



Rapport de stage - TN10

Intelligence Artificielle et Science de Données

A23

Redéfinition des régions

Pengyu JI
17 décembre 2023

Entreprise : Amazon EU S.A.R.L
Adresse : 5 Rue Plaetis,
L-2338 Luxembourg

Suiveur UTC : Ghada JABER
Suiveur Entreprise : Noah BERLANDI

Ce document synthétise le travail et la réflexion sur le stage que
j'ai effectué chez Amazon EU S.A.R.L. du 28 août 2023 au 9
février 2024.

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué au succès de mon stage et qui m'ont aidé lors de la rédaction de ce rapport.

Je souhaite tout d'abord remercier l'UTC et Amazon pour m'avoir permis de réaliser ce stage Projet de fin d'études. Ces 6 mois enrichissants m'ont permis d'avoir un contact approfondi avec le monde du travail, de connaître le déroulement de travail au sein d'une entreprise internationale et de constituer une expérience professionnelle solide en gestion des datas.

Je remercie également mon maître de stage, Noah BERLANDI, pour avoir accepté ma candidature, ainsi que pour son suivi régulier de mes progrès pendant le stage. Son expertise analytique acquise au fil des années a enrichi mon expérience en m'apportant de précieuses idées et des perspectives commerciales éclairantes. Je suis reconnaissant du temps qu'il a consacré au développement de mes compétences professionnelles, m'offrant ainsi une vision approfondie du monde des affaires.

Je voudrais aussi remercier Karan NIMRANI, le directeur de notre équipe, pour la reconnaissance du travail, le temps et les conseils qu'il m'a accordés. Son parcours personnel marqué par des défis surmontés et des succès obtenus m'a énormément inspiré et a constitué une source significative de motivation. Ses expériences de lutte et de réussite personnelle ont été une leçon précieuse pour moi, renforçant mon engagement envers mes objectifs professionnels.

Je remercie au même titre Francesca CRISTOFORETTI, Sr BIE dans l'équipe, pour son aide précieuse et ses conseils avisés. Son expérience professionnelle et son expertise techniques m'ont permis de mieux comprendre les enjeux de mon travail et de mener à bien mes missions. Je suis reconnaissant de son soutien et de son aide tout au long de mon stage.

Je tiens également à exprimer ma gratitude envers Matthieu BALY, Senior Analyst, pour sa contribution unique à l'équipe. Son sens de l'humour a apporté une note légère et positive, contribuant à créer un environnement de travail agréable et propice à la collaboration. Sa persistance sur le choix des chemises bleues, bien que légère et amusante, a ajouté une dimension ludique à notre quotidien.

Table des matières

Résumé

Lors de mon TN10, j'ai effectué un stage de 6 mois chez Amazon EU S.A.R.L. en tant que Business Intelligence Engineer. Le projet sur lequel j'ai travaillé consistait à redéfinir les régions de vente des produits Amazon à la place des régions géographiques existantes définies par les pays. L'équipe *inventory*, dans laquelle j'ai évolué, est chargé de la mise en place de ce projet.

Bien que le sujet paraisse simple, il s'agit d'un projet complexe qui nécessite non seulement une base technique solide, mais aussi une compréhension approfondie des données et des processus métier, et des compétences analytiques.

Afin de mener à bien ce projet, j'ai dû interroger les bases de données, donner des analyses profondes sur les données, et créer des outils qui servaient à bien les présenter et à aider à prendre des décisions au sein de l'équipe.

Les techniques que j'ai utilisées sont principalement SQL, Python, Excel, ainsi que des outils internes chez Amazon. Grâce à ces outils internes, j'ai pu accomplir des tâches qui auraient été très difficiles à réaliser avec des outils externes ou par moi-même. Par exemple, j'ai pu créer des applications web qui permettent de prendre des fichiers d'entrée, de manipuler les données, de les visualiser sans intervention humaine, et de générer des fichiers de sortie.

Grâce à ce projet, j'ai pu améliorer mes compétences techniques et analytiques, et j'ai pu acquérir une expérience professionnelle solide en gestion des données. Dans ce document, je vous présente l'entreprise, le stage, mes réalisations et mes sentiments personnels.

1 Présentation générale

1.1 Entreprise

1.1.1 Histoire d'Amazon

Amazon (NASDAQ : AMZN) est une multinationale américaine basée à Seattle.

