

QU'EST-CE QUE DATAGENCES ?

Datagences est un portail de données multi-sources, multi-échelles et multi-thématiques basé sur une plateforme décisionnelle Open Source composée :

- d'un entrepôt de données,
- d'un outil d'analyse de données,
- d'un outil de conception automatisée de rapports thématiques ou multi-thématiques.

Un entrepôt de données

Un entrepôt de données est un **espace de stockage centralisé** regroupant, dans un format homogène, les données issues de différentes sources, **transformées ou non** et **capitalisées dans le temps**.

Les données sont structurées selon un modèle multidimensionnel et stockées sous forme d'indicateurs ou « **mesures** » pouvant être analysées selon un ensemble d'axes appelés « **dimensions** ».

Cette modélisation permet donc d'effectuer des requêtes **efficacement** croisant **donnée, échelle géographique** et **temporalité**.

Ainsi l'entrepôt de Datagences regroupe l'essentiel des données open data **nécessaires aux besoins d'études** de l'agence à travers plus de 130 cubes de données intégrant près de 500 indicateurs simples ou calculés.

Les principales sources de données intégrées sont ACOSS, AGRESTE, BPE, CAF, les chiffres-clés et le recensement de l'INSEE, SIRENE, Sit@del2 et RPLS.

Par exemple on retrouve des données sur les thématiques suivantes :

Économie : effectifs salariés privés, caractéristiques de l'emploi, démographie des entreprises et établissements, etc.

Démographie – Population : évolution et structure de la population (par âge, sexe et catégorie socio-professionnelle), nombre et taille des ménages, etc.

Éducation : effectifs étudiants inscrits dans les établissements de l'enseignement supérieur, nombre de personnes diplômées, etc.

Habitat - Logement : nombre de logements autorisés et commencés, nombre de permis de construire par type, données relatives aux résidences principales, etc.

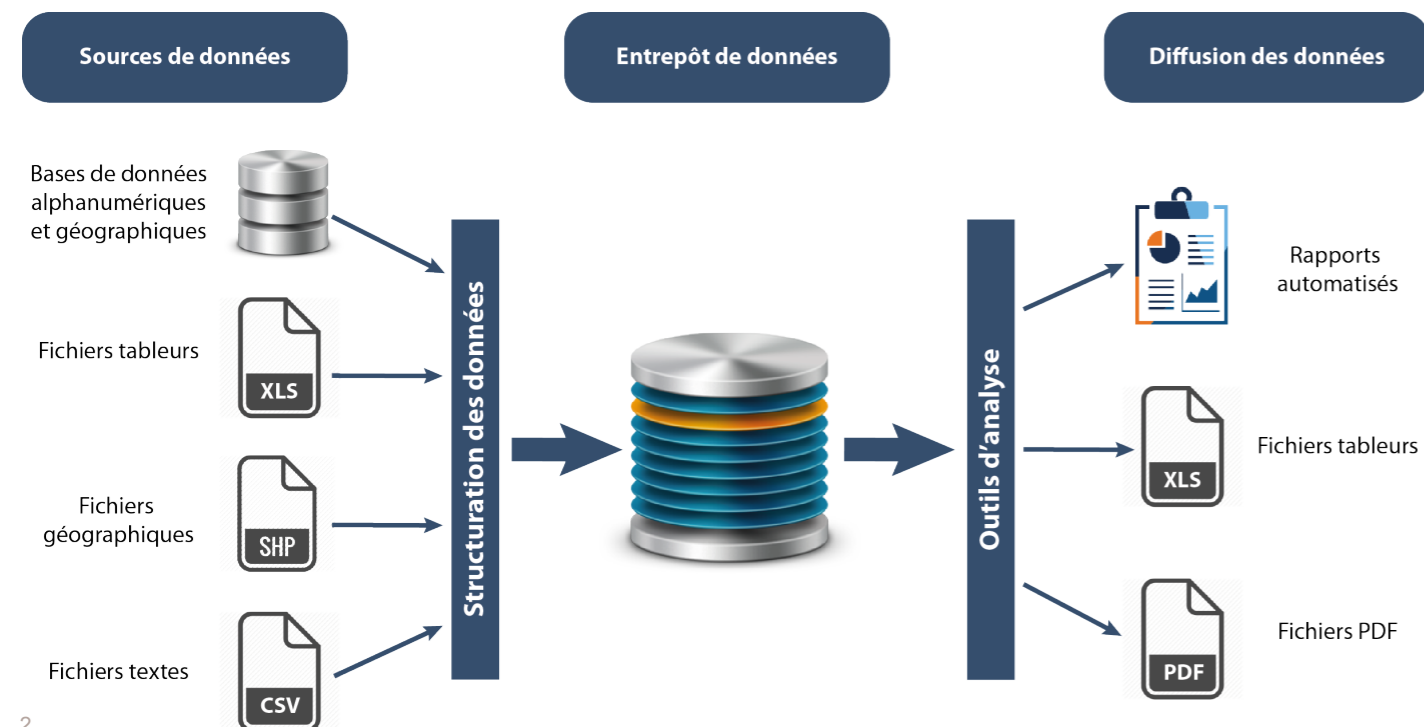
Mobilité : mobilités professionnelles et scolaires.

