

Université IBM i 2018

16 et 17 mai

IBM Client Center Paris



S28 - IA sur vos données DB2 avec Watson Analytics

Christophe LALEVEE - Architect
IBM Client Center Montpellier, France

@ : lalevee@fr.ibm.com  : [linkedin.com/in/clalevee](https://www.linkedin.com/in/clalevee)  : [@chris_lalevee](https://twitter.com/@chris_lalevee)

IBM Watson
Analytics



Présentation



- Le but de cette session est de vous faire :
 - Déployer et configurer les composants nécessaires à l'importation des données Db2 for i dans Watson Analytics : IBM Watson Knowledge Catalog
 - Prendre en main l'interface de Watson Analytics pour créer vos premiers rapports et tableaux de bord

IBM Watson
Analytics



Agenda

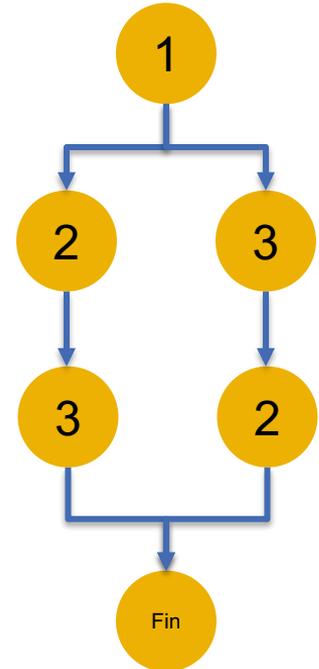


Présentation du Hands-on

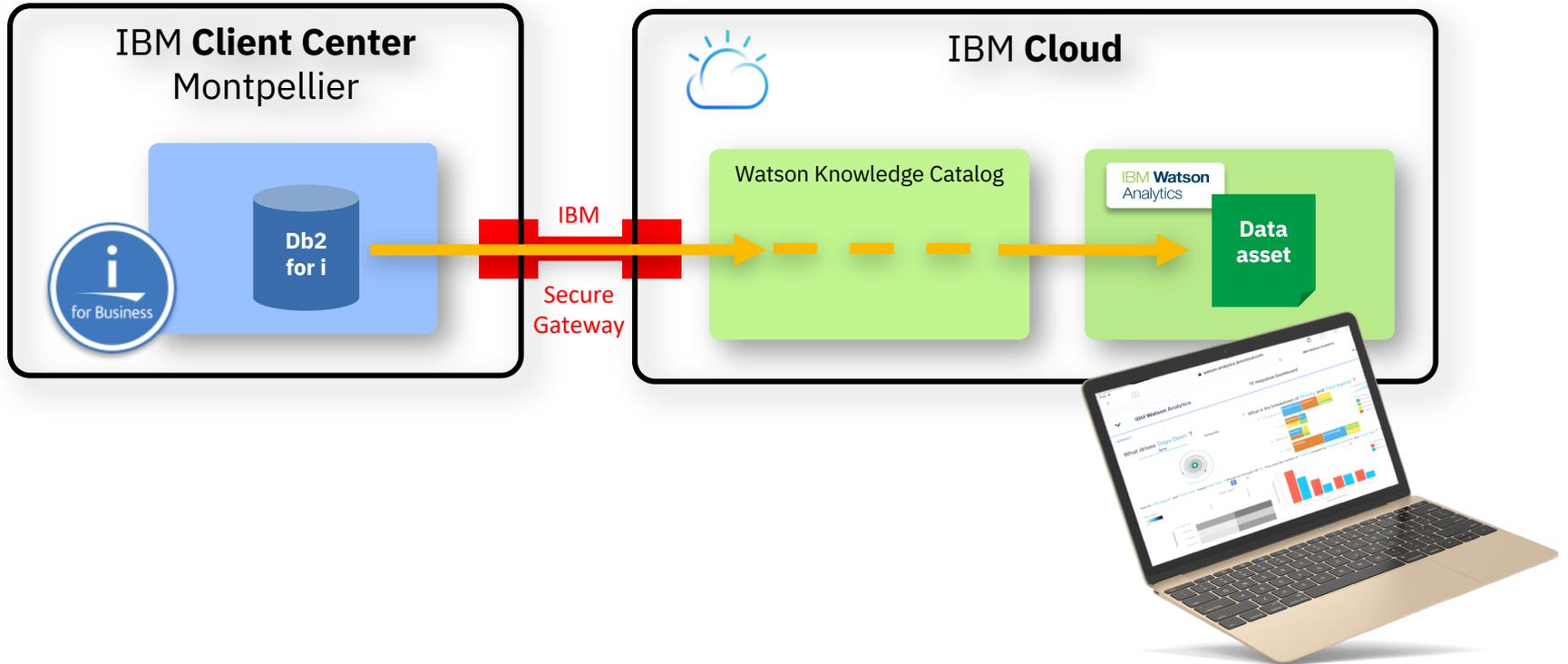
1. Création d'un compte Watson Analytics
2. Importation des données à partir de Db2 for i
 1. Connexion au module Data Refinery
 2. Création d'un projet
 3. Création des connexions
 4. Création et exécution du flux de données
3. Analyse des données importées
 1. Importation des données à partir d'un fichier Excel (option)
 2. Prise en main de l'environnement
 3. Analyse des données
 4. Création d'un tableau de bord
 5. Partage du tableau de bord

Agenda : attention !!

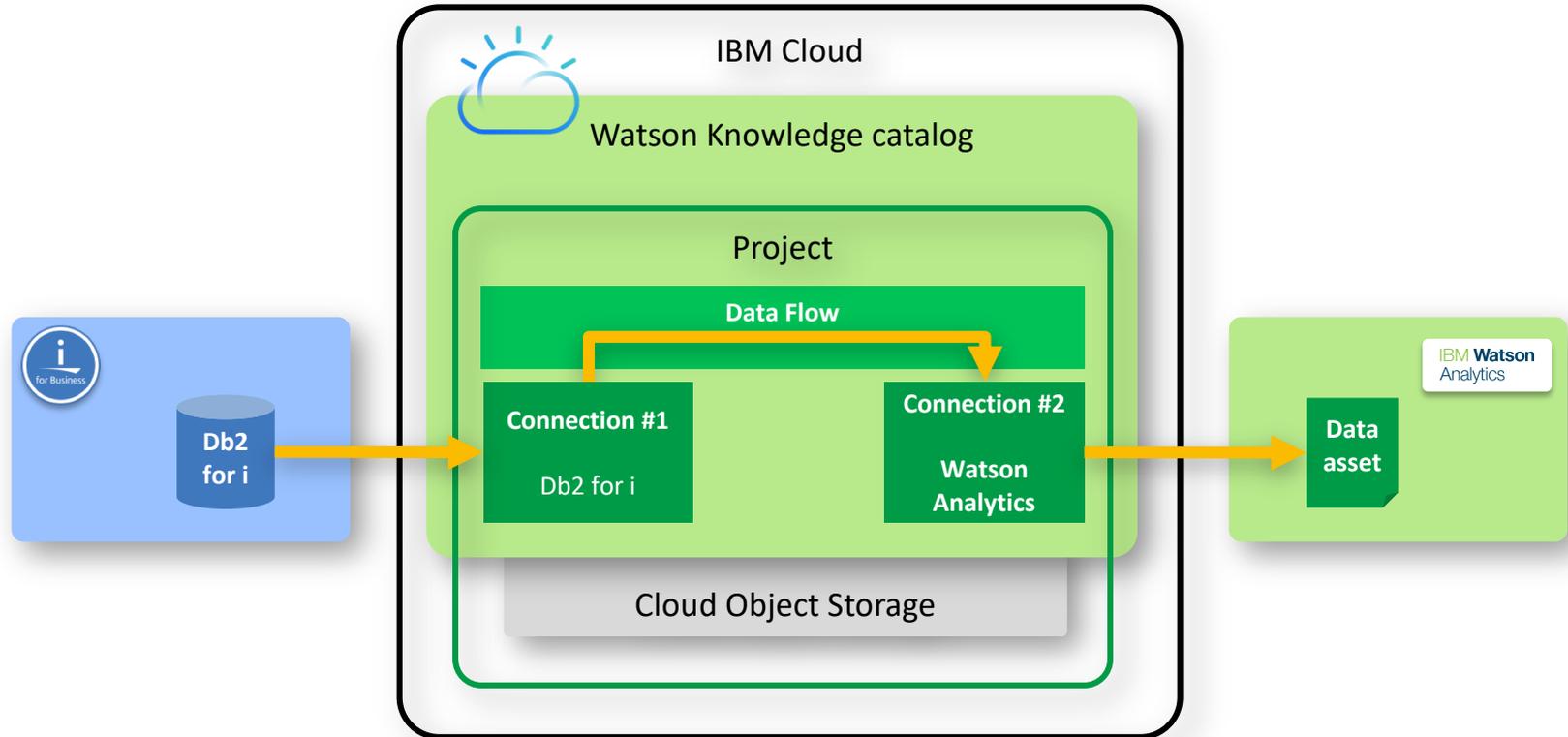
- Le lab est très long car il vous permet d'implémenter un cas d'usage de bout en bout
- Commencez par la partie qui vous intéresse le plus
- Après l'exercice 1, vous pouvez commencer par
 - Importation des données de l'IBM i (Exercice 2)
 - ou
 - Visualisation des données (Exercice 3)
- Vous pourrez faire la suite à partir de chez vous
L'infrastructure du lab restera en place jusqu'à la fin du mois de juin



Principe général



Principe général : Watson Knowledge catalog



Liens

- Guide d'exercice :

<http://ibm.biz/univi-s28>

- Données Tickets IT Helpdesk (Microsoft Excel) :

<http://ibm.biz/univi-s28-data>

{ “hands-on” }

S28 - Watson Analytic & IBM i



Université IBM i

16 et 17 mai 2018

IBM Client Center - Bois-Colombes

Agenda

Présentation du Hands-on.....	4
0. Connection au poste de travail	5
1. Création d'un compte Watson Analytics.....	6
2. Importation des données à partir de Db2 for i	9
Section 1. Connexion au module Data Refinery	10
Section 2. Création d'un projet	15
Section 3. Création des connexions	18
Section 4. Création et exécution du flux de données.....	23
3. Analyse des données importées.....	28
Section 0. Importation des données à partir d'un fichier Excel (option).....	29
Section 1. Prise en main de l'environnement.....	31
Section 2. Analyse des données	34
Section 3. Création d'un tableau de bord	41
Section 4. Partage du tableau de bord	46

Présentation du Hands-on

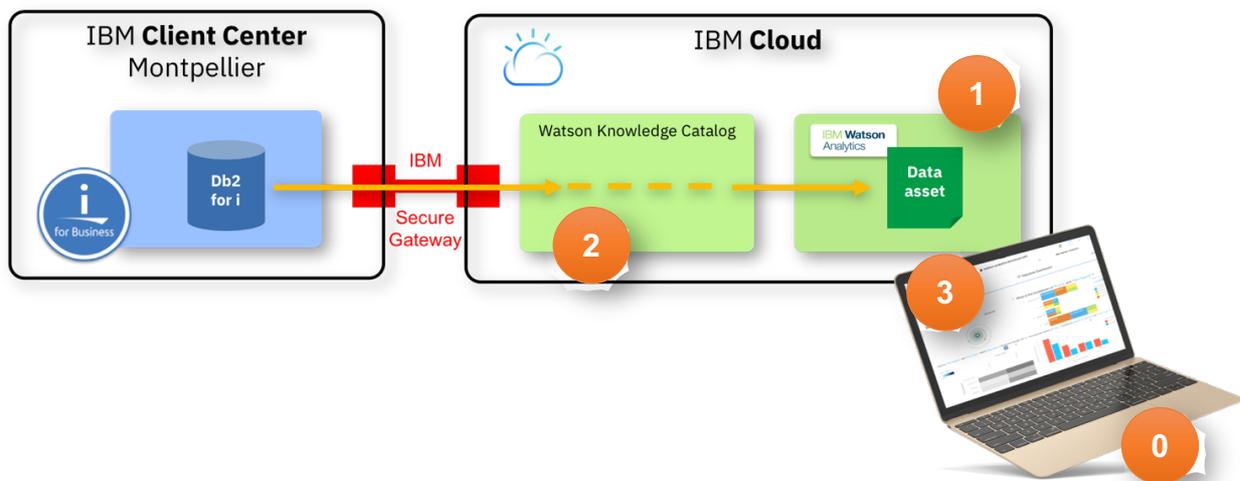
IBM Watson Analytics est une solution Cloud vous apportant des fonctionnalités d'exploration intelligente de données et d'analyse prédictive, et ce, de manière conversationnelle.

Combiné à un service IBM d'intégration de données, Watson Analytics peut également se connecter directement à DB2 for i, afin de télécharger et d'affiner vos données en toute sécurité. Ainsi, si vous avez besoin de repérer rapidement une tendance ou si votre équipe souhaite afficher des informations synthétiques dans un tableau de bord, Watson Analytics est l'outil idéal et complémentaire de DB2 Web Query.

Le but de cette session est de vous faire déployer et configurer les composants nécessaires à l'importation des données Db2 for i dans Watson Analytics, puis de prendre en main l'interface de Watson Analytics pour créer vos premiers rapports (« Discovery ») et tableaux de bord (néanmoins l'exercice 3 peut être réalisé sans avoir fait l'exercice 2).

Durant cette session pratique, vous allez donc mettre en œuvre les différents composants de ce cas d'usage :

0. Utiliser le poste de travail mis à votre disposition dans la salle
1. Créer les comptes d'accès à Watson Analytics et IBM Cloud (si nécessaire)
2. Dans IBM Watson Knowledge Catalog (module Data Refinery), créer un projet d'importation de données à partir de Db2 for i
3. Explorer les données importées et générer rapport et tableau de bord



(Les numéros indiqués sur le schéma ci-dessus correspondent à la numérotation des exercices)

L'IBM i utilisé est hébergé dans un Datacenter IBM à Montpellier, France. Pour pouvoir y accéder à partir du Cloud IBM, un service d'intégration, IBM Secure Gateway, est nécessaire.

Le déploiement et la configuration de la Secure Gateway ne font pas partie de ce hands-on.

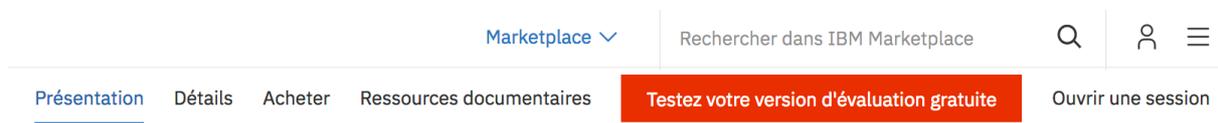
Vous utiliserez une instance Secure Gateway préconfigurée pour vous. Elle vous donnera accès à la base de données Db2 for i, au travers d'un nom de host et d'un numéro de port TCP publics (translation de l'adresse IP privée et du numéro de port DRDA de l'IBM i).

0. Connection au poste de travail

Connectez-vous à votre poste de travail comme indiqué par l'instructeur.

1. Création d'un compte Watson Analytics

1. A l'aide de votre navigateur, connectez-vous à l'adresse suivante : <https://www.ibm.com/watson-analytics>.
2. Si vous possédez déjà un compte, cliquez sur le bouton **Ouvrir une session** et passer à la session 3.



3. Dans le cas contraire, cliquez sur le bouton **Testez votre version d'évaluation gratuite**.

4. Si vous avez déjà un IBM ID, vous pouvez le rattacher à Watson Analytics (bouton Connexion). Sinon, saisissez l'ensemble des champs demandés, puis cliquez sur le bouton Continuer. Un email avec un code de validation vous sera envoyé à l'adresse email que vous avez indiqué. Vous devrez le saisir pour finaliser l'inscription

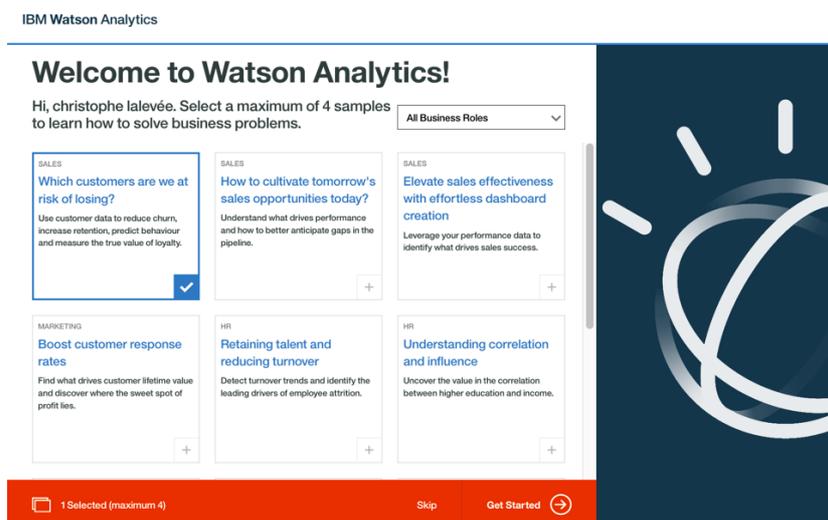
5. Vous recevrez deux autres emails : le premier vous indiquant que votre IBM ID a bien été créé, le deuxième, un peu plus tard, vous indiquant que votre service Watson Analytics est prêt à être utilisé.



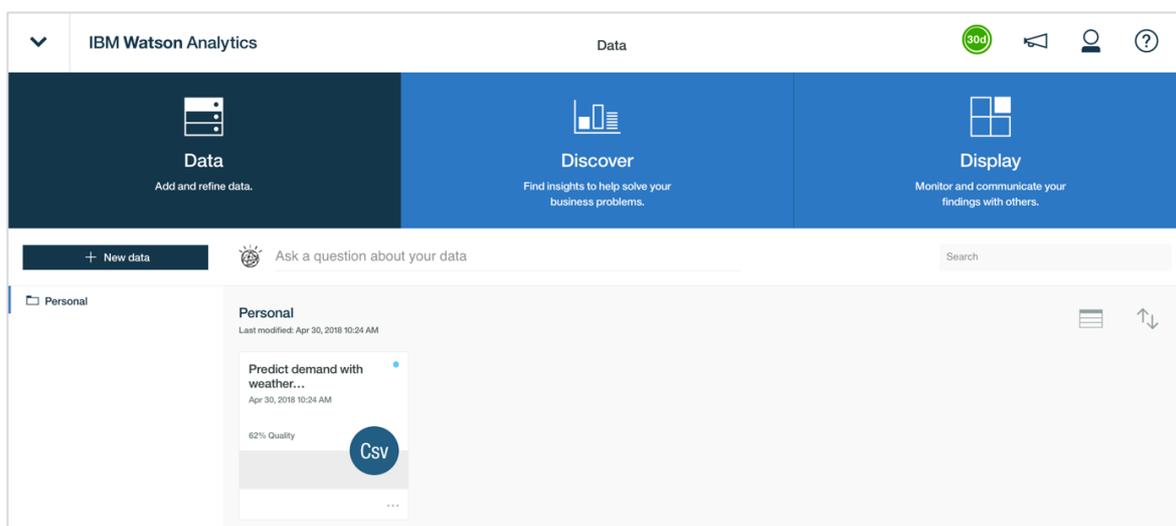
Information

Cette souscription vous donne un accès illimité à un compte Watson Analytics version Freemium, mais aussi, pendant 30 jours, à la version professionnelle et l'option Social Media. Pour plus d'information sur les différentes éditions, consultez le lien suivant : <https://www.ibm.com/fr-fr/marketplace/watson-analytics/purchase#product-header-top>

6. Après une petite vidéo vous permettant de patienter pendant la création de votre compte Watson Analytics, vous arrivez sur une page vous permettant de sélectionner jusqu'à 4 jeux de données pour apprendre à utiliser Watson Analytics par l'exemple. Cliquez sur **Skip**, pour sauter cette étape : vous pourrez accéder ultérieurement à ces exemples.

