



PUBLIC

# Data Synthesizer Operator in SAP Data Intelligence

## Usage Guide

# TABLE OF CONTENTS

<b>LAUNCH SAP DATA INTELLIGENCE .....</b>	<b>3</b>
<b>DOWNLOAD AND INSTALLATION .....</b>	<b>3</b>
<b>RUN THE GRAPHS.....</b>	<b>6</b>

[www.sap.com/contactsap](http://www.sap.com/contactsap)

© 2018 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company.

The information contained herein may be changed without prior notice. Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, SAP SE or its affiliated companies have no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and SAP SE's or its affiliated companies' strategy and possible future developments, products, and/or platform directions and functionality are all subject to change and may be changed by SAP SE or its affiliated companies at any time for any reason without notice. The information in this document is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, and they should not be relied upon in making purchasing decisions.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies. See [www.sap.com/copyright](http://www.sap.com/copyright) for additional trademark information and notices.

**THE BEST RUN**



Usage Guide for Data Synthesizer Operator provides detail steps from download, import, to run graph in in SAP Data Intelligence.

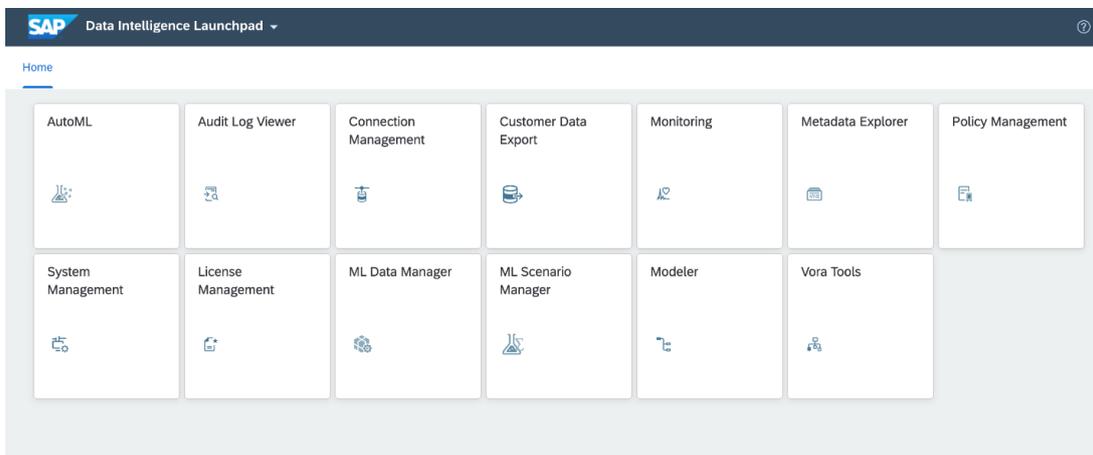
Before you start using this guide, please make sure that you are familiar with the basic concepts of SAP Data Intelligence modeling, such as Operators, and Graphs. For more information, please refer to the Modeling Guide for SAP Data Intelligence ([https://help.sap.com/viewer/product/SAP\\_DATA\\_INTELLIGENCE/Cloud/en-US](https://help.sap.com/viewer/product/SAP_DATA_INTELLIGENCE/Cloud/en-US)).

## LAUNCH SAP DATA INTELLIGENCE

Launch SAP Data Intelligence in a browser using the URL of your instance. And log on to the SAP Data Intelligence Launchpad application using the following information:

