

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.

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

Instructions

Follow these steps to implement the industry accelerator:

1. Navigate to the **Assets** tab and scroll to the **Notebooks** section.
2. 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 step through the execution. Alternatively, you can run the notebook from the Jobs tab by executing **1-model-training-notebook-job**.
3. Edit and run the **2-model-scoring-and-app-deployment** notebook. This notebook deploys data assets, a model scoring function and the r-shiny dashboard. It also generates a URL to launch the r-shiny dashboard. Alternatively, you can run the notebook from the Jobs tab by executing **2-model-scoring-notebook-job**.
4. Launch the r-shiny dashboard from one of the following ways.
 - Open the URL generated by deploying r-shiny dashboard in **2-model-scoring-and-app-deployment** notebook.
 - Navigate to **Deployments -> Spaces -> Utilities Customer Micro-Segmentation Space -> Deployments -> Utilities-Customer-Micro-Segmentation-Shiny-App** to find the URL of deployed r-shiny dashboard and open it in a new tab.
 - Run the dashboard from RStudio console by completing these steps:
 - i. 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.
 - ii. Click **Launch IDE > RStudio** on the menu bar.
 - iii. In the **Files** pane, select the **Upload** toolbar button and upload the `utilities-customer-micro-segmentation-analytics-dashboard.zip` file into RStudio.
 - iv. 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 username and password on the **Client View** tab.
 6. Optional. To connect the data assets used in this accelerator to the business terms in Watson Knowledge Catalog, you can edit and run the **0-map-business-terms-to-data-headers** notebook. Enter the authentication details required in the first few cells.

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 notebooks to step through the execution.

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-and-app-deployment: 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
- Store and deploy R Shiny app
- Generate URL to view the app.

0-map-business-terms-to-data-headers: This optional notebook performs the following functions:

- Publish the **customer micro-segmentation input.csv** file into a specified catalog.
- Read mappings from **utilities-customer-micro-segmentation-map-terms.csv** and applies business terms to the published dataset headers.

Jobs

Navigate to **Jobs** tab to execute following jobs.

1-model-training-notebook-job: Runs the 1-model-training notebook end to end.

2-model-scoring-notebook-job: Runs the 2-model-scoring-and-app-deployment notebook end to end.

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.

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 using the below files available in the project tar file.

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.

Once the glossary is imported into Watson Knowledge Catalog, Navigate to **Governance > Categories > Industry Accelerator > Utilities Customer Micro-Segmentation** to explore the glossary contents.

Terms and Conditions

This project contains Sample Materials, provided under 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.