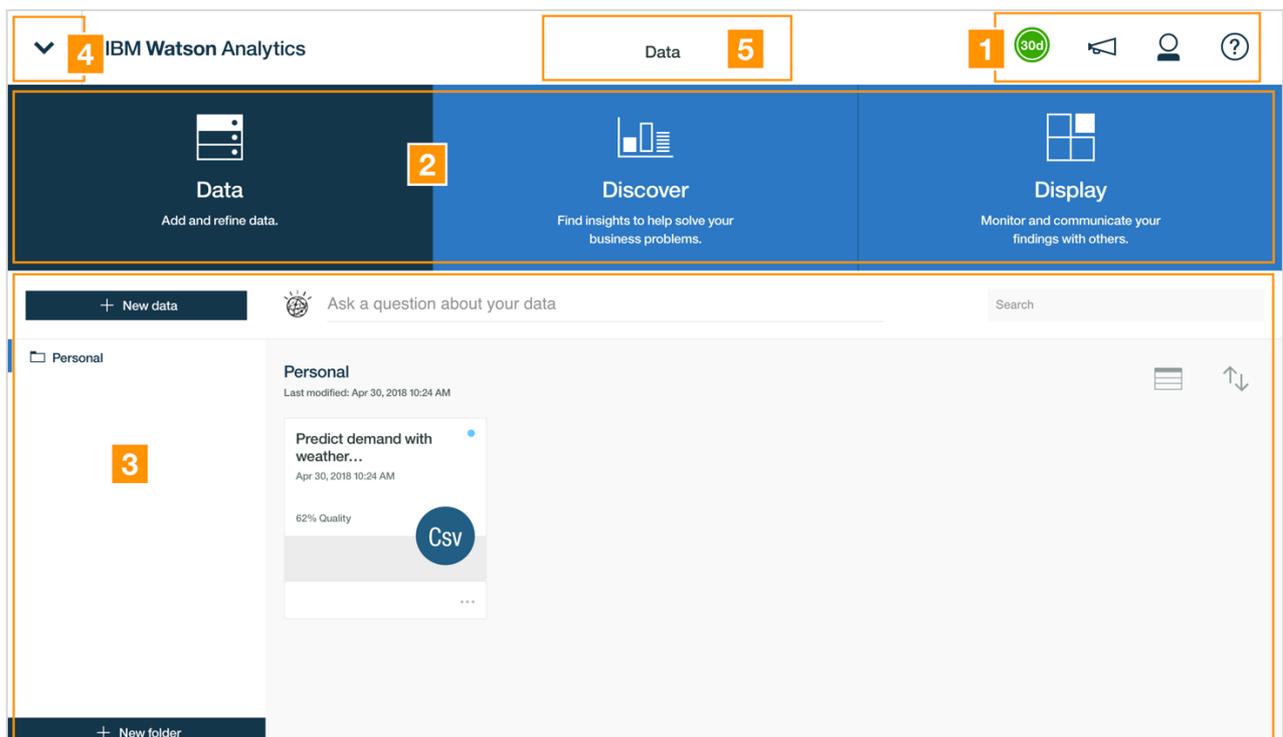


7. L'environnement de travail de Watson Analytics s'ouvre. Vous êtes maintenant prêt à l'utiliser. Notez le rond vert indiquant le nombre de jours restants pour l'évaluation de la version professionnelle.



___ 8. L'interface s'organise de la manière suivante :

1. L'accès aux menus relatifs à votre compte, ses options, et aux sources d'aide et d'assistance.
2. L'accès aux grandes catégories d'action pour manipuler vos données : ajouter des données, les explorer, les afficher.
3. Les données importées et les dossiers permettant de les organiser.
4. L'accès aux modules de Watson Analytics : Watson Analytics lui-même, Watson Analytics for Social Media, ...
5. La liste déroulante permettant de naviguer dans les différentes fenêtres que l'on ouvre.



2. Importation des données à partir de Db2 for i

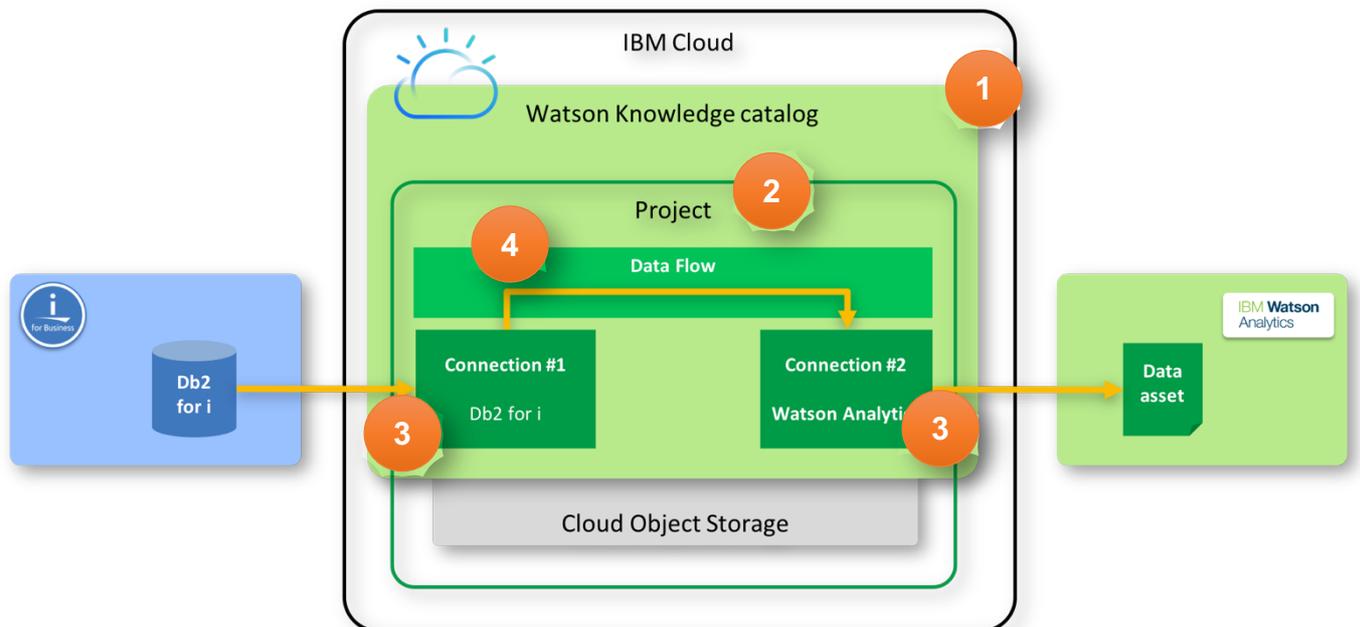
L'importation des données à partir d'une base de données externe se fait à partir du module **Data Refinery** de **Watson Knowledge Catalog** (depuis le début de 2018).

Le service Watson Knowledge Catalog fournit une plate-forme de gestion sécurisée de catalogue de données d'entreprise. Un catalogue relie les données et les connaissances aux personnes qui en ont besoin. Pour plus d'information :

<https://dataplatform.ibm.com/docs/content/catalog/overview-wkc.html?linkInPage=true>

L'outil Data Refinery, disponible via les services IBM Cloud Watson Studio et Watson Knowledge Catalog, permet d'économiser du temps de préparation des données en transformant rapidement de grandes quantités de données brutes en informations de qualité, prêtes pour l'Analytics. Il s'interface avec toutes sortes de bases de données, comme Db2 for i, et avec Watson Analytics.

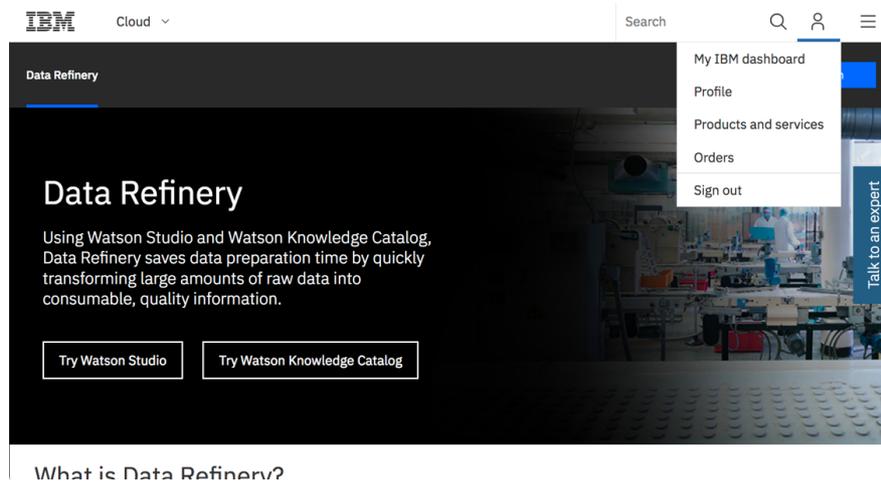
Le schéma ci-dessous présente l'architecture de ce que vous allez mettre en œuvre durant cet exercice.



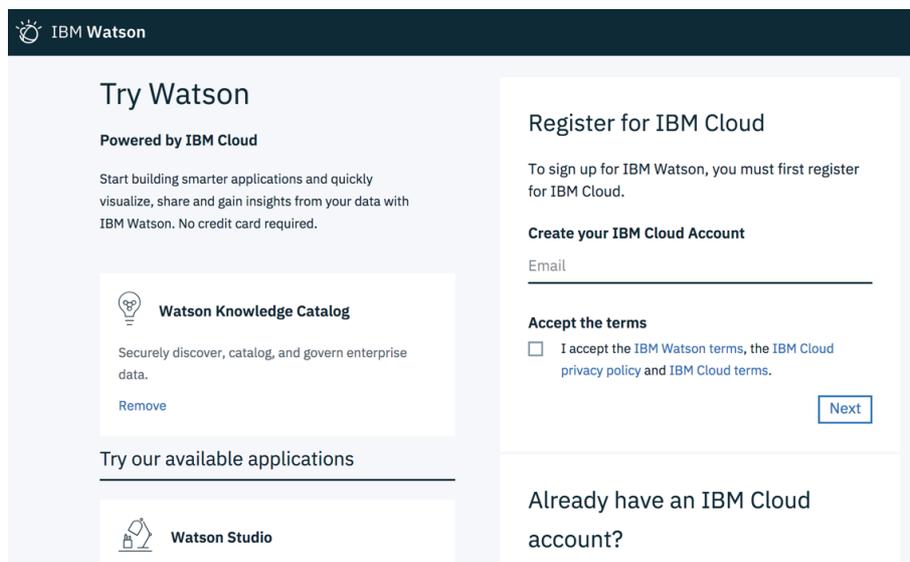
1. Création des comptes nécessaires et déploiements du service Watson Knowledge Catalog
2. Création d'un projet d'importation de données
3. Création des connexions
4. Création du flux de données et exécution

Section 1. Connexion au module Data Refinery

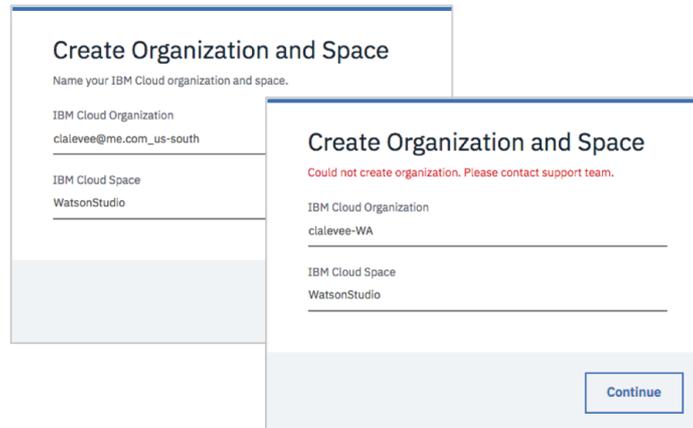
1. A l'aide votre navigateur, connectez-vous à l'adresse <https://www.ibm.com/cloud/data-refinery> et connectez-vous, si nécessaire, avec les identifiants et mot de passe utilisés pour Watson Analytics (Il se peut que vous soyez déjà identifiés si vous êtes connectés avec le même navigateur sur Watson Analytics ; cf. image ci-dessous). Sélectionnez Watson Knowledge Catalog (bouton **Try Watson Knowledge Catalog**)



2. Connectez-vous à IBM Cloud pour pouvoir utiliser ce service Watson. Différents choix sont proposés selon les comptes que vous avez ou que vous n'avez pas :
 - **Register for IBM Cloud** : vous n'avez pas de compte IBM Cloud, et pas de compte IBM Watson
 - **Already have an IBM Cloud account?** : vous avez un compte IBM Cloud, mais pas de compte IBM Watson
 - **Already signed up?** : vous avez un compte IBM Cloud, et un compte IBM Watson



- ___ 3. Cliquez sur l'option correspondant à votre situation et suivez la procédure proposée pour utiliser un compte IBM Cloud (et Watson).
- ___ 4. A la fin de la procédure, il vous sera demandé de créer une organisation et un espace. Saisissez les informations demandées, puis cliquez sur **Continuer**.



- ___ 5. Selon la procédure et les configurations préalablement réalisées, il est possible que vous obteniez le message d'erreur « **Could not create organization. Please contact support team.** ». Cela est probablement dû au fait que vous avez un compte IBM Cloud Freemium, possédant déjà une organisation et/ou un espace (les comptes Freemium sont limités à 1).

Dans ce cas, suivez la procédure ci-dessous.

Sinon, suivez les étapes qui vous sont proposées et rendez-vous à l'étape 13

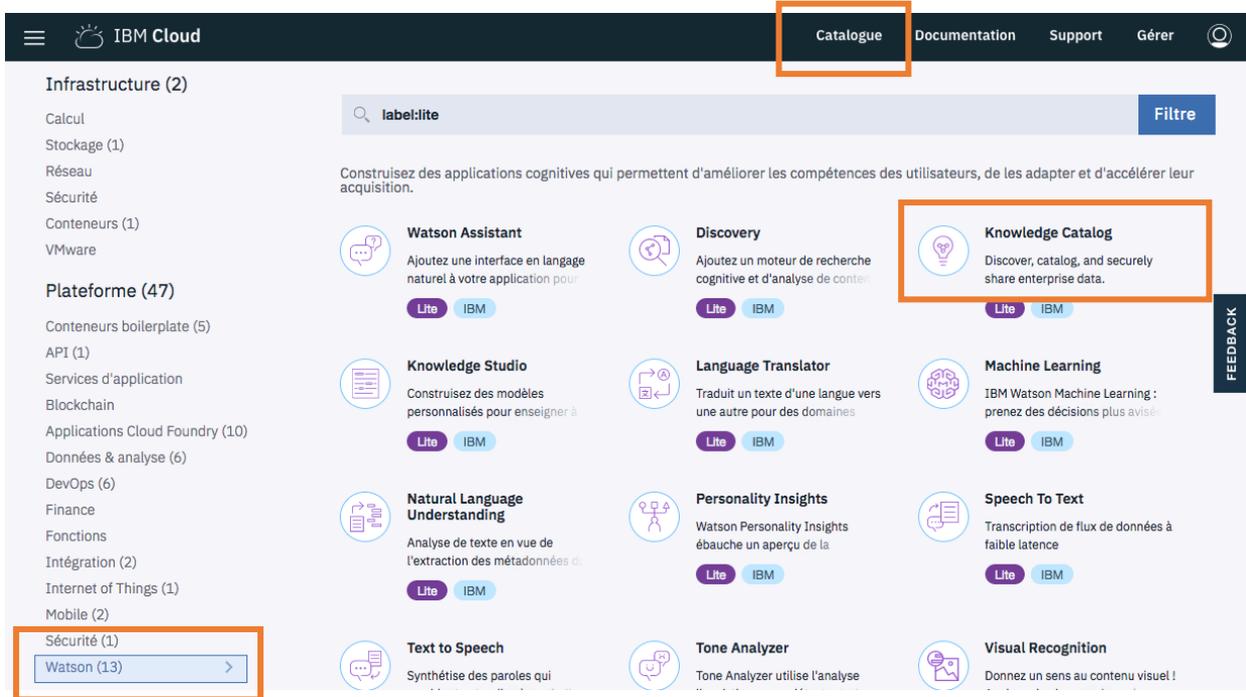
- ___ 6. Connectez-vous à IBM Cloud en utilisant l'adresse : <https://console.bluemix.net>.
- ___ 7. A partir du menu « burger » en haut, à gauche, sélectionnez **Tableau de bord**.



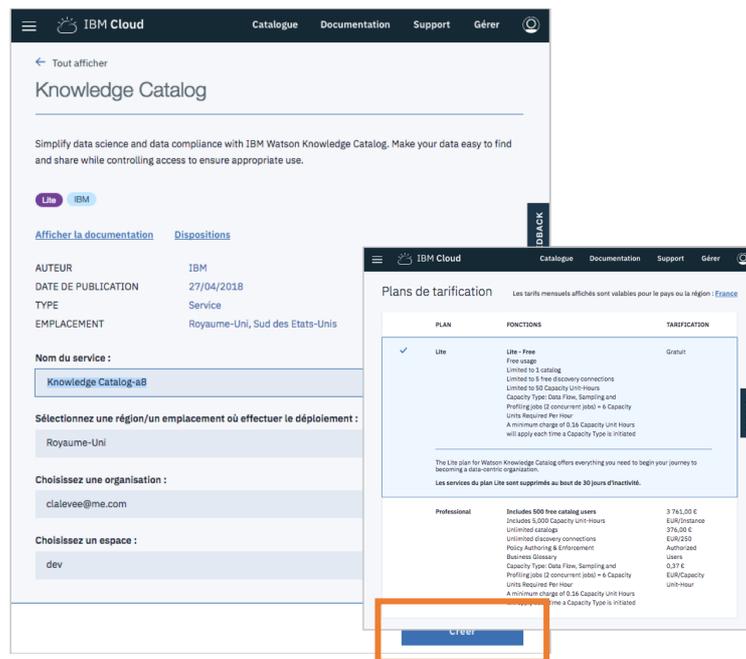
- ___ 8. En utilisant les listes déroulantes du tableau de bord, vous allez voir les organisations et espaces existants



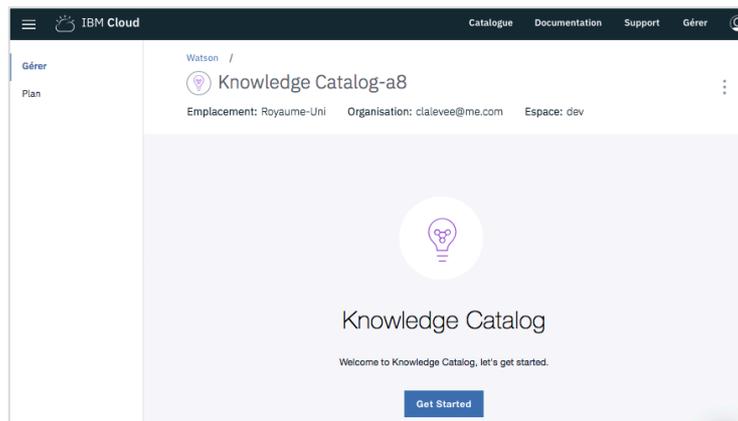
9. Sélectionnez un espace existant pour créer manuellement le service **Watson Knowledge Catalog**.
 Dans la barre de menu, cliquez sur **Catalogue**, puis, dans le panneau de gauche, sélectionnez **Watson** dans la catégorie **Plateforme**.
 Cliquez sur le service **Knowledge Catalog** :



Saisissez un nom de service et sélectionnez l'organisation, l'espace et la région.
 Vérifiez également en bas de page, que le plan lite (gratuit) est sélectionné.
 Cliquez sur le bouton **Créer**.

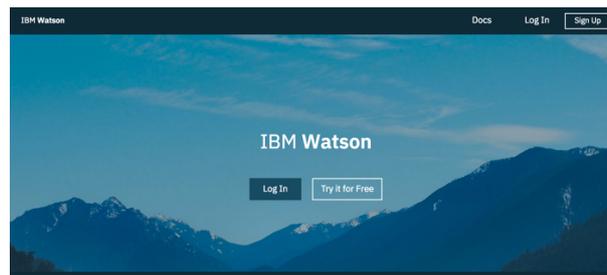


Votre service Watson Knowledge Catalog est créé.

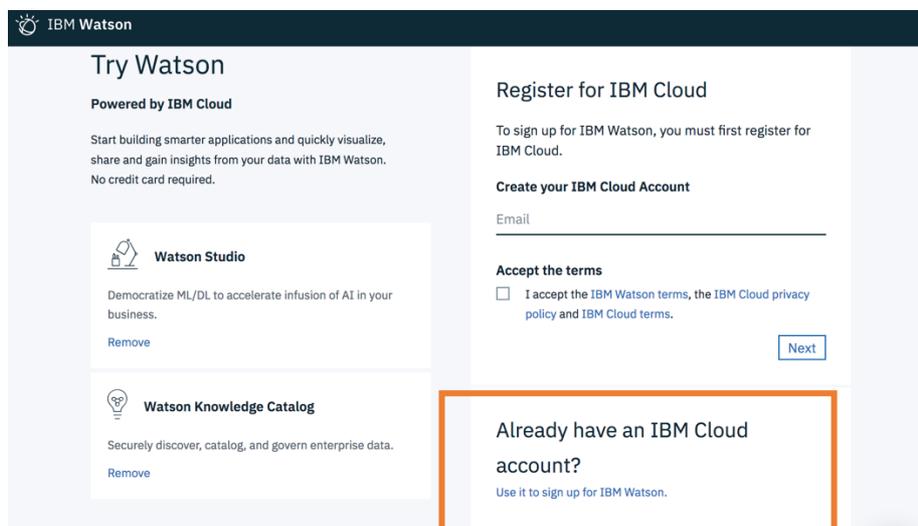


Il faut maintenant l'activer sur la plateforme IBM Watson Studio (Watson Data Platform).