Sociale : nombre de foyers allocataires du RSA ou du RSO, nombre d'enfants couverts par l'allocation de rentrée scolaire, etc.

Agriculture : nombre d'exploitations par taille et par type, nombre de cheptels selon la taille du troupeau, etc.

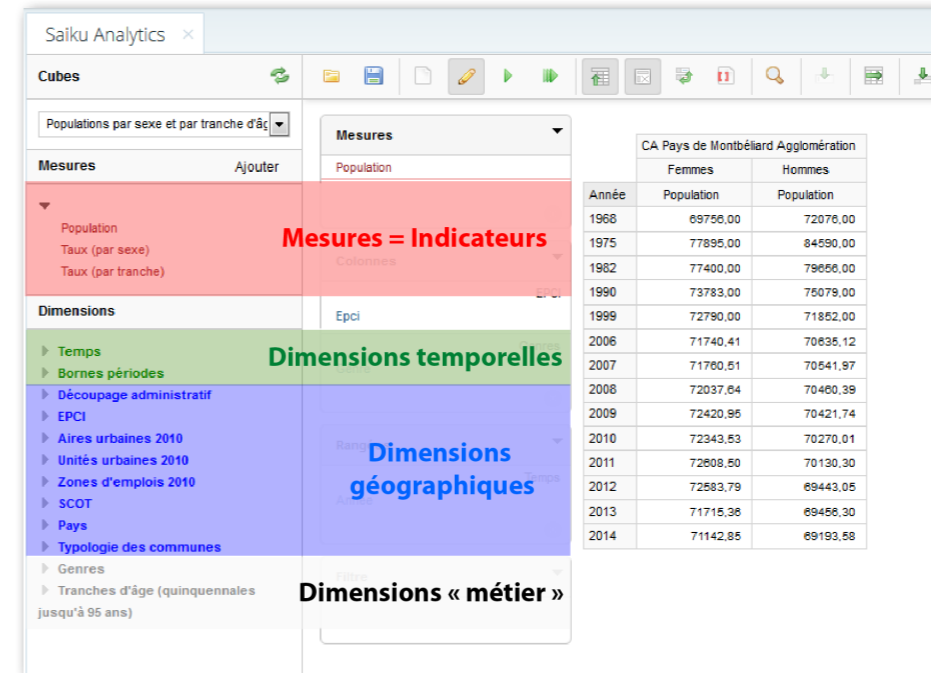
Équipement : nombre d'équipements par type, etc.

Tourisme : hébergement touristique, nombre d'hôtels et de campings par classement (nombre d'étoiles), etc.



Un outil d'analyse de données

Un outil d'analyse, Saiku analytics, permet via une interface **épurée** et **ergonomique** de **visualiser**, **d'agréger**, **d'explorer** et **d'exporter** les données de l'entrepôt pour la réalisation de diagnostics, d'études ou de notes de conjoncture.



Cet outil permet de filtrer les données selon 3 types de dimensions d'analyse : **temporelles**, **géographiques** ou **métiers** propres à chaque jeu de données.

Il est donc aisé d'observer les données selon une maille géographique (nationale, régionale, départementale, cantonale, intercommunale, communale et infra-communale pour les jeux de données le permettant), des années et/ou des critères plus spécifiques aux données visualisées comme les codes d'activités principales des entreprises, l'âge, le sexe, les catégories socio-professionnelles, le statut ou le type d'équipement.

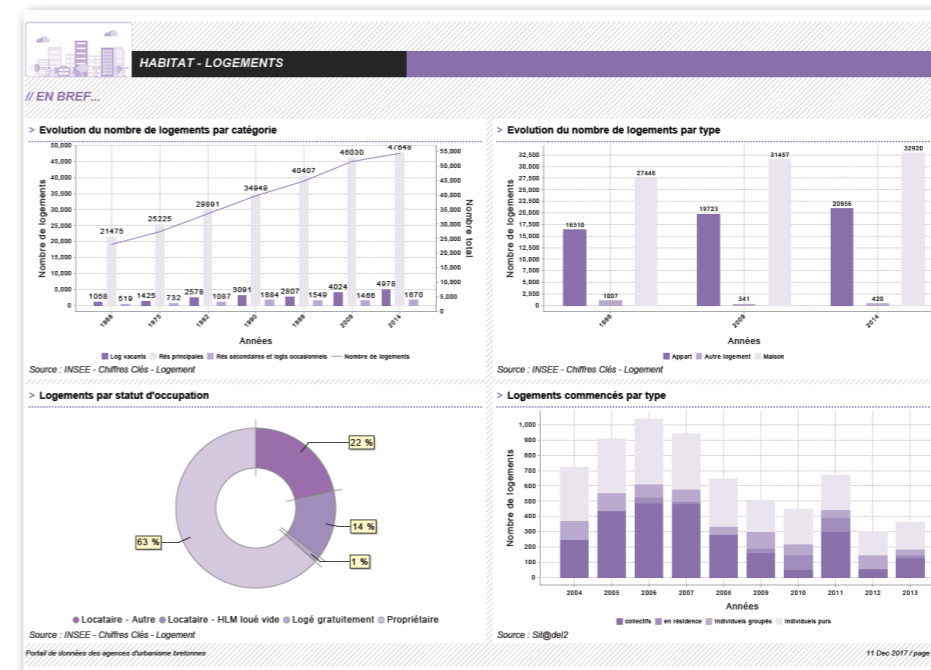
Les données étant à l'échelle de la France, la comparaison au.x territoire.s voisin.s ou au.x territoire.s ayant des caractéristiques similaires sur l'ensemble des thématiques présentes dans l'entrepôt de données est facilitée.