Fondée par Jeff Bezos en 01 juillet 1994, l'entreprise a été introduite en Bourse au NASDAQ en 01 mai 1997.

L'activité initiale de la société Amazon concernait la vente à distance de livres, avant que la société ne se diversifie dans la vente en ligne de produits culturels et marchands, profitant de l'explosion d'internet.

Progressivement, le groupe se lance agressivement dans de nombreux domaines comme l'informatique, la grande distribution, la production cinématographique, les objets connectés, les cliniques de santé, ou l'industrie militaire.

C'est un des géants du Web, regroupés sous l'acronyme GAFAM. [1]

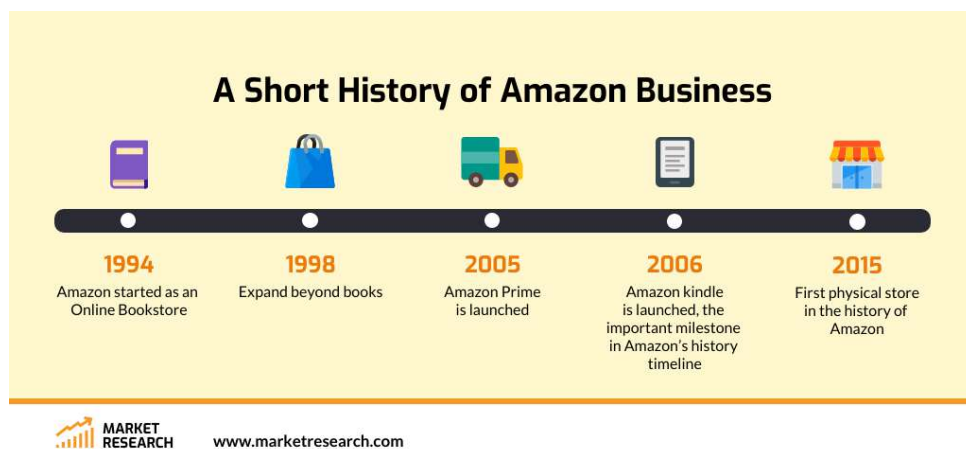


FIGURE 1 – L'Histoire d'Amazon [2]

1.1.2 Domaines d'activité et ses compétences

Amazon, initialement fondé comme une simple plateforme de commerce électronique, a rapidement évolué bien au-delà de ses débuts modestes.

- Commerce Électronique :
 - Plateforme exhaustive offrant une vaste gamme de produits.
 - Interface conviviale facilitant la navigation et les achats en ligne.
- Logistique Efficace :
 - Réseau mondial d'entrepôts et de centres de distribution.
 - Livraison rapide et fiable grâce à une chaîne d'approvisionnement bien organisée.
- Amazon Web Services (AWS) :
 - Leader mondial du cloud computing.
 - Offre des solutions de pointe en matière de stockage, de traitement de données et d'intelligence artificielle.
- Services d'Abonnement (Amazon Prime) :
 - Avantages exclusifs tels que la livraison express gratuite.
 - Accès à une variété de contenus multimédias, y compris des séries, des films et de la musique.
- Marketplace pour Vendeurs Tiers :
 - Plateforme permettant à des milliers de petites entreprises de vendre leurs produits.
 - Favorise une communauté diversifiée d'entreprises et d'entrepreneurs.

En résumé, la structure diversifiée des services d'Amazon reflète ses compétences exceptionnelles dans la gestion de multiples facettes du commerce et de la technologie.

1.1.3 Ses filiales en Europe

Amazon a solidement ancré ses activités à travers toute l'Europe, établissant des centres opérationnels, des entrepôts et des bureaux dans des pays clés tels que le Royaume-Uni, l'Allemagne, la France, l'Italie, l'Espagne et d'autres.

Amazon a diversifié ses activités en Europe à travers différentes filiales spécialisées. Par exemple, Amazon Web Services (AWS) offre des services cloud, tandis que des entités comme Amazon Prime se concentrent sur les services de streaming et la livraison rapide.

L'Europe a également été le terrain d'innovations majeures pour Amazon, notamment dans le domaine de l'intelligence artificielle, de l'automatisation des entrepôts et du développement de solutions technologiques avancées pour améliorer l'expérience client.