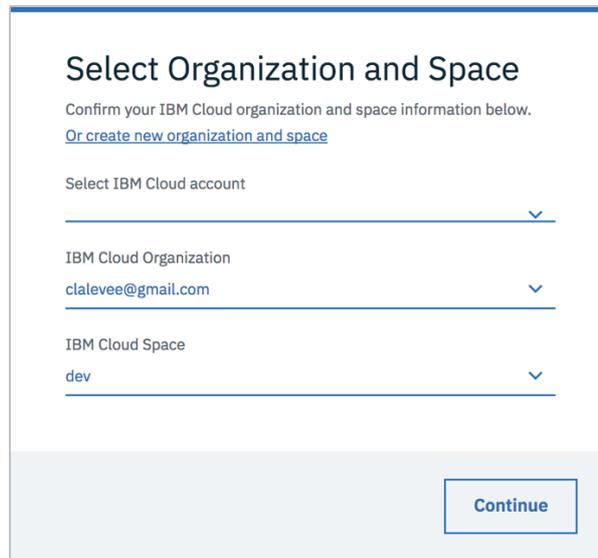
- 10. Connectez-vous à l'adresse suivante pour activer et utiliser le service Watson créé : <https://dataplatform.ibm.com> (ou <https://eu-gb.dataplatform.ibm.com> selon la région IBM Cloud que vous utilisez) et cliquez sur **Try it for Free**.



- 11. Comme vous avez déjà un compte IBM Cloud, cliquez sur **Use it to sign up for IBM Watson** dans la section **Already have an IBM Cloud account?**



- ___ 12. Sélectionnez votre compte, l'organisation et l'espace contenant votre service Watson Knowledge Catalog.



Select Organization and Space

Confirm your IBM Cloud organization and space information below.
[Or create new organization and space](#)

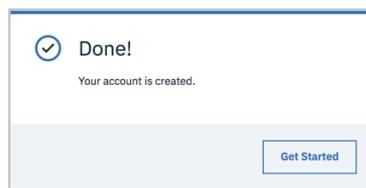
Select IBM Cloud account

IBM Cloud Organization
clalevee@gmail.com

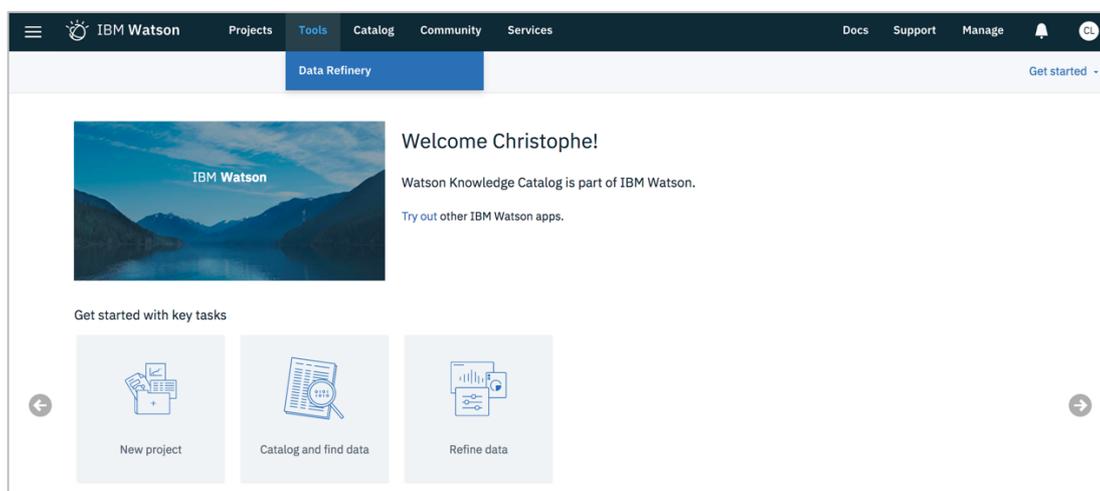
IBM Cloud Space
dev

Continue

- ___ 13. Votre compte IBM Watson Studio est créé, et contient le service Watson Knowledge Catalog.

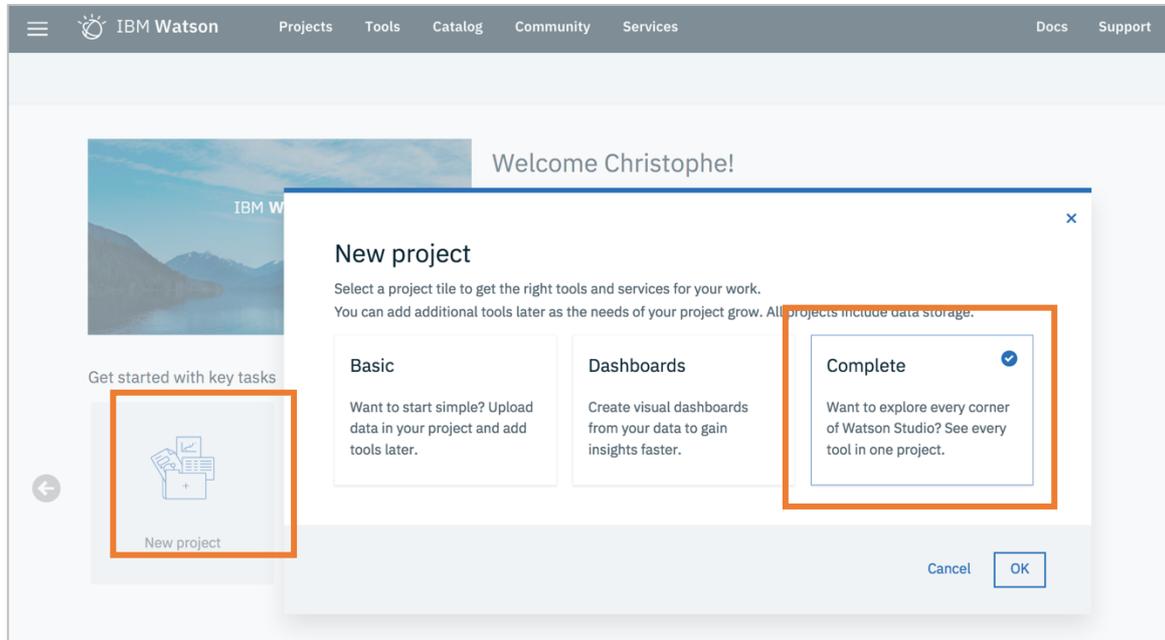


- ___ 14. Vérifiez que vous avez bien l'option **Data Refinery** dans le menu Tools.

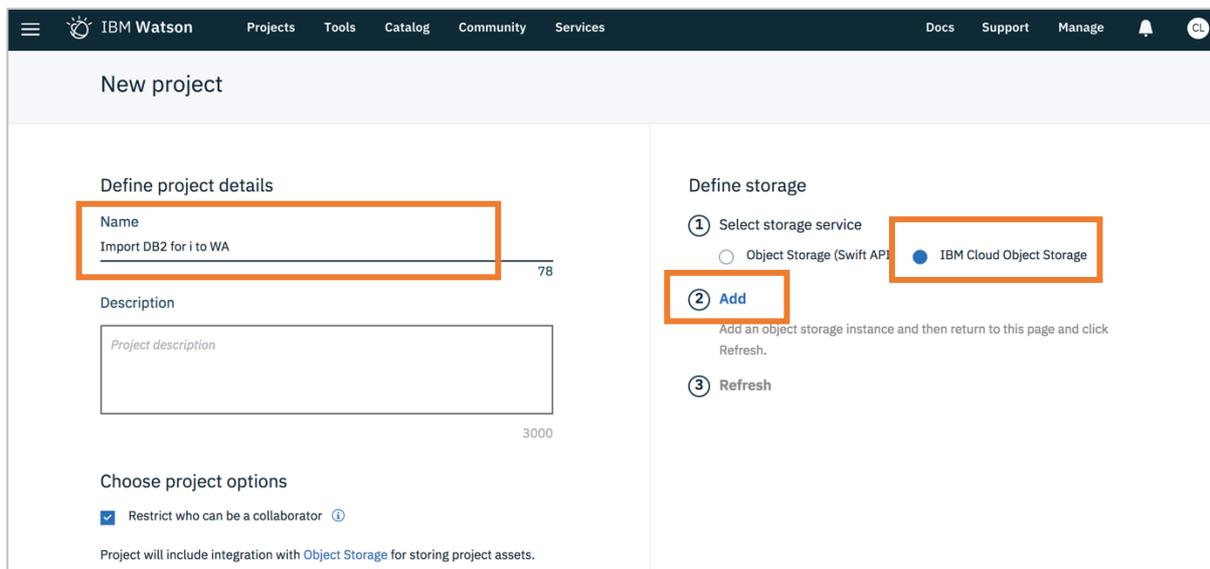


Section 2. Création d'un projet

1. Vous allez maintenant créer un nouveau projet d'importation de données, pour transférer les données de Db2 for i dans Watson Analytics. Cliquez sur **New Project.**, sélectionnez **Complete**, puis cliquez sur **OK**.

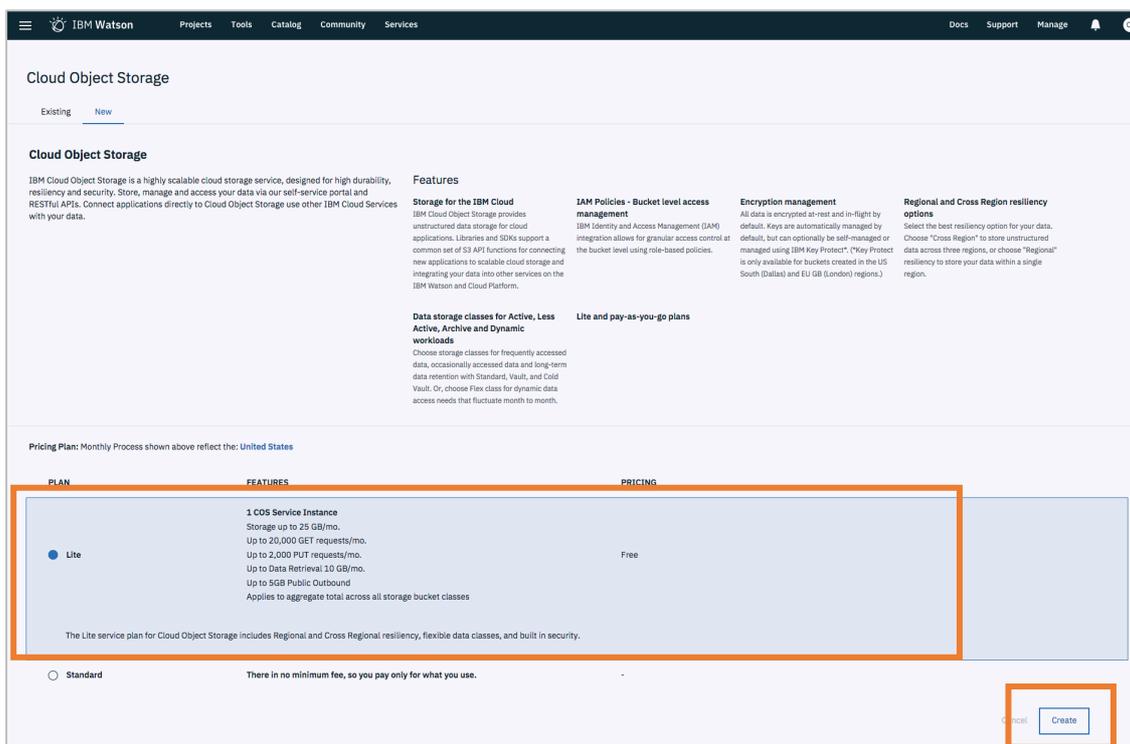


2. Nommez le projet et cliquez sur **Add** pour ajouter un service de stockage de type Cloud Object Storage (du stockage est nécessaire pour le fonctionnement du service).



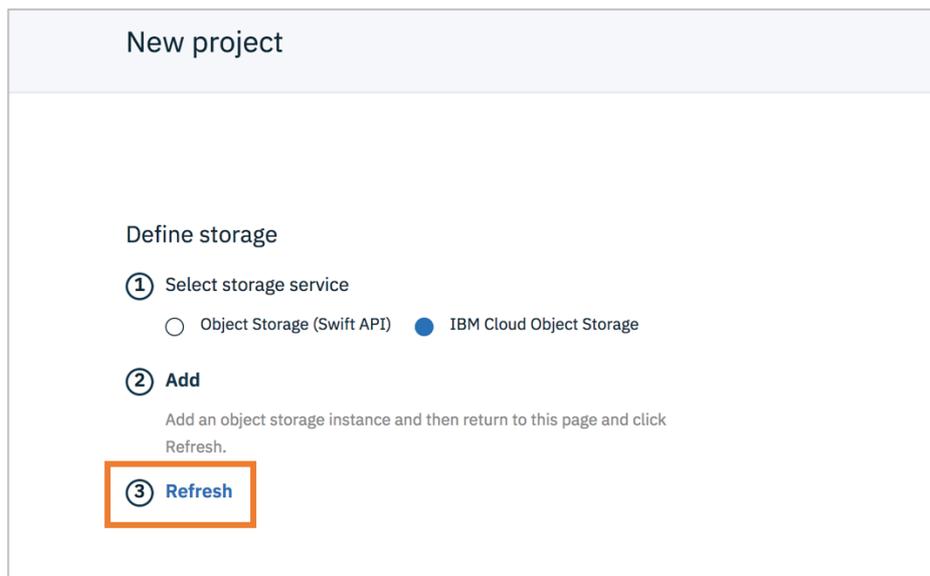
Un nouvel onglet s'ouvre dans votre navigateur sur la page de création du service stockage Cloud Object Storage.

3. Vérifiez que le plan **Lite** est sélectionné et cliquez sur **Create**.



4. Confirmez les paramètres et cliquez sur **Confirm** pour lancer la création du stockage.

5. De retour sur la page de création du projet, cliquez sur **Refresh** (bas de page) pour faire apparaître le service de stockage.



___ 6. Puis cliquez sur **Create** pour finalement créer le projet.

The screenshot shows the 'New project' form in the IBM Watson interface. The form is titled 'New project' and is divided into three main sections: 'Define project details', 'Choose project options', and 'Storage'.
1. **Define project details:** This section contains a 'Name' field with the text 'Import Db2 for i to WA' and a character count of 78. Below it is a 'Description' field with a placeholder 'Project description' and a character count of 3000.
2. **Choose project options:** This section has a checked checkbox labeled 'Restrict who can be a collaborator' with an information icon. Below this, it states 'Project will include integration with Object Storage for storing project assets.'
3. **Storage:** This section shows a dropdown menu with the selected option 'cloud-object-storage-ub'.
At the bottom right of the form, there are two buttons: 'Cancel' and 'Create'.

Votre projet est créé.

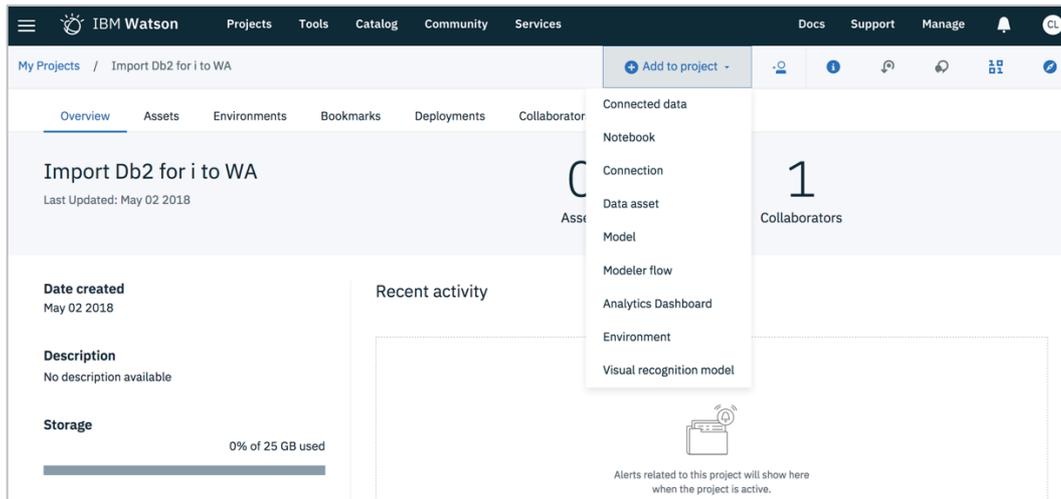
The screenshot shows the 'Overview' page for the project 'Import Db2 for i to WA'. The page header includes the IBM Watson logo and navigation tabs: 'Projects', 'Tools', 'Catalog', 'Community', 'Services', 'Docs', 'Support', and 'Manage'. Below the header, there are tabs for 'Overview', 'Assets', 'Environments', 'Bookmarks', 'Deployments', 'Collaborators', and 'Settings'.
The main content area displays the project name 'Import Db2 for i to WA' and 'Last Updated: May 02 2018'. To the right, there are three large numbers: '0' for Assets, '0' for Bookmarks, and '1' for Collaborators.
Below this, there are several sections:
- **Date created:** May 02 2018
- **Description:** No description available
- **Storage:** A progress bar showing '0% of 25 GB used'.
- **Collaborators:** A list showing one collaborator, 'Christophe Lalevee' (Admin), with a 'View all (1)' link.
- **Bookmarks:** A section showing 'You currently have 0 bookmarks' with a 'View all (0)' link.
On the right side, there is a 'Recent activity' section with a placeholder box and a message: 'Alerts related to this project will show here when the project is active.' The page also features a dark navigation bar at the top with various icons and a user profile icon.

Section 3. Création des connexions

Les sources de données de type base de données et la connexion à Watson Analytics sont configurées comme des « **Connections** ».

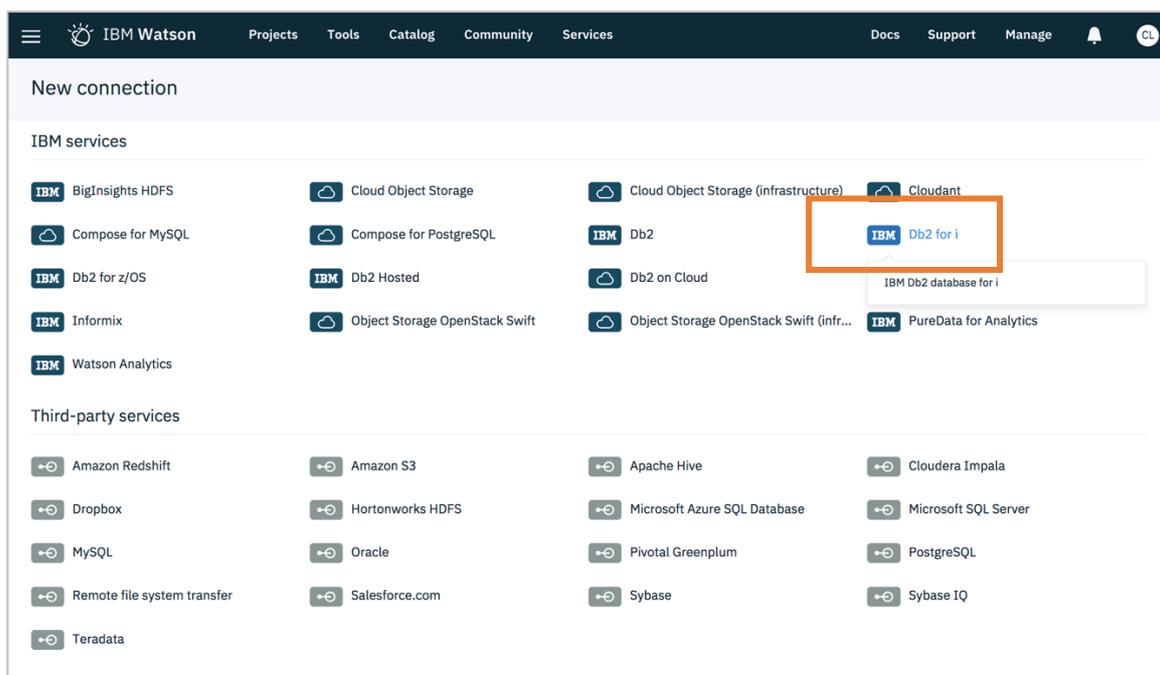
Vous allez donc maintenant configurer la connexion à une source de données externe (Db2 for i). Par la suite, vous configurerez la connexion nécessaire à Watson Analytics.

