

A propos de l'objet datagen

Description

Spectre et Visual Integrator disposent tous deux d'entrées capables de générer un ensemble de valeurs. Dans Spectre 7.2, la balise `input - datagen` n'est disponible qu'en mode éditeur de texte, mais elle est prise en charge en tant qu'objet dans Spectre 8.0 Flow.

Les entrées de génération de données (*Datagen*) présentent des propriétés similaires dans les deux outils, bien que l'objet Visual Integrator en propose quelques-unes de plus. Il est important de noter qu'aucun paramètre ne garantit l'unicité des valeurs dans votre jeu de données. En effet, des doublons sont possibles et peuvent survenir à grande échelle, en particulier lorsque le nombre de valeurs générées dépasse le nombre de valeurs internes prédéfinies.

Les éléments « prénom » et « nom de famille », par exemple, ne disposent que de 1 000 valeurs internes. Par conséquent, les doublons seront bien plus fréquents pour ces éléments que dans un jeu de données réel. Lors de la création de jeux de données de démonstration, il est plus réaliste d'utiliser des listes plus vastes de prénoms et de noms trouvés sur le Web.

Il convient également de noter que le type de données « word » (mot) dans *datagen* ne génère pas nécessairement un mot reconnaissable en anglais ou dans une autre langue.

Référence de la documentation concernant les attributs et les propriétés :

<https://www.dimins.com/online-help/workbench/72/spectre/spectre-datagen-properties.html>

Fonctionnalités supplémentaires de Visual Integrator

Le type de données ZIP génère une chaîne de 5 chiffres comprise entre 00001 et 99999. Cette chaîne peut correspondre ou non à un véritable code postal.

Distribution : linéaire, standard ou 80-20. Cet attribut définit la concentration des valeurs.

Le mode linéaire est la valeur par défaut ; chaque valeur a alors la même probabilité d'apparition.

Le mode standard augmente la fréquence des valeurs situées au milieu de la plage du jeu de données, créant ainsi une distribution rappelant la « courbe en cloche ».

Le mode 80-20 génère un ensemble de valeurs plus asymétrique, où 80 % du jeu de données se concentre sur 20 % des valeurs possibles. Ce mode consommant davantage de mémoire que les deux autres, il convient de faire preuve de prudence lors de sa sélection. (Remarque : le mode 80-20 n'est pas disponible pour la génération de données d'écrites.)

Tags

1. Datagen
2. Spectre
3. Visual Integrator