

Utilities Customer Micro-Segmentation Industry Accelerator

The Utilities Customer Micro-Segmentation accelerator includes a structured set of sample data science assets. The glossary provides the information architecture that you need to segment your customers based on commonalities in client lifestyle and their engagement behaviors. Your data scientists can use the sample notebooks, segmentation and predictive models, and dashboard to accelerate data preparation, machine learning modelling and data reporting. Since there aren't any predefined segments we will need to discover those segments within the underlying data structure. Once that structure is found the clients will be assigned to their respective clusters. The segmentation helps you understand how groups of customers differ from one another.

Tip: Download the PDF of these instructions from the Data assets section on the **Assets** page so you can keep these instructions open while you work.

- [Instructions](#)
- [Sample data assets](#)
- [Notebooks](#)
- [R Shiny dashboard](#)
- [Sample business glossary](#)

Instructions

Follow these steps to implement the industry accelerator:

1. Create a deployment space, which you'll need when you run the notebooks. From the navigation menu, select **Deployments > View All Spaces > New deployment space** and then complete these steps:
 1. Enter the name **Utilities Customer Micro-Segmentation Space**.
 2. Select your Watson Machine Learning service. If you don't have one, from the **Select machine learning service** pull-down menu, select **Create a new machine learning service**, provision one, and then click the **Create** button.
 3. Click **Create** to create the deployment space.
2. Navigate back to the accelerator project, select the **Assets** tab and scroll to the **Notebooks** section.
3. Edit the **1-model-training** notebook by clicking the edit icon that looks like a tiny pencil next to the notebook name. This notebook prepares the data, builds ML models, and deploys the model. Follow the instructions in the notebook to configure and step through running it.
4. Edit and run the **2-model-scoring** notebook. This notebook deploys data assets and a model scoring function.
5. Run the dashboard from RStudio console by completing these steps:
 1. Download the `utilities-customer-micro-segmentation-analytics-dashboard.zip` file from the Data assets section of the **Assets** page. If you don't see the file, click **View All** to display the full list of assets.
 2. Click **Launch IDE > RStudio** on the menu bar.
 3. In the **Files** pane, select the **Upload** toolbar button and upload the `utilities-customer-micro-segmentation-analytics-dashboard.zip` file into RStudio.

4. Navigate to `utilities-customer-micro-segmentation-analytics-dashboard`, select the `app.R` file, and click the **Run App** toolbar button to launch the dashboard. If you see a warning message that certain packages are not installed, you can ignore it because the packages will be installed first time you run the app.
5. Once the app has launched, you can perform model scoring in real time by entering your API key and selecting your region on the **Client View** tab.
6. Optional. You can import the glossary of business terms into Watson Knowledge Catalog to get started on data governance. With the **Lite plan** only 5 Business terms can be imported. You must have [the permission to create governance artifacts](#).
 1. Download the `utilities-customer-micro-segmentation-glossary-categories.csv` and `utilities-customer-micro-segmentation-glossary-terms.csv` files from the Data assets section of the **Assets** page.
 2. Navigate to **Governance > Categories**.
 3. Click **Add Category > Import From File**.
 4. Import the `utilities-customer-micro-segmentation-glossary-categories.csv` file. Select **Replace all values** as your merge option.
 5. Navigate to **Governance > Business Terms**.
 6. Click **Add Business Term > Import From File**.
 7. Import the `utilities-customer-micro-segmentation-glossary-terms.csv` file. Select **Replace all values** as your merge option.
 8. Once the Import completes successfully, click on **Go to task** and then click **Publish** in the next page.
 9. Navigate to **Governance > Categories > Industry Accelerators** to explore the business terms.

Note: Importing the accelerator and running the notebooks and dashboard consumes approximately 1.5 - 2 Watson Studio CUH and 3 - 3.5 Watson Machine Learning CUH.

Sample data assets

These sample data files that act as dimensional and fact tables are included in the project on the **Assets** page:

'Customer Micro-Segmentation Input.csv' : Customer demographic data, historical energy usage data and the answers provided to a survey on lifestyle & sustainability questionnaire.

Additionally, there is another dataset created via the analytics project :

'model output summary.csv' : Consolidated prepped data after cluster Mapping for Exploratory data analysis and data visualization in the R shiny dashboard.

Notebooks

Follow the instructions in the first few cells of the notebook to configure the project token and API key. Then run the notebooks step-by-step.

- **1-model-training:** This notebook performs the following functions:
 - Load data
 - Prepare and clean data
 - Build clusters for lifestyle and customer engagement, analyze and visualize the data

- Create a Watson Machine Learning based deployment space
- Store the models in the deployment space and deploy the models
- **2-model-scoring:** This notebook performs the following functions:
 - Get the deployment space and deployments
 - Deploy the data assets
 - Create and deploy a function for assigning new data to lifestyle and customer engagement clusters
 - Call the function and return cluster assignments

R Shiny dashboard

The R Shiny dashboard displays model insights, customer summaries and scores new data. The dashboard has the following tabs:

- **Model Insights :** Displays the overall customer segments based on lifestyle and customer engagement and charts that show the distribution of each feature within each segment.
- **Client View :** Targets individual client information, displays segmentation related attributes for the selected customer. It provides the option to run the model scoring webservice which assigns the customer to a segment.

Sample business glossary for use with Watson Knowledge Catalog

Optionally, you can import the glossary of business terms into Watson Knowledge Catalog to get started on data governance.

The `utilities-customer-micro-segmentation-glossary-categories.csv` file defines the main and sub categories for the business terms.

The `utilities-customer-micro-segmentation-glossary-terms.csv` file defines the business terms, category of the business terms and their Related Terms/Part of Terms, if applicable.

Note: This optional step requires the Watson Knowledge Catalog service instance.

Terms and Conditions

This project contains Sample Materials, provided under this [license](#).

Licensed Materials - Property of IBM.

© Copyright IBM Corp. 2020, 2021. All Rights Reserved.

US Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.