



Rapport de stage

Comment les outils de Microsoft Data Factory
permettent-ils d'améliorer l'efficacité de la
récupération, de la transformation et de l'analyse de la
donnée.

CMI Ingénierie de la Statistique et Informatique, L1

JULIENNE Maxence

02/07/2025 – 31/07/2025

Technomade, CITE NUMERIQUE-BAT C-4EME ETAGE 406 BOULEVARD JEAN-JACQUES BOSC 33130 BEGLES

Tuteur de stage : Yannick GABILLE, directeur des opérations

Remerciements

En premier lieu, je souhaite remercier Yannick GABILLE pour la confiance qu'il m'a accordée ainsi que pour m'avoir permis de m'intégrer pleinement dans l'entreprise. Je suis également reconnaissant pour ses conseils avisés et la pédagogie dont il a fait preuve tout au long du stage.

Je remercie chaleureusement toute l'équipe de Technomade qui m'a accompagné avec bienveillance et a toujours été disponible pour m'aider.

Je remercie aussi Hélène MENNEGUERRE qui a géré les procédures administratives.

Enfin, je remercie grandement Bruno REAL d'avoir rendu ce stage possible.

Sommaire

- I. Présentation de l'entreprise
 - Présentation globale
 - Motifs de sollicitation par les clients
- II. Présentation du stage et des attentes
 - Objectif pédagogique et personnel
 - Mission
- III. Découverte des outils de Data Factory
 - Méthodologie adoptée
 - Déroulement des opérations
- IV. Sécurité
- V. Projet réalisé
 - Stratégie de mise en œuvre
 - Plan
- VI. Présentation Orale
 - Préparation
 - Présentation
- VII. Après l'oral
 - Amélioration du rapport
 - Recherches supplémentaires
- VIII. Bilan du stage
 - Difficultés
 - Points positifs professionnels
 - Points positifs personnels
- IX. Annexes
 - Rapport
 - Support oral

L'ensemble des termes techniques ou spécifiques nouvellement rencontrés feront l'objet d'une définition lors de leur première apparition dans le rapport.

Présentation de l'entreprise

Technomade est une entreprise à taille humaine fondée en 2004. On y compte une vingtaine d'employés. En tant que Microsoft gold Partner depuis 2016, elle a développé une expertise pointue dans l'intégration et le développement des solutions CRM et ERP via des outils de Microsoft Dynamics 365.

Les clients font appel à Technomade afin d'identifier leur besoin métier et bénéficier de solutions parfaitement adaptées, grâce à la mise en place de systèmes CRM ou ERP sur mesure.

Ils font confiance à Technomade pour :

- Une expertise reconnue, avec plus de 20 ans de pratique et une certification Microsoft gold Partner
- Proposer des solutions à des problèmes métier complexes à travers des méthodes fiables et innovantes
- Un accompagnement personnalisé tout au long du projet
- L'engagement d'un partenariat humain et responsable tout au long de la collaboration

CRM : Customer Relationship Management. Système de gestion de la relation client, il permet de suivre les relations avec les clients, d'optimiser les ventes, de personnaliser le marketing et de fidéliser les utilisateurs.

ERP : Entreprise Resource Planning. Système qui centralise et automatise les processus interne d'une entreprise tels que la comptabilité, les achats, la gestion des stocks...

Présentation du stage et des attentes

Les objectifs pédagogiques de ce stage obligatoire, réalisé dans le cadre du parcours CMI ont convergé vers mes objectifs personnels : explorer le monde professionnel et mieux comprendre le fonctionnement d'une entreprise dans le secteur technologique.

J'ai réalisé ce stage d'une durée de 4 semaines, du 2 au 31 juillet 2025 à Technomade sous la tutelle de monsieur Yannick GABILLE, directeur des opérations. J'avais des horaires de bureaux classiques (9h-17h) qui m'ont convenu car ils me permettaient une régularité et une prévision plus aisée du temps nécessaire à la réalisation d'une tâche.

