Analyze Data with Explain in Oracle Analytics

Before you Begin

This lab shows you how to automatically generate insights from your data using the Explain machine learning capability in Oracle Analytics.

Background

The machine learning Explain algorithm generates data insights and patterns about the selected data element in the context of the whole dataset.

Selecting Explain generates visualizations that describe basic facts, key drivers of the column values, segments (hidden groupings), and anomalies. Segments are leaf nodes containing data classification rules for the node's data value in the data that correlate and predict outcome values for the selected attribute. Anomalies are outliers or unexpected results for the data model used to Explain the attribute column. The Explain-Anomalies algorithm is not the same algorithm used for Advanced Analytics Outliers.

What Do You Need?

- Access to Oracle Analytics Cloud or Oracle Analytics Desktop
- Download terminations_workers_details.xlsx to your computer

Add the Dataset and Prepare the Data

Before using the Explain algorithm, you can review the data profiling results in the Prepare page, and implement changes to standardize the data.

1. Sign in to Oracle Analytics.

   When using Oracle Analytics Desktop, you must install machine learning (DVML) to use Diagnostics Analytics (Explain), Machine Learning Studio, or advanced analytics.

2. On the Home page, click **Create**, and then click **Dataset**.
3. In Create Dataset, click **Drop data file here or click to browse**, select the terminations_workers_details.xlsx file, and then click **Open**.
4. In Create Dataset Table from terminations_workers_details.xlsx, click **OK**.
5. Click the **terminations_workers_details** tab. Select the **Tenure in Months** column, click **Measure** #, and then click **Attribute**.
6. In Tenure in Months properties, click **Number Format** #. In the Number Format row, click **Auto**, and then select **Number**. In the Decimal Places row, click **2**, and then select **1**.
7. Select the **Terminations Week** column, click **Measure** #, and then select **Attribute**.
8. Click **Save** 🖫. In Save Dataset As, enter terminations_workers_details in **Name**, and then click **OK**.
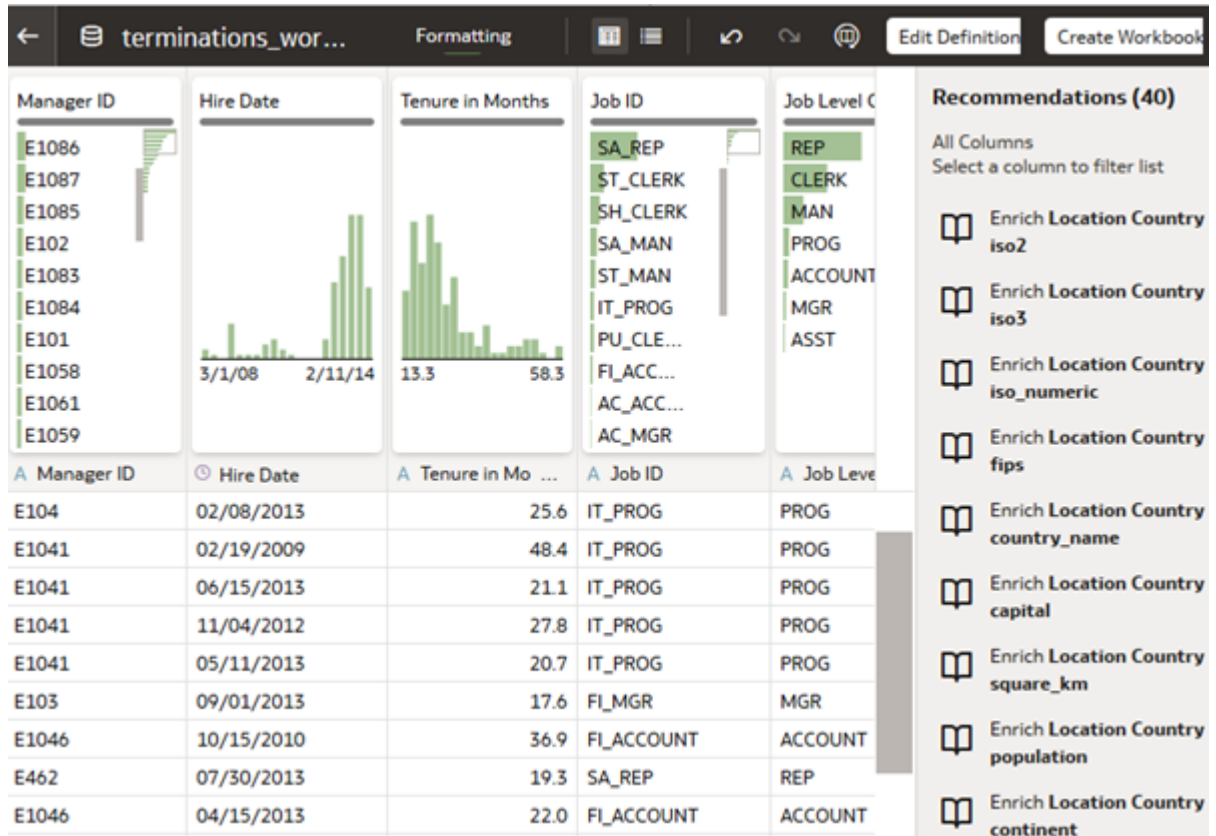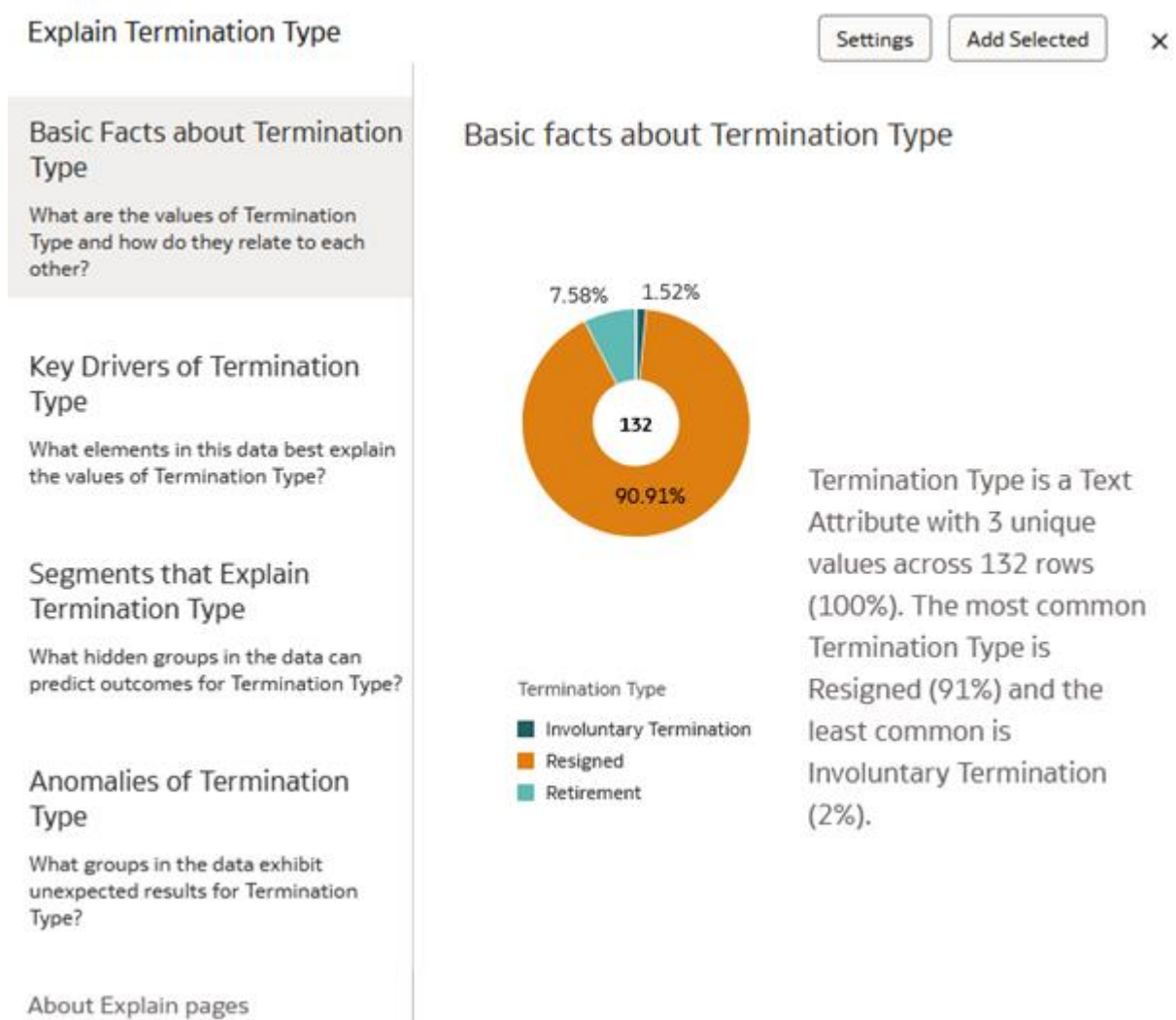
Visualize the Data