- Tenant name
- Your username
- Your password

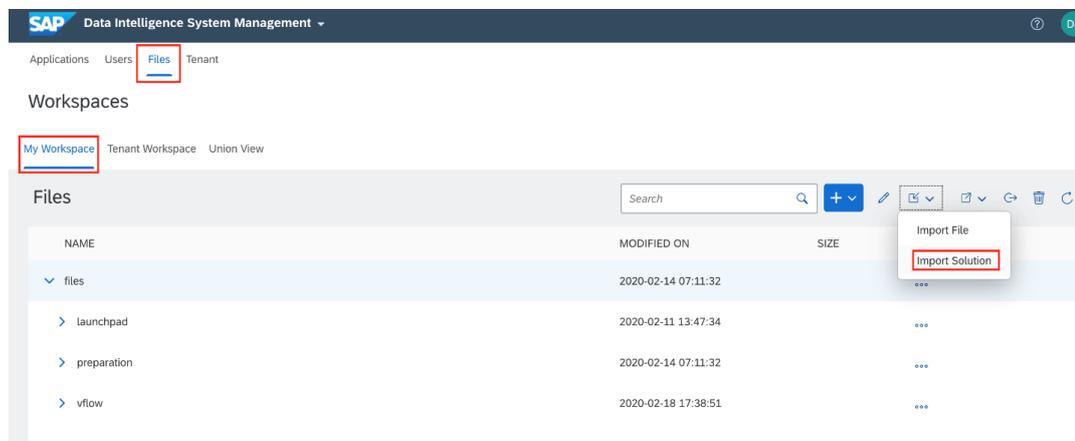
The SAP Data Intelligence Launchpad opens and displays the initial home page.



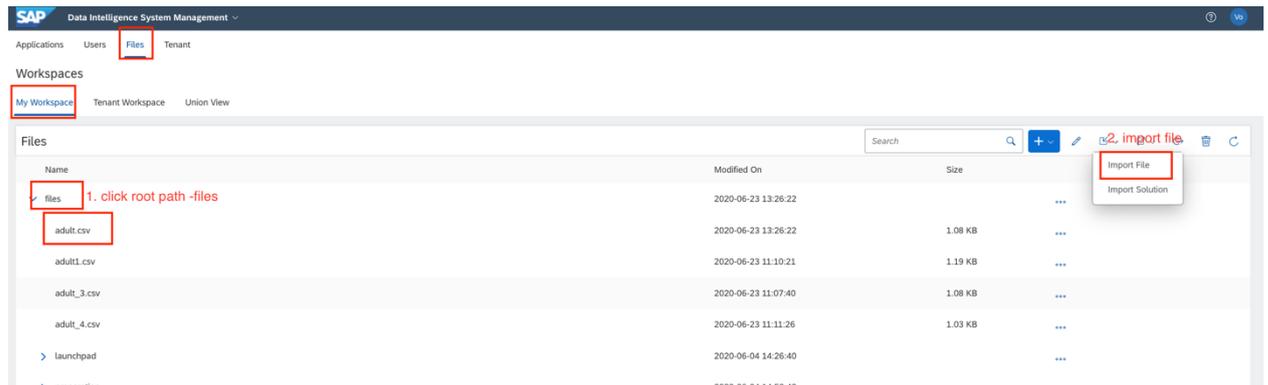
## DOWNLOAD AND INSTALLATION

After downloading the Data Synthesizer Operator attached to this package, you import the solution archive into your SAP Data Intelligence user workspace by following below steps.

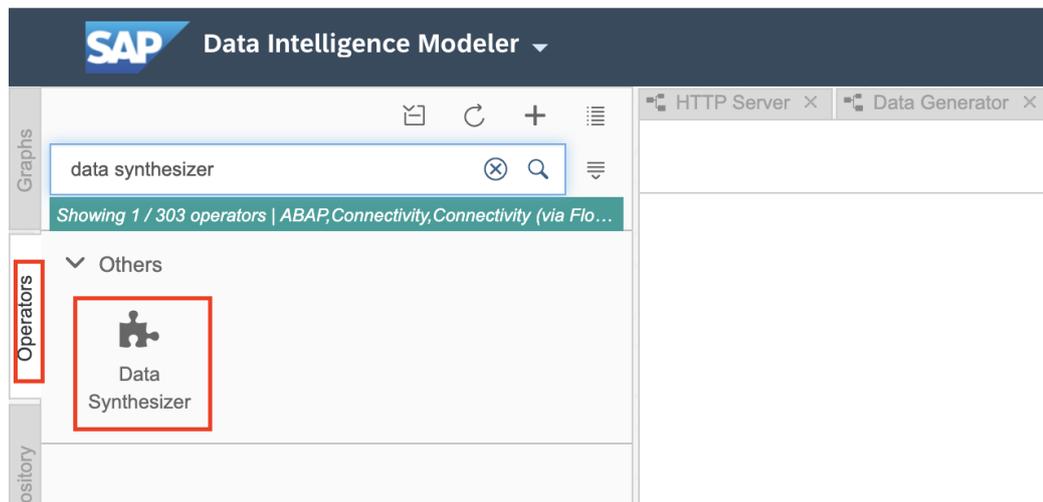
- Click the **"System Management"** tile on the launchpad page. Click **"Files"**, select **"Import Solution"** in **"My Workspace"** and import **data-synthesizer-operator.tgz**. Wait until it shows import success.