Dès le début de mon stage, il m'a été confié la mission d'explorer l'outil Data Factory, intégré à l'environnement Microsoft Fabric.

Ce service est composé de plusieurs modules orientés vers le traitement et l'analyse de données. Il n'avait pas été encore exploité par l'entreprise par faute de temps et d'intérêt apparent à son utilisation.

Ma mission consistait donc à mener une étude approfondie des fonctionnalités offertes par Data Factory, à identifier les apports potentiels pour l'entreprise puis à retranscrire ces observations dans un rapport écrit, accompagné d'une présentation orale devant les membres de l'équipe.

Les premiers jours, l'idée de faire une présentation orale m'a beaucoup impressionné et intimidé. Mais dès lors que je me suis plongé dans le sujet, je l'ai tout de suite trouvé très intéressant, au point de me dire que le vrai défi serait d'avoir assez de temps pour partager toutes mes découvertes.

Découverte des outils de Data Factory

Avant de débiter ce stage, mes connaissances en matière de traitement de données se limitaient principalement à ce que j'avais acquis en cours, une base utile mais assez éloignée des réalités du monde professionnel.

Durant l'intégralité de mon expérience, j'ai pris des notes des différentes étapes de mon avancement pour la création du rapport final. Voici comment je me suis organisé :

Tâches	02	03	04	07	08	09	10	11	15	16	17	18	21	22	23	24	25	28	29	30	31	
Découverte Fabric																						
Recherches Data Factory																						
Rédaction rapport																						
Préparation Oral																						
Présentation Orale																						
Amélioration rapport + recherches supplémentaires																						
Rédaction rapport de stage																						

J'ai choisi de découvrir dans un premier temps l'environnement de Microsoft Fabric et de Power BI, j'ai rapidement été impressionné par la puissance de cette plateforme et la diversité de possibilités qu'elle offre. J'ai utilisé la plateforme Microsoft Learn qui recense une multitude de tutoriels très utiles.

Une fois la vision d'ensemble du système acquise, j'ai entamé une exploration approfondie de l'outil Data Factory en gardant en tête l'objectif principal : « la reprise, l'interfaçage, la transformation et la synthèse de données ».

Pour cela, j'ai pensé qu'il valait mieux partir d'un cas concret afin de bien comprendre l'ensemble des capacités de ces outils. J'ai donc pris l'exemple du propriétaire d'un château :

« Prenons l'exemple générique d'un propriétaire d'un domaine viticole. Il y fait de la production de vin et propose aussi la location de son domaine pour des événements. Depuis 15 ans, il utilise plusieurs tableaux Excel pour gérer toutes ses données (clients, réservations, production de vin...). Ces fichiers ne sont pas optimaux et trop compliqués à modifier pour lui.

Son CRM développé par Technomade correspond à ses besoins mais nécessite l'import manuel de ses fichiers Excel ce qui a plusieurs désavantages :

- Tâche répétitive
- Risque d'erreur dans la saisie ou le choix du format
- Chronophage

Le propriétaire, peu familier avec les outils informatiques récents ne souhaite pas s'y former ni modifier ses habitudes avec Excel.

L'objectif est donc d'automatiser ce processus sans changer ses habitudes. »

Extrait de l'introduction de l'exemple de mon rapport

Je me suis donc renseigné et ai découvert des outils qui pourraient me permettre de mettre cela en œuvre simplement :

- Dataflow Gen2
 - **Besoin** : interfaçage et transformation des données sans coder
 - **Description** : Interface visuelle qui permet de filtrer, fusionner, corriger et transformer des données de manière intuitive
 - **Exemple** : Récupérer un fichier Excel, enlever les lignes vides, corriger les noms et charger ces données propres dans un Lakehouse
- Lakehouse
 - **Besoin** : Stocker des données brutes ou transformées dans un format structuré, consultable et optimisé pour l'analyse
 - **Description** : conteneur qui combine la flexibilité du Datalake (stockage massif de données) avec la structure et les performances d'un data Warehouse (requête SQL, intégration avec Power BI)
 - **Exemple** : Garder les fichiers de stock mensuels d'une entreprise sur le Lakehouse, en faire des analyses avec Power BI qui y est intégré

