Deployments on KNIME Business Hub Guide

KNIME AG, Zurich, Switzerland

Version 1.15 (last updated on)

Table of Contents

Introduction

The KNIME Business Hub has significantly expanded the range of options for deploying workflows. In this guide, we will explore the full spectrum of opportunities provided by this innovative infrastructure to streamline data science production with KNIME. We have prepared a variety of deployment examples that serve both as demonstrations and real-world use cases, offering a comprehensive showcase of the platform's latest features. To configure these examples within the KNIME Business Hub, the first step is to initiate an installation workflow.

> ℹ️ Find all the workflows relative to this guide here.

Deployment Examples Installation

The deployment examples installation workflow is a versatile data application that can be executed both directly on KNIME Business Hub and within a local environment. This folder data area contains the complete Repository Examples folder. After inputting the Application Password and the Hub URL into the "Hub connection" component, this folder is promptly uploaded to the Hub as soon as the next component is executed. After successfully establishing the connection, the user will be presented with a list of all available teams. Subsequently, the content will be uploaded to all the teams that have been selected by the user.
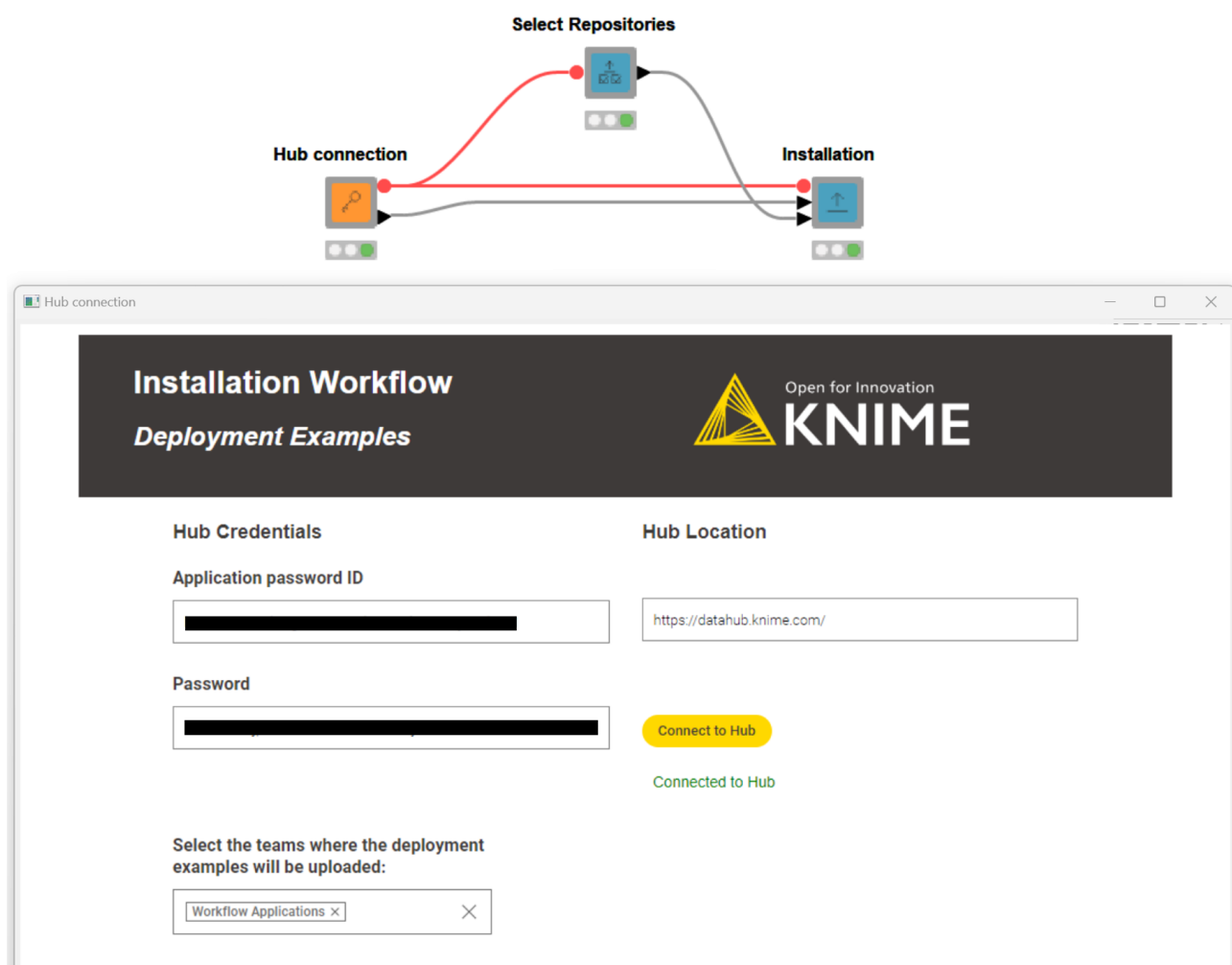
In the next step, the Installation component will initially create a space named "KNIME Business Hub Deployment Examples" within the selected teams on the Hub, followed by the upload of workflows and files to this newly created space.
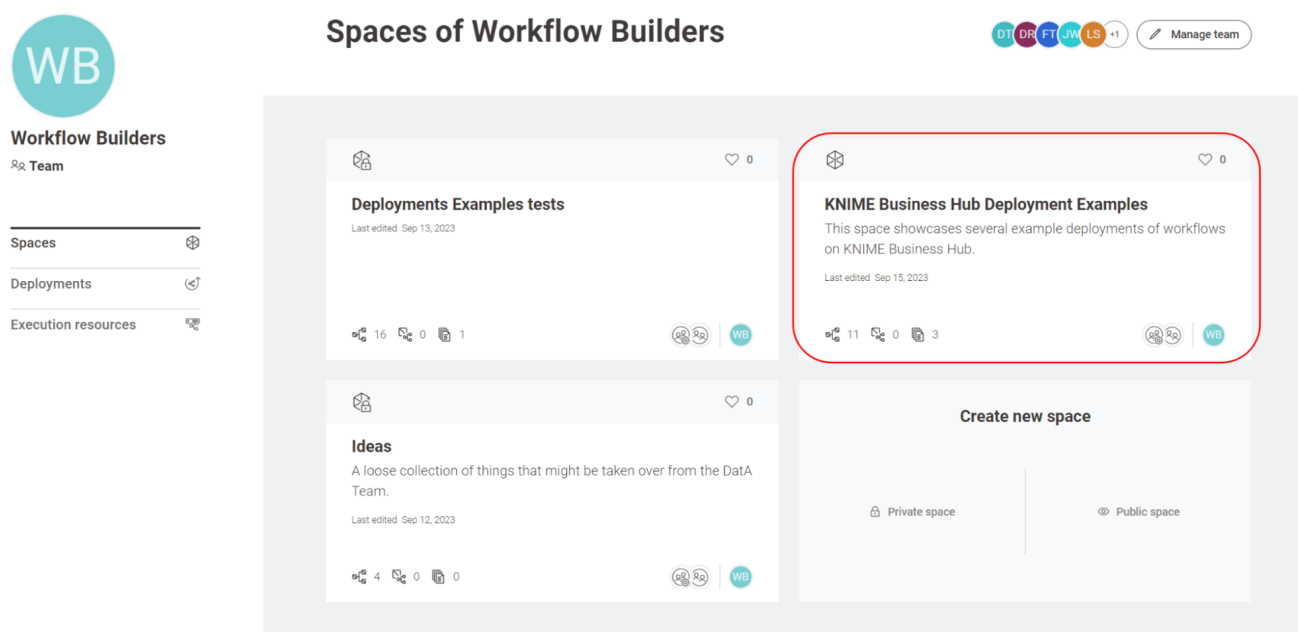
In this space, you will encounter the following organized structure:

- "Deployment Examples" Folder: This folder contains four distinct workflows, each tailored for a specific deployment type, as indicated by the subfolder names.

- "Use Cases" Folder: In contrast, the "Use Cases" folder comprises three projects with a slightly elevated level of complexity.

The workflows situated within the "Deployment Examples" folder are designed for manual deployment by the user, providing a hands-on experience. In contrast, the applications within the "Use Cases" folder are configured for fully automated deployment, offering an effortless experience for users. After the successful execution of the Installation workflow, you can expect to find the content neatly organized within the previously mentioned space. The "Installation" component's view will provide you with a comprehensive overview of the results pertaining to the automatic deployment creation for the projects contained within the "Use Cases" folder.

For an introduction to all available deployment types on KNIME Business Hub, please refer to the linked documentation. In summary, there are four primary deployment options:

- Data App: You can deploy a workflow as a Data App.
- Schedule: Schedule workflows for automatic execution at a specified time.
- Service: Create an API Service to facilitate workflow access via external calls.
- Trigger: Implement trigger-based deployments for specific events.

Deployment Examples

Item Versioning

Creating a deployment of a workflow, regardless of the deployment type, necessitates the mandatory creation of a version of the workflow. This requirement stems from the fact that users must select a specific version of the item they intend to deploy. For additional information, you can refer to the official documentation.

To create a version of a workflow, uploaded to a Hub space, follow these steps:

- Click on the "History" button, which will trigger the appearance of a "side drawer" on the left-hand side

- Within this "side drawer," locate and click the "Create Version" button

- You will be prompted to provide a name (mandatory) and an optional description for the version.