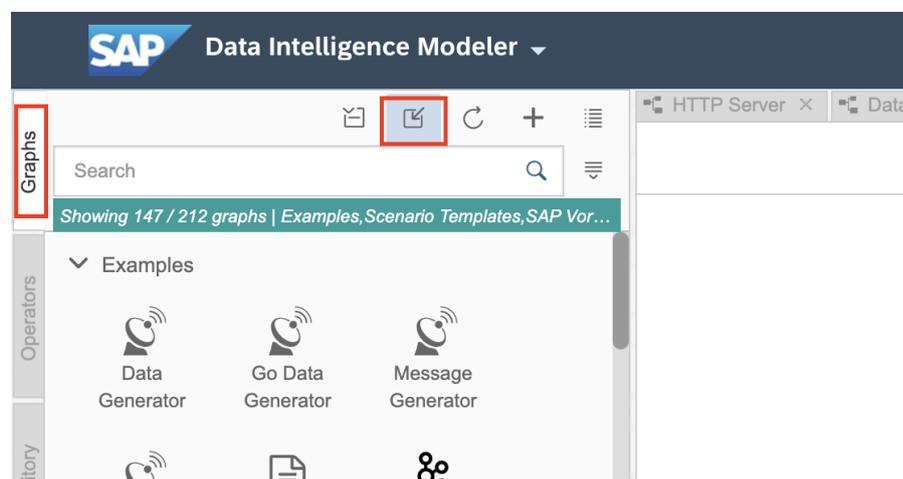
- Click "**Files**", select "**Import File**" in "**My Workspace**" and import **adult.csv**, wait until it shows import success. Please download the sample csv file from [SAP Jam](#), or create your own csv file.



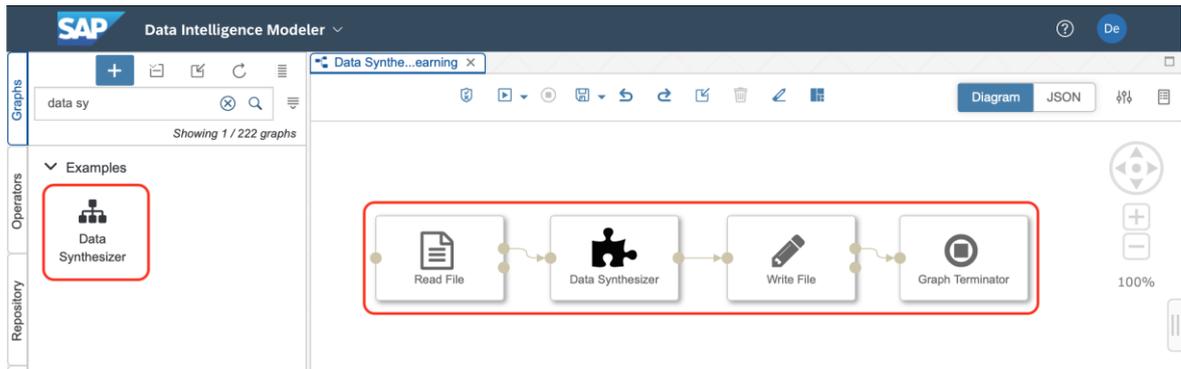
- After import success, click into "**Modeler**" tile on the Home page. Select "**Operators**" on the left panel, and search "**data synthesizer**" to ensure **Data Synthesizer Operator** has been imported successfully. (Our released .tgz package works as one of the many operators in Data Intelligence to handle metadata processing.)