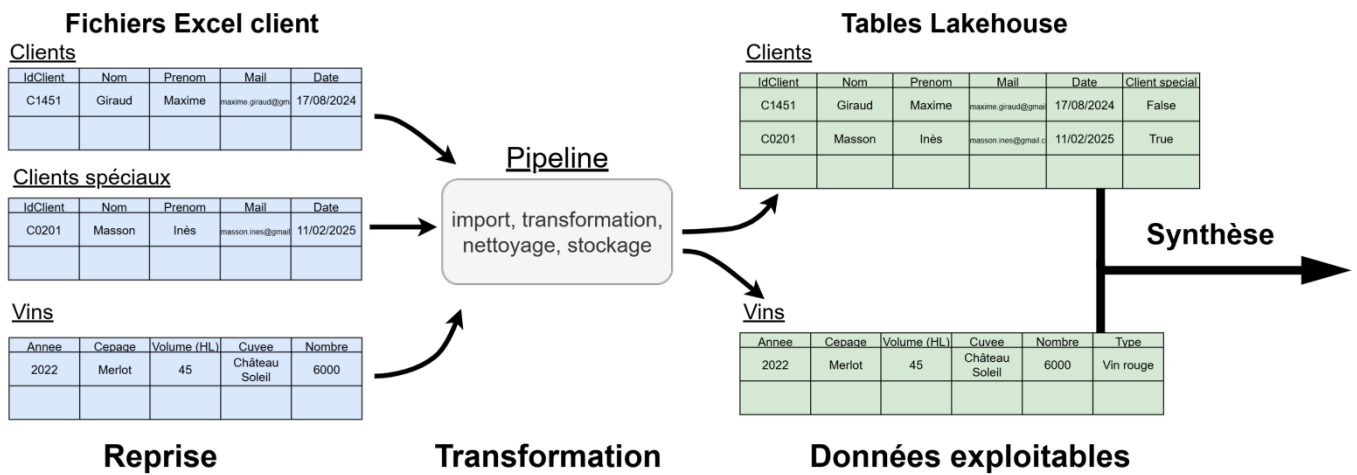
Data Pipeline

- **Besoin** : Automatiser les tâches de récupération et de transformation des données

- **Description** : Suite d'étapes programmées qui s'exécutent automatiquement à intervalle régulier selon les déclencheurs
- **Exemple** : Tous les matins, le pipeline prend les ventes de la veille, les nettoie et les envoie dans un Lakehouse et met à jour le rapport Power BI.

Ainsi, grâce à ces outils, il est possible de :

- Tous les jours, à heure fixe, charger les données de ses tableurs
- Nettoyer et transformer les données afin qu'elles soient utilisables
- Les charger dans une base de données



Tout cela permet de débiter avec des données brutes et non analysables et d'aboutir à des données nettoyées sur lesquelles on peut faire des rapports dynamiques et fiables directement dans le CRM.

Ce que j'ai particulièrement apprécié avec Power BI est l'interactivité des rapports. En effet, c'est l'un de ses atouts majeurs. Chaque élément visuel est cliquable et permet à l'utilisateur de filtrer dynamiquement les données et d'obtenir instantanément un rapport plus ciblé et pertinent.



Ici, j'ai juste sélectionné un type de cépage et tout le rapport s'est adapté

J'ai ensuite été chargé d'approfondir cette méthodologie dans un cas plus récurrent avec leur collaborateur, les données sont stockées dans une base de données en local.

Cette mission m'a particulièrement plu car elle m'a permis de découvrir le fonctionnement concret des serveurs hébergeant des bases SQL.

La solution était d'utiliser une passerelle. Sa mise en œuvre a été difficile pour moi, la complexité venait de la mise en place d'une passerelle de données qui devait se connecter à un port spécifique d'une instance particulière et répliquer l'identité d'un certain utilisateur qui avait les permissions d'accéder aux données. Une fois configurée, cette passerelle permettait le transfert des données vers le CRM.

