separate tab is also included for each sub-table, color coded to identify the header and data rows.

If Data Interpreter does not provide the expected results, clear the **Cleaned with Data Interpreter** check box to use the original data source.



1. To replace the current table with any of the found tables, drag the current table off the canvas and then drag the found table that you want to use to the canvas.

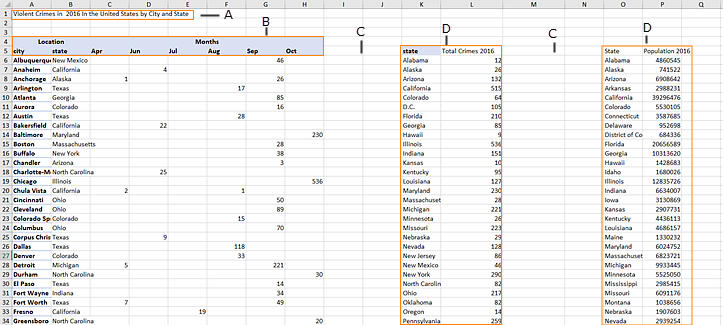
If Data interpreter has misidentified the range of the found table, after you drag the found table to the canvas, click the drop-down arrow on that table, and then select **Edit Found Table** to adjust the corners of the found table (the top-left cell and bottom-right cell of the table).



1. After you have the data that you want to work with, you can apply any additional cleaning operations to your data so that you can analyze it.

### Data Interpreter Example

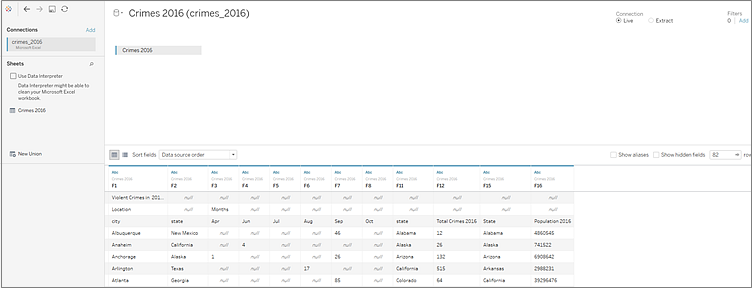
In this example we are connecting to an Excel spreadsheet with violent crime data by city and state for the year 2016. This spreadsheet includes multiple tables on one sheet and some extra formatting.



1. Title
2. Merged header cells
3. Extra white space
4. Sub-tables

The extra formatting in this spreadsheet makes it difficult for Tableau to determine what the field headers and values are.

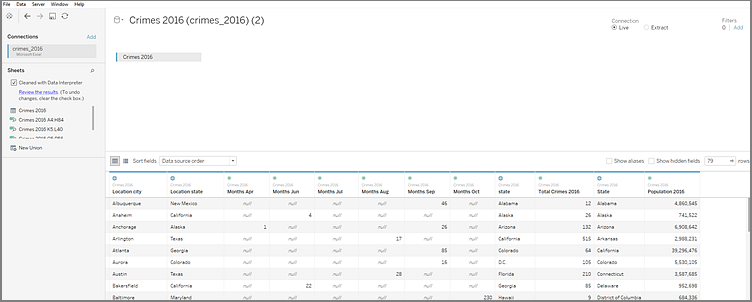
Instead, it reads the data vertically and assigns each column the default value F1, F2, F3 (Field 1, Field 2, Field 3) and so on. Blank cells are read as null values.



To see if Data Interpreter can help clean this data set, we select **Use Data Interpreter**.

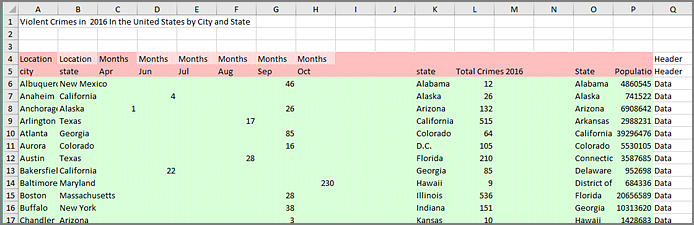
Data Interpreter detected the proper headings for the fields, removed the extra formatting and found several sub-tables. The sub-tables are listed in the **Sheets** section in the Data pane and are named using the original sheet name and the cell ranges for each sub-table.