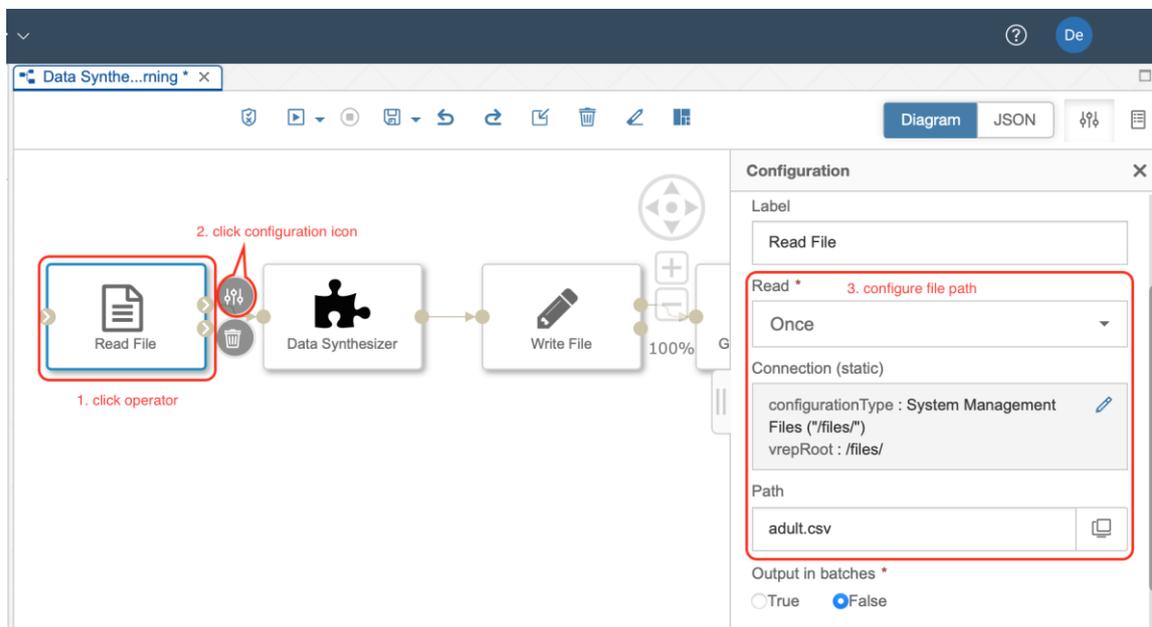
- Click into "**Modeler**" on the Home page, select "**Graphs**" on the left panel and click "**import**" icon to import **graph.json**. This is a predefined data process workflow which applies data synthesizer operator to synthesize data from the "Read File" operator and transport metadata to the "Write File" operator. Please download this sample graph from [SAP Jam](#), or create your own graph.



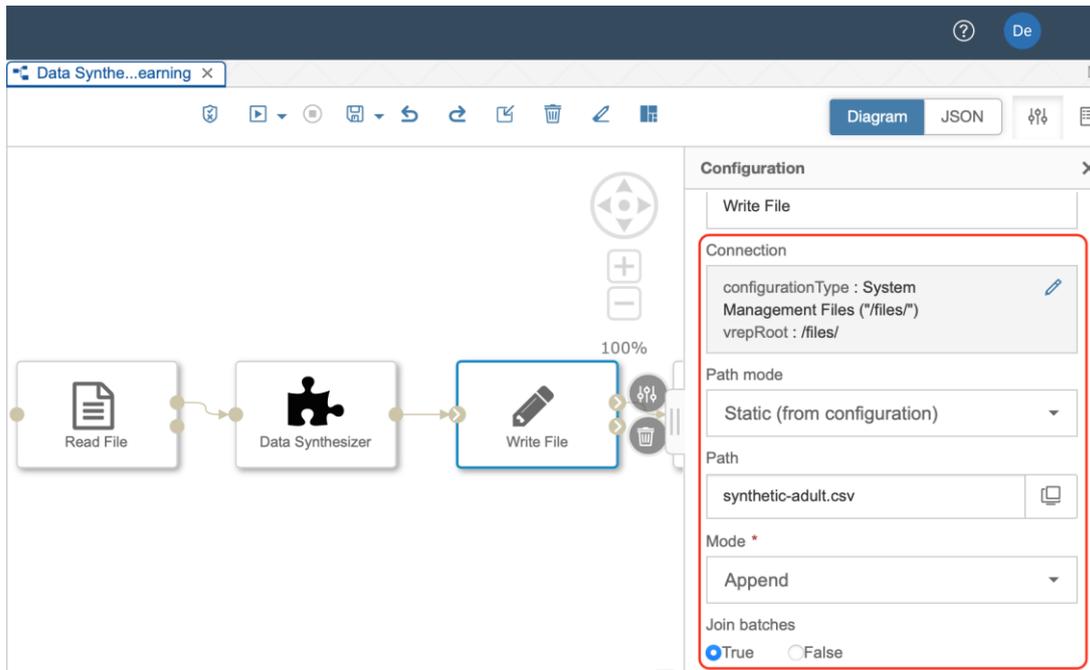
- After import success, the Data Synthesizer graph will be opened directly as follows (please kindly ignore the minor UI difference caused by functionality change of Data Intelligence).