Enfin les données analysées peuvent être exportées sous un format interopérable permettant de les représenter sous différentes formes (cartes, graphiques, tableaux, etc.)

Un outil de conception automatisée de rapports

La solution Datagences intègre un outil de conception automatisée de rapports appelé PRD (Pentaho Report Designer).

Il permet de structurer des rapports thématiques ou multi-thématiques complexes et multi-source automatisés ou semi-automatisés à partir des données de l'entrepôt mais également à partir d'autres sources de données capitalisés dans le système d'information de l'ADU.



Ces rapports sont ensuite hébergés sur la plateforme décisionnelle afin d'être accessibles à l'ensemble des utilisateurs qui peuvent alors les générer **à la volée** selon le périmètre géographique souhaité. Ces rapports peuvent être exportés au format texte, tableur ou web pour faciliter leur diffusion.

L'image ci-contre montre un exemple de rapport sur la thématique du logement et de l'habitat avec l'utilisation de 4 indicateurs issus de deux sources de données et représentés sous forme de graphiques. Ce rapport peut être généré à l'échelle d'une commune, d'une intercommunalité, d'un département, etc.

Datagences permet de **rationaliser les tâches chronophages d'intégration** des données au profit de temps consacrés à l'analyse. Il **facilite** aussi **l'accès à des données open-data structurées** nécessaires à la réalisation des missions de l'ADU.

L'outil offre également la possibilité de **créer ses propres cubes de données** à partir de données locales pour suivre plus efficacement les projets de territoire comme le Schéma de Cohérence Territoriale (SCoT) ou le Nouveau Programme National de Renouveau Urbain (NPNRU).

Il permet également d'**acquérir un socle de connaissances et de données** sur les territoires d'intervention actuels et futurs de l'ADU.

Face à l'augmentation des données à capitaliser et à traiter, cet outil permet ainsi à l'ADU de se focaliser sur sa capacité à les analyser, les problématiser et les mettre en perspective selon le contexte local.

Le portail de données Datagences est né **de l'initiative des agences d'urbanisme de Lorraine** (Longwy, Metz et Nancy) avec la collaboration d'une entreprise spécialisée pour traiter, qualifier et structurer les données publiques.

Depuis, il a été déployé au sein **des agences d'urbanisme de Bretagne** (Brest, Lorient, Rennes et Quimper), de **Mulhouse**, de **Saint-Nazaire**, de **Sud Bourgogne** et de **Montbéliard**.

Ce déploiement progressif dans les agences d'urbanisme a permis la création d'**une communauté d'utilisateurs experts** ayant pour but de **partager leurs expériences**, de **mutualiser leurs savoir-faire** et d'**échanger des indicateurs, des tableaux de bords ou des rapports automatisés**.

ZOOM SUR...

L'**open data** ou **données ouvertes** est un mouvement visant à mettre à disposition, facilement et gratuitement, du public ou des professionnels des données d'intérêt général.

Ainsi l'open data désigne les données qu'un organisme met à la disposition de tous sous forme de fichiers numériques interopérables afin de permettre facilement leur réutilisation. Toutefois ces données ne doivent pas avoir un caractère personnel et leur réutilisation peut être soumise à conditions.

Ces données sous licence ouverte peuvent être d'origine publique (émanant des services publics ou collectivités) ou privée (provenant d'entreprises et institutions dont les données concourent à des projets d'utilité publique).

Encadré par des réglementations légiférées au Parlement Européen dès 2003, l'open data prend un nouvel essor avec la loi pour une République Numérique promulguée le 7 octobre 2016 mettant en place plusieurs mesures **favorisant une politique d'ouverture des données et des connaissances**. Dans ces dernières, on peut noter l'ouverture et la gratuité des données de l'INSEE ou l'ouverture par défaut des données des collectivités territoriales de plus de 3500 habitants.

Le **big data** ou **données massives** est une notion apparue à la fin des années 1990 pour désigner « *l'ensemble des données numériques produites par l'utilisation des nouvelles technologies à des fins personnelles ou professionnelles*² ». Elles nécessitent de nouvelles approches technologiques pour leur stockage, leur traitement et leur utilisation. Ces approches doivent donc prendre en compte leur volume, leur vélocité (produites, récoltées et analysées en temps réel) et leur variété (brutes, non structurées ou semi-structurées).

² Source : Futura Tech - <https://www.futura-sciences.com/>

