Smart Discovery as a Story

Based on version 2018.16

TABLE OF CONTENTS

www.sap.com/contactsap

ABOUT THIS DOCUMENT

This document provides instructions for how to use the new version of Smart Discovery in SAP Analytics Cloud.

ABOUT SMART DISCOVERIES

A Smart Discovery is the result generated by running a Machine Learning algorithm to uncover new or unknown relationships between columns within a dataset. Smart Discovery also gives an overview of your data by automatically building charts to begin discovering more about your data.

In SAP Analytics Cloud, you can choose to run a discovery against a measure or dimension within a model to determine the influencers on the focus of the discovery, how they relate to one another, and key members or value ranges. The guided analysis will also alert you if no significant or insightful relationships are present in your data.

Smart Discoveries are driven by users working on stories. You can exclude particular dimensions or measures from the analysis, focus on particular dimension members, and display relationships from a list of influencers. The results can be saved as story pages which you can share within your organization.

To narrow down the scope of the analysis, or to optimize running discoveries on large datasets, you can create filters to exclude specific records from dimensions in your model.

Note: Currently, you cannot run a discovery analysis on more than one million cells. By filtering the data first, you can reduce the number of cells included in your dataset. The number of cells is calculated by multiplying the number of measures by the number of records.

After a running a discovery, you can simulate numeric targets, using the results of the key influencer analysis. You can also view data records which are highlighted by the predictive model as being unexpected.

RUNNING A SMART DISCOVERY

You can run a Smart Discovery to gain insights on how underlying variables influence a selected dimension or a measure within a dataset in a story.

Context

You can launch a Smart Discovery when you start a new story or open a story containing a dataset you want to explore.

Procedure

1. Launch a Smart Discovery:

    a) If you are creating a new story: go to Main Menu> Create> Story and then select **Run a Smart Discovery**.

    b) If you want to run a Smart Discovery on a data set in an existing story: go to Main Menu> Browse> Files and select the story. Once the story is loaded, change the mode to **Edit**. From the **Tools** menu select **Smart Discovery**.

2. Configure settings in the **Smart Discovery** side panel.

    a) Select the target dimension or measure on which to run the discovery from the **Select Measure/Dimension** list under **Discovery Settings**.

    Note: If you selected a dimension, the Set Classification Groups dialog is displayed with all the members of the dimension listed under Baseline Group.
    Drag all the members you want to serve as the focus of the Smart Discovery to the Comparison Group column and select OK to continue.

    b) Specify **Advanced Options** settings:

    *   Determine which **Version** you want to use for the discovery.

    *   Under **Each record refers to** provide a singular and plural identifier for each row of data, such as *Employees* or *Sales Transactions*. *Record* and *Records* are the default values.

    *   Use the **Included Columns** area to specifically determine which measures and dimensions to include in the discovery. By default, all measures and dimensions in the data set are included in the discovery scope.

        a) Select **Measures** to display the **Set Measures to Include in Discovery** dialog. Choose the measures you want to include.

        b) Select **Dimensions** to display the **Set Dimensions to Include in Discovery** dialog. Choose the dimensions you want to include.

Select **OK** to submit your selections.

The number of selected measures and dimensions is indicated as well as a listing of the selected items under **Measures** and **Dimensions**.

- Under **Page Filters** select **+ Add Filters** to further refine the discovery scope on any of the included dimensions or measures.

An entry will appear under **Page Filter** for every filter you create.
Note: Discovery analysis can handle a finite data size, use filters to reduce the data size.

c) When you are ready to launch the discovery analysis, select **RUN**.
Once the process is complete the discovery results are displayed under separate pages:

- **Overview**: This page provides visualizations to summarize the results for the target dimension or measure. Even if the discovery's insight quality is considered to be poor or low, this page is created, as it is just an overview of your data.

Note: The Adjust Measure widget is displayed in the Overview when the target column is a dimension. Select a measure from the widget analyze against the target dimension.

- **Key Influencers**: This page lists (ranked from highest to lowest) up to ten dimensions and measures that significantly impact the target of the discovery. A summary and insight quality is provided. For every listed influencer there are specific visualizations displayed to show the relationship between the influencer and the target. This page will not display if the insight quality is considered to be too low.

- **Unexpected Values**: This page includes a table listing existing values and the predicted value is displayed along with the other corresponding dimensions. To add or remove columns from the displayed table select **Edit Columns**. This page will not display if the insight quality is considered to be too low or there are no unexpected values.

Note: For regression analysis, two additional interactive charts (scatterplot and bar/column) are provided to help visualize the difference between the predicted and existing values listed in the table. Use the Search or Edit Columns options to focus or filter out columns from the table. When you sort the table on a particular column, the columns in the bar/column chart will reflect this sort order. Any changes in the table will be automatically reflected in the charts. Note: Unexpected Values are available only for measures

- **Simulation**: This page displays a waterfall chart containing the influencers' relative contributions. To the right, a listing of the key influencers and their corresponding values is displayed. To modify a value and simulate its impact, select the value and use either the displayed slider or radio buttons to specify a new value. The impact of the new value is reflected in the chart on the left. The value of the smart discovery target column also changes. This page will display when the insight quality is sufficient and there are variables to analyze. Note: Simulations are available only for measures

Results

After you have finished analyzing the results from your discovery, you can:

- Save the discovery as a part of the new or existing story.

- Share the story with other users in your organization.

You can change the discovery target column, classification groups, and what columns to include in the analysis. However, these changes cannot be applied directly to your current discovery results. From the ▦ **Tools** menu select **Smart Discovery.** You will have to run the discovery again.

Import / Export Functionality

Note: When importing a .tgz file that contains pages generated by Smart Discovery, not all Smart Discovery charts will be imported. This is a temporary limitation that will be resolved within future versions.