Lors de mes recherches, j'ai trouvé une fonctionnalité innovante de Data Factory : Le Mirroring :

- **Besoin** : accéder aux données d'une base SQL en temps réel dans Fabric sans les dupliquer
- **Description** Révolution dans le traitement de données, chaque modification de la source est visible instantanément dans Power BI, cela remplace le schéma Dataflow + Pipeline + Power BI par une connexion directe et en temps réel.
- **Exemple** : Le client ajoute une commande dans sa base de données, quelques secondes plus tard elle apparaît dans le rapport power BI sans toucher à rien

Cette technologie est basée sur des capacités de Change Data Capture (CDC) et de réplication en streaming. Cela permet à Fabric d'obtenir un reflet des données instantané et sans duplication.

CDC : méthode qui permet de détecter et transmettre uniquement les modifications apportées aux données d'une base : ajouts, suppression ou mise à jour.

La mise en place du mirroring n'a pas été instinctive pour moi. Dans un souci de productivité, j'ai d'abord utilisé Copilot à outrance pour m'accompagner, sans effectuer au préalable mes propres recherches approfondies. Cette erreur évitable m'a mené vers des techniques complexes et détournées de l'objectif initial me faisant perdre presque deux journées de travail. Finalement, une simple vidéo tutorielle de 10 minutes m'a permis de débloquer la situation.

Une fois configuré correctement, Mirroring est un outil très puissant et utile, notamment pour garantir la fraîcheur et la pertinence des données affichées sur le rapport.

Voici les principales étapes de mes recherches sur cet aspect :

1. Création d'une base de données très complète qu'un client pourrait utiliser
2. Paramétrage de celle-ci afin d'accepter le CDC
3. Branchement de la passerelle sur le bon port
4. Configuration de la base de données miroir sur Fabric
5. Création d'un rapport interactif.

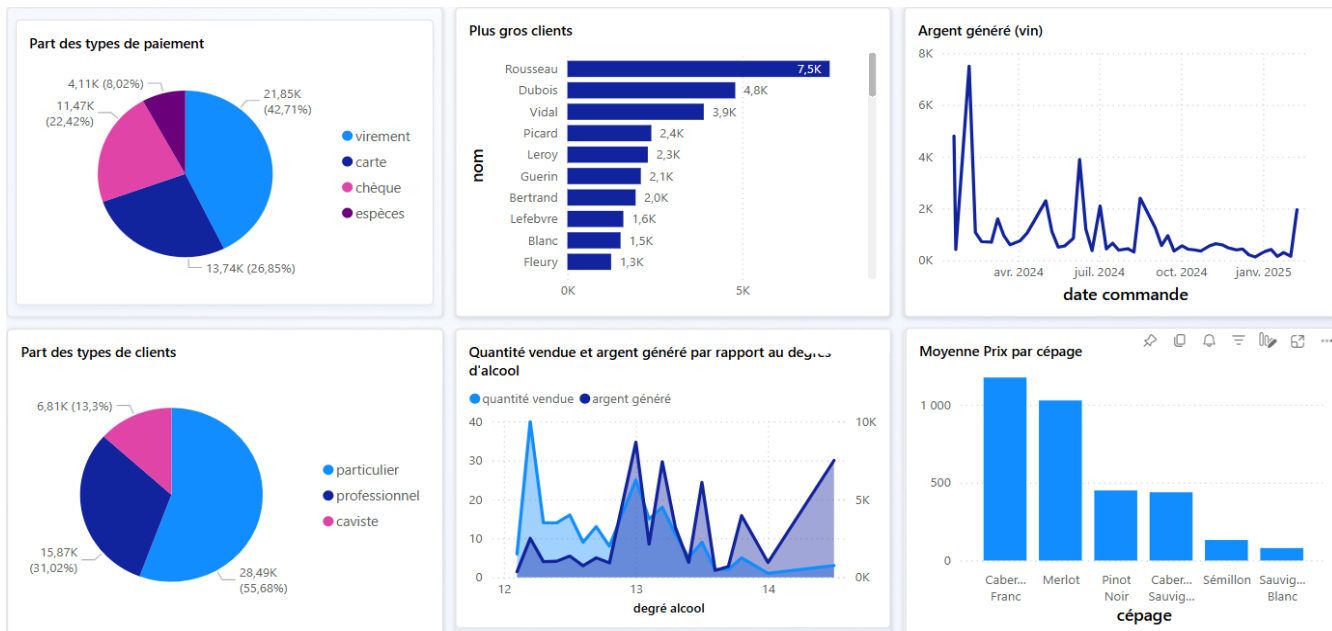
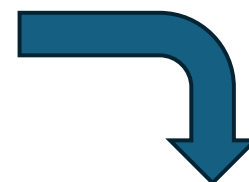
Ainsi, le client obtient un rapport à jour et fiable presque instantanément après avoir rentré de nouvelles données.