In this section, you select the basic facts generated by the Explain machine learning algorithm. You can also view the key drivers, segments, and anomalies for the termination type column.

Oracle Analytics enables Auto Insights as the default behavior.

1. In the terminations_workers_details dataset, click **Create Workbook**.

2. Close the Auto Insights panel.

3. In the Data pane, right-click **Termination Type**, and then select **Explain Termination Type**.

Explain Termination Type

Settings   Add Selected   ✕

Basic Facts about Termination Type

What are the values of Termination Type and how do they relate to each other?

Key Drivers of Termination Type

What elements in this data best explain the values of Termination Type?

Segments that Explain Termination Type

What hidden groups in the data can predict outcomes for Termination Type?

Anomalies of Termination Type

What groups in the data exhibit unexpected results for Termination Type?

About Explain pages

Basic facts about Termination Type

7.58%   1.52%

132

90.91%

Termination Type

■ Involuntary Termination
■ Resigned
■ Retirement

Termination Type is a Text Attribute with 3 unique values across 132 rows (100%). The most common Termination Type is Resigned (91%) and the least common is Involuntary Termination (2%).

4. In Explain Termination Type, hover your cursor over the upper right side of the canvas, and then click **Select for Canvas** ✔. When the check mark ✔ changes to green, click **Add Selected**.
5. In the Data pane, right-click **Termination Type**, select **Explain**. In Explain Termination Type, click **Key Drivers of Termination Type**.

   You could select the visualizations generated for key drivers of termination type data to explore the data further. These visualizations aren't used in this tutorial.

Key Drivers of Termination Type

Based on Termination Type: All Values the 5 attributes that are most strongly correlated are: **Job ID, Tenure in Months, Job Level Code, Termination Department, Termination Reason**