1. L'ajout des composants d'un projet, se fait à l'aide du menu **Add to project**. Ouvrez le menu **Add to project** et cliquez sur **Connection**.



Vous pouvez voir toutes les connexions disponibles (IBM ou non).

2. Cliquez sur **Db2 for i**.



Remarque

L'IBM i utilisé est hébergé dans un Datacenter IBM à Montpellier, France. Pour que vous puissiez y accéder à partir de Watson Knowledge Catalog, nous avons configuré pour vous le service d'intégration : IBM Secure Gateway.

Le déploiement et la configuration de ce service ne fait pas partie de ce hands-on. Vous utiliserez une instance préconfigurée pour vous.

L'IBM i aura donc un hostname de type *.ibmcloud.com, et Db2 sera joignable, via une translation de port, sur un port différent du port DRDA par défaut (446 ou 447) : vous accèderez à cet IBM i comme s'il était sur Internet, et non sur un réseau privé.

Pour configurer vous-même ce service, vous pouvez consulter la documentation en ligne comme, par exemple :

https://console.ibmcloud.com/docs/services/SecureGateway/secure_gateway.html

___ 3. Renseignez les paramètres de la connexion avec les valeurs suivantes :

- **Name** : Db2 for i MOP
- **Connection details**
 - **Hostname or IP address** : caplongprd-3.integration.ibmcloud.com
 - **Port** : 15060
 - **Location** : T01FF3B4
 - **Username** : LABIWA
 - **Password** : trentears
 - **Secure Gateway** : non coché

The screenshot shows the 'New connection (Db2 for i MOP - Db2 for i)' configuration page in the IBM Watson console. The page is divided into two main sections: 'Connection overview' and 'Connection details'.

Connection overview:

- Name:** Db2 for i MOP
- Description:** IBM Db2 database for i

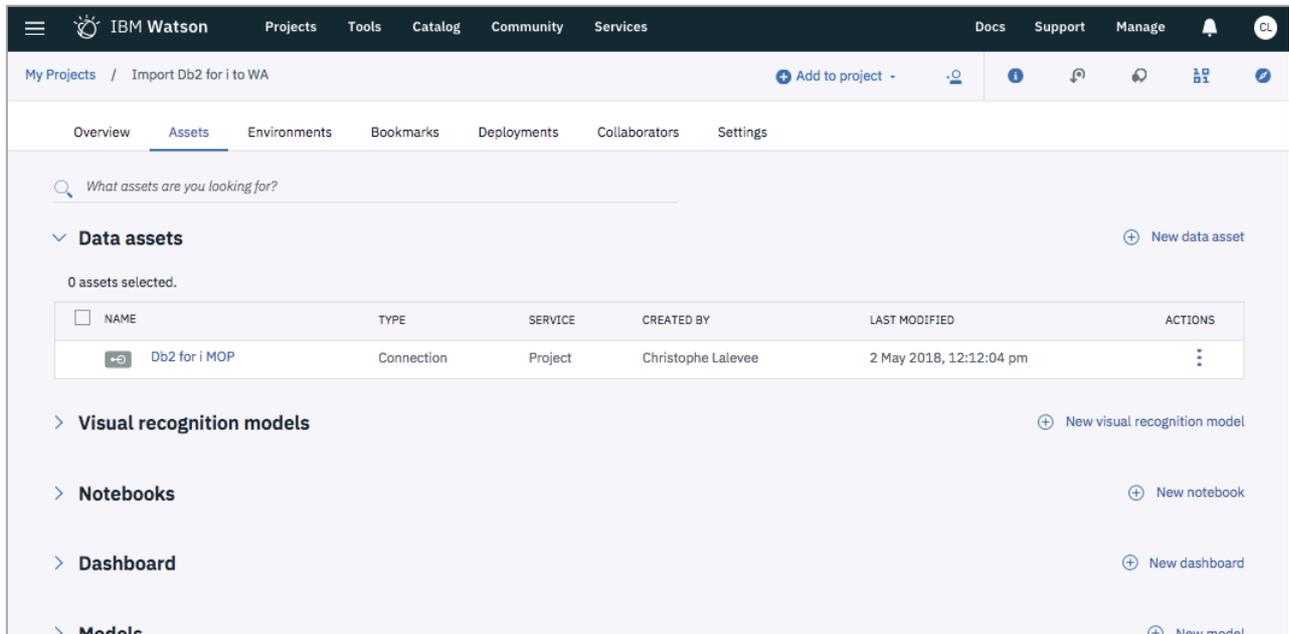
Connection details:

- Hostname or IP Address ***: onsgprd-3.integration.ibmcloud.com
- Location ***: T01FF3B4
- Secure Gateway**: Use a secure gateway
- Port ***: 15060
- Username ***: labiwa
- Port is SSL-enabled**: The port is configured to accept SSL connections
- Password ***:

At the bottom right, there are 'Cancel' and 'Create' buttons, along with a blue circular icon containing a white speech bubble.

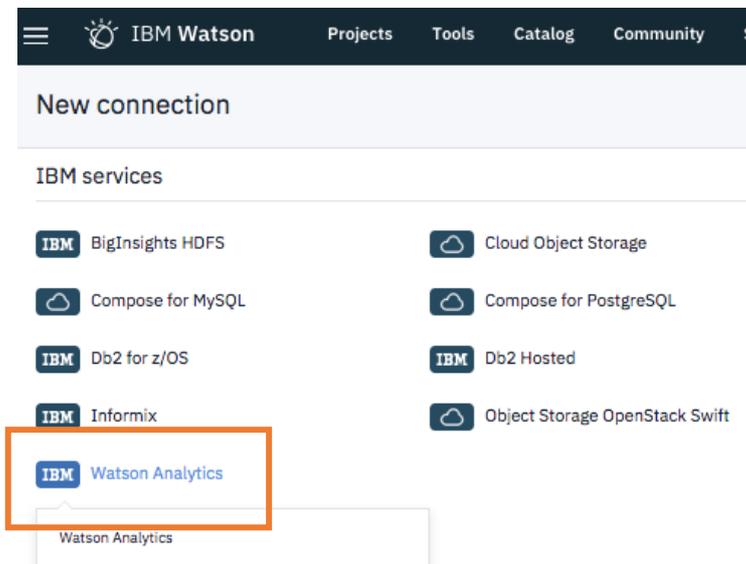
4. Cliquez sur **Create**.

Vous obtenez alors la vue globale du projet, avec la nouvelle connexion présente dans la section **Data Asset**.



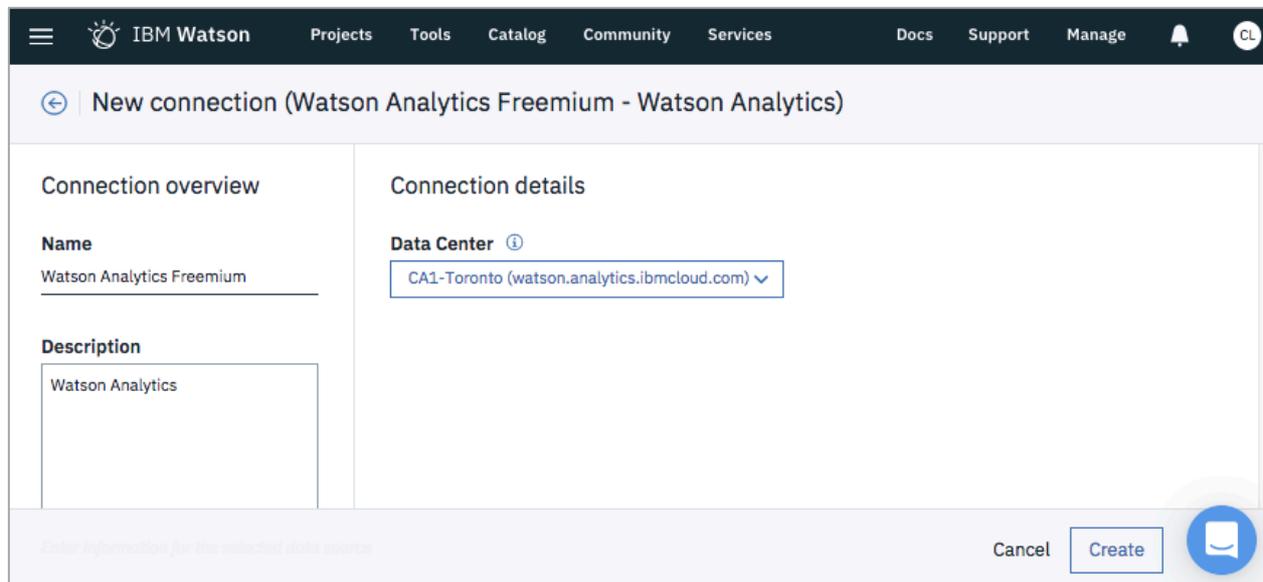
5. Vous allez maintenant créer la connexion pour Watson Analytics.

Comme précédemment, l'ajout de la connexion se fait par le menu **Add to project** puis **Connection**. Sélectionnez **Watson Analytics**.



Donnez un nom à la connexion, puis sélectionnez le Datacenter **CA1-Toronto**.

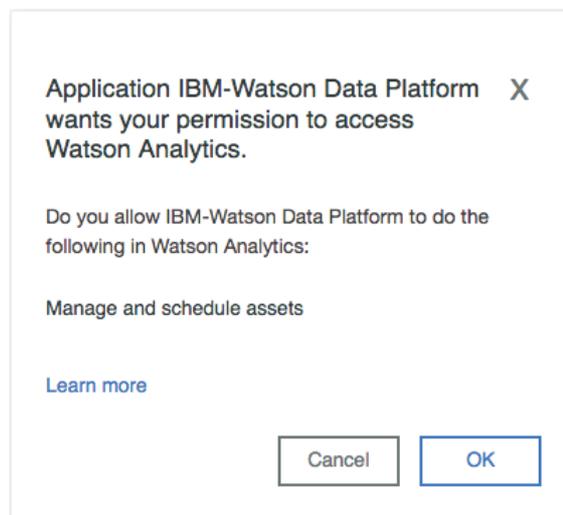
Ce Datacenter héberge les comptes Freemium de Watson Analytics, mais vous pouvez voir que d'autres Datacenters sont également disponibles pour les comptes non-freemium, notamment en Europe (Amsterdam).



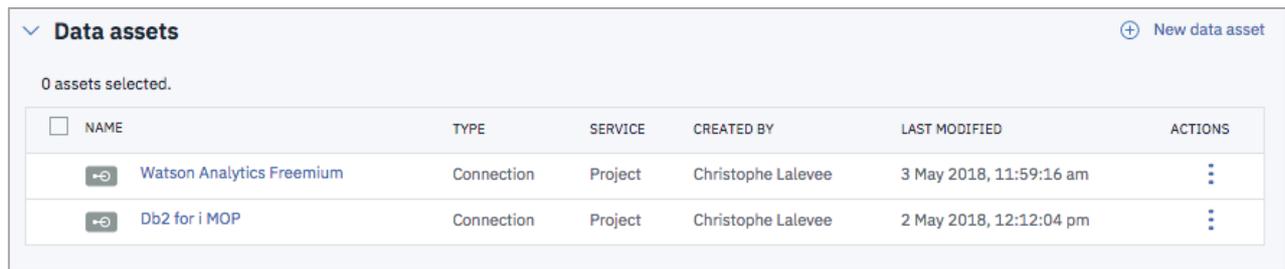
6. Cliquez sur **Create**.

Afin de valider la connexion à Watson Analytics, vous devez autoriser la connexion de Watson Knowledge Catalog (dans Watson Data Platform) à Watson Analytics (si vous n'êtes pas encore connecté à Watson Analytics, une phase d'authentification préalable aura lieu).

Pour cela, cliquez sur **OK**.



- ___ 7. Vous obtenez alors la vue globale du projet, avec la nouvelle connexion Watson Analytics présente dans la section **Data Asset**.



▼ **Data assets** ⊕ New data asset

0 assets selected.

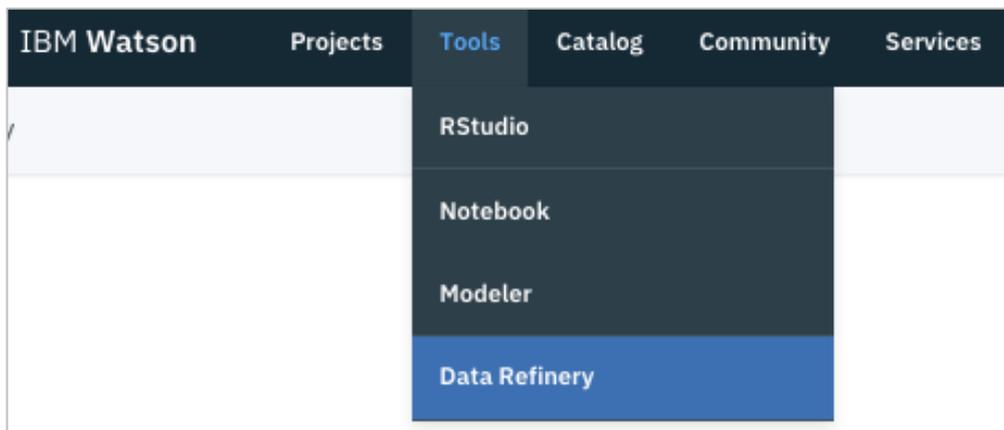
<input type="checkbox"/>	NAME	TYPE	SERVICE	CREATED BY	LAST MODIFIED	ACTIONS
<input checked="" type="checkbox"/>	Watson Analytics Freemium	Connection	Project	Christophe Lalevee	3 May 2018, 11:59:16 am	⋮
<input checked="" type="checkbox"/>	Db2 for i MOP	Connection	Project	Christophe Lalevee	2 May 2018, 12:12:04 pm	⋮

Vous allez maintenant créer le Data Flow entre la connexion à la base de données Db2 for i et la connexion à Watson Analytics.

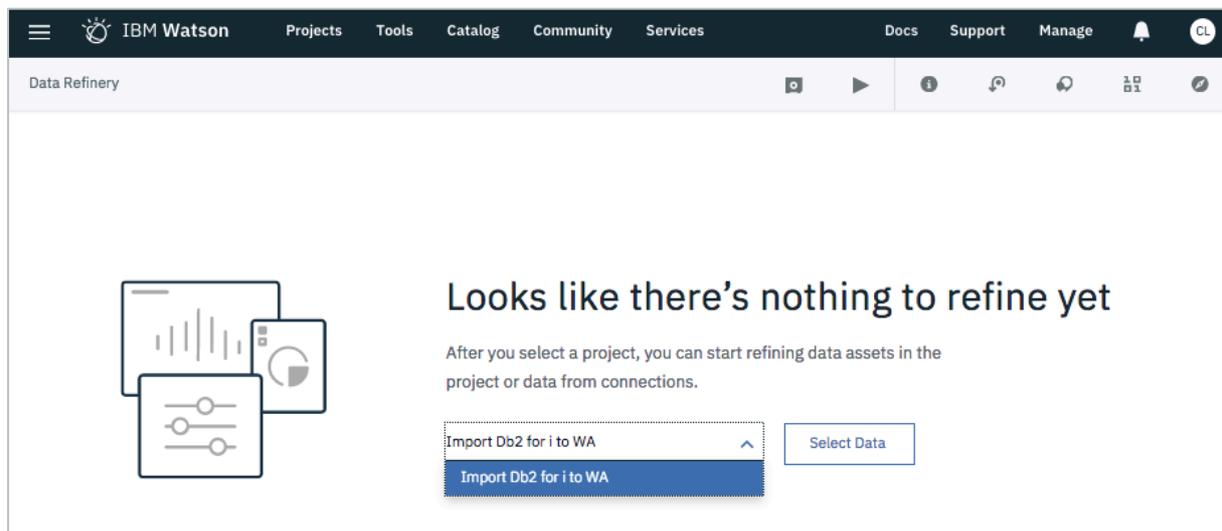
Section 4. Création et exécution du flux de données

Vous allez maintenant créer le Data Flow entre la connexion à la base de données Db2 for i et la connexion à Watson Analytics.

1. Dans le menu **Tools**, cliquez sur **Data Refinery**.



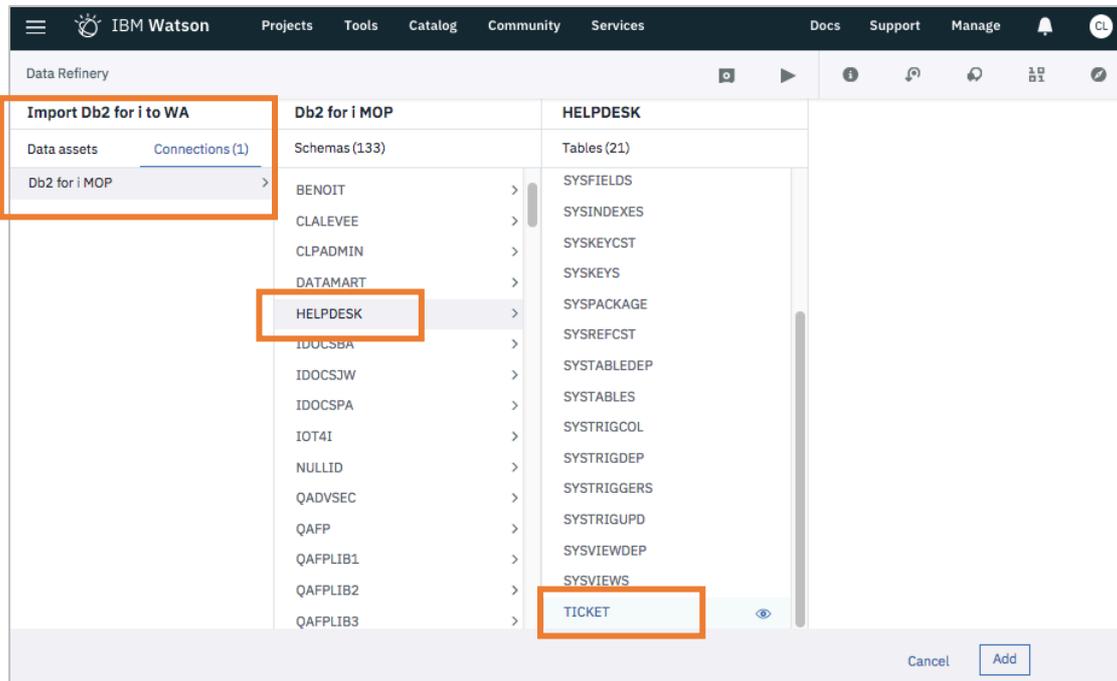
2. Sélectionnez le projet pour lequel vous voulez créer un flow et cliquez sur **Select Data**.



Vous allez maintenant définir la source de donnée du flow.