Cette automatisation est invisible pour lui mais fait gagner un temps précieux et permet d'éviter toute erreur humaine.

	id_vente	id_client	id_vin	date_commande	quantité	prix_unitaire	prix_total	date_livraison	mode_paiement	notes
1	1	1	9	2024-01-20	10	480.00	4800.00	2024-01-25	carte	Livraison rapide
2	2	2	10	2024-01-22	1	420.00	420.00	2024-01-27	virement	NULL
3	3	3	6	2024-02-06	3	2500.00	7500.00	2024-02-10	virement	Commande exceptionnelle
4	4	4	15	2024-02-13	6	180.00	1080.00	2024-02-18	carte	NULL
5	5	5	28	2024-02-20	4	180.00	720.00	2024-02-25	espèces	Client fidèle
6	6	6	12	2024-03-02	2	350.00	700.00	2024-03-05	carte	NULL
7	7	7	13	2024-03-09	5	320.00	1600.00	2024-03-14	virement	Commande pro
8	8	8	4	2024-03-16	1	950.00	950.00	2024-03-21	chèque	NULL
9	9	9	14	2024-03-23	2	300.00	600.00	2024-03-28	carte	Livraison express
10	10	10	33	2024-04-03	10	75.00	750.00	2024-04-08	virement	NULL
11	11	11	26	2024-04-11	3	350.00	1050.00	2024-04-14	carte	NULL
12	12	12	11	2024-04-19	4	380.00	1520.00	2024-04-24	virement	Client entreprise
13	13	13	8	2024-05-02	2	1150.00	2300.00	2024-05-07	chèque	NULL
14	14	14	3	2024-05-09	1	1100.00	1100.00	2024-05-14	espèces	NULL
15	15	15	24	2024-05-16	6	85.00	510.00	2024-05-20	carte	NULL
16	16	16	35	2024-05-23	10	55.00	550.00	2024-05-28	virement	NULL
17	17	17	1	2024-06-02	1	850.00	850.00	2024-06-07	virement	Mariage
18	18	18	7	2024-06-09	3	1300.00	3900.00	2024-06-14	chèque	NULL
19	19	19	2	2024-06-16	1	1200.00	1200.00	2024-06-21	virement	NULL
20	20	20	25	2024-06-23	5	75.00	375.00	2024-06-28	carte	NULL
21	21	21	5	2024-07-02	2	1050.00	2100.00	2024-07-07	virement	NULL
22	22	22	38	2024-07-09	4	110.00	440.00	2024-07-13	chèque	NULL
23	23	23	23	2024-07-16	6	110.00	660.00	2024-07-21	carte	NULL