It contains “Read File”, “Data Synthesizer” and “Write File” operators. You can define your own file path in the **Configuration** of “Read File” Operator (as follows). (**NOTE:** The default root path is **System Management Files** where you upload the files in above steps.)

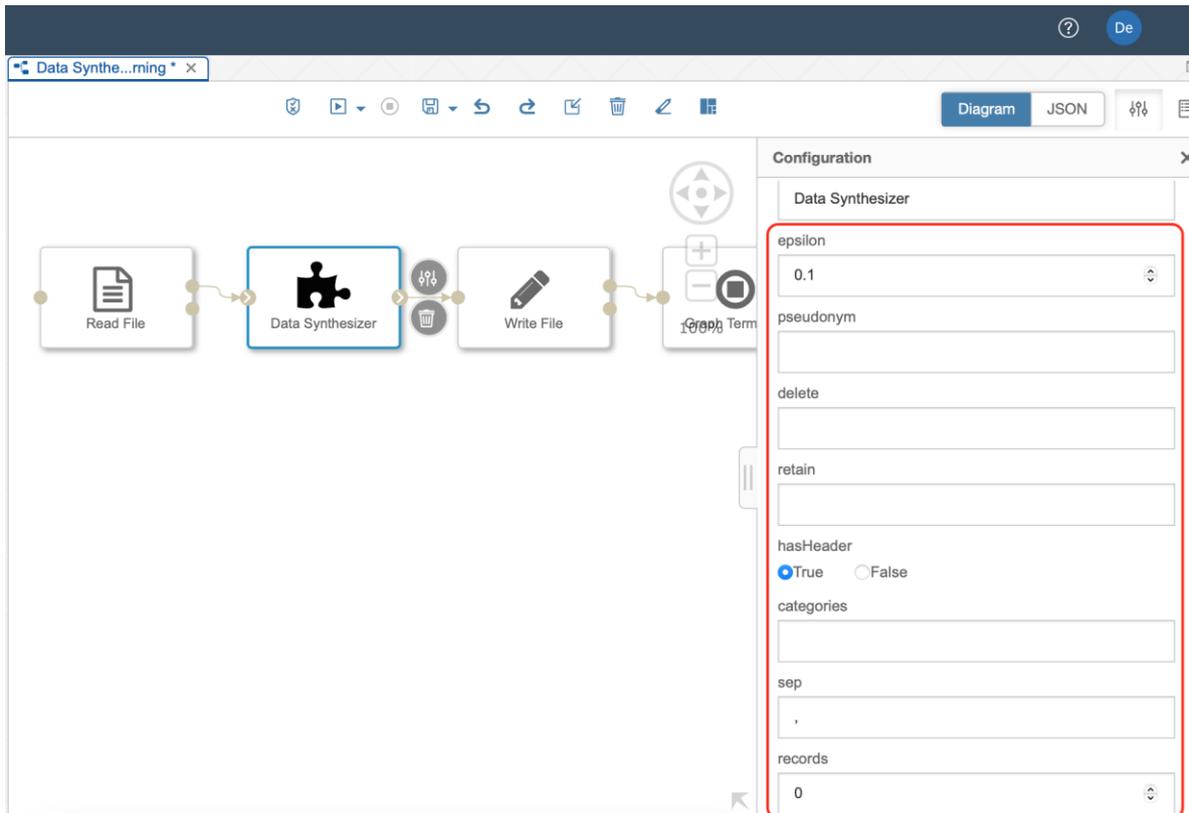


And similarly, define the file path for the synthetic dataset in the **Configuration** of “Write File” operator (as follows).