By following these steps, you ensure that a version is created and ready for deployment
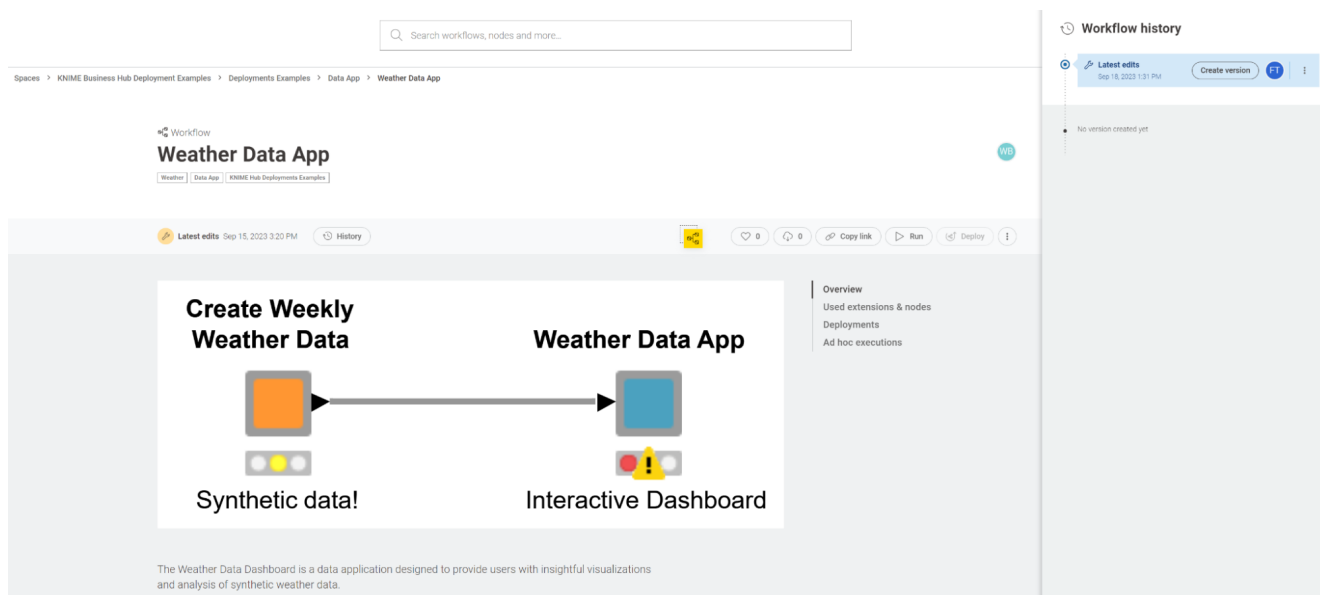
From this point onward, the step-by-step guide will assume that you have already created an item version before proceeding with the workflow deployment.
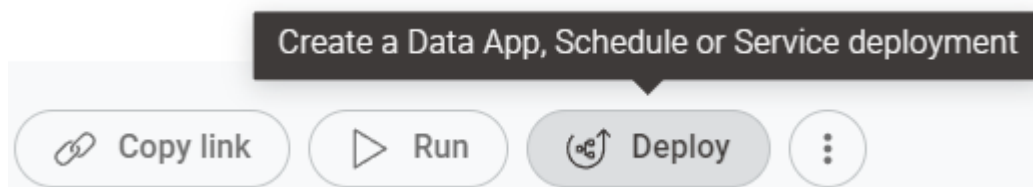
Data App - Weather Data App

Workflow explanation

The Weather Data Dashboard is a data application designed to provide users with insightful visualizations and analysis of synthetic weather data.

This user-friendly app generates synthetic weather data for various parameters, including Date and time, Sunlight Duration (minutes/hour), Rain Probability, Relative Humidity, and Temperature (°C). It presents this data in an intuitive dashboard featuring dynamic graphs (Line Plot and Bar Chart), a table view, and a Date Filter widget.

Creating a Data App Deployment: Step-by-Step Guide

Upon successful execution of the installation workflow, you will find the Weather Data App conveniently located within the Data App folder of the newly created space. After successfully creating the version, proceed by clicking the "Deploy" button, which becomes enabled once there is at least one version of the item.
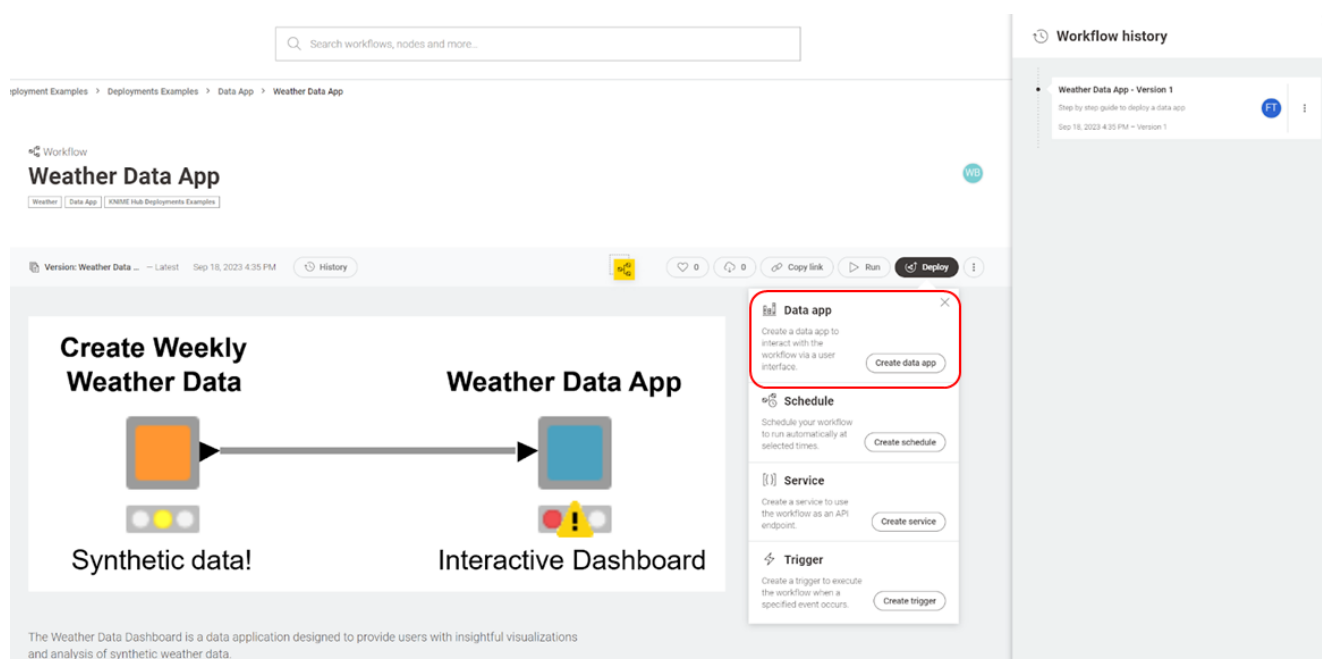
A panel will appear, facilitating the configuration of essential deployment details. Here, you can specify information, such as selecting a deployment name and description, choosing an execution context, and categorizing the data app on the DataApps Portal.

⬛ Create data app

Create a data app to interact with the workflow via a user interface.

Deployment name

Weather Data App

Version

Weather Data App - Version 1 - Sep 18, 2023 4:35 PM

Select execution context

5.1.1 General Purpose ⌄

Category

Deployment Examples

Description

The Weather Data Dashboard is a data application designed to provide users with insightful visualizations and analysis of synthetic weather data.

Workflow actions

Control basic actions of the workflow execution.

☐ Enable workflow actions

Advanced settings
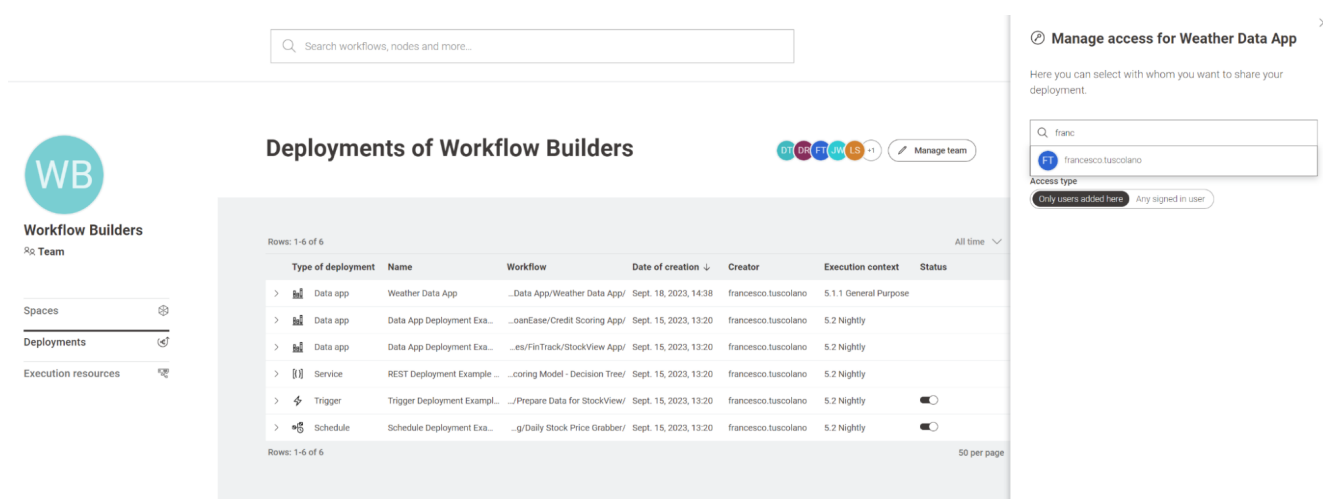
Change job lifecycle behavior, timeouts, ...   Set ＞

Once the Data App has been successfully created, you can access it via the "Deployments" page within your team's workspace. To share the Data App with other users, follow these steps:

- Click the ⋮ icon next to the Data App
- Select "Manage Access" from the menu
- Configure access permissions.