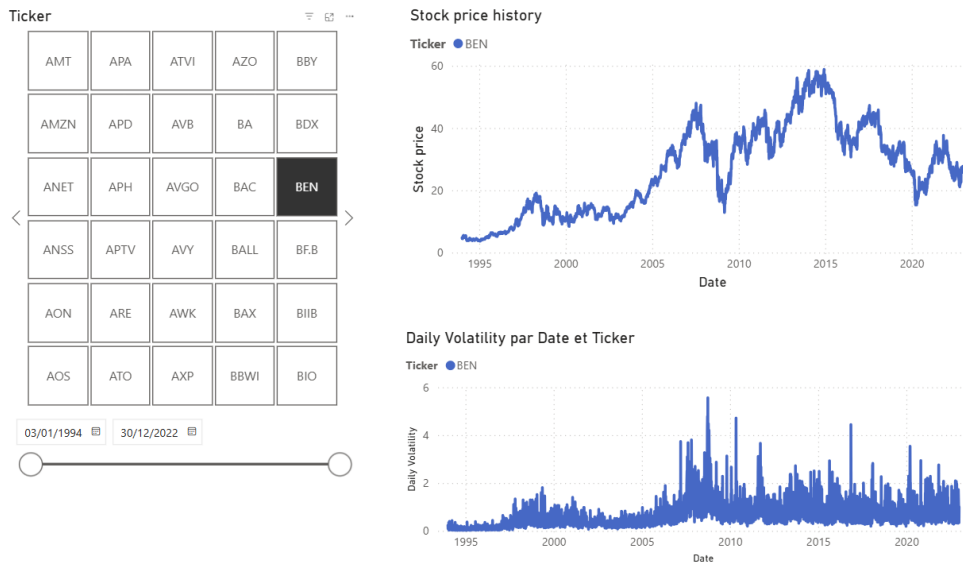
Données initiales

≈ 1 min



Rapport final, disponible sur le CRM

Voici un autre exemple d'un rapport produit avec Power BI via les outils de Data Factory



Exemple d'utilisation de Power BI intégré à Fabric : Stocks_sample.

Sécurité

Après avoir exploré l'intégralité des outils de Datafactory, on m'a chargé de découvrir comment améliorer la sécurité des données. Il est important de savoir que Microsoft Fabric bénéficie de la sécurité native de l'active directory de Microsoft, ce qui lui donne déjà un niveau de sécurité élevé.

Cependant, après analyse, j'ai découvert qu'il était possible d'améliorer la sécurité des données en utilisant un outil complémentaire : Pureview. N'ayant pas accès à une licence, je n'ai pas pu tester les fonctionnalités proposées.

Ce thème n'étant pas le sujet principal de mon stage, je fais le choix de ne pas m'y attarder dans ce rapport. Néanmoins, la synthèse des recherches est tout de même disponible dans le rapport remis à l'entreprise ([ici](#)).

Découvrir quelles stratégies de protection de la donnée peuvent être adoptées dans des contextes critiques ne m'a pas particulièrement captivé, bien qu'il présente des enjeux importants.

Projet réalisé

Tout au long de mon stage, j'ai rédigé à la demande de Technomade un rapport qui leur était destiné. Ce document avait pour objectif de centraliser la synthèse de l'ensemble des recherches que j'ai effectuées. Pour cela, j'ai procédé méthodiquement en établissant tout d'abord un plan détaillé que j'ai fait évoluer au fil de mes recherches. J'ai également tenu à jour des documents de prise de notes avec les informations pertinentes et les captures d'écran qui pourraient s'avérer nécessaires pour la suite.

La rédaction du rapport s'est faite en gardant en tête la problématique posée ce qui m'a permis de rester cohérent dans mes analyses. Voici le plan du rapport :

- I. Introduction
 - Problématique
 - Objectif du rapport
- II. Les outils de Datafactory
 - Dataflow Gen2
 - Lakehouse
 - Data Pipeline
 - Tache Apache Airflow
 - Copy Data
 - Miroiring
 - Bibliothèque de variables
- III. Exemples complets
 - Vigneron (Excel)
 - Vigneron (SQLEXPRESS)
- IV. Sécurité
 - Problèmes de sécurité de Datafactory
 - Solution avec Pureview
- V. Bilan
- VI. Documentation/tutoriels
 - Lier un fichier Excel à un Lakehouse via un Dataflow
 - Ajouter une table dans ce Dataflow

Ecrire ce rapport m'a permis de développer une compétence que je considère essentielle pour la poursuite de mes études : la capacité à rechercher une information, à l'analyser et en à faire une synthèse.