The charts below show the distribution of Termination Type values across each of the key drivers (sorted by descending row count). Click the checkmarks above any of the visuals to add them to your workbook when done.
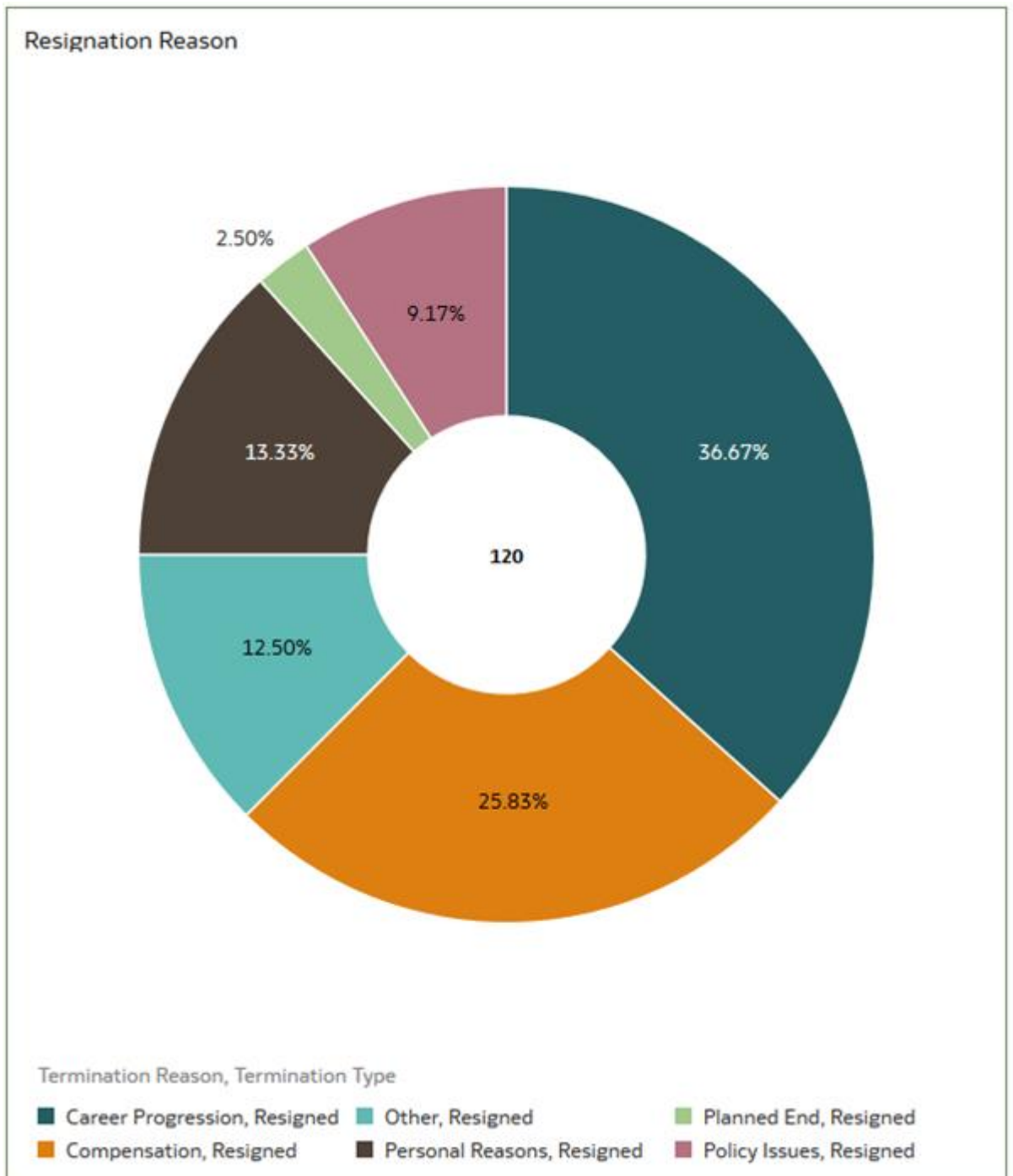

6. Click **X** to close Explain Termination Type.
7. Click **Save**, and select **Save As**. In Save Workbook, enter HR Attrition in the **Name** field, and then click **Save** 🖫.
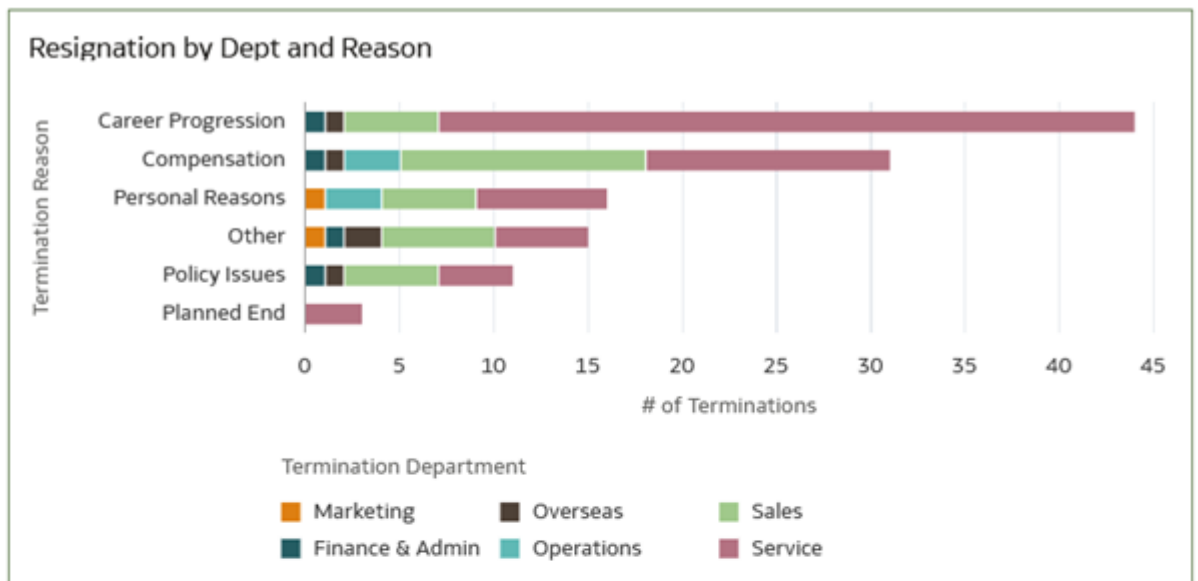
Explain Resignations

In this section, you analyze resignations to uncover the reason why employees voluntarily resign. You also identify the departments that had the most resignations.