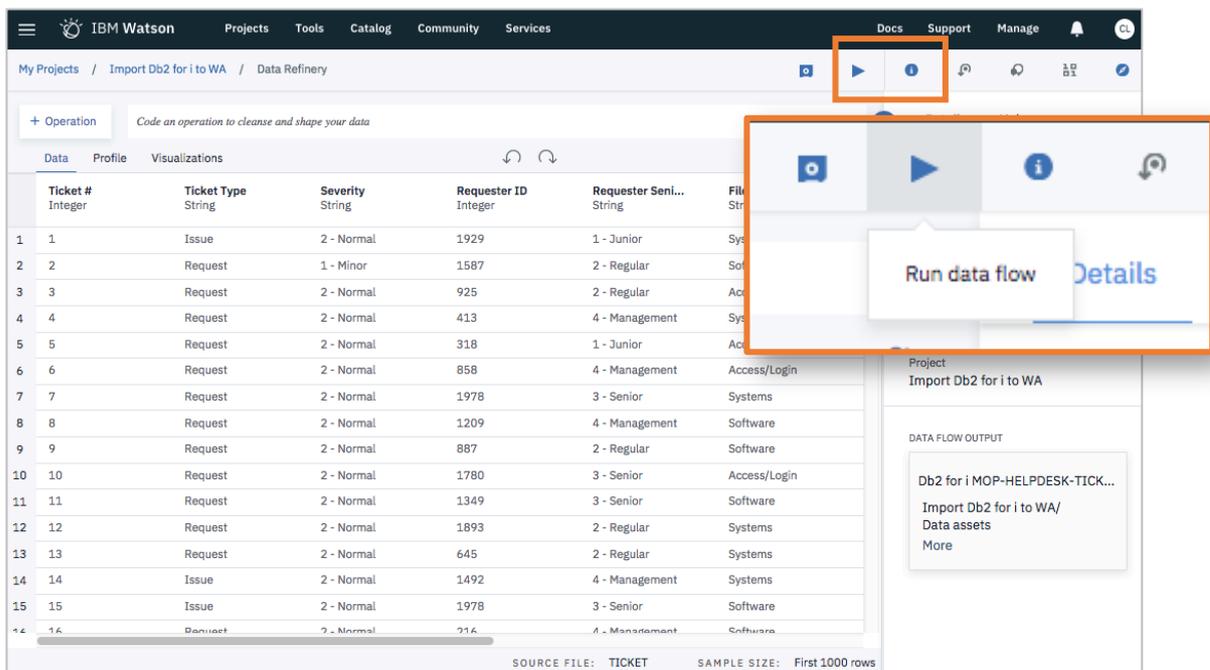
3. Sélectionnez l'onglet **Connection**, puis votre connexion à Db2 for i. Watson Knowledge Catalog se connecte alors à la base de données Db2 for i et vous permet de sélectionner, la ou les tables que vous voulez utiliser.

Dans la bibliothèque HELPDESK, sélectionnez la table TICKET. Cliquez sur **Add**.



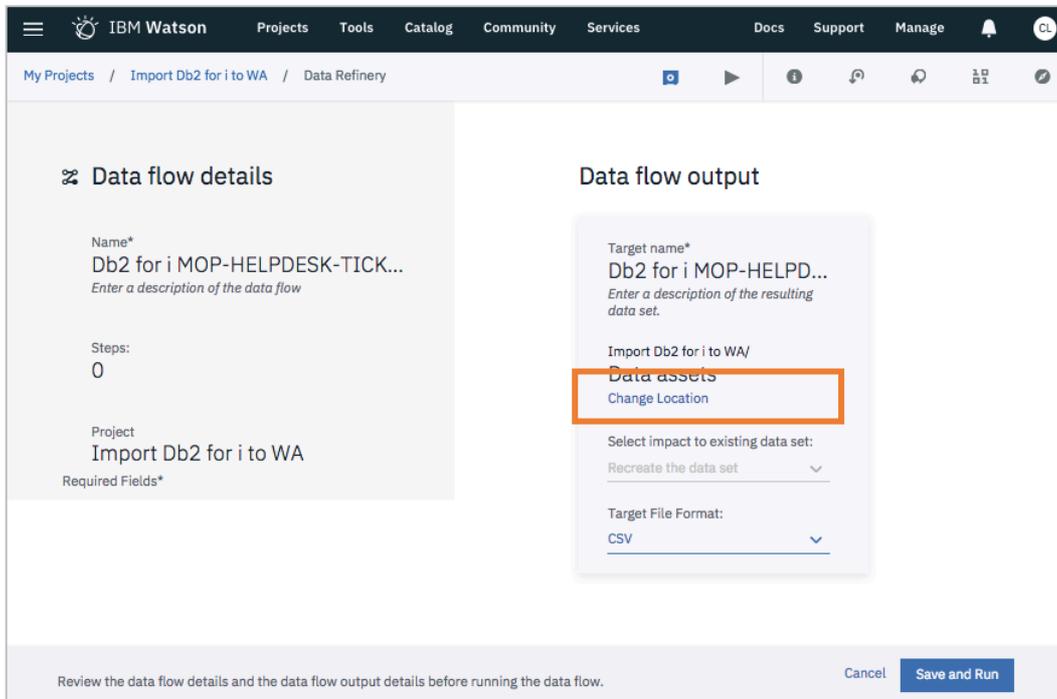
Un échantillon des données est chargé vous permettant de définir les actions de nettoyage et préparation des données. Les données étant « propres », nous n'avons pas besoin de raffiner la donnée ici.

4. Cliquez sur le menu flèche puis **Run Data Flow**.



Un résumé du flow est affiché. Si vous regardez la sortie définie (**Data Flow Output**), vous pouvez constater que les données seront extraites de l'IBM i et stockées dans un fichier CSV (Data Asset).

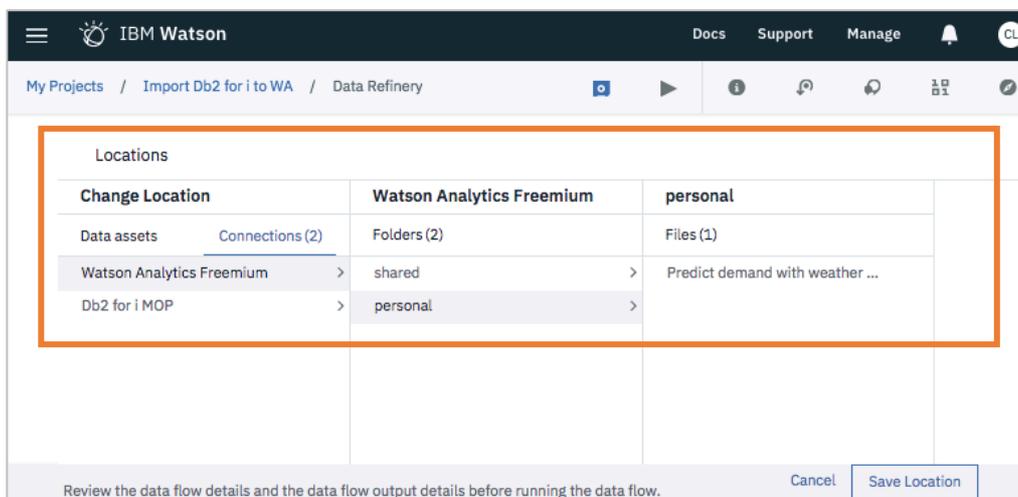
5. Comme vous voulez les envoyer directement dans Watson Analytics, vous allez modifier la sortie du flow. Cliquez sur **Change Location**.



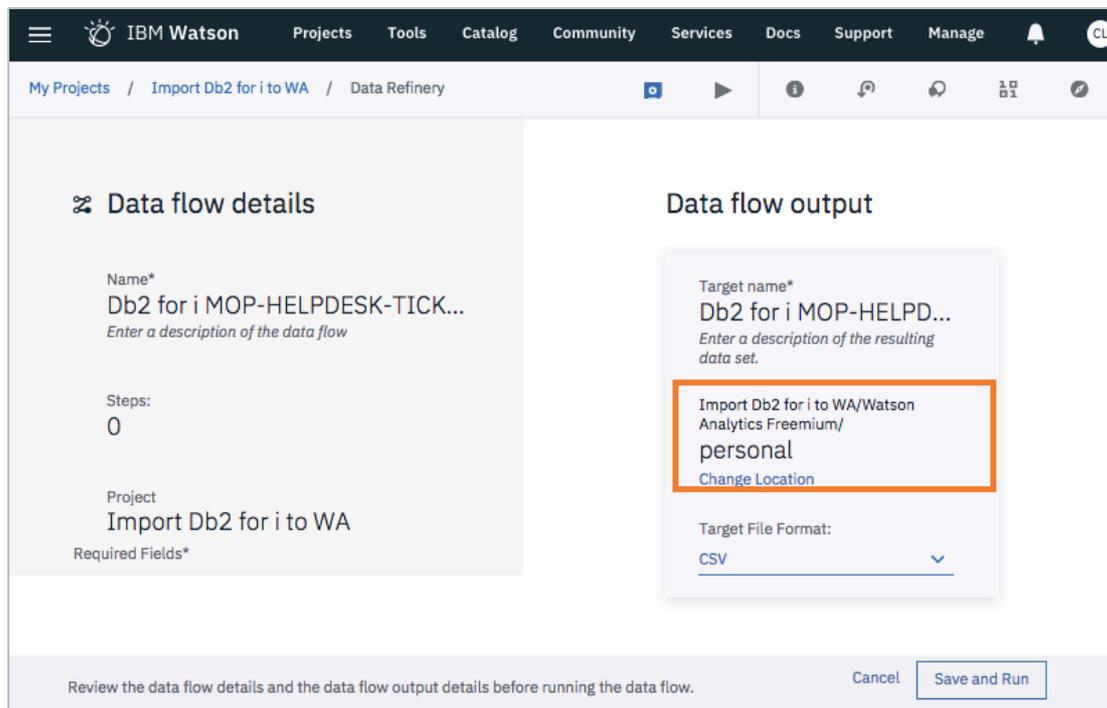
6. Sélectionnez l'onglet **Connections**, puis la connexion Watson Analytics que vous avez précédemment définie.

Watson Studio se connecte à votre instance Watson Analytics (vous l'avez autorisé précédemment) et vous montre le contenu de votre espace de travail.

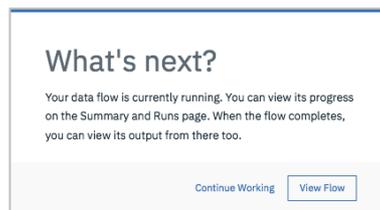
7. Sélectionnez le dossier **Personnal** et cliquez sur **Save Location**.



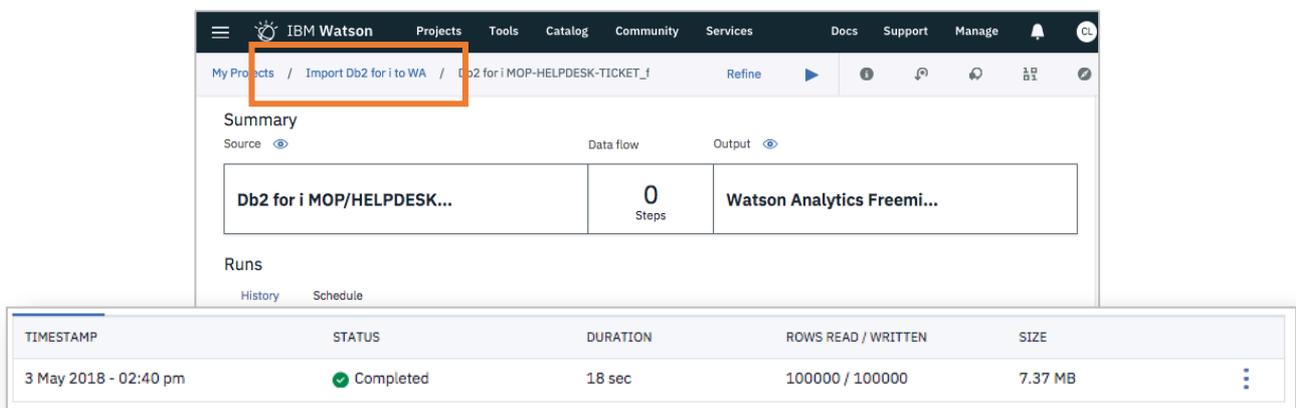
La destination du flow a changé : les données provenant de l'IBM i seront directement envoyées dans Watson Analytics.



- 8. Cliquez sur **Save and Run** pour exécuter manuellement ce flow.
L'exécution du flow est en cours. Cliquez sur **View Flow** pour visualiser sa progression.



Le flow prendra différents statuts avant de se terminer (**Completed**).



9. Cliquez sur le nom de votre projet (ici Import Db2 for i to WA) pour revenir au tableau de bord du projet.

Vous pouvez voir que vous avez créé les 3 composants nécessaires du projet : 2 connexions dans la section **Data Assets**, 1 flow dans la section **Flow**.

The screenshot displays the IBM Watson project interface for 'Import Db2 for i to WA'. The 'Assets' tab is active, showing a search bar and a list of assets. Two 'Data assets' are highlighted with an orange box:

NAME	TYPE	SERVICE	CREATED BY	LAST MODIFIED	ACTIONS
Watson Analytics Freemium	Connection	Project	Christophe Lalevee	3 May 2018, 11:59:16 am	
Db2 for i MOP	Connection	Project	Christophe Lalevee	2 May 2018, 12:12:04 pm	

Below the 'Data assets' section, there are expandable sections for 'Visual recognition models', 'Notebooks', 'Dashboard', 'Models', and 'Modeler flows'. The 'Data flows' section is also highlighted with an orange box, showing one data flow:

NAME	TYPE	CREATED BY	LAST MODIFIED	ACTIONS
Db2 for i MOP-HELPDESK-TICKET_flow	Data flow	Christophe Lalevee	3 May 2018, 2:27:57 pm	

Le flow créé a été exécuté une fois manuellement.

Vous pourrez le re-exécuter manuellement quand vous le souhaitez, ou, si besoin, programmer son exécution.

Retournez sur Watson Analytics pour voir les données importées et les explorer.

3. Analyse des données importées

Retournez sur la page de votre navigateur ouverte sur Watson Analytics.

Le cas d'usage sur lequel vous allez travailler est le suivant :

Le taux de satisfaction des utilisateurs de votre service support informatique (« Helpdesk ») a baissé pendant 3 trimestres consécutifs et, vous devez inverser cette tendance.

On vous a fourni des données sur des milliers de tickets fermés (100 000 tickets) avec des informations sur :

- les utilisateurs qui ont ouvert ces tickets,
- les agents de support à qui les tickets ont été attribués,
- le type de problème remonté,
- le domaine technique,
- la sévérité et la priorité des tickets,
- la durée pendant laquelle les tickets sont restés ouverts,
- et la satisfaction de l'utilisateur après qu'ils aient été résolus.

Voici un exemple des données disponibles :

Filed Against	IT Owner ID	Priority	Requester ID	Requester Se...	Severity	Ticket #	Ticket Type	Satisfaction	Days Open
Showing 1000 rows. Not all rows can be shown.									
Systems	28	3 - High	605	3 - Senior	2 - Normal	15004	Request	1 - Unsatisfied	14
Systems	47	0 - Unassigned	595	4 - Managem...	2 - Normal	15005	Request	0 - Unknown	8
Systems	3	0 - Unassigned	863	1 - Junior	2 - Normal	15006	Request	1 - Unsatisfied	22
Systems	26	0 - Unassigned	1594	3 - Senior	2 - Normal	15007	Request	0 - Unknown	8
Systems	14	3 - High	1955	3 - Senior	2 - Normal	15008	Request	0 - Unknown	6
Access/Login	7	2 - Medium	101	1 - Junior	3 - Major	15009	Request	1 - Unsatisfied	0
Systems	18	3 - High	1536	2 - Regular	2 - Normal	15010	Request	0 - Unknown	6
Systems	32	0 - Unassigned	741	1 - Junior	2 - Normal	15011	Request	2 - Satisfied	15
Access/Login	25	1 - Low	1957	2 - Regular	2 - Normal	15012	Request	2 - Satisfied	0
Hardware	35	0 - Unassigned	503	3 - Senior	2 - Normal	15013	Request	1 - Unsatisfied	12
Access/Login	29	3 - High	1013	4 - Managem...	2 - Normal	15014	Issue	3 - Highly sat...	0
Access/Login	32	1 - Low	1134	2 - Regular	2 - Normal	15015	Request	3 - Highly sat...	0
Access/Login	23	3 - High	1374	3 - Senior	2 - Normal	15016	Issue	2 - Satisfied	0
Systems	36	2 - Medium	621	3 - Senior	2 - Normal	15017	Request	3 - Highly sat...	5
Systems	21	3 - High	431	4 - Managem...	2 - Normal	15018	Request	3 - Highly sat...	5
Systems	21	2 - Medium	1168	2 - Regular	2 - Normal	15019	Request	3 - Highly sat...	8
Systems	38	0 - Unassigned	1761	2 - Regular	2 - Normal	15020	Request	0 - Unknown	6
Software	11	2 - Medium	1782	4 - Managem...	2 - Normal	15021	Request	0 - Unknown	11
Software	28	0 - Unassigned	1404	2 - Regular	2 - Normal	15022	Request	2 - Satisfied	14
Systems	40	0 - Unassigned	1746	3 - Senior	2 - Normal	15023	Request	2 - Satisfied	6
Hardware	44	2 - Medium	1373	3 - Senior	2 - Normal	15024	Issue	0 - Unknown	16

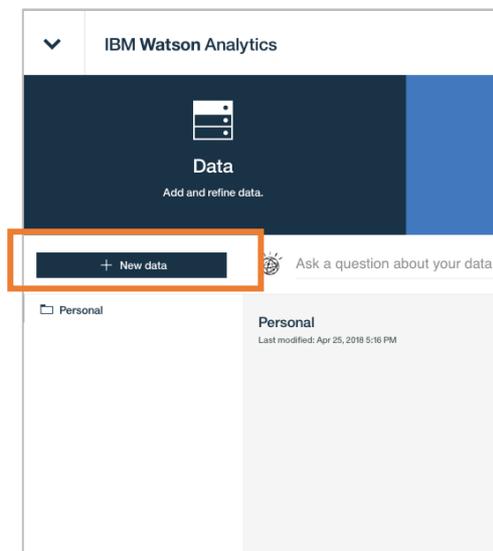
Vous allez maintenant tirer parti des capacités cognitives, prédictives et analytiques de Watson Analytics pour étudier ce qui détermine réellement les délais de résolution des tickets et les moyens de les améliorer pour renforcer la satisfaction de vos utilisateurs.

Section 0. Importation des données à partir d'un fichier Excel (option)

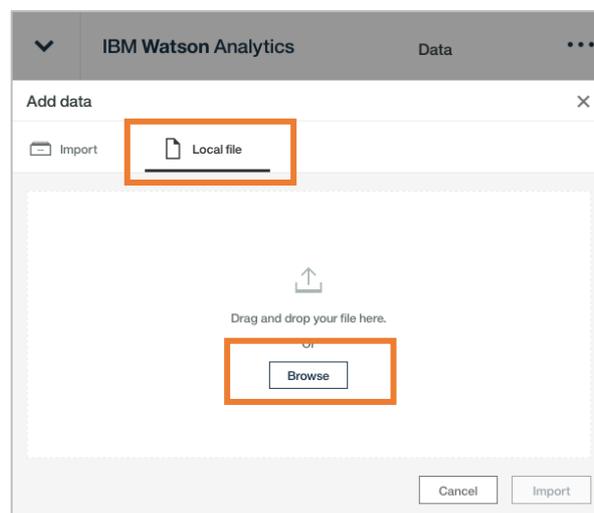
Si vous n'avez pas réalisé l'exercice 2, vous n'avez pas, dans Watson Analytics, le jeu de données nécessaire à l'exécution de ce nouvel exercice.

Vous allez donc devoir d'abord importer manuellement les données à partir d'un fichier Microsoft Excel.

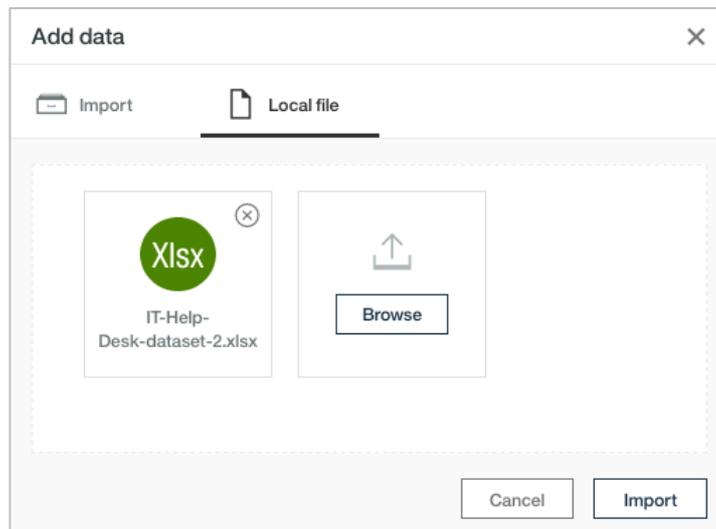
- ___ 1. Demandez à votre instructeur l'adresse du fichier **IT-Help-Desk-dataset-2.xlsx**
- ___ 2. Téléchargez-le sur votre poste de travail
- ___ 3. Sur l'interface de Watson Analytics, dans le module **Data**, cliquez sur le bouton **+ New data**.