Upon completing these actions, the added users will gain visibility and access to the Data App within the Data App Portal, organized under the assigned category. For more comprehensive details regarding the Data Apps Portal, please refer to the provided documentation.


Now that the Data App has been successfully deployed, you have two convenient methods for execution:

- Data Apps Portal: Access and execute the Data App via the Data Apps Portal, especially when shared with other users. (`https://apps.my-business-hub.com/`)
- Deployment Page: Alternatively, you can run the Data App directly from the Deployment page by clicking the ⋮ icon and selecting "Run."
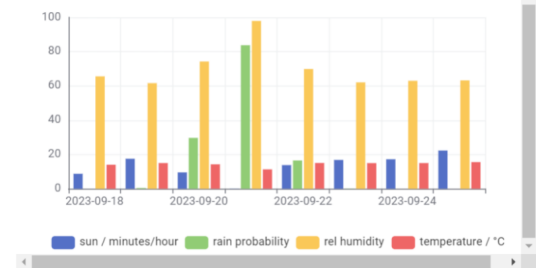
Weather Data App
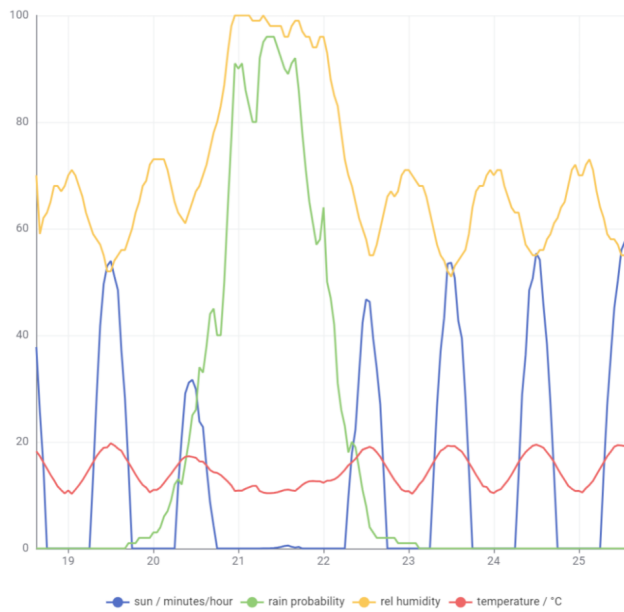
Data App - Flags Games Data App

Workflow explanation

This workflow details the creation of a captivating and interactive game data app that engages users in a flag-guessing challenge.

In this app, users are presented with flags from different nations, and their task is to identify the nation to which each flag belongs correctly.

After a series of flag guessing rounds, a score is computed, and users are offered the opportunity to submit their scores. This engaging and educational app combines entertainment with data collection.

Furthermore, this workflow includes various tips and tricks that aid in the creation of high-quality data apps that deliver an exceptional user experience.

To create a data app deployment for this workflow, you need to follow the steps outlined in the previous section, Creating a Data App Deployment.

Welcome!

Player: account:user:9eb1ae19-76bd-413a-8bd6-d5e9a1ac05d0

GuessTheCountry is an interactive game to test your knowledge around the world.
The game runs three times: in each round it shows you an image and three countries as options.
Watch out: points are given and taken!

- 3 points for a *correct* answer
- -1 point for a *wrong* answer

How many points will you be able to score?
Click **Next** and start playing!

Upload a file or use the default dataset

Select file   countries-flags-png.table

Round 1/3

Current Score: 3

Make your choice!

○ South Ossetia   ● Estonia   ○ Colombia

Estonia is correct!

Click Next to continue

← Back   Cancel   Next

Round	Country	Number of attempts	Points
1	Antigua and Barbuda	1	-1
1	Mali	1	-1
2	Singapore	1	3
3	Sweden	1	3

First Name

first name

Last name

last name

Age

25

Country

Albania

Email Address

example@email.com

☑ I want to subscribe to the newsletter

From 1 to 5, how much would your recommend our game to a friend?

○ 1   ○ 2   ○ 3   ○ 4   ● 5
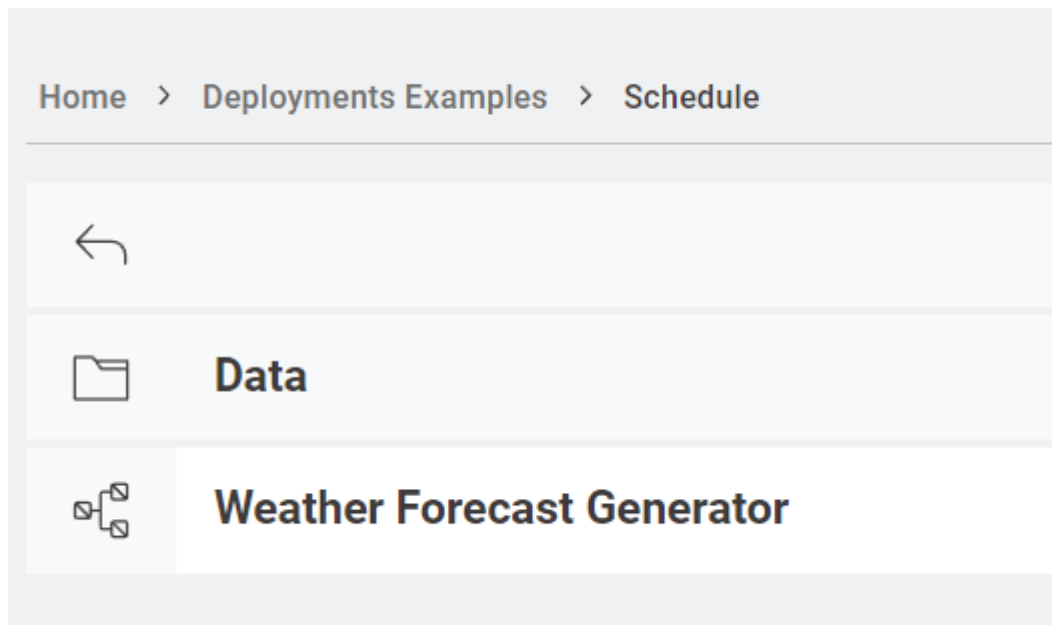
Schedule - Weather Forecast Generator

Workflow explanation

The workflow is an automated solution designed to generate synthetic weather data on a weekly basis and store it in a designated folder named "Data."

It generates synthetic weather data for various parameters, including Date and time, Sunlight Duration (minutes/hour), Rain Probability, Relative Humidity, and Temperature (°C).
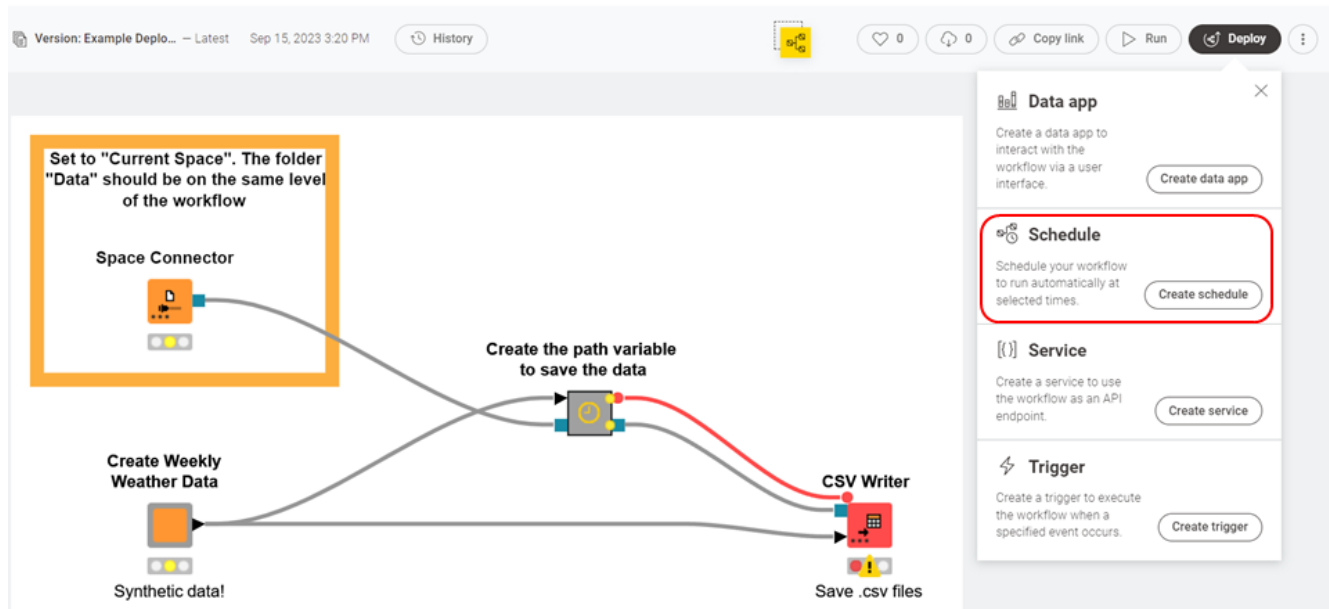
By scheduling this workflow, you ensure a steady supply of weather data without manual intervention.