FIGURE 2 – Les domaines d’activité d’Amazon [3]

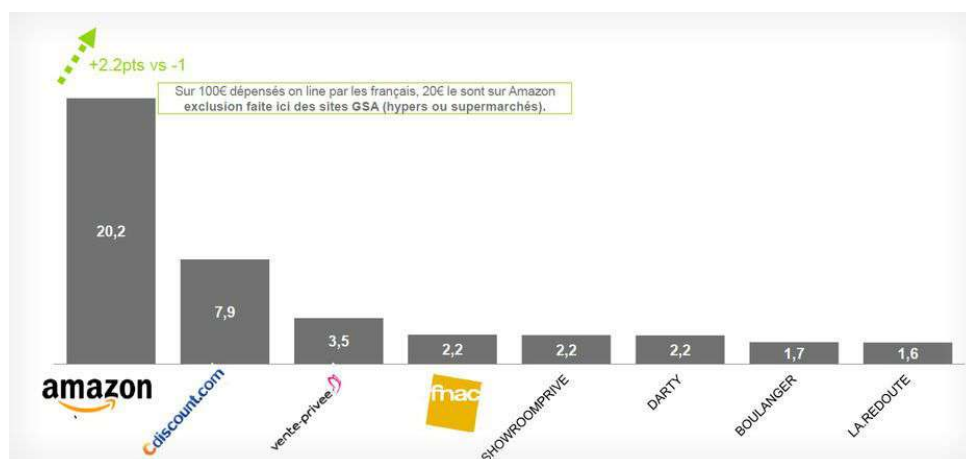


FIGURE 3 – La comparaison entre Amazon et les autres plateformes en France [4]

La filiale française du groupe est ouverte en 2000. En 2016, Amazon devient le premier distributeur non alimentaire en France par le chiffre d'affaires.

La présence d'Amazon en Europe a eu un impact significatif sur l'économie de la région, générant des emplois directs et indirects, stimulant le commerce local. Amazon a également pris des engagements envers la durabilité environnementale en Europe, en investissant dans des initiatives d'énergies renouvelables et en travaillant à atteindre des objectifs de neutralité carbone.

1.1.4 Son statut au marché

Amazon occupe actuellement une position prédominante sur le marché mondial, définissant les normes dans divers secteurs économiques. En tant que leader incontesté du commerce électronique, l'entreprise détient une part de marché significative dans de nombreux pays, offrant une vaste gamme de produits et de services à des millions de clients.

Amazon Web Services (AWS) domine le marché du cloud computing, fournissant des solutions technologiques à des entreprises du monde entier. En matière de streaming et de divertissement, Amazon Prime Video rivalise avec les meilleures plates-formes, créant du contenu original et attirant un nombre croissant d'abonnés. L'acquisition de Whole Foods a consolidé sa présence dans le secteur alimentaire et de la distribution physique. La plateforme Marketplace continue de prospérer, permettant à des vendeurs tiers de profiter de l'infrastructure et de la clientèle d'Amazon.

Son statut au marché est également renforcé par sa capacité à anticiper et à s'adapter aux tendances, à investir dans l'innovation, et à maintenir une réputation solide en matière de service client. Dans l'ensemble, Amazon reste un acteur incontournable, influençant considérablement l'économie mondiale et redéfinissant continuellement les normes de l'industrie.

L'évolution des chiffres d'affaires d'Amazon au fil des années est remarquable, reflétant une croissance soutenue et une expansion constante de ses activités. Depuis ses débuts en tant que librairie en ligne dans les années 1990, Amazon a connu une augmentation exponentielle de ses revenus, passant d'une entreprise de commerce électronique à une entité diversifiée offrant une multitude de services. Les chiffres d'affaires ont connu une croissance significative, alimentée par l'expansion de son catalogue de produits. En 2019, Amazon a réalisé un chiffre d'affaires de 280,5 milliards de dollars, contre 232,9 milliards de dollars en 2018.