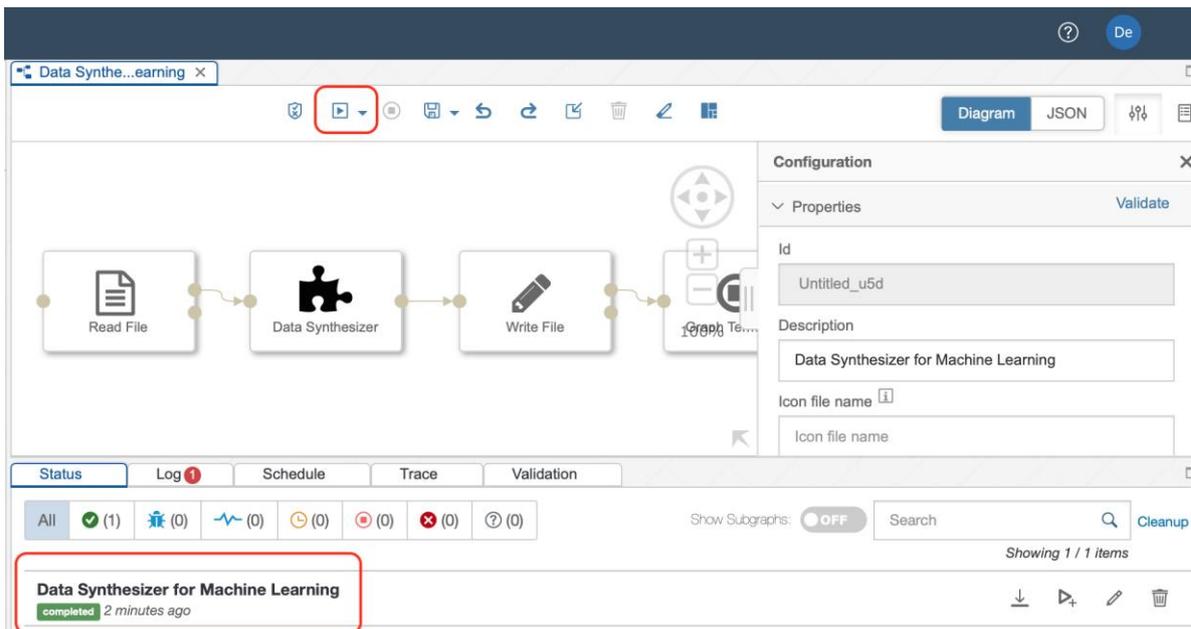
## RUN THE GRAPHS

There are several flexible configurations for users to set from for Data Synthesizer operator. All these configurations have default value, you can just run it directly.



Configuration	Description
epsilon	The privacy budget in Differential Privacy, which can be used to balance between data privacy and utility.
pseudonym	The columns to synthesize with pseudonym algorithm in the synthetic dataset. Its value is separated by comma.
retain	The columns to retain in the synthetic dataset. Its value is separated by comma.
delete	The columns to delete in the synthetic dataset. Its value is separated by comma.
hasHeader	Indicates if there is a header specified in the input CSV file. If there is no header in source csv, operator will generate new headers for it.
categories	The categorical columns user can set. Its value is separated by a comma. If not set, the operator will check in the code by itself.
sep	The field separator in the input CSV file (default: ',')
records	The records the end user want to generate in the synthetic dataset; the default value is the same with original dataset.

- Click "run" button as follows. After the **Status** shows completed, go to the workspace to check the synthesized data.



- Click into "**System Management**" on the Home Page. Navigate to **Files** tab, find the synthesized csv data, you can view the data or export data as you like.

SAP Data Intelligence System Management

Applications Users **Files** Tenant

Workspaces

My Workspace Tenant Workspace Union View

Files

Search

Name	Modified	Size
> launchpad	2020-02-11 16:02:03	
manifest.json	2021-03-16 16:02:03	
> preparation	2020-02-11 16:02:03	
synthetic-adult.csv	2021-03-16 16:02:03	1.00 KB

- Open File
- Export File
- Export as Solution
- Export as solution to solution repository
- Copy Path
- Move to Tenant Workspace
- Rename
- Delete

/files/synthetic-adult.csv