



Free Training Transcript:
Explain Data

Introduction

Welcome to this video on Explain Data. Explain Data automatically provides AI-driven explanations for the value of a data point with a single click. Based on advanced statistical models, explanations are relevant and integrated, saving us time and enabling us to uncover insights we may not have found otherwise.

Explain Data works when authoring in Tableau Desktop, Tableau Server, Tableau Online and Tableau Public. It is available to users with Creator and Explorer access. Use the workbook below to follow along.

Scenario

Our workbook contains data on London House Sales. The viz displays the average duration of home ownership broken down by borough. Most boroughs have an average home ownership duration greater than 5 years. However, the borough of Newham is an outlier with an average home ownership duration of only 4.8 years. We will use Explain Data to help us explore why homes in Newham are selling more frequently than in other boroughs.

Launching Explain Data

To use Explain Data, we must select a single mark to analyze while editing the view. Our viz must contain a measure that is aggregated with Sum, Avg, Count, or Count Distinct. In our scenario, we click on the mark that represents the average duration of home ownership for Newham. We then click on the tooltip icon for Explain Data. We can also launch Explain Data from the Analysis menu or by right-clicking on the mark.

Explanations Window

The explanations window displays the name of the mark we selected, the measure in use, and an expected value summary. This information confirms that the average duration of home ownership for Newham is lower than expected and is not within the natural range of variation. Below that, on the left, we see several possible explanations for the expected value summary. These explanations are listed in order based on how informative they

are. The explanations are based on all the dimensions and measures of our data source, not just those displayed in the viz we are exploring. To the right, we see the explanation's description displayed as a combination of Tableau-generated natural language sentences and Tableau vizzes.

Explanations

Our first and most informative explanation is that Newham has a distinctive ratio of records with Residential Group equal to Young Ethnic Communities. In the explanation description, the blue bars in the chart represent Newham while the gray bars represent all boroughs in the map. The Y-axis represents the percentage of Total Number of House Sales. The gray bars tell us that across all London boroughs a lower proportion of house sales occurred in boroughs with young ethnic communities. Meanwhile, the blue bars tell us that in the London borough of Newham, a higher proportion of house sales occurred in young ethnic communities. As we scroll down, we see that the average duration of home ownership throughout London is slightly lower in young ethnic communities than in all other categories. However, this average is still not as low as the 4.8 average for Newham. Let's explore one of the other explanations for additional insights into the outlier.

The second explanation is regarding housing affordability groups. Again, the Y-axis represents the percentage of Total Number of House Sales. The blue bars tell us that Newham has a high proportion of house sales in the extremely low housing affordability group in contrast to London overall.

The third explanation is regarding median age. At the median age of 32, a higher proportion of house sales occur in Newham than in London overall. At other median ages, Newham has a lower proportion of house sales than London overall.

The fourth explanation indicates that almost all the house sales in Newham occur among those who are not retired.

Flow of Analysis

Explain Data is a jumping-off point for further exploration of our data. Each viz within the explanation description can be opened as a new worksheet for further analysis. Let's open our "Distribution of Values for % of Population that is Retired" viz as a new worksheet to examine it further.

Let's swap out “% of Population that is Retired”, replacing it with “Median Age”. We see that all the house sales in Newham are among relatively young people. Young people are generally quite mobile and move more frequently than older adults. This may help explain why the average home ownership duration in Newham is lower than the rest of London.

Recap

We started our analysis by asking Explain Data why homes in Newham sold more frequently than in the rest of London. Explain Data analyzed our entire dataset to quickly uncover relationships within our data. We then used those insights to help guide our deeper analysis of the data.

Additional Help

For additional help on the use of Explain Data, please visit help.tableau.com.

Closing

Thank you for watching this video on Explain Data. We invite you to continue with the free training videos to learn more about using Tableau.