

## Nouveautés de la technologie SIG

Bulletin d'information de l'OSIG n°4/2019

### Apache Superset et autres outils de visualisation et de publication de données

Des applications d'un genre nouveau font parler d'elles ces temps-ci. Elles sont regroupées sous l'appellation d'*outils de visualisation et de publication de données*, mais sont aussi fréquemment dénommées *outils de business intelligence*.

Une bonne visualisation est importante pour donner une traduction concrète à des données issues d'enquêtes et de capteurs ou encore à des données spatiales. Elles sont comprises plus facilement et donc mieux acceptées. De plus, on ne décèle souvent la présence de liens de dépendance que sur des diagrammes et des cartes appropriés. Cela signifie que les visualisations peuvent également servir à l'analyse des données. En outre, les données préparées pour l'établissement de listes, de diagrammes ou de cartes peuvent être publiées en interne, au sein d'une organisation ou partagées via Internet avec des clients externes. Ces applications majoritairement à base web ont le potentiel requis pour prendre en charge certaines fonctions de systèmes d'information géographique (SIG).

Les deux produits les plus largement répandus sont les applications commerciales Tableau et Microsoft PowerBI. Insights for ArcGIS et CARTO existent dans le domaine des SIG. En y regardant de plus près, on découvre encore de nombreux autres produits – en particulier si l'on étend la recherche à des produits servant à surveiller des systèmes, cas de Kibana par exemple. Il semble assez difficile de tracer des frontières précises.

Ces outils ont tous les fonctions suivantes en commun:

- Analyser: sélectionner, filtrer et préparer des données.
- Visualiser: établir des listes, des diagrammes et des cartes interactifs.
- Présenter: réunir des listes, des diagrammes et des cartes au sein d'un «cockpit» (appelé «dashboard» (*tableau de bord*) dans Superset).
- Publier: partager ces cockpits/tableaux de bord via Internet avec des tiers.

Tous les outils permettent d'établir des diagrammes simplement. A cela s'ajoutent des fonctions de gestion des utilisateurs comme la création de comptes et de groupes de comptes, en incluant l'attribution de droits définissant les sources de données et les fonctions auxquelles un utilisateur peut accéder.

Un projet open source est également disponible depuis 2015 pour Apache Superset. L'auteur principal du présent article s'est donc intéressé d'un peu plus près à cette application et a brièvement récapitulé ses impressions dans la suite, en commençant par les trois points suivants:

- L'analyse avec Apache Superset est étendue aux sources de données les plus diverses, allant de fichiers au format CSV ou GeoJSON jusqu'à des services en ligne en passant par des systèmes de banques de données.
- Quelques clics suffisent pour réaliser les visualisations (diagrammes et cartes).
- Les résultats se laissent généralement manier et publier avec une grande facilité.

Il existe quelques possibilités d'analyse des jeux de données. Et avec 30 types de diagrammes et douze types de cartes, c'est un large éventail de variantes de visualisation qui est proposé.

On n'insistera jamais assez sur le fait qu'il est inutile de disposer de connaissances particulières (ni en matière de programmation ni en SIG) pour des outils de cet ordre. Et si les fonc-

tions existantes sont insuffisantes, on peut (faire) intégrer des requêtes SQL qui sont alors disponibles en tant que nouvelles sources de données.

Le projet open source Apache Superset n'a pas encore atteint le stade de développement des outils commerciaux précités et l'installation de l'application à base web reste encore très compliquée – sauf si l'on peut s'appuyer sur quelqu'un qui s'y connaît un tant soit peu. Ces insuffisances s'expliquent surtout par le fait que le projet a douze ans de moins que la société Tableau par exemple, fondée en 2003.

Quoi qu'il en soit, les fonctionnalités de Superset ont déjà atteint un degré de maturité suffisamment élevé pour des applications intéressantes en interne, si bien que des entreprises telles que Airbnb ou Twitter y recourent.

Stefan Keller et le groupe de travail 4 Technologie SIG de l'OSIG

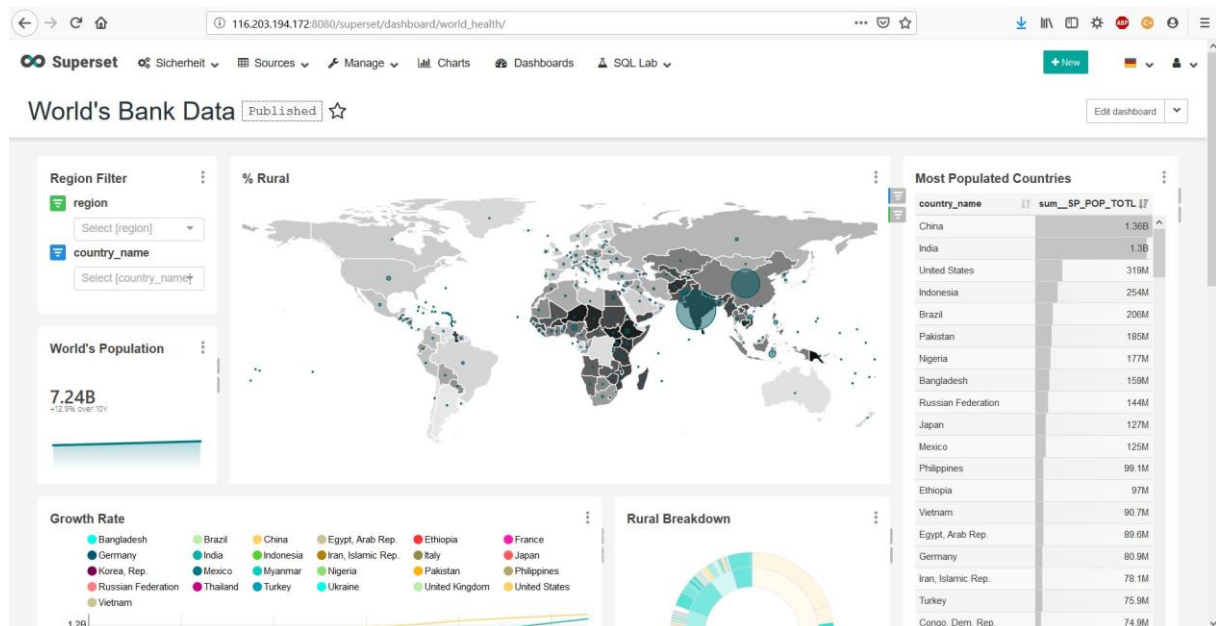


Figure: copie d'écran de diagrammes basés sur des données nationales provenant de la Banque mondiale et visualisés sous forme de tableau de bord (dashboard) à l'aide d'Apache Superset.