Data is saved in `.csv` format after each execution.
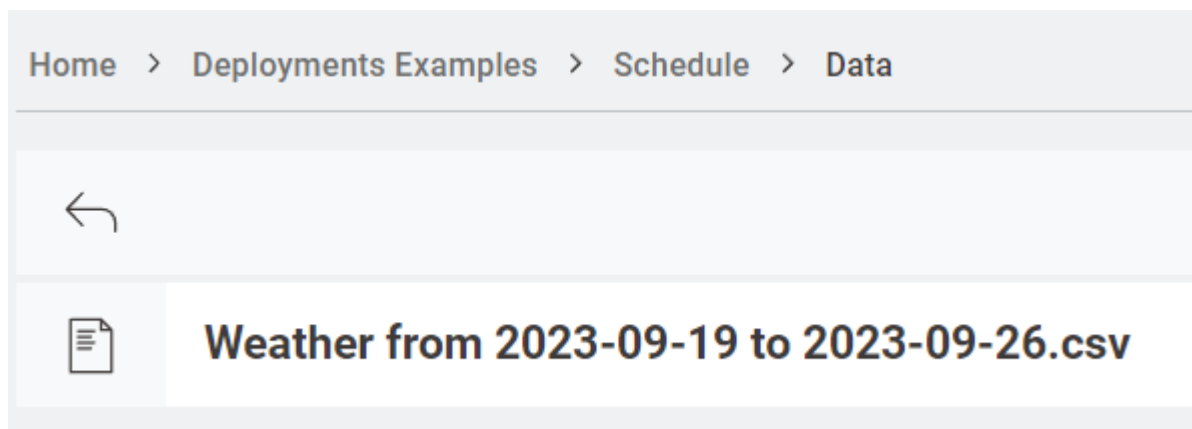
Creating a Schedule Deployment: Step-by-Step Guide

By using a schedule, users can specify a time for workflows to run automatically, eliminating the need for manual intervention.

We offer various options for specifying when the workflow will run. For instance, you can configure the workflow to run every seven days from the current date and continue running for a year.

Schedule options

Define when the workflow should be executed.

Initial execution

| 2023-09-19 📅 | 11 ⌃⌄ | : | 02 ⌃⌄ | : | 00 ⌃⌄ |

☑ Repeat execution

(**Every**   Start times)

| 7 ⌃⌄ | Days ⌄ |

Schedule ends date

(Never   **Date**)

| 2024-09-20 📅 | 11 ⌃⌄ | : | 02 ⌃⌄ | : | 00 ⌃⌄ |

Clicking on "Set Schedule Details" provides additional capabilities, allowing you to include or exclude specific months or days, as well as specifying precise execution timeframes.

Recurring on

Days of the week

| M ⌄ | T ⌄ | W ⌄ | T ⌄ | F ⌄ | S ⌄ | S ⌄ |

Invert selection of days of the week

Days

1 ⌄	2 ⌄	3 ⌄	4 ⌄	5 ⌄	6 ⌄	7 ⌄
8 ⌄	9 ⌄	10 ⌄	11 ⌄	12 ⌄	13 ⌄	14 ⌄
15 ⌄	16 ⌄	17 ⌄	18 ⌄	19 ⌄	20 ⌄	21 ⌄
22 ⌄	23 ⌄	24 ⌄	25 ⌄	26 ⌄	27 ⌄	28 ⌄
29 ⌄	30 ⌄	31 ⌄	Last day of the month ⌄

Invert selection of days

Months

| Jan ⌄ | Feb ⌄ | Mar ⌄ | Apr ⌄ | May ⌄ | Jun ⌄ |
| Jul ⌄ | Aug ⌄ | Sep ⌄ | Oct ⌄ | Nov ⌄ | Dec ⌄ |

Invert selection of months

Timeframes

Between   and

| 00 ⌃⌄ | : | 00 ⌃⌄ | | 23 ⌃⌄ | : | 59 ⌃⌄ |

You can always edit the schedule by accessing it on the "Deployment" page.

Service - Prime Numbers Data App

Workflow explanation

To achieve the goal of integrating the service into other workflows, data apps, or clients, we will showcase the use of two workflows in this demonstration.

The first workflow, named "Prime Number," will be deployed as a service. Simultaneously, the "Prime Number Data App" will invoke this workflow during its execution to retrieve information about the prompted prime number.

Creating a Service Deployment: Step-by-Step Guide

To initiate the service deployment process, click the "Deploy" button and select "Create Service."

The configuration process only necessitates a few specified parameters, such as the deployment name, execution context, and version selection if multiple versions are available.

After the deployment has been created, you can open it and read the API specifications for invoking the service on the Deployments page within the Hub team.

It's crucial to remember that the deployment ID is key information for making REST API calls to the specific deployment.

It is important to configure access permissions for the service; otherwise, users may be unable to access it. To accomplish this, you can click on "Manage access" and add all the users who require access to the service. Alternatively, you can choose to permit access for any signed-in user. This is the same procedure to grant access to a data app.

Once the Service deployment has been created, you can proceed to run the Prime Number Data App, which can also be deployed. This data app requires the Deployment ID as input to fetch the service for invocation. Once the Service has been successfully connected to the data app, you can proceed to the next page.

Prime Numbers Data App

Hub Credentials

Application password ID

bLIQR9tKVtCdwzgCiTRzOWvlAYw91hHvU65joQ3bsIM

Password

Br6BkxCzXjpWtYOTXJman8Nnbm_ZyKKCf-5h52NMvHbLfB22FI8Zs8izp_Zuc_0AR12Bk0o1efS

Deployment ID

rest:511dca10-0b0d-488c-aec7-d4050a70382e

Fetch REST deployment

Hub Location

https://datahub.knime.com/

Connect to Hub

Connected to Hub

You fetched the correct deployment.

This straightforward data app demonstrates how the user interface collects input, sends it to the service, and receives the determination of whether the number is prime or not in response.

Prime Numbers Data App

Test if the input number is a prime number

Input number:

3

Prime Number?

Results from the Service Deployment

3

is a prime number

Showing 1 to 1 of 1 entries

Service - ETL and Integrated Deployment

Workflow explanation

This workflow outlines the process of using the Integrated Deployment extension to create a brand-new production workflow that can be deployed as a service.

The workflow retrieves data from the data area and conducts essential preprocessing tasks. Thanks to the Integrated Deployment Extension, these preprocessing steps are encapsulated and replicated within a production workflow, making them readily available for external use when deployed as a REST endpoint. The resulting production workflow is saved within the

"Production Folder (Integrated Deployment)" which is located at the same level as the original workflow.

The goal is to create a deployable service that can effectively process a data table when it's invoked correctly.
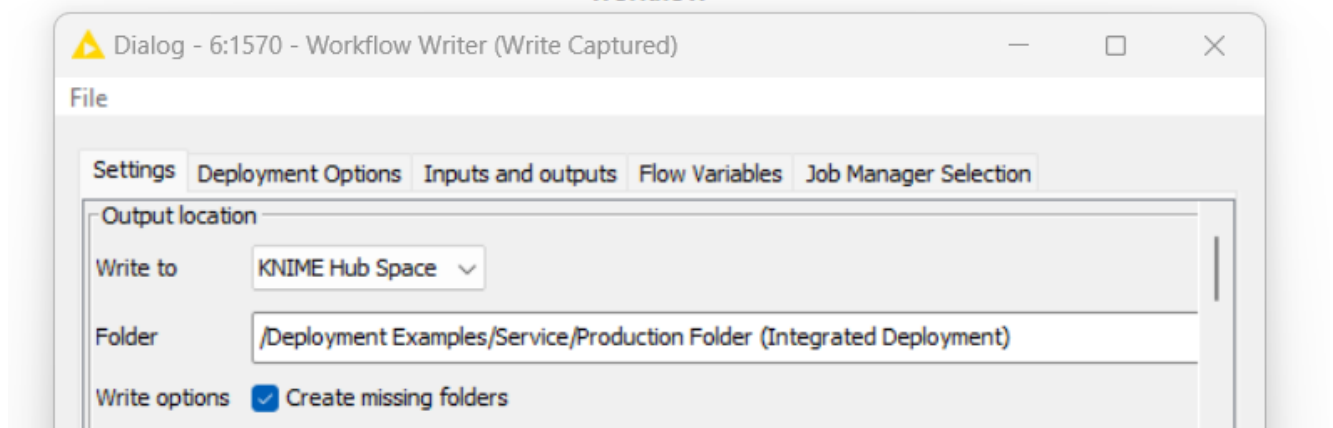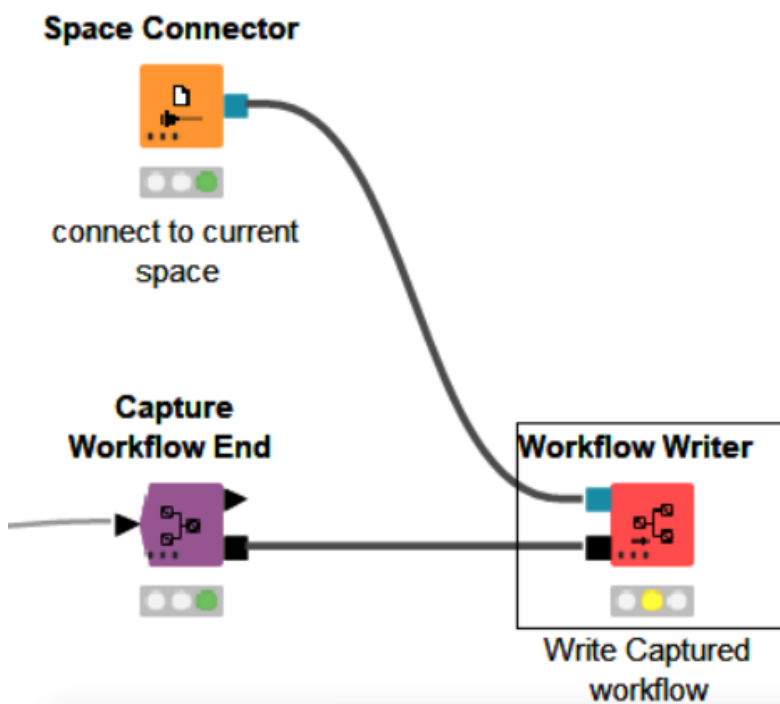
To start, execute the "ETL and Integrated Deployment" workflow. You will observe that the

Workflow Writer node will save the relevant portion of this workflow in the Production Folder.