1.1.5 Ses impacts

Amazon est aujourd'hui devenu un acteur et un employeur engagé dans les territoires français. Dès sa création, Amazon n'a cessé d'innover en faveur d'une économie responsable, qui profite aux PME comme aux grandes entreprises sur l'ensemble du territoire. Amazon se mobilise également afin d'apporter sa contribution aux communautés locales proches de ses sites logistiques, à travers un soutien humain, matériel et financier, en particulier pour les personnes les plus démunies et durement touchées par la crise sanitaire actuelle. [6]

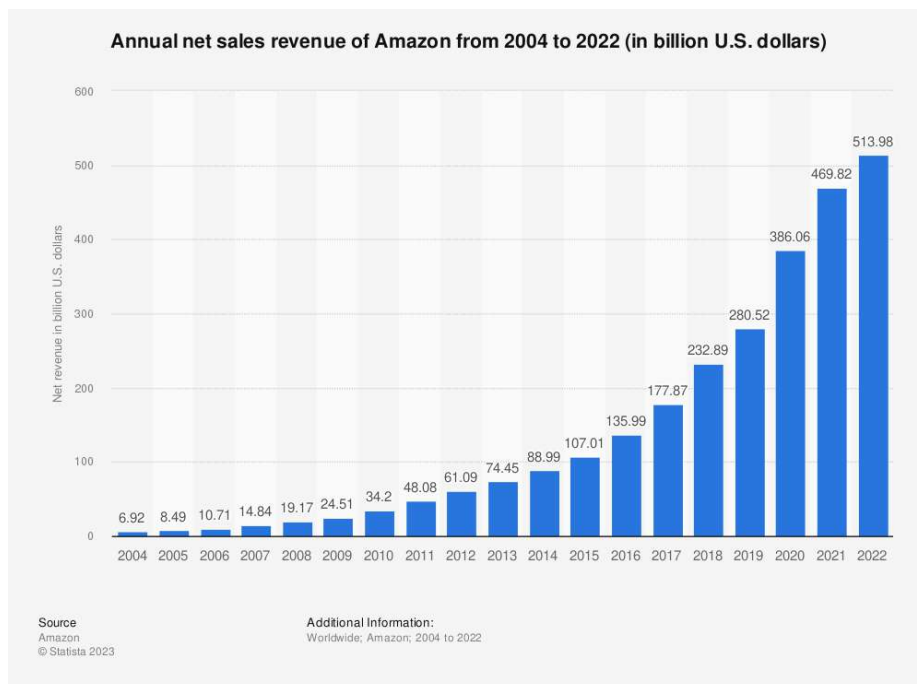


FIGURE 4 – L'évolution des chiffres d'affaires d'Amazon [5]

Croissance économique

Depuis deux décennies, Amazon a noué une relation forte avec des dizaines de millions de clients qui la font confiance et apprécient l'apport d'Amazon à leur qualité de vie et à leur pouvoir d'achat. Pour mieux les servir, Amazon a investi plus de 11 milliards d'euros dans nos activités françaises entre 2010 et 2020. Grâce à ces investissements, Amazon est devenu l'un des principaux créateurs d'emplois en France et emploie 15 500 salariés en CDI en France.

Soutien à l'entrepreneuriat

Amazon propose des outils et des ressources pour aider les petites et moyennes entreprises, les auteurs, les clients AWS, et les partenaires de livraison à développer leur activité. Ils peuvent ainsi être au service de leur territoire, tout en offrant plus de choix aux clients du monde entier.

En plus, Amazon accompagne le succès des petites et moyennes entreprises françaises : elles ont créé plus de 25 000 emplois en France pour soutenir leur activité en ligne.

L'an dernier, les petites et moyennes entreprises (PME) françaises ont vendu plus de 55 millions de produits dans les boutiques Amazon. En moyenne, elles vendent plus de 100 produits par minute dans les boutiques.

Impact à l'échelle locale

Amazon est aujourd'hui devenu un acteur et un employeur engagé dans les territoires français. En effet, Amazon se mobilise afin d'apporter sa contribution aux territoires et communautés dans lesquels les sites logistiques sont implantés, en mobilisant sa capacité à innover rapidement mais également à travers un soutien humain, matériel et financier.

Amazon travaille en étroite collaboration avec des partenaires pour trouver des solutions aux enjeux mondiaux les plus pressants, en créant notamment des programmes innovants, qui laisseront une empreinte durable et positive. Partout en France, Amazon accompagne les populations locales en veillant notamment à ce que les jeunes, dans le cadre du programme Amazon Future Engineer, soient sensibilisés aux opportunités et métiers du numérique et ce dès le plus jeune âge. Amazon se mobilise également pour que leurs besoins essentiels soient couverts.

1.1.6 Ses concurrents

Amazon opère dans un environnement commercial concurrentiel où plusieurs entreprises cherchent à rivaliser dans des secteurs spécifiques.