1. In the visualization, hover your cursor over **Resigned**, right-click and select **Keep Selected**.
2. From the Data pane, select **Termination Reason**, and then drag it to the canvas. Select **Termination Reason**, drag it to **Color** to switch places with **Termination Type**.
3. Click **Properties** ⸬. In the Title row, click **Auto**, select **Custom**, and then enter Resignation Reason as the title of the visualization.

Resignation Reason

2.50%

9.17%

13.33%

36.67%

120

12.50%

25.83%

Termination Reason, Termination Type

- Career Progression, Resigned
- Other, Resigned
- Planned End, Resigned
- Compensation, Resigned
- Personal Reasons, Resigned
- Policy Issues, Resigned

4. In the Data pane, hold down the **Ctrl** key and select the following:
   o **Termination Department**
   o **Termination Reason**
   o **# of Terminations**
5. Right-click, select **Pick Visualization**, and then select **Horizontal Stacked** ▤.
6. Click the visualization **Menu** ⋮ , select **Sort by**, select **# of Terminations**, and then select **High to Low**.
7. Click **Properties** ⫴, click **Auto** in the Title row, and then select **Custom**. In **Title**, enter Resignations by Dept and Reason.
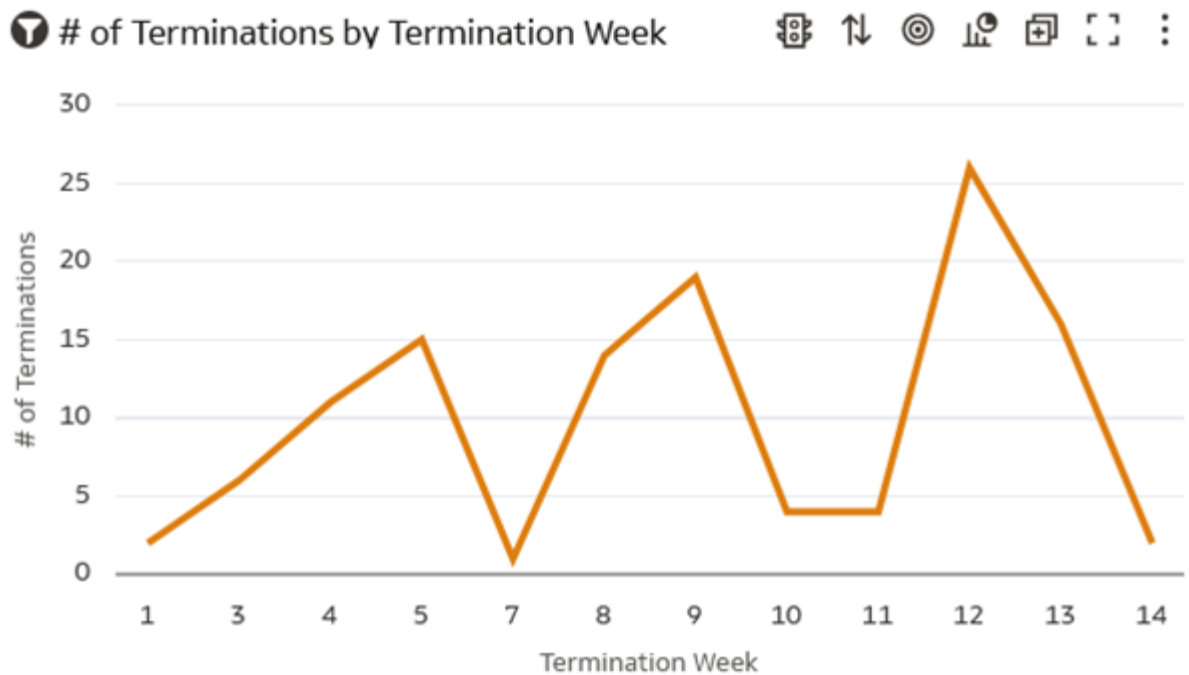8. Click **Save** 🖫.