The Workflow Writer node is configured to use the Workflow Service nodes as input and output.
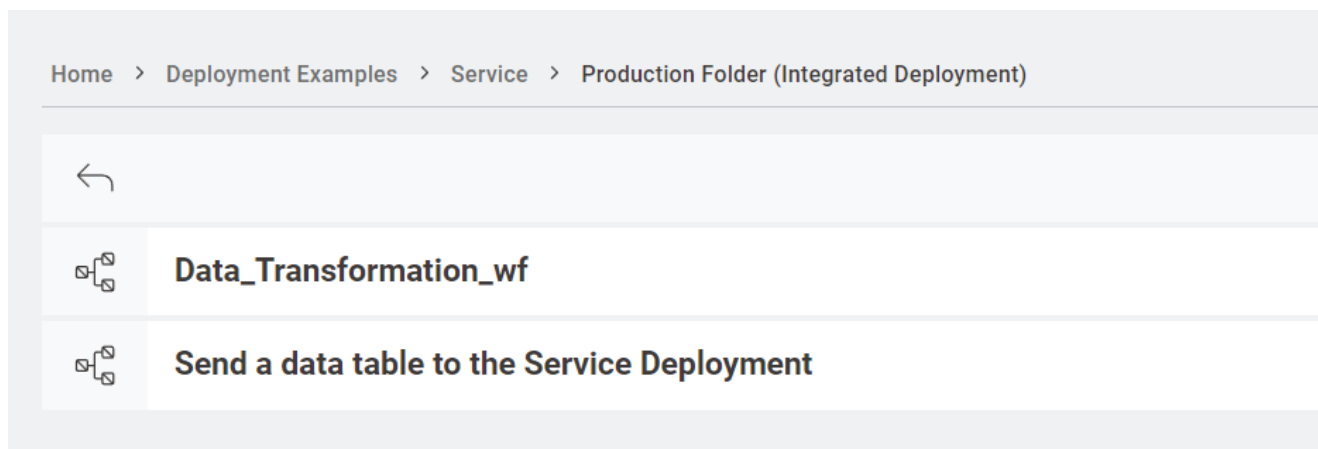
Transmit a table to the Service Deployment

If you want to process a new table using this ETL workflow, once you've created a Service Deployment for the "Data_Transformation_wf," the most straightforward approach is to utilize the Call Workflow Service node along with the KNIME Hub Authenticator node.

Open the "Send a data table to the Service Deployment" workflow directly on the Hub, then browse the available deployments and select the one that's connected to the ETL workflow generated using Integrated Deployment.
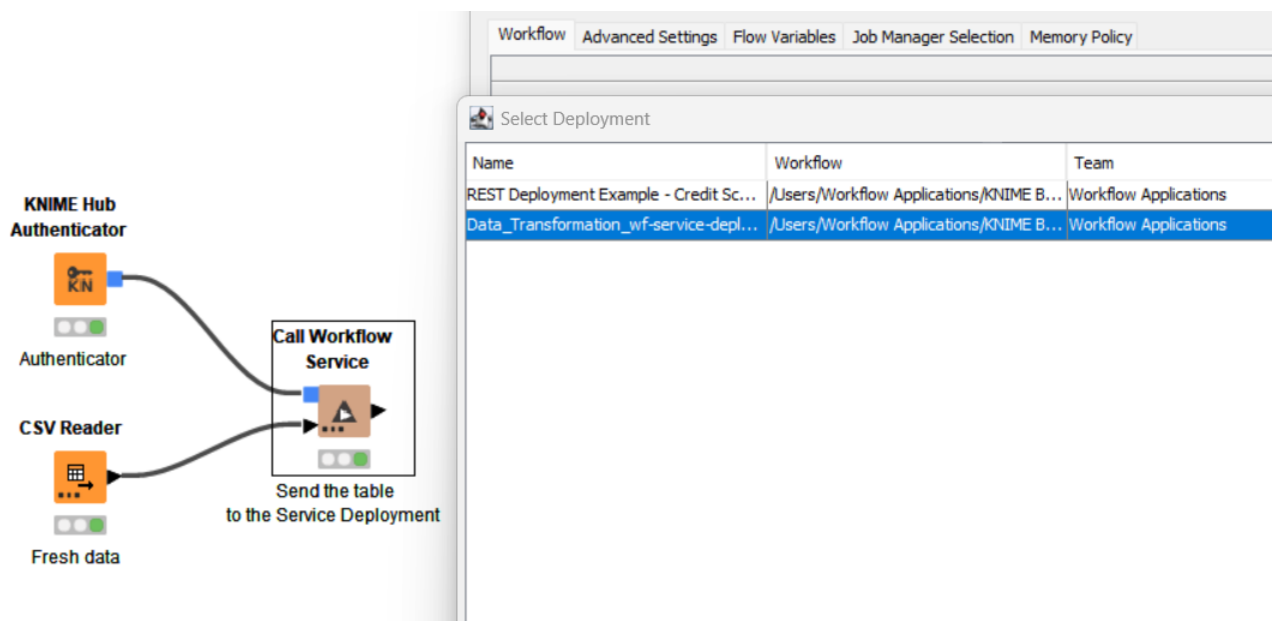
After making your selection, you should click the "Adjust node ports" button to create the input port that sends the data.
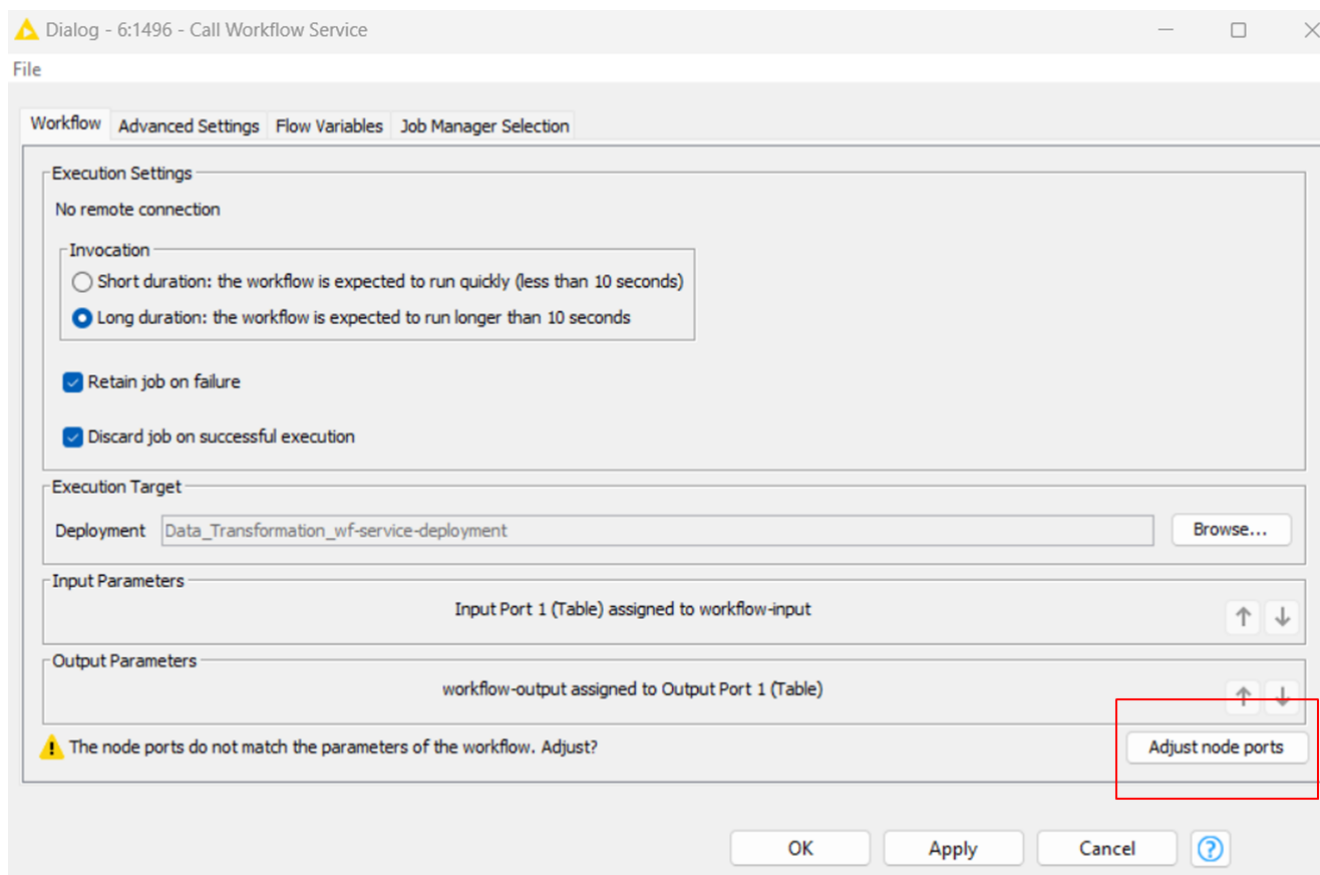
Trigger - Automatic Version Creator

Workflow explanation

This KNIME workflow is designed to automate the creation of item versions when a file is added, modified, or updated to a designated space on KNIME Business Hub.