In this example there are three sub-tables: **Crimes 2016 A4:H84**, **Crimes 2016 K5:L40**, and **Crimes 2016 O5:P56**.

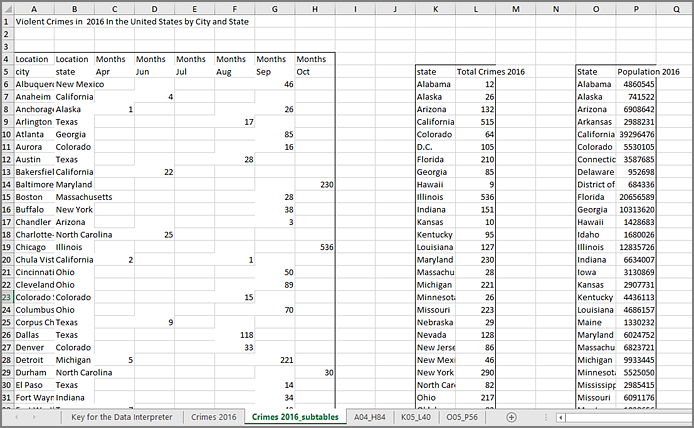


To examine the results of the Data Interpreter more closely, we click the **Review the results** link in the Data pane to view an annotated copy of the spreadsheet.

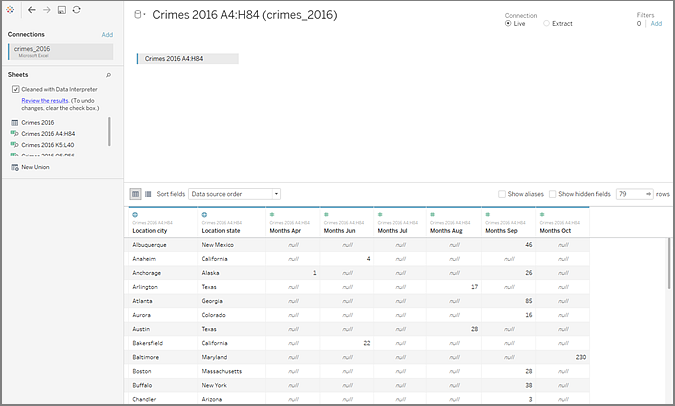
Here we see a copy of the original data, color coded to identify which data was identified as header data and which data was identified as field values.



The next tab shows us the sub-tables that Data Interpreter found, outlined by the cell ranges.



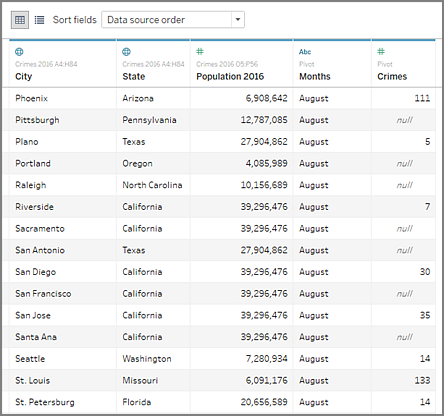
In this example the first sub-table, **Crimes 2016 A4:H84**, has the main data that we want to work with. To use this table as our data table, we can simply drag the original table off the canvas and then drag the new table to the canvas.



Once we have the data that we want to work with in the canvas, we can do some additional clean up on the data. For example we can:

* Change the field names so that they represent city, state, and month names.
* Pivot the months fields.
* Drag in the third sub-table **Crimes 2016 o5:P56** and join it to our first sub-table on the **State** field to include state populations for our analysis.
* Hide any duplicate fields that were added as a result of the join.

The results might look something like this:



Now we are ready to start analyzing our data in Tableau.

## When Data Interpreter is not available

The Data Interpreter option might not be available for the following reasons:

* **The data source is already in a format that Tableau can interpret:** If Tableau Desktop doesn't need extra help from Data Interpreter to handle unique formatting or extraneous information, the Data Interpreter option is not available.
* **Many rows or many columns:** The Data Interpreter option is not be available when your data has the following attributes:
  + Data contains more than 2000 columns.
  + Data contains more than 3000 rows and more than 150 columns.