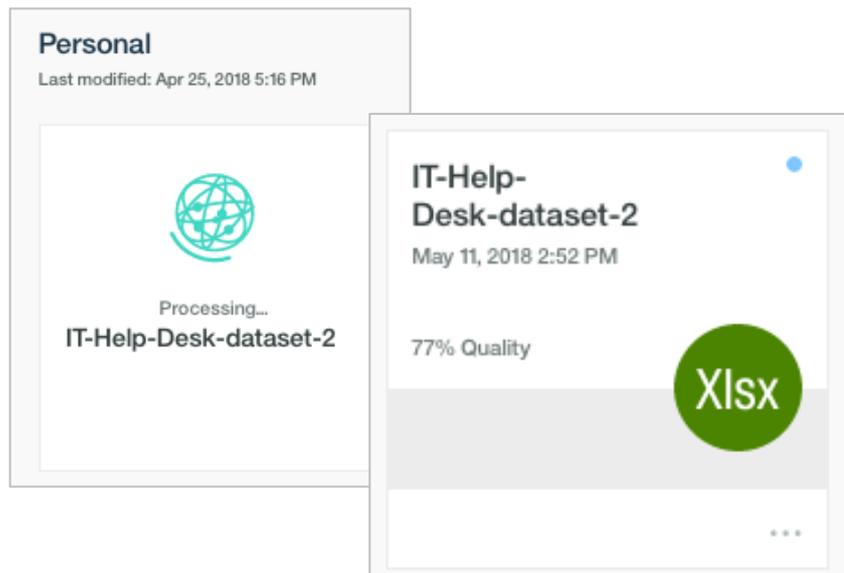
- ___ 4. Sélectionnez l'onglet **Local file** pour pouvoir charger le fichier à partir du disque de votre poste de travail. Cliquez sur **Browse**.



- ___ 5. Retrouvez le fichier sur votre disque dur, sélectionnez-le et cliquez sur **Open**. Cliquez sur **Import**



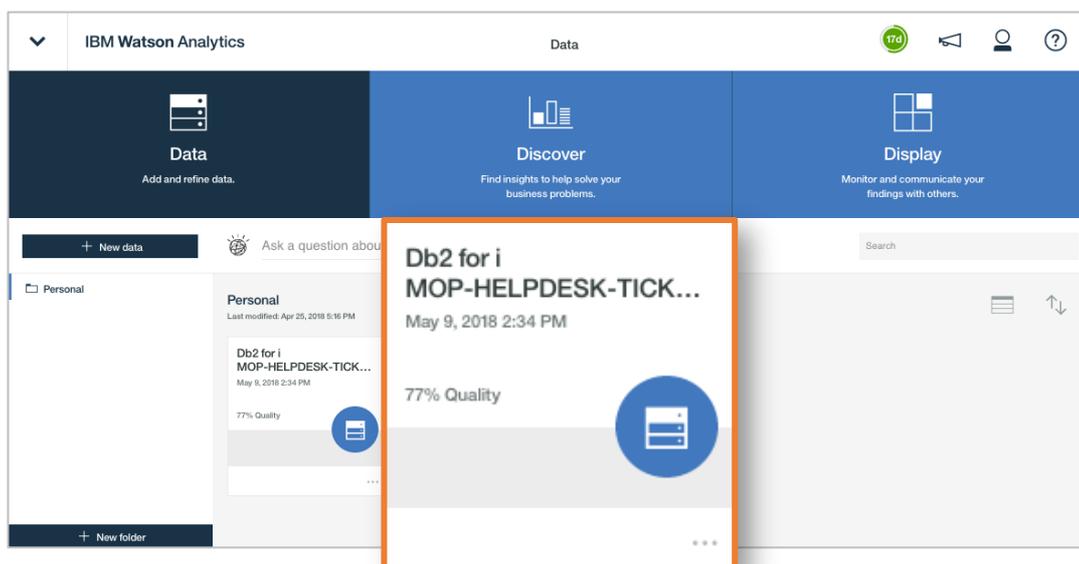
- ___ 6. Le fichier est téléchargé dans Watson Analytics, puis analysé (qualité des données). A la fin du traitement, vous pouvez disposer des données pour la suite de l'exercice.



Section 1. Prise en main de l'environnement

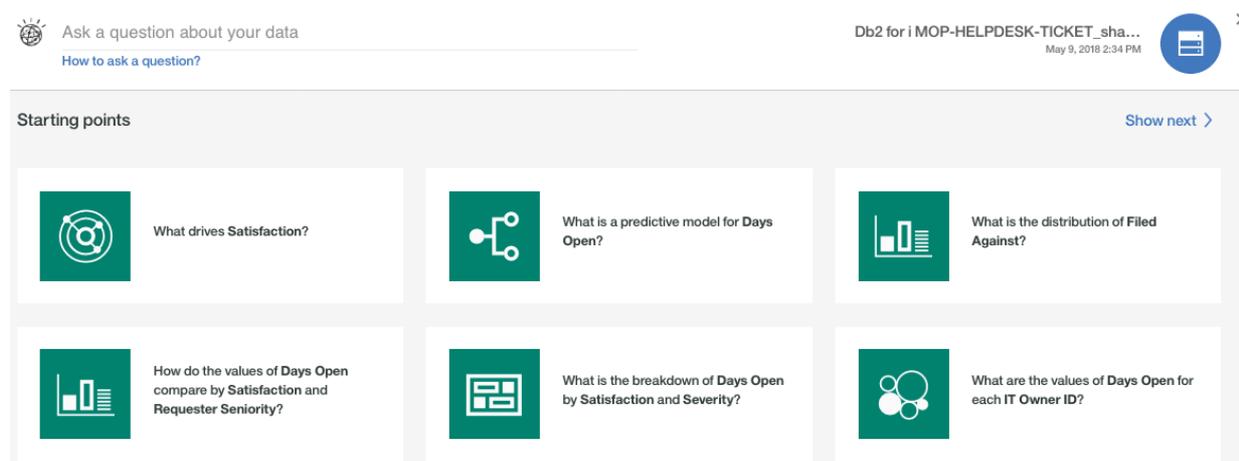
Pour prendre en main l'environnement, vous allez maintenant faire une première interrogation simple des données importées.

- ___ 7. Sur la page d'accueil, dans le dossier Personnel, vous retrouvez les données importées.



Watson Analytics a ingéré et analysé les données sous-jacentes et un score de qualité lui a été attribué (ici 77%) : cela vous aide à comprendre s'il y a des irrégularités dans vos données. Si ce score n'était pas suffisant, le module **Data - Add and refine data**, vous permettrait d'améliorer la qualité des données. Ici, dans notre exemple, le score de 77% est suffisant.

- ___ 8. Vous allez commencer votre investigation en cliquant simplement sur le jeu de données importé. Même sans avoir à taper une question, Watson Analytics vous présente des « points de départ » intuitivement découverts sur cet ensemble de données.



- ___ 9. Bien que vous puissiez commencer par l'un ou l'autre de ces points de départ, vous allez entrer votre propre question pour connaître la durée moyenne d'un ticket ouvert (colonne Days Open) : « **What is average of days open for hardware request?** », puis appuyez sur la touche **Entrée**.

Bien entendu, vous pouvez entrer une autre phrase ayant la même signification. Cependant, n'oubliez pas que la compréhension de Watson Analytics se base sur les entêtes de colonne et les données du Data set.

The screenshot shows the Watson Analytics interface. At the top, the search query is "What is average of days open for hardware request?". Below the search bar, there are six "Starting points" suggested by the system, each with an icon and a description:

- Most relevant**: What is the summary of Days Open? (Filtered by Filed Against: Hardware, Ticket Type: Request)
- Most relevant**: What is the distribution of Days Open? (Filtered by Filed Against: Hardware, Ticket Type: Request)
- Most relevant**: What is the breakdown of Days Open by Ticket Type? (Filtered by Filed Against: Hardware, Ticket Type: Request)
- Most relevant**: What are the most relevant values of Ticket Type for Days Open? (Filtered by Filed Against: Hardware, Ticket Type: Request)
- Most relevant**: What are the values of Days Open for each Ticket Type? (Filtered by Ticket Type: Request, Filed Against: Hardware)
- Most relevant**: What are the values of Days Open for Filed Against and Ticket Type? (Filtered by Ticket Type: Request, Filed Against: Hardware)

Watson Analytics vous présente les points d'entrée qu'il a déterminé en comprenant la requête que vous avez exprimé en langage naturel.

Notez qu'il donne pour chacun de ces points d'entrée un niveau de confiance dans la pertinence de sa proposition : **Most relevant** ou **Somewhat relevant** ou **Not really relevant**.

- ___ 10. La proposition qui correspond le plus à ce que nous cherchons est « **What is the summary of Days Open** ». Cliquez sur ce point d'entrée.

Vous obtenez la durée moyenne d'un ticket : Watson Analytics a bien compris votre demande.

6.84

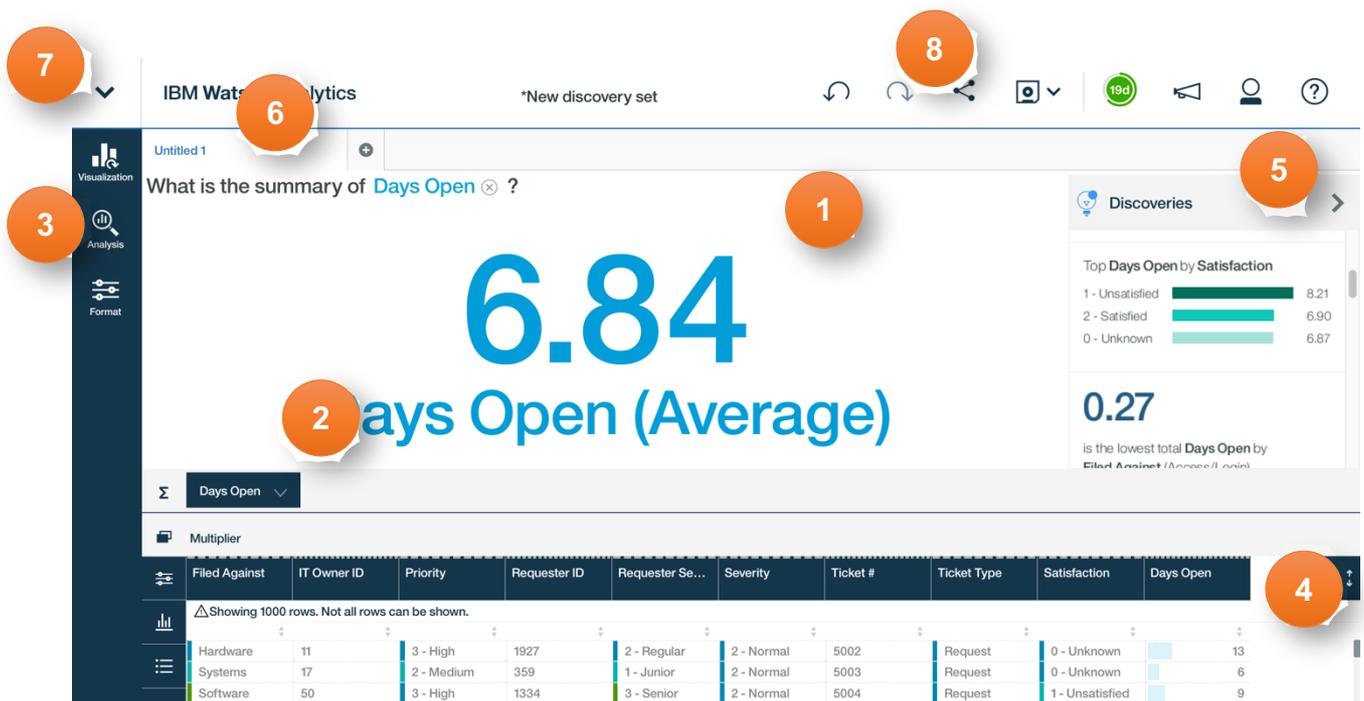
Days Open (Average)

Il a, par exemple, compris que

- le mot **Average** devait être utilisé comme une fonction d'agrégation,
- le mot **Hardware** était une valeur de la colonne **Filed Against**,
- le mot **Request** était une valeur de la colonne **Ticket Type**.

Cela a donc été traduit par une fonction d'agrégation (type GROUP BY en SQL) et 2 filtres (type WHERE en SQL).

__ 11. Étudions l'interface obtenue :



1. La présentation du résultat de la requête avec la méthode de visualisation déterminée par Watson Analytics
2. Une liste déroulante vous permettant de modifier la requête
 - a. Un clic gauche pour changer les colonnes
 - b. Un clic droit pour changer la fonction d'agrégation (moyenne, nombre total, ...)
3. Un menu vous permettant de changer la présentation, comme le type de graphique, les couleurs, l'affichage des légendes, ...
4. L'accès aux données « brutes » en cliquant sur les 2 petites flèches . Le menu de gauche de cette section vous permet d'obtenir des informations diverses sur les données, leur qualité, leur format, ...
5. Des propositions d'analyses ou de recherches que Watson Analytics a jugé pertinentes pour compléter la requête en cours. Par exemple, ci-dessus, il vous propose de croiser la durée d'un ticket (son temps de résolution) avec la satisfaction finale de l'utilisateur.
6. Le nom de ce rapport (« Untitled 1 ») : un clic sur le nom vous permettra de le remplacer par quelque chose de plus parlant
7. Une liste déroulante (▼) vous permettant de naviguer dans les différents rapports que vous avez ouverts ou de revenir à la page d'accueil de Watson Analytics. Dans cette liste, un clic sur la croix (✕) permet de fermer le rapport ouvert.
8. Le menu général vous donnant accès notamment, à l'annulation de la dernière action ou à la sauvegarde du rapport en cours.

__ 12. Fermez ce premier rapport sans le sauvegarder.



Section 2. Analyse des données

- ___ 1. Cliquez de nouveau sur votre jeu de données importées.

Notre intention étant de comprendre ce qui détermine réellement les délais de résolution des tickets, tapez la requête en langage naturel « **What drives days open** » (= « qu'est ce qui détermine le nombre de jours ouverts »).

- ___ 2. Watson Analytics répond immédiatement avec un ensemble de points d'entrée, classés par ordre de pertinence.
Cliquez sur la visualisation du diagramme en spirale la plus pertinente pour commencer votre analyse : « **What drives day open** ».

The screenshot shows the IBM Watson Analytics interface. At the top, there's a search bar with the query 'What drives days open' and a 'How to ask a question?' link. Below the search bar, there's a 'Starting points' section with a 'Show next >' link. The starting points are arranged in a grid of six cards. The first card, 'Most relevant: What drives Days Open?', is highlighted with an orange border. The other cards include: 'Most relevant: What is a predictive model for Days Open?', 'Most relevant: What is the summary of Days Open?', 'Most relevant: What is the distribution of Days Open?', 'Somewhat relevant: What is the breakdown of Days Open by Severity?', and 'Somewhat relevant: What are the values of Days Open for each IT Owner ID?'.

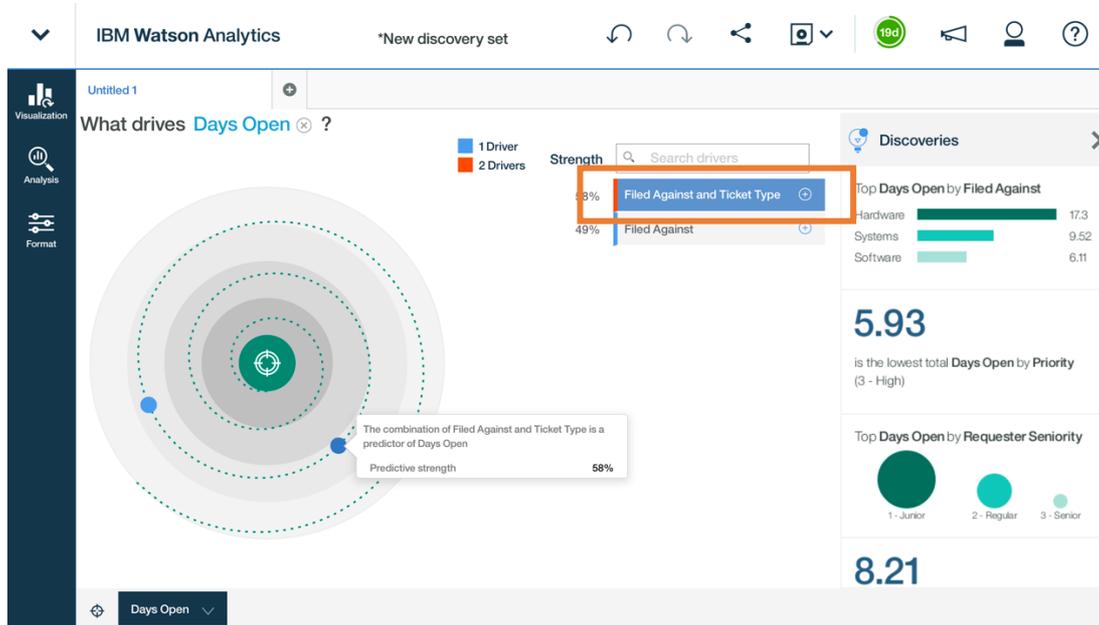
- ___ 3. Le diagramme en spirale vous montre tous les facteurs qui déterminent le nombre de jours d'ouverture d'un ticket ou, en d'autres termes, le temps de résolution du ticket. Plus un facteur est proche du centre de la spirale, plus la confiance dans la prédiction est forte.

Notez que la question que vous avez posée, n'est pas une simple interrogation de données, mais une analyse prédictive.

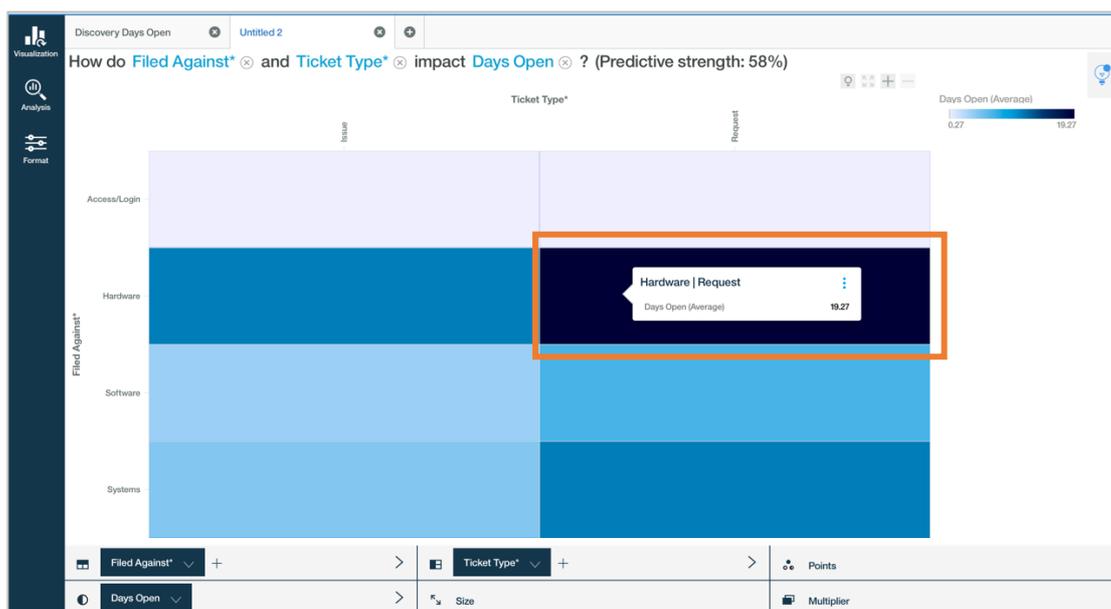
- ___ 4. Changez le type d'agrégation pour prendre en compte les temps moyens.

The screenshot shows a dropdown menu for aggregation options. The menu is titled 'Summarize' and has a 'Back' button. The options are: 'Sum (Default)', 'Average' (selected with a blue checkmark), 'Minimum', 'Maximum', 'Count', and 'Count distinct'. There is a 'Reset' button at the bottom of the menu. The background shows a partial view of a spiral diagram and a table with columns like 'Days Open', 'Filed Against', 'Requester Se...', and 'Severity'.

- ___ 5. Regardez la liste des « Drivers » (= « déterminants ») proposés. Vous pouvez vous rendre compte que la combinaison de **Filed Against**, la catégorie à laquelle le ticket a été assigné, et de **Ticket Type**, le type du ticket, est la raison la plus probable dans la détermination du temps de résolution de ticket.