It leverages the power of KNIME's automation capabilities to streamline the process of managing and versioning files within a shared space. The workflow is made to be deployed on KNIME Business Hub and is triggered whenever a new file is uploaded to the space where the workflow is stored.

The container Input (Repository Event) conveys information about the new file to the workflow. These parameters are then utilized to automatically generate a version of the file by means of the Version Creator node.

Creating a Trigger Deployment: Step-by-Step Guide

To initiate the service deployment process, click the "Deploy" button and select "Create Trigger."

In addition to the standard deployment parameters like name and execution context, the critical aspect here is to specify the type of event that serves as the trigger for the workflow. These events can encompass actions such as file addition, deletion, update, or the creation of a new version for any data file. Furthermore, users have the option to specify a sub-path within the chosen target space.

⚡ Create trigger

Create a trigger to execute the workflow when a specified event occurs.

Deployment name

Automatic Version Creator

Version

Example Deployment Version - Automatic Version Creator - Sep 15, 2023 3:20 PM

Execution context

5.1.1 General Purpose ⌄

Trigger options

Define events that will trigger execution.

Select space

KNIME Business Hub Deployment Examples ⌄

Item type

(**File** Folder)

File type

☑ Workflow

☑ Component

☑ Data file

Action

☑ Added

☑ Modified

☐ Deleted

☐ New version created

Additional filtering

(**No** By subfolder)

With these settings in place, a version will be automatically created each time there is a modification or a new upload within the KNIME Business Hub Deployment Examples space.
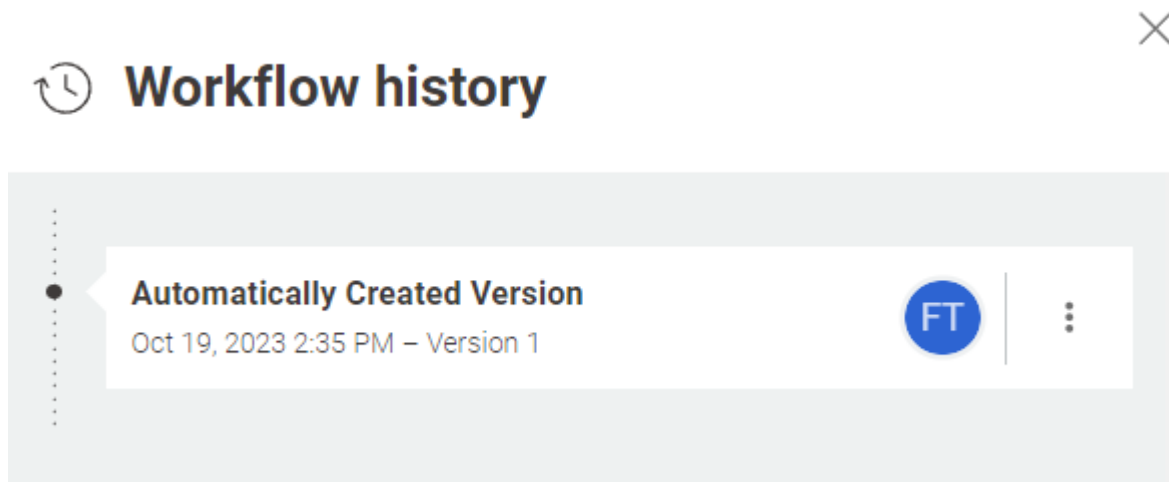
Use Cases

Use cases are already ready to use as the installation workflow takes care of the setup.

Loan Ease

This use case consists of two workflows that allow the achievement of different goals:

- The Credit Score App
- The Credit Score Model - Decision Tree

The "Credit Scoring App" is designed to determine the creditworthiness of an individual applying for a loan. It takes a set of input variables and uses a deployed "Credit Scoring Model- Decision Tree" workflow as a service. This workflow is deployed on KNIME Business Hub as a data app, providing an easy-to-use interface for users to assess their credit eligibility.

The "Credit Score Model - Decision Tree" serves as a service deployed to provide credit scoring predictions to the Credit Scoring Data App.
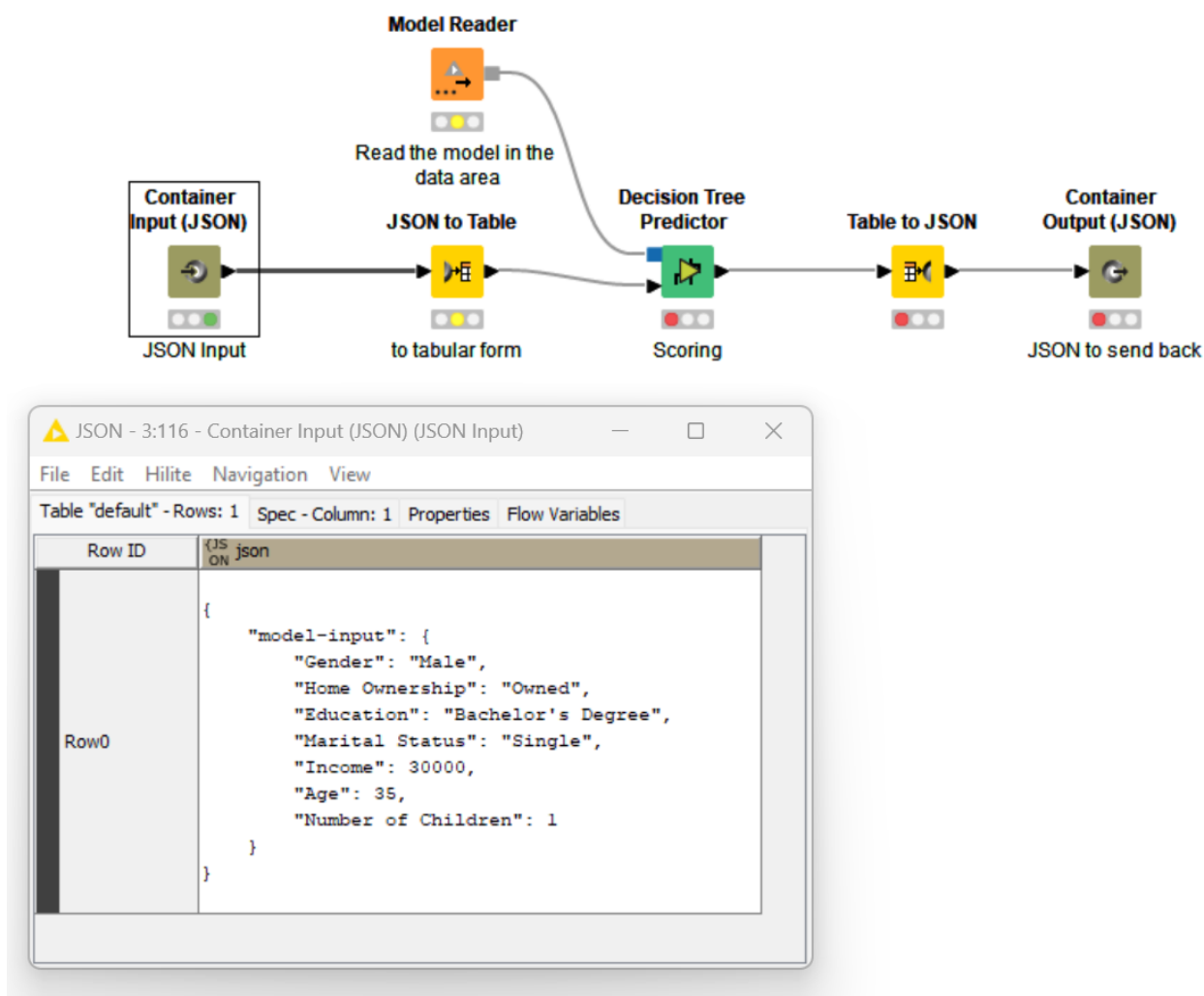
It utilizes a decision tree model stored in the data area to score incoming data, generating predictions and estimated probabilities for creditworthiness.

The results are sent back to the data app for display and decision-making.

The workflow begins when a user accesses the Credit Scoring App through the KNIME Business Hub.

The user is presented with a user-friendly interface to input their information.

The data app collects input variables from the user, which are required for credit scoring through seven widgets:

1. Gender

2. Home Ownership

3. Education

4. Marital Status

5. Income

6. Age

7. Number of Children

After, it constructs a JSON body with the user-provided input variables, that is sent to the predictive model for the scoring.