Resignation by Dept and Reason
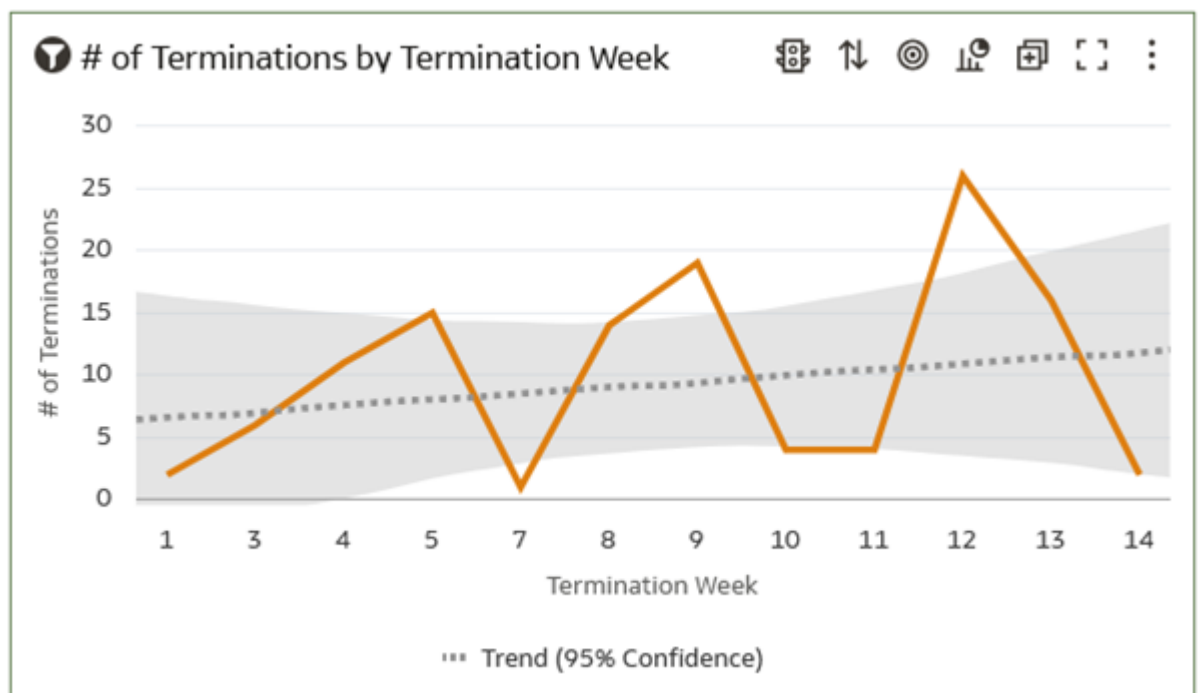

Resignation Reason
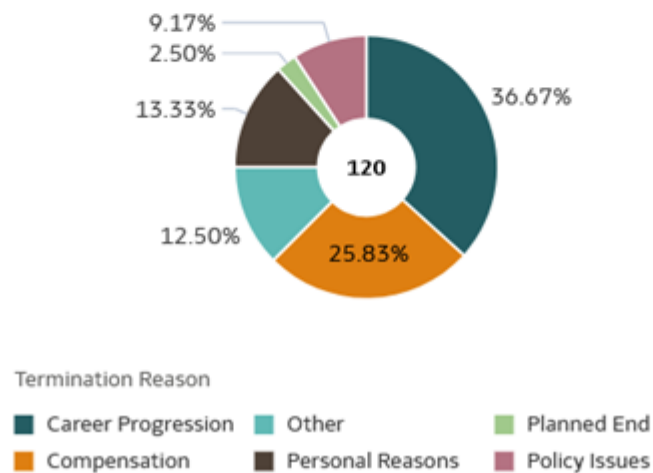
Analyze Trends in Resignations

In this section, you complete the picture of the termination data.

1. In the Data pane, hold-down the **Ctrl** key, select **Termination Week** and **# of Terminations**. Right-click, select **Pick Visualization**, and then select **Line** ≣.

2. In the line visualization, hover over the line, right-click and select **Add Statistics**, and then select **Add Trend Line**.