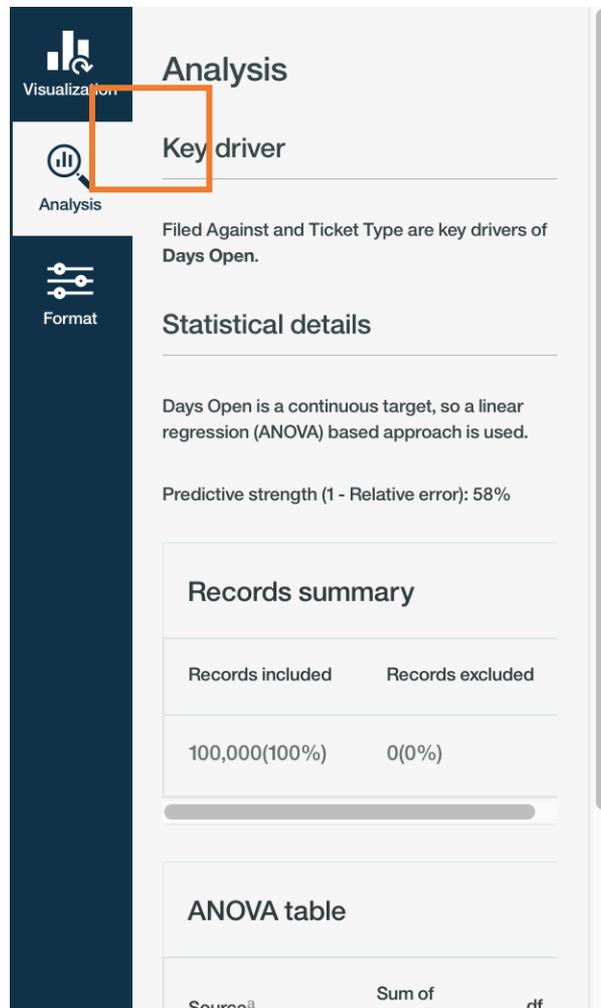


- ___ 6. Donnez le nom **Discovery Days Open** à cette découverte.
- ___ 7. Poursuivez votre analyse en examinant les détails de la combinaison de facteurs la plus importante. Dans la liste des « drivers », cliquez sur **Filed Against and Ticket Type** pour approfondir cette piste.
- ___ 8. Un nouvel onglet s'ouvre présentant les interactions entre ces 2 colonnes.



Il est évident que les demandes relatives au matériel sont celles qui prennent le plus de temps à résoudre.

- ___ 9. Notez que Watson Analytics vous a suggéré cette combinaison en se basant sur des modèles statistiques pertinents, confrontés à votre jeu de données. Les informations sur les modèles statistiques utilisés sont disponibles dans le menu **Analysis** à gauche de l'interface (approche de type régression linéaire ANOVA, dans cet exemple).

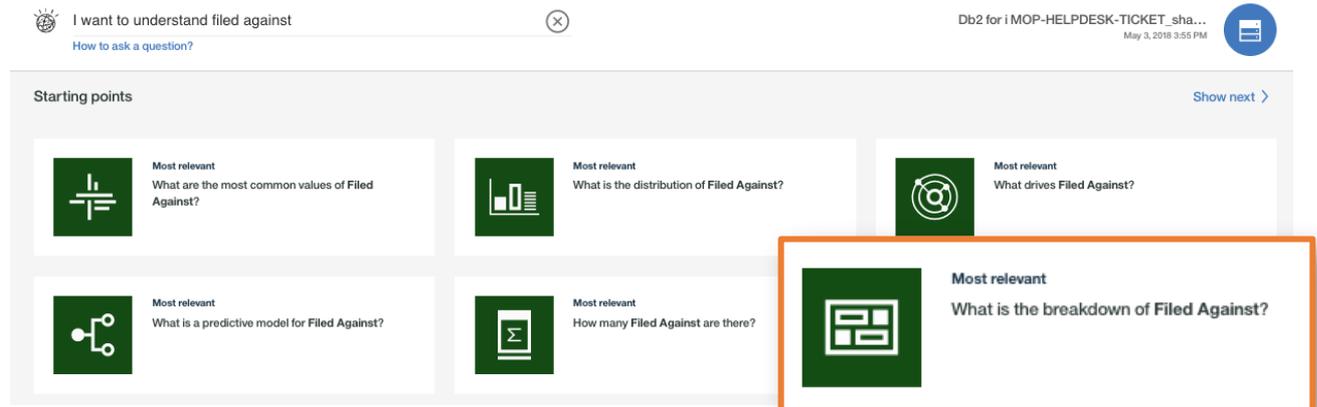


- ___ 10. Donnez le nom **Top Driver Combination** à ce rapport et ouvrez un nouvel onglet.



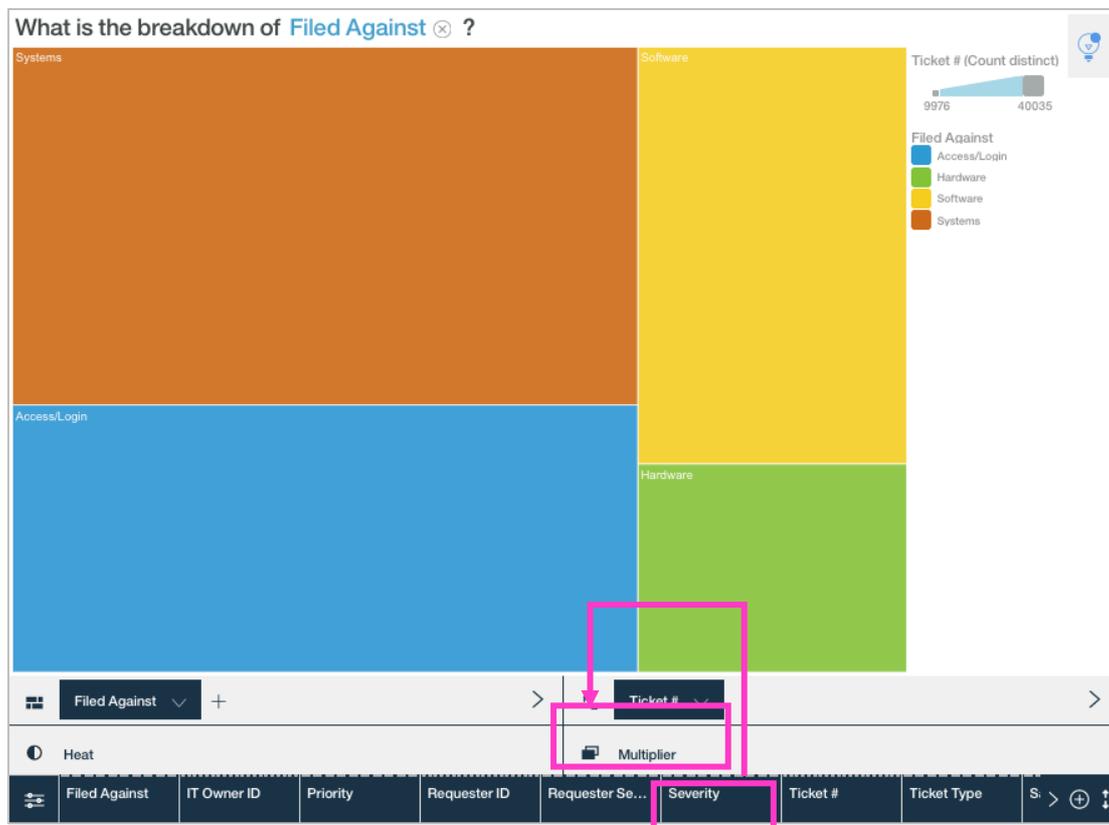
- ___ 11. Tapez une nouvelle question pour en savoir plus sur les catégories de ticket (colonne « Filed Against ») : **I want to understand filed against.**

___ 12. Sélectionnez la visualisation de type Treemap la plus pertinente : **What is the breakdown of Filed Against?**



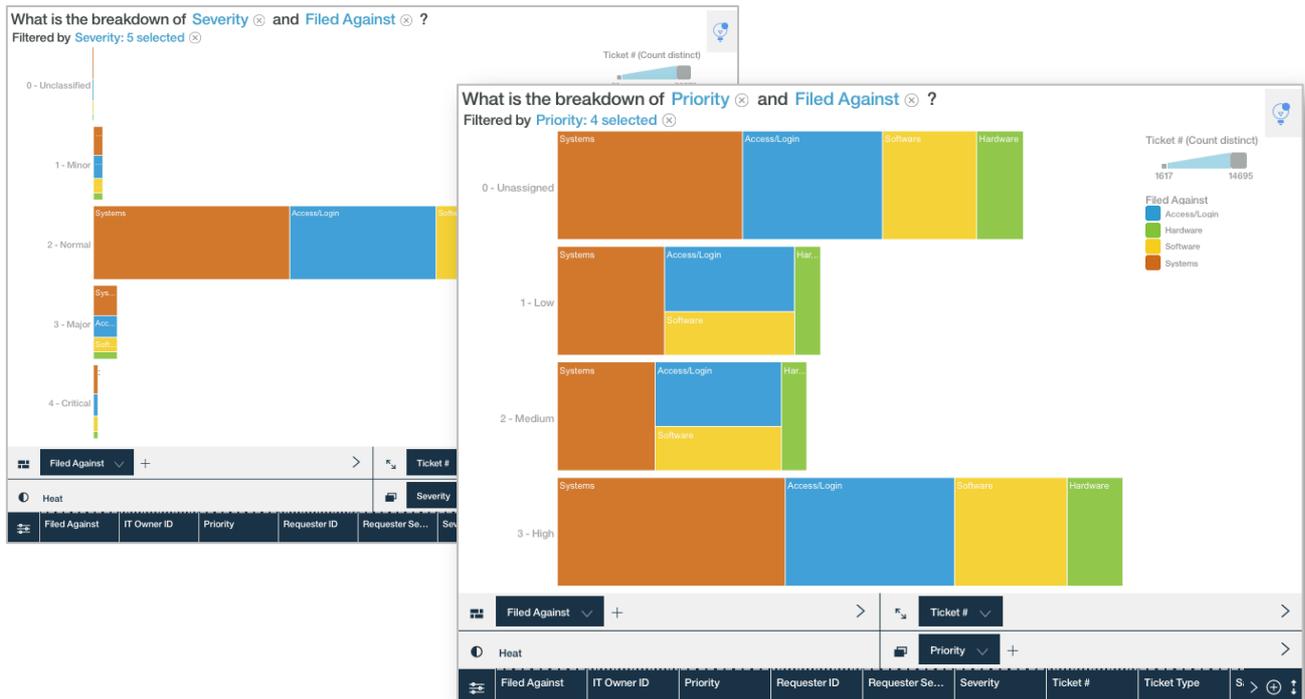
Vous obtenez la visualisation ci-dessous qui vous permet de voir la répartition du nombre de tickets selon leur catégorie (colonne **Filed Against**).

___ 13. Vous allez maintenant croiser cette information avec la sévérité (colonne **Severity**), puis avec la priorité (colonne **Priority**), allouées aux tickets. Pour cela, glissez le champ **Severity** (bandeau au bas de la fenêtre) sur la case **Mutiplier**.



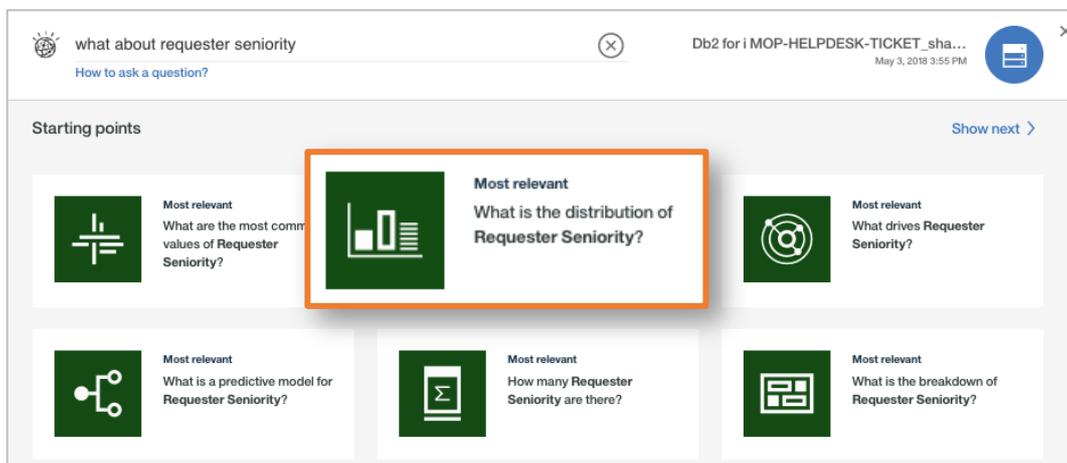
Interprétez le résultat.

Puis, faites la même manipulation, pour croiser la catégorie des tickets avec leur priorité.



La catégorie **System** affiche le plus grand nombre de tickets dans tous les niveaux de priorité.

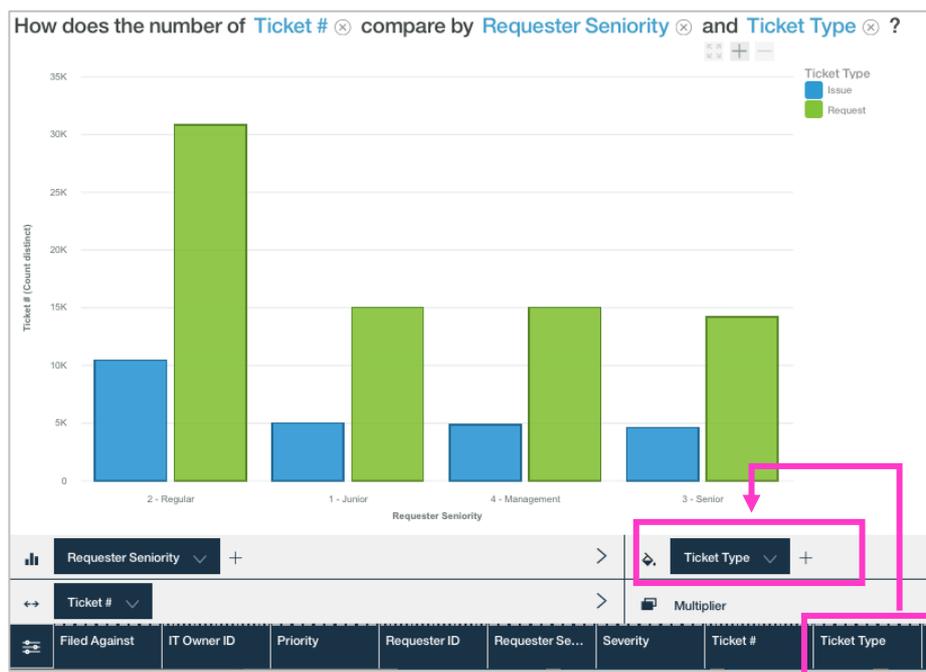
- ___ 14. Renommez cette analyse **Category & Priority** et sauvegardez.
- ___ 15. Jusqu'à présent, nous avons découvert ce qui détermine le temps de résolution des tickets et la répartition du volume de tickets sur les différentes catégories et priorités. Regardons maintenant si l'ancienneté de l'utilisateur a un impact sur la résolution des tickets. Créez un nouvel onglet et posez la question suivante : **What about requester seniority ?**
- ___ 16. Cliquez sur la visualisation en barre la plus pertinente.



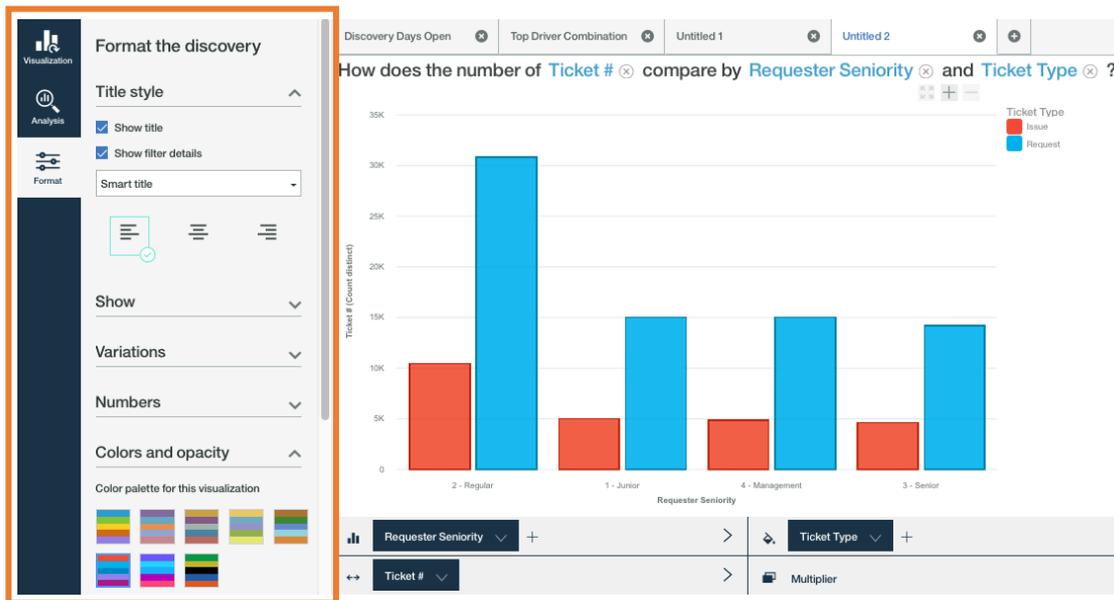
___ 17. Triez la série **Requester Seniority** par nombre de tickets descendant (clic bouton droit sur le titre de la série, puis **Sort, Descending by value**).



___ 18. Croisez l'ancienneté avec le type de ticket ouvert : glissez-déposez la colonne **Ticket Type** sur la case **Color** (cf ci-dessous).

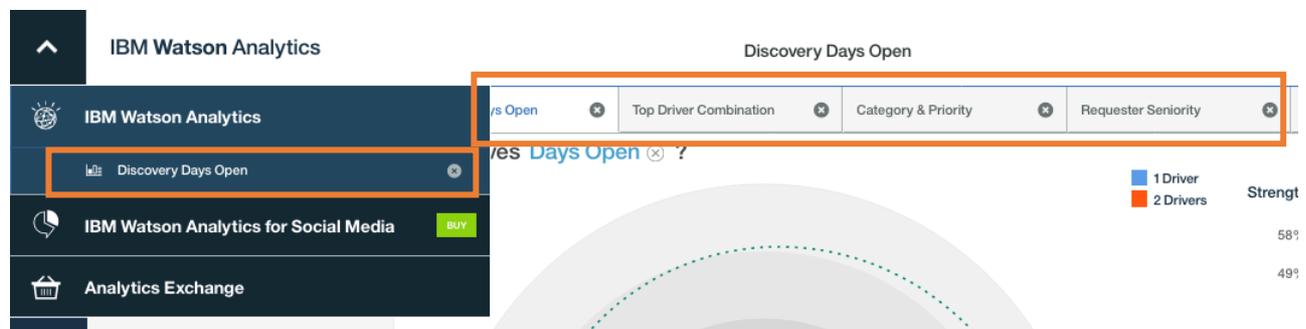


___ 19. Utilisez le menu **Format** du bandeau de gauche pour modifier la couleur de votre graphique.



___ 20. Renommez cette analyse **Requester Seniority** et sauvegardez.

___ 21. A ce stade de l'exercice, vous avez un ensemble de découvertes (un « Discovery Set », ci-dessous nommé **Discovery Days Open**) contenant 4 rapports ou « discoveries » (4 onglets).



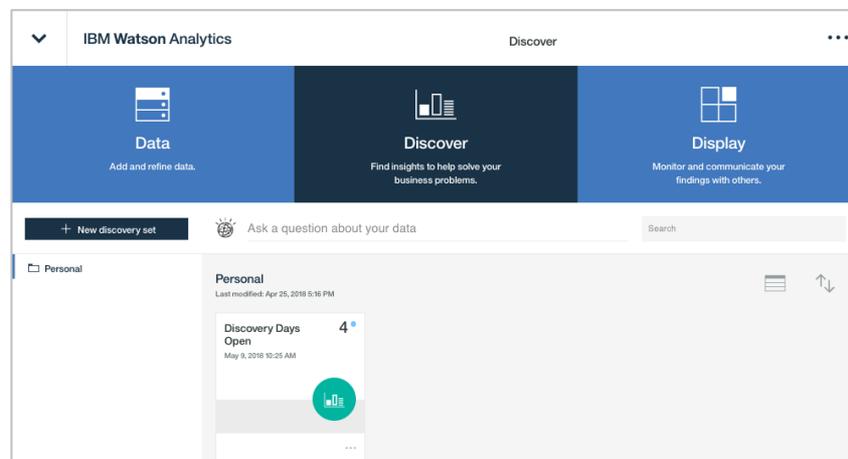
Vous allez maintenant créer un tableau de bord pour partager ces informations.