FinTrack

This data app encompasses three deployment types, each serving distinct objectives.

The 'Daily Stock Price Grabber' downloads stock price data and runs as a scheduled execution every day at 08:00 a.m., excluding weekends.

Home > Use Cases > FinTrack > Data Processing

↩

📁 **Data Validation**

📁 **Staging Area**

📁 **StockView Data Input**

🔗 **Daily Stock Price Grabber**
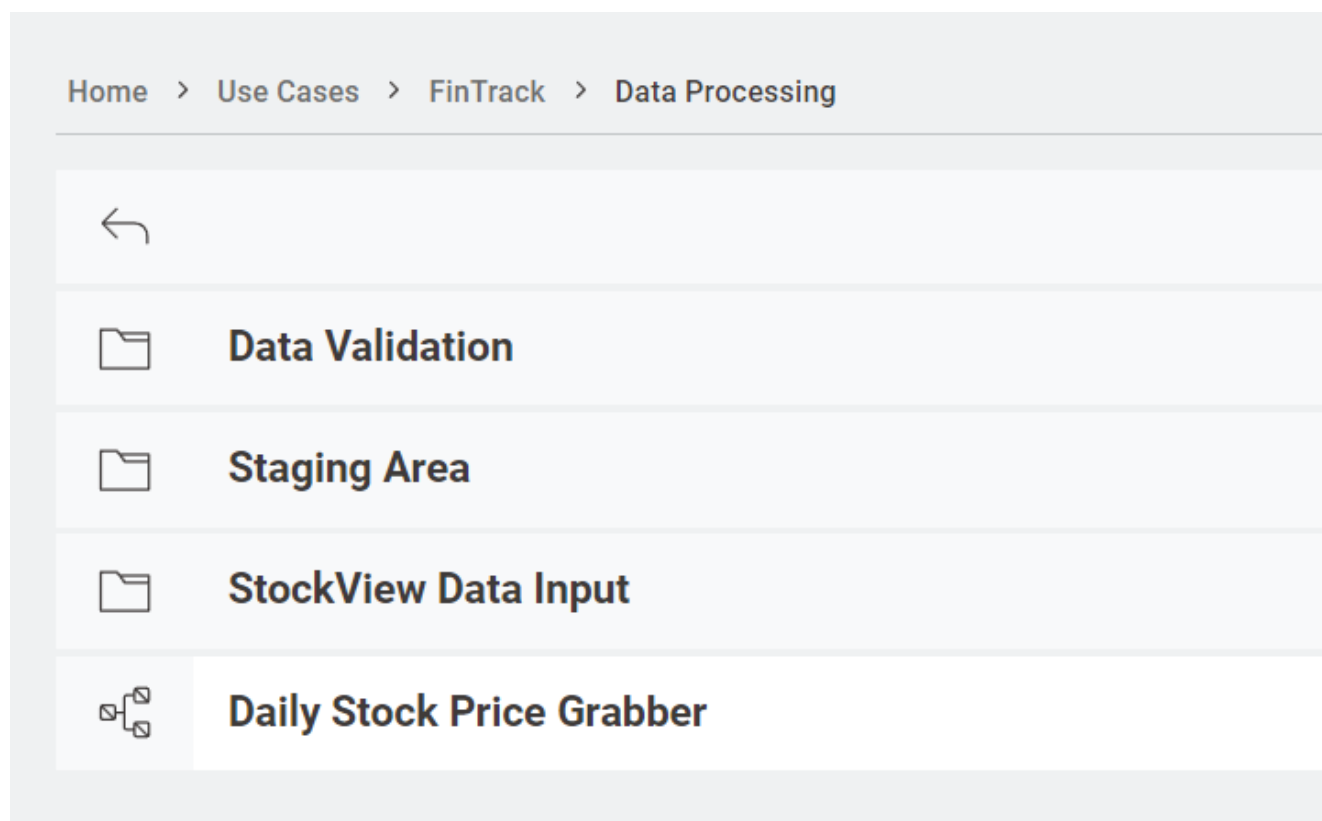
You can select the stocks to download by entering their tickers separated by commas in the workflow configuration. If you wish to add more stocks, you will need to create a new schedule.

In the Staging Area, the app stores historical prices in separate data files

Home > Use Cases > FinTrack > Data Processing > Staging Area

AAPL.csv

GOOG.csv

The data files are processed to be formatted in a way that is compatible with the StockView App. Before this processed data file is transferred to the StockView Data Input folder, a trigger deployment initiates a basic validation process. If this validation process is successfully completed, the data is then moved to the designated location (The StockView Data Input folder).

Home > Use Cases > FinTrack > Data Processing > Data Validation

Validate File

Home > Use Cases > FinTrack > Data Processing > StockView Data Input

Processed-data.csv

The dashboard reads data from the 'StockView Data Input' folder.
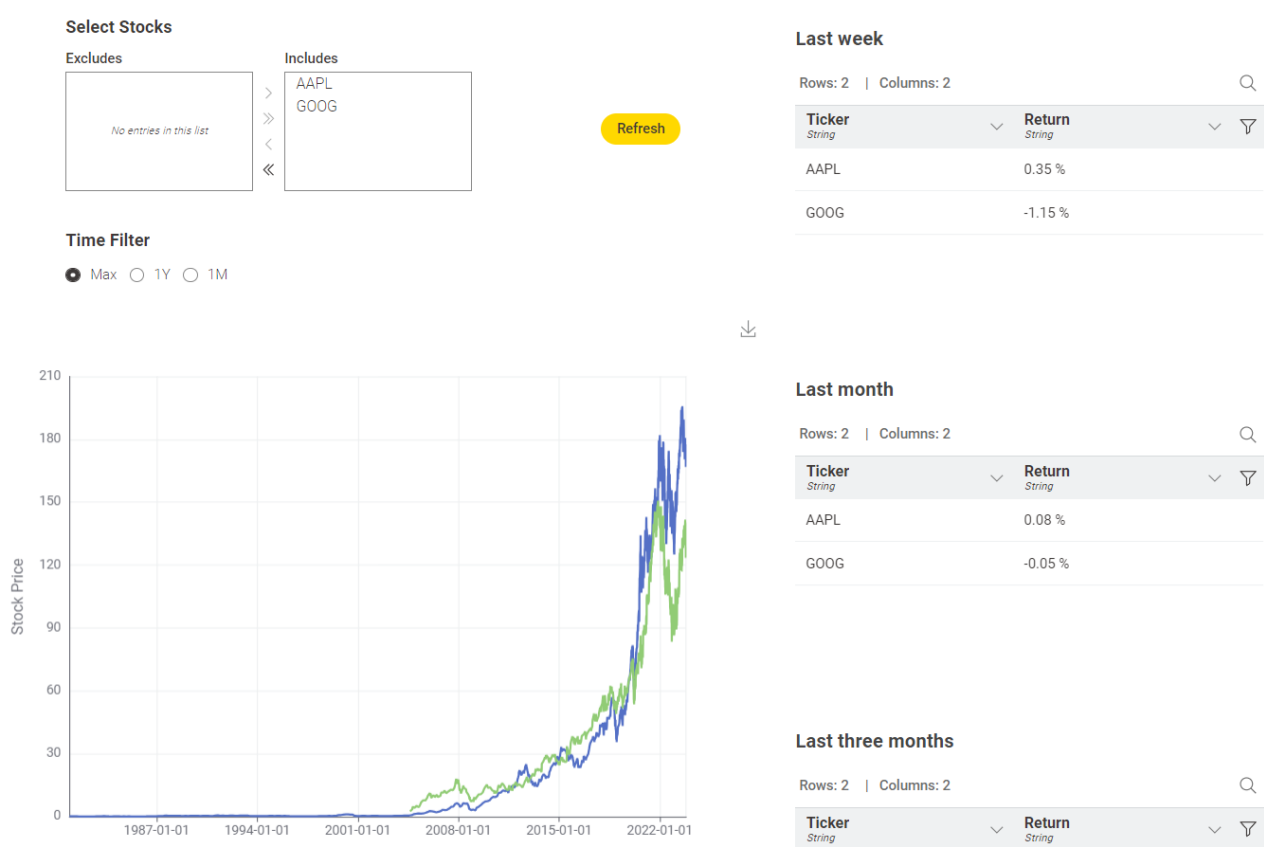
Market Sizing