L'intégralité du rapport délivré à Technomade est disponible dans les annexes.

Présentation orale

Afin de compléter mon rapport écrit, j'ai également réalisé une présentation orale devant des membres de l'équipe. Comme mentionné précédemment, cet exercice m'a particulièrement impressionné. Un créneau d'une heure m'a été réservé le mercredi 23/07 durant lequel j'ai présenté et valorisé le fruit de mon travail devant eux.

Pour me préparer, j'ai conçu un support visuel et structuré, reprenant les éléments clés de mon rapport. Il a été enrichi avec des illustrations parlantes afin de rendre la présentation plus accessible et percutante. Ce support est joint en annexe du rapport.

La présentation s'est bien déroulée, le public a été très compréhensif et m'a partagé de précieux conseils.

Ce qui m'a particulièrement marqué est le fait de produire une présentation avec un véritable objectif. En effet, en cours nous sommes habitués à faire des présentations dans le but de nous évaluer à l'oral ou dans nos recherches. Ici, il s'agissait de présenter mes recherches devant des professionnels attentifs à mon contenu, prenant des notes sur ce qui pouvait leur être utile.

Dès le début, j'ai précisé que les interruptions étaient les bienvenues si des questions ou des remarques s'avéraient nécessaires. Après une vingtaine de minutes d'exposé, j'ai réalisé une démonstration en direct des outils présentés, suivie d'une phase d'échange marquée par des questions ciblées et des retours sur ma performance.

Je suis globalement satisfait de ma prestation. Les principaux points à améliorer qui ont été relevés sont : une élocution parfois trop rapide et le fait d'avoir lu une ou deux diapositives.

Cette présentation m'a fait prendre conscience que contrairement aux travaux scolaires destinés à développer nos compétences, les missions d'entreprise s'effectuent dans un cadre réel avec de véritables enjeux temporels ou économiques.

Après l'oral

Après mon passage à l'oral et la remise de mon rapport, ma tâche principale était terminée. On m'a donc donné de nouvelles missions de courte durée pour améliorer le rapport ou pour ma culture personnelle. Voici donc les différentes tâches qui m'ont occupé du 24/07 jusqu'à la fin de mon stage.

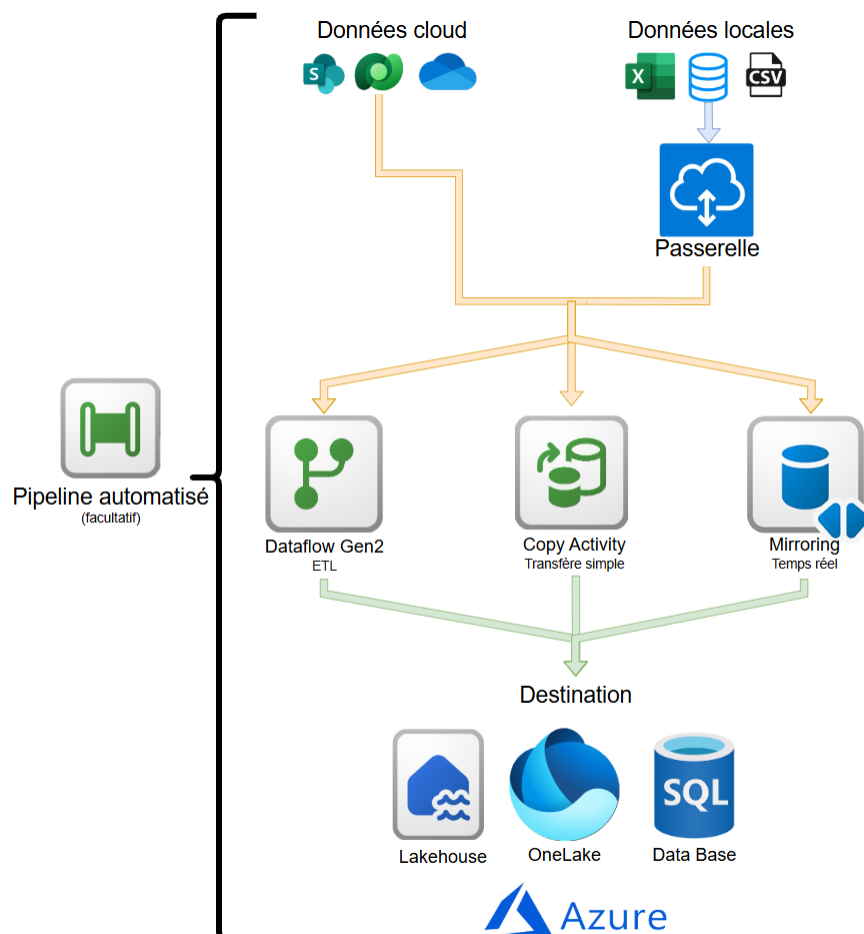
Tarifcation :