Section 3. Création d'un tableau de bord

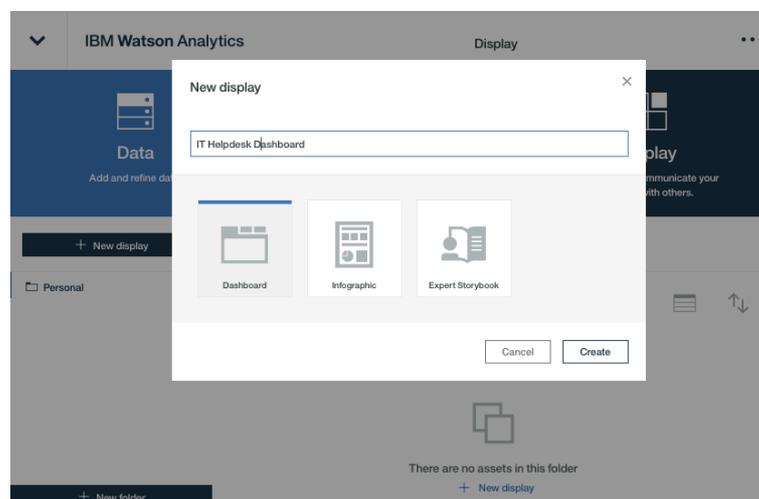
1. En utilisant le menu déroulant en haut à gauche de l'interface, revenez sur le menu principal de Watson Analytics.



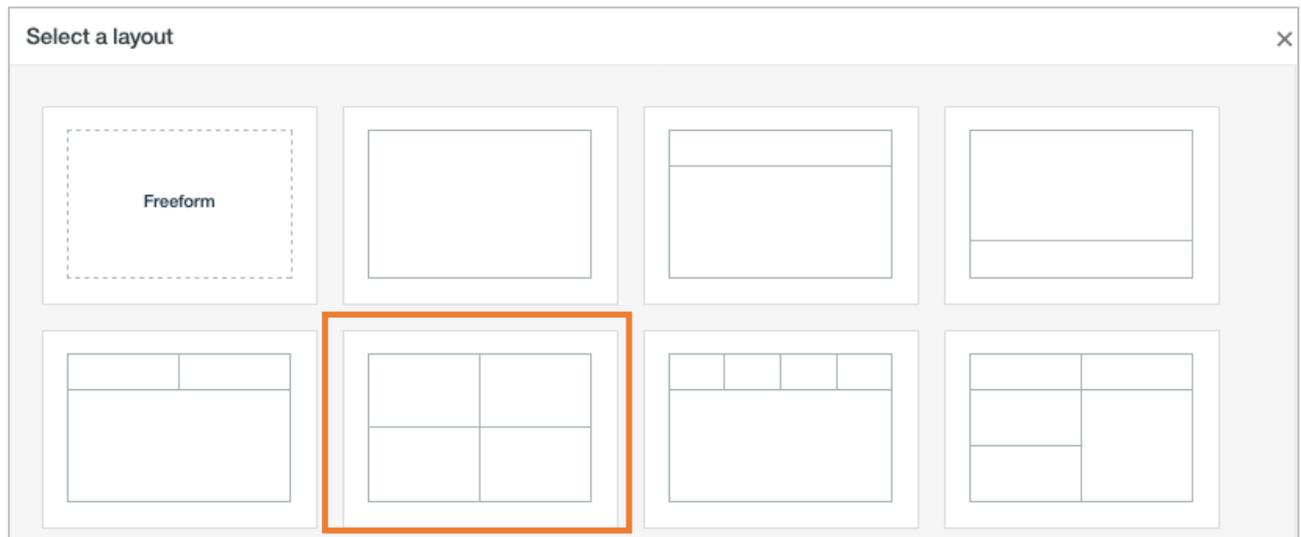
Dans la section précédente, vous venez d'utiliser le module **Discover** et vous pouvez maintenant voir votre **Discovery Set** dans votre dossier **Personnel**.



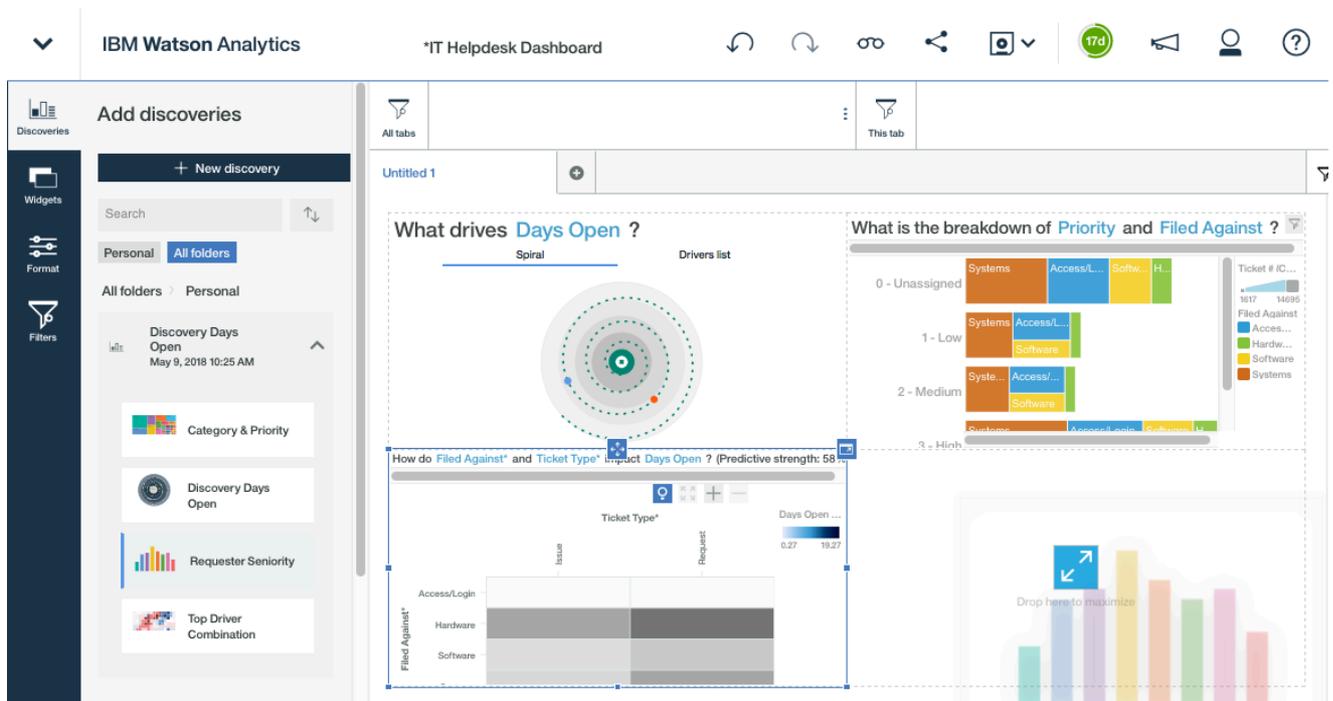
2. Cliquez sur le module **Display**. Le dossier est vide.
3. Cliquez sur **New Display** et donnez un nom à votre tableau de bord.
4. Cliquez sur **Create**.



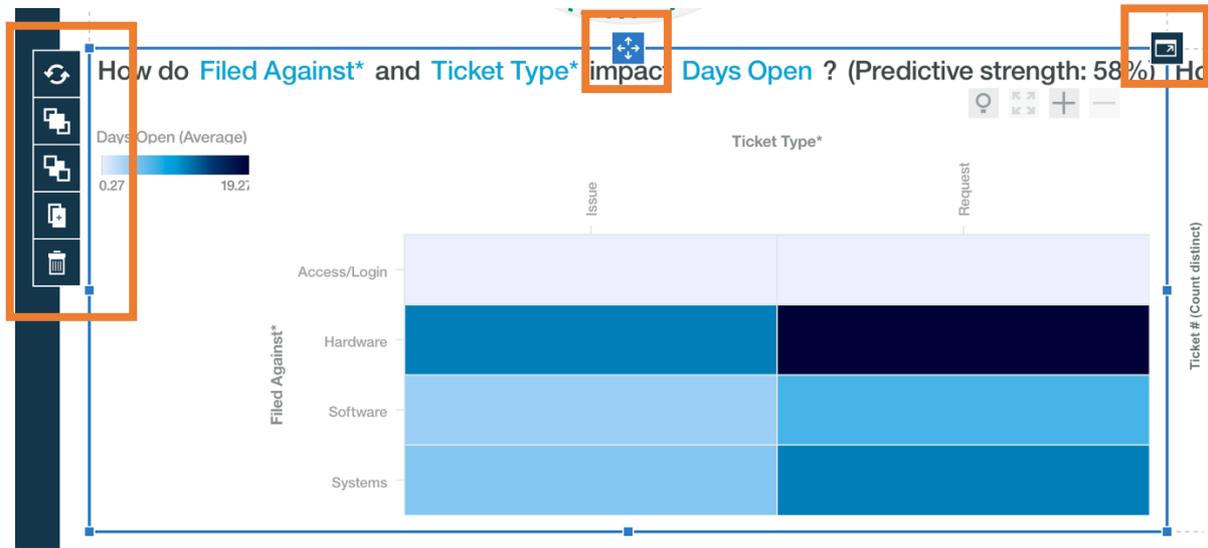
__ 5. Choisissez une mise en page (« Layout ») pour accueillir vos 4 rapports.



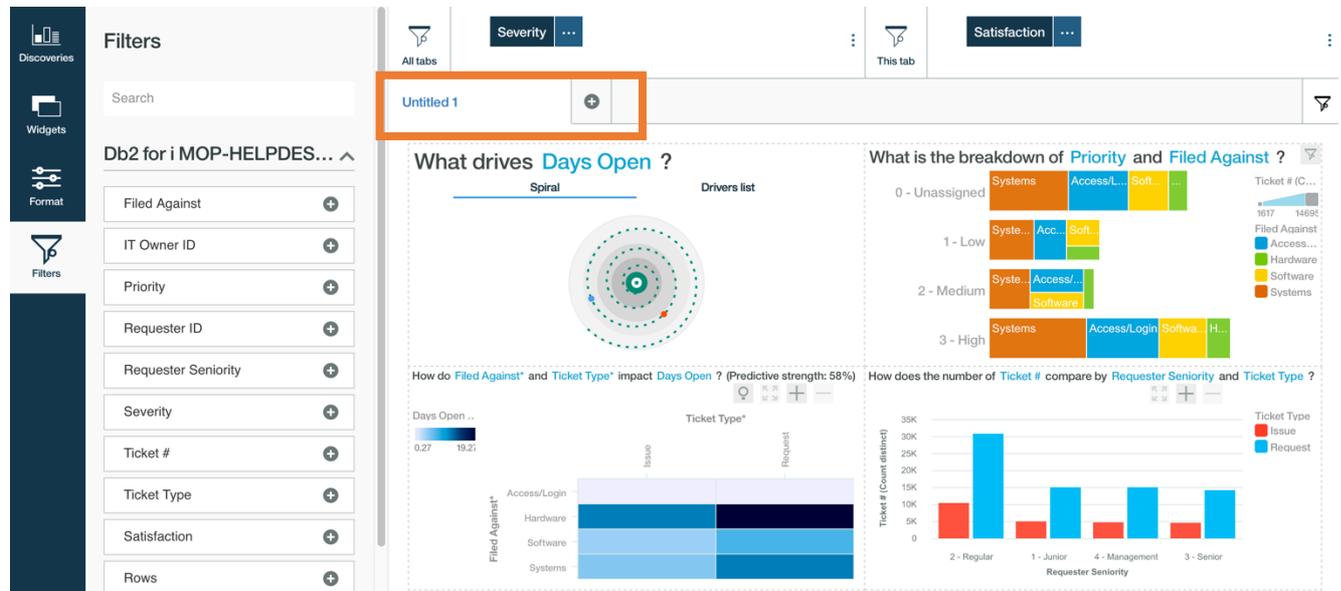
__ 6. Dans le bandeau de gauche, sélectionnez **Discoveries**. Puis, à partir de votre dossier **Personnel**, en sélectionnant votre ensemble de découvertes, glissez-déposez vos 4 découvertes dans les différentes parties du tableau de bord (visez l'icône ↗ pour obtenir un redimensionnement automatique).



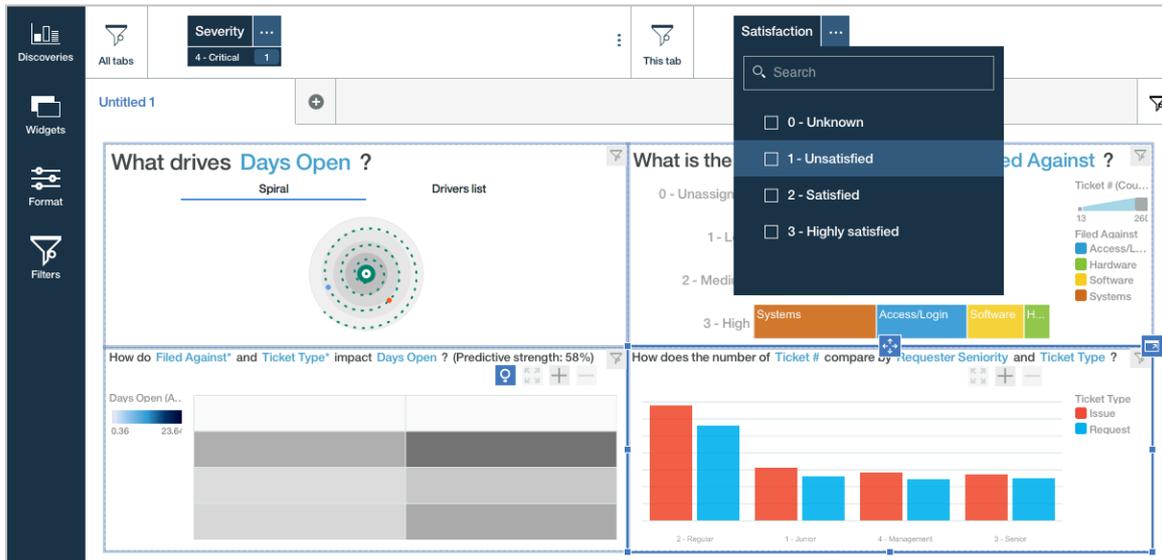
Pour chaque rapport, vous pouvez faire apparaître des menus permettant de modifier sa mise en forme, de le déplacer, voir même, de passer en mode édition sur ce rapport pour le modifier.



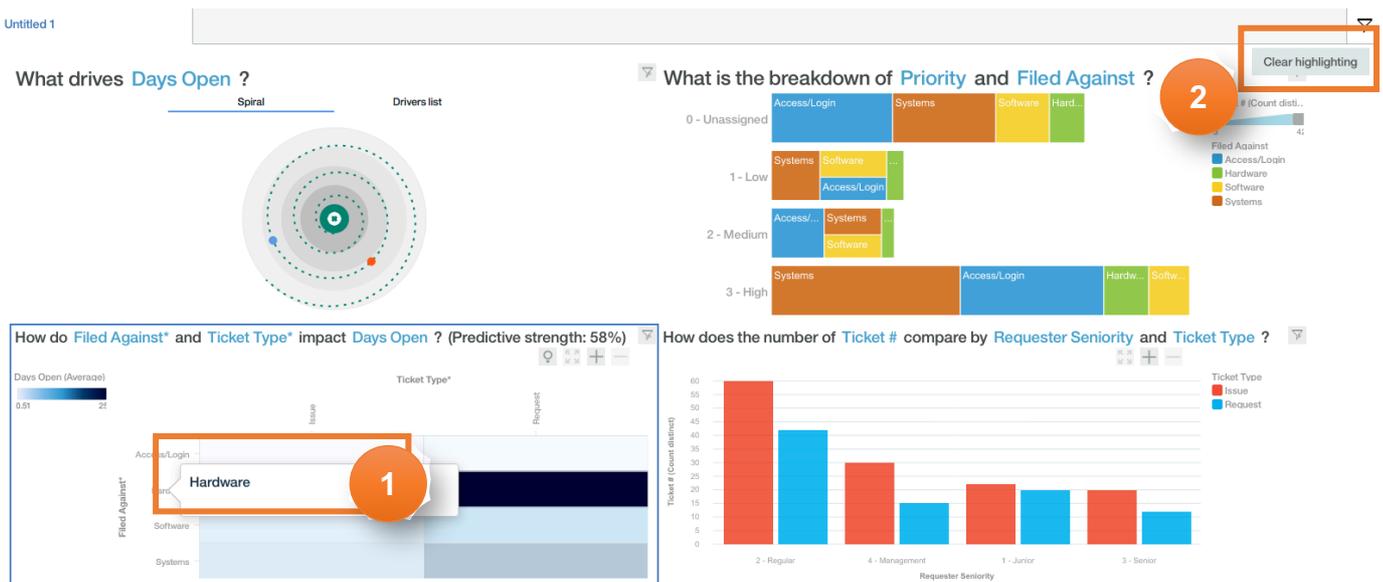
7. Vous allez maintenant ajouter des filtres permettant de restreindre le tableau de bord à la vision de certaines données uniquement. A partir du bandeau de gauche, sélectionnez **Filters**, puis glissez-déposez la colonne **Security** dans la zone **All Tabs**, et la colonne **Satisfaction** dans la zone **This Tabs**. Ceci sert à illustrer le fait que vous pouvez filtrer simultanément toutes les données de tous les onglets de votre tableau de bord, ou les données d'un seul onglet (même si dans cet exemple, vous n'avez qu'un seul onglet « Untitled 1 »).



- 8. En cliquant sur le nom de la colonne, vous pouvez utiliser ces filtres et voir immédiatement leur résultat sur les graphiques. Les tickets qui font l'objet de toute notre attention étant les tickets critiques ayant généré de l'insatisfaction, filtrez les données pour n'afficher que ces tickets.



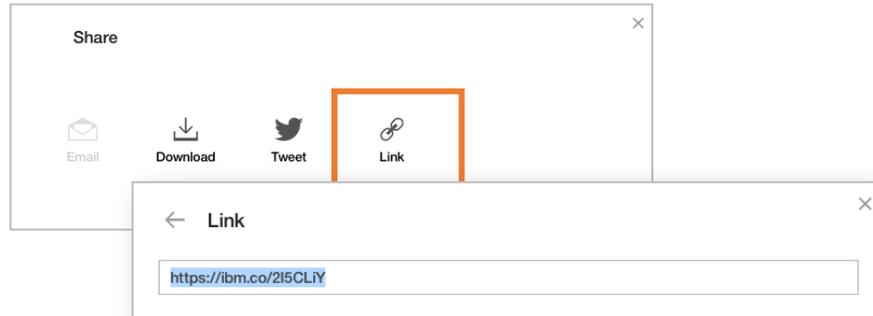
- 9. Sortez du mode édition en utilisant le bouton  : vous pouvez utiliser votre tableau de bord pour naviguer dans vos découvertes.
- 10. Remarquez que les différentes découvertes sont « liées » en elles : si, en cliquant dessus, vous sélectionnez la catégorie **Hardware** dans le rapport **Top Driver Combinaton** (numéro 1, ci-dessous), les 4 rapports seront automatiquement mis à jour pour prendre en compte ce filtre dynamique. En cliquant sur **Clear Highlight** (numéro 2, ci-dessous), vous supprimerez toutes ces sélections.



La dernière étape va maintenant consister à partager ce tableau de bord.

Section 4. Partage du tableau de bord

- ___ 1. Dans le menu du haut, cliquez sur l'icône  pour partager le tableau de bord, puis sur **Link**.



- ___ 2. Copiez ce lien et ouvrez la page liée à partir d'un autre navigateur, ou, envoyez-vous ce lien sur votre tablette (l'écran d'un smartphone est un peu petit pour ce genre de tableau de bord...).
- Ouvert sur une tablette (après authentification) cela vous donnera :



Bien entendu, vous pouvez également télécharger le tableau de bord au format image, PDF ou Microsoft Powerpoint. Mais dans ce cas, vous perdrez les fonctionnalités interactives et dynamiques du tableau de bord.

Félicitations ! Vous avez fini ce hands-on.

Maintenant, à vous de découvrir le sens vos données Db2 for i !

_____ End of Lab. _____