In addition to the deployment options mentioned earlier, another valuable KNIME feature for seamlessly integrating different workflows is the "Call Workflow" node. In this particular use case, we deploy a data app that leverages another workflow. The data app invokes this workflow to process data using the "Call Workflow (Table Based)" node and retrieve specific data into the data app.
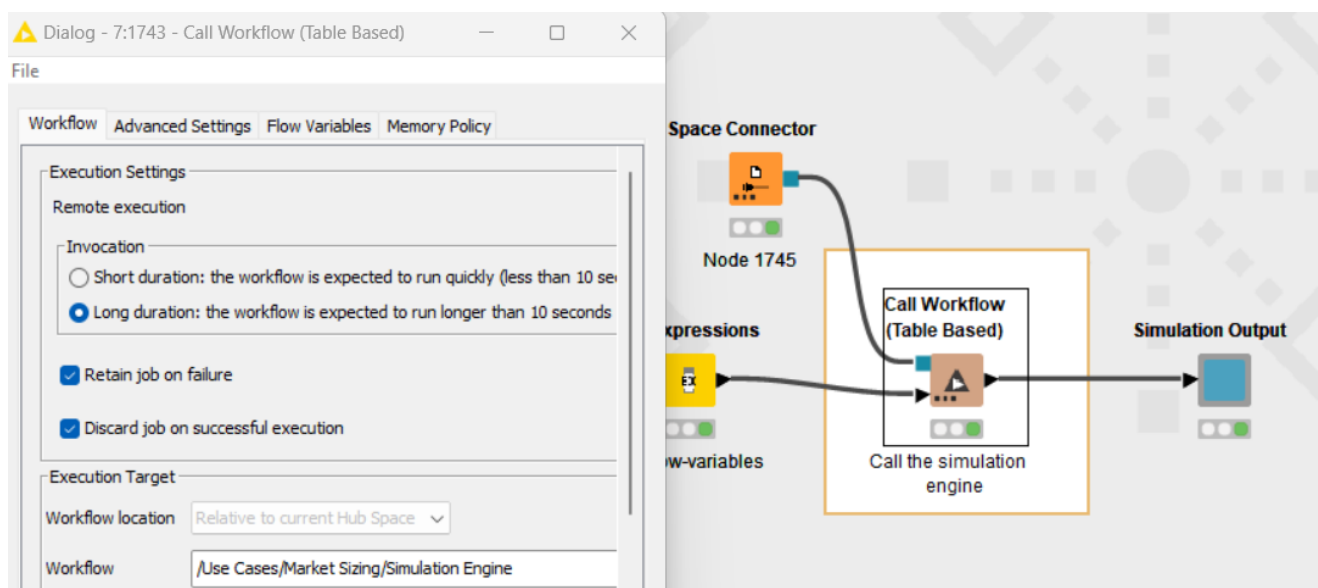
This workflow is designed to estimate market size and potential revenues for a product or service by harnessing the power of Monte Carlo simulation techniques. The data app uses some input parameters including product or service unit price, awareness, consideration, purchase intention, and the size of the target population, culminating in the generation of reasonable market size projections.

It all begins with the conceptualization of a theoretical purchase process undertaken by potential customers. After defining the size of the target population, which represents the group of customers most likely to embrace the product or service, the simulation makes an assumption regarding the number of customers.

1. Awareness is the first stage of the marketing funnel. At this stage, potential customers become aware of your product or service's existence. They may have encountered your brand through various channels, such as advertising, social media, word-of-mouth, or online search.

2. Consideration is the second stage of the marketing funnel. Once people are aware of your product or service, they start considering it as a potential solution to their needs or problems. They begin to research and gather information about your offering and may compare it with alternatives.

3. Purchase intention is the final stage of the marketing funnel before a potential customer becomes an actual customer. At this point, the individual has a strong interest in your product or service and is leaning toward making a purchase.

The formula for estimating the potential number of customers is as follows:

```
Number of customers = Target Population size x % Awareness x %
Consideration x % Purchase Intention.
```

This formula closely follows the logical progression of a market funnel.

In the subsequent step, users are prompted to specify the range of variability for the marketing factors that exert influence on the ultimate outcome. These defined ranges serve as the parameters for the Triangular Distribution, a probabilistic process employed to generate random combinations of these marketing factors. This is accomplished through a series of repeated sampling experiments, closely aligned with the fundamental principles underlying Monte Carlo simulations.

You have the ability to set the boundaries for the generated values using the parameters of the triangular distribution.
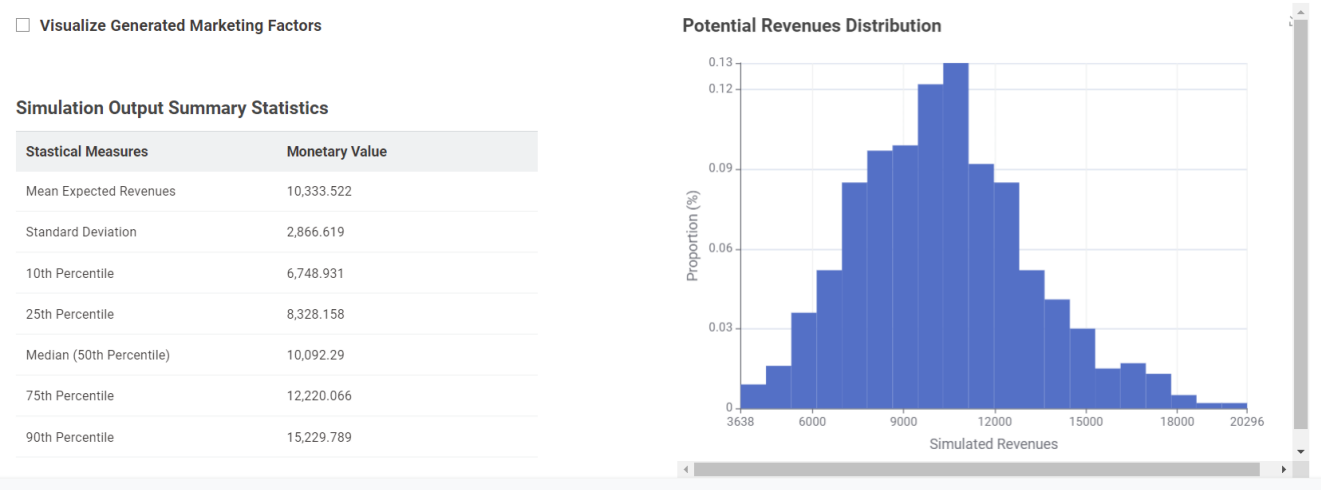

Random values for awareness, consideration, and purchase intention will be generated through a triangular distribution process.

The simulation engine conducts 1000 iterations, randomly combining potential outputs using the following generative process:
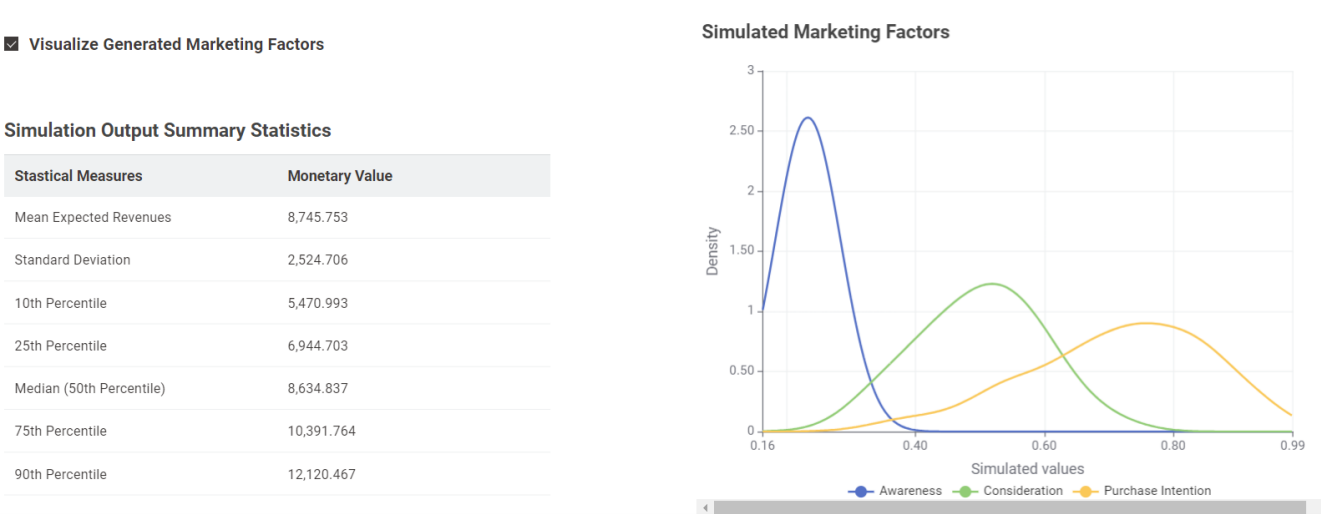
- Awareness ~ Triangular(a, b, c)

- Consideration ~ Triangular(a, b, c)

- Purchase Intention ~ Triangular(a, b, c)

The final formula for computing the simulated revenues is as follows:

```
Revenues = Target Population x (Awareness ~ Triangular(a, b, c)) x (Consideration ~
Triangular(a, b, c)) x (Purchase Intention ~ Triangular(a, b, c)) x Unit Price
```

☐ **Visualize Generated Marketing Factors**

Simulation Output Summary Statistics

Stastical Measures	Monetary Value
Mean Expected Revenues	10,333.522
Standard Deviation	2,866.619
10th Percentile	6,748.931
25th Percentile	8,328.158
Median (50th Percentile)	10,092.29
75th Percentile	12,220.066
90th Percentile	15,229.789

Potential Revenues Distribution

You can examine the distribution of simulated values defined by your input, which serves as the basis for revenue calculations in each iteration, by selecting the checkbox.

☑ **Visualize Generated Marketing Factors**

Simulation Output Summary Statistics

Stastical Measures	Monetary Value
Mean Expected Revenues	8,745.753
Standard Deviation	2,524.706
10th Percentile	5,470.993
25th Percentile	6,944.703
Median (50th Percentile)	8,634.837
75th Percentile	10,391.764
90th Percentile	12,120.467

Simulated Marketing Factors

KNIME AG
Talacker 50
8001 Zurich, Switzerland
www.knime.com
info@knime.com