3. In **Properties** ⁣ . In the Title row, click **Auto** and select **Custom**. In **Title**, enter Terminations by Week Trend.
4. Click **Note** 🖳 and select **Add Note**. In the **Note** field, enter Career progression is the primary reason for resignations. Click **Save**.
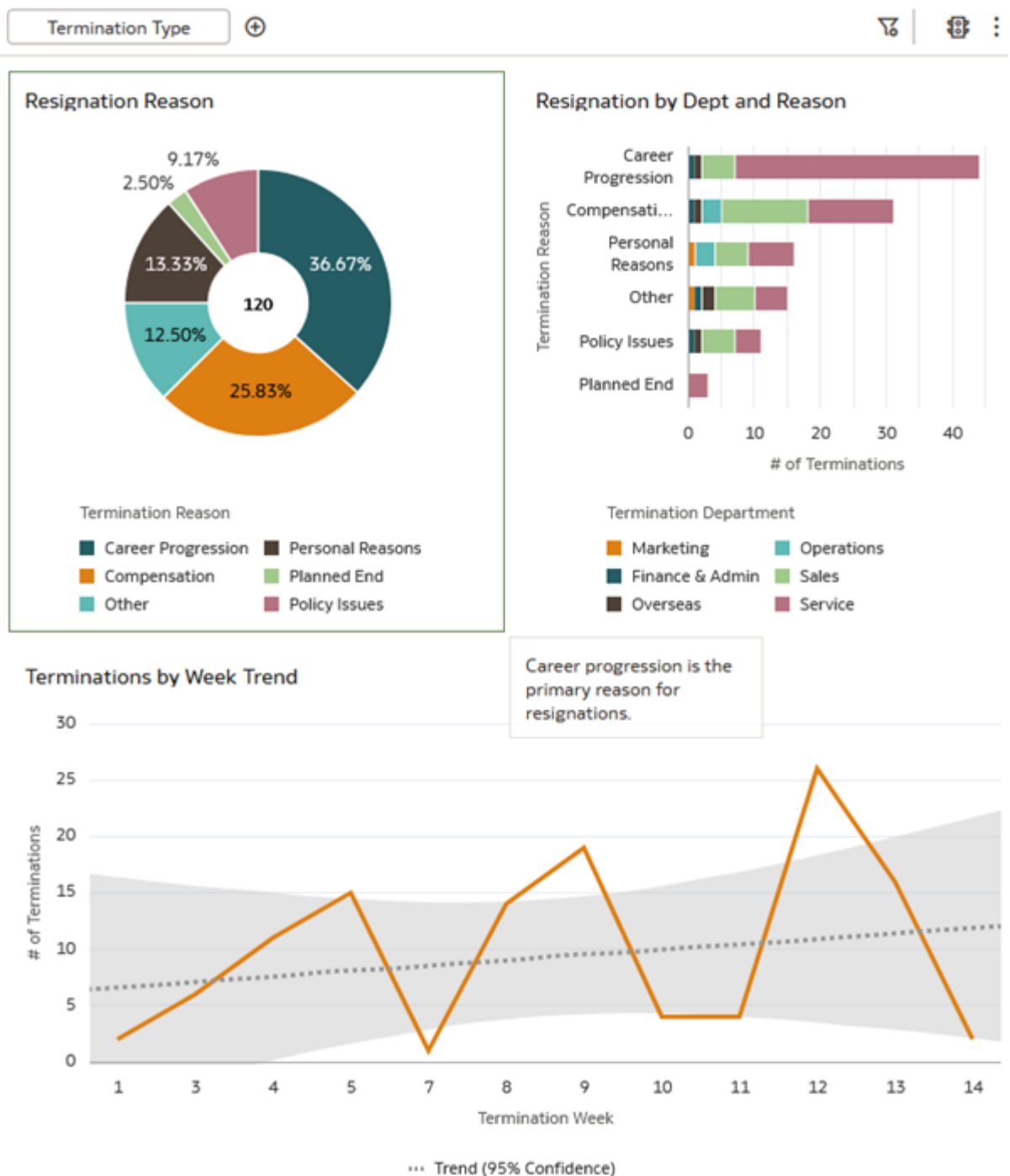
Create a Presentation

In this section, you create the view your users see when the workbook is opened.