Le principal point à améliorer de mon rapport était le manque d'indication sur la tarification. En effet, au moment de l'écriture, je n'avais pas encore réalisé que mon rapport serait vraiment utile et devrait donc être le plus complet possible.

Je me suis donc initié au fonctionnement de la facturation de Microsoft. C'est un système complexe avec énormément de variables qui, en résumé fonctionne avec des CU (Capacity Units) que l'on paye.

Illustration du circuit des données :

Il manquait un schéma des principaux outils utiles présents dans mon rapport. J'en ai donc créé et intégré un.



Découverte des mesures Power BI :

On m'a informé qu'il serait intéressant pour moi de découvrir le fonctionnement des mesures sur le rapport Power BI. Les mesures sont des nouvelles colonnes ou variables dans les tables branchées directement sur des fonctions. Elles sont codées en langage DAX (dont je n'avais jamais entendu parler). J'ai trouvé ça très intéressant et utile car elles permettent une analyse plus poussée et un rendu plus parlant.

Par exemple, il est possible de créer une mesure calculant le total des revenus générés par les commandes. Il est aussi possible de rajouter une colonne fictive qui concatène la colonne nom et prénom pour plus de clarté dans le rapport.

Bilan du stage

Pour conclure ce rapport, voici un bilan objectif de ce stage.

Difficultés :

Durant cette expérience, je n'ai pas rencontré de véritables difficultés. Celles auxquelles j'ai été confronté m'ont permis de progresser et de découvrir de nouvelles méthodes de travail.

Par exemple, en buttant plusieurs fois sur la mise en place d'une connexion pour un transfert de données, j'ai appris le fonctionnement d'un serveur SQL et comment en assurer la gestion.

Points positifs :

Les retours positifs sur mon rapport et ma présentation m'ont fait particulièrement plaisir. Être félicité sur un travail qui a demandé du temps et qui pourra leur être utile est très gratifiant.

J'espère que mon rapport et mon diaporama leur permettront soit de constater que certains outils ne répondent pas entièrement à leurs attentes, soit de les intégrer à leur méthode de travail.

Cette expérience professionnelle en quasi-total autonomie a été pour moi une véritable opportunité, elle m'a permis de découvrir le monde professionnel et de confirmer que je suis dans une voie qui me plaît. J'aimerais exercer plus tard un métier similaire à ceux que j'ai observé mais plus orientés sur l'analyse de la donnée. Cela m'a motivé à persévérer dans mes études afin de pouvoir prétendre à ce type d'emplois.

Ce stage m'a permis de découvrir les processus de récupération et de transformation de données à l'échelle d'une entreprise.

Pour une prochaine expérience, j'aimerais approfondir le côté analytique du métier en découvrant comment les données sont interprétées et comment elles peuvent être utilisées pour faire des choix stratégiques. Cela me permettrait d'obtenir une vision plus complète du cycle de vie des données.

Annexe

Voici sur les pages suivantes le rapport sur mes recherches que j'ai rendu à l'entreprise et le support qui m'a permis la présentation orale.