1. Click **Present**. From Canvases, drag **Explain Termination Type** to **Drag a Canvas Here to Begin**.
2. In the Explain Termination Type canvas, click the **menu** ▼, and then select **Rename Page**.
3. In the **Page Title** field, enter Attrition Reasons.

4. Click **Active Canvas**. In the Filter Bar group, remove the check for **Termination Type**.
5. Click **Save**.
6. Click **Preview** ▷.
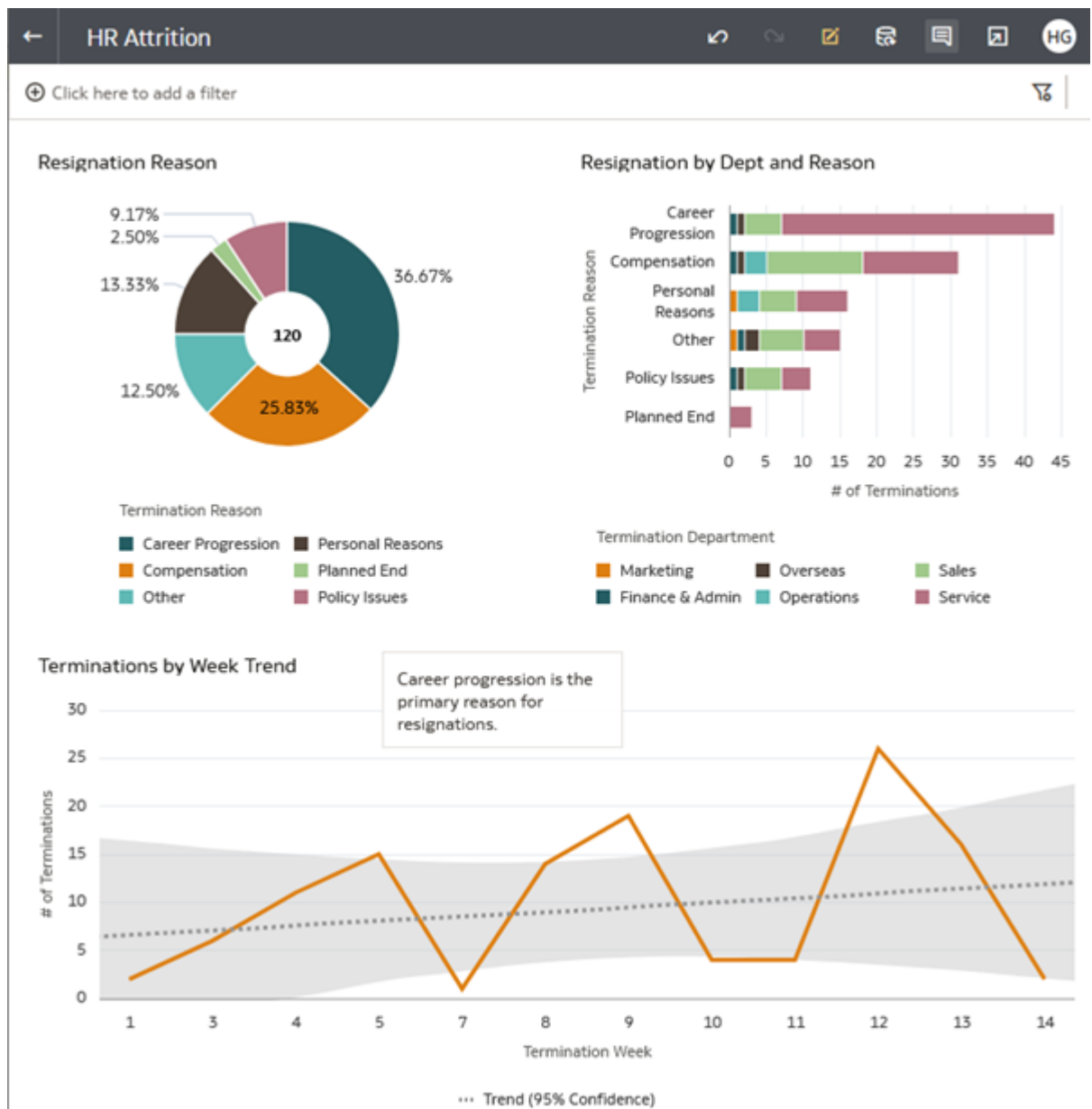
7. Click **Edit** 🗹 return to the Present page.