Dans le domaine du commerce électronique, des concurrents tels qu'eBay et Alibaba se positionnent en tant que forces importantes, chacun avec des approches distinctes. Alibaba, par exemple, domine le marché chinois avec une présence mondiale croissante. En ce qui concerne le cloud computing, Microsoft Azure et Google Cloud sont des concurrents majeurs d'Amazon Web Services (AWS). La concurrence dans le secteur du streaming vidéo provient de sociétés telles que Netflix et Disney+. Dans le commerce alimentaire, des concurrents traditionnels tels que Walmart et des chaînes de supermarchés européennes jouent un rôle significatif.

Malgré cette concurrence féroce, Amazon maintient sa position dominante en raison de sa capacité à innover rapidement, à investir massivement dans la technologie, et à offrir une gamme diversifiée de services. La rivalité constante dans ces différents secteurs stimule l'innovation et incite Amazon à rester à l'avant-garde de l'évolution des marchés mondiaux.

1.1.7 Son succès pendant la pandémie

La pandémie de coronavirus a entraîné la mort de plus d'un million de personnes dans le monde et a eu des répercussions dévastatrices sur l'économie mondiale, mettant certaines industries à l'arrêt, provoquant des licenciements massifs et accélérant le déclin des chaînes de grands magasins déjà en difficulté.

Amazon, en tant que géant du commerce électronique, s'est démarqué comme l'une des rares exceptions. Offrant une vaste sélection et une recherche constante de commodité et de prix bas, il est devenu le détaillant par défaut et un service essentiel pour de nombreux consommateurs pendant la crise du coronavirus.

Face à la fermeture des magasins et aux rayons vides, les consommateurs se sont tournés en masse vers Amazon pour des produits de protection tels que des désinfectants pour les mains, des masques et des désinfectants. Les ventes de produits ménagers et d'épicerie ont explosé, propulsant Amazon vers des ventes records au deuxième trimestre.

Malgré des défis opérationnels initiaux, tels que des retards de livraison et des problèmes d'inventaire, Amazon a continué à embaucher massivement pour répondre à la

demande croissante. Le PDG Jeff Bezos a souligné les difficultés rencontrées, mais Amazon prévoit de maintenir sa croissance pendant la saison des fêtes.

Cependant, la pandémie a révélé des vulnérabilités dans la chaîne d’approvisionnement d’Amazon, perturbant son fonctionnement. La société continue de faire face à des problèmes d’inventaire, restreignant la quantité de marchandises stockées dans ses entrepôts.

Malgré ces défis, Amazon prévoit de dépasser les 100 milliards de dollars de revenus trimestriels pour la première fois au quatrième trimestre, ce qui en ferait l’une des rares entreprises américaines à atteindre ce seuil aux côtés de Walmart et Exxon.

1.2 Equipe

1.2.1 Mission

Notre équipe s’appelle *Speed-Long term planning*. Notre objectif est de prendre les meilleures décisions pour nos clients et pour l’ensemble d’Amazon en nous basant sur les principaux indicateurs de performance (KPI) de l’entreprise : **Speed** (la promesse de livraison que nous voyons sur le site), **Delivery promise accuracy** (DEA, notre capacité à livrer à temps) et **Cost**.

Nous optimisons l’ensemble des opérations d’Amazon et les interactions entre ces métriques pour prendre les meilleures décisions de bout en bout. Notre vision est de réduire les coûts tout en maintenant une excellente expérience client grâce à la rapidité et à la DEA. Pour concrétiser cette vision, nous participons à la planification à long terme sur un horizon de 3 ans et à la planification opérationnelle à un an. Notre équipe est organisée en fonction des différentes activités d’Amazon (approvisionnement, expédition, placement, conception du réseau, ...) et nous travaillons en étroite collaboration avec les équipes opérationnelles pour définir les objectifs et les plans d’action. Nous travaillons également avec les équipes techniques pour développer des outils et des modèles qui nous permettent de prendre des décisions basées sur des données et de les mettre en œuvre à grande échelle.

Au sein de cette équipe, je rejoins la sous-équipe *Invenory* qui s’occupe des projets de placement des stocks (où et comment nous disposons les articles). Notre sous-équipe est composée de Marc, Matthieu, Francesca, Joy, Abhishek, Vinayak, et nous sommes tous sous la direction de Karan.

Dans notre équipe, la bienveillance est au cœur de notre culture. Nous encourageons l’échange ouvert d’idées, favorisons un esprit d’équipe solide et sommes là pour soutenir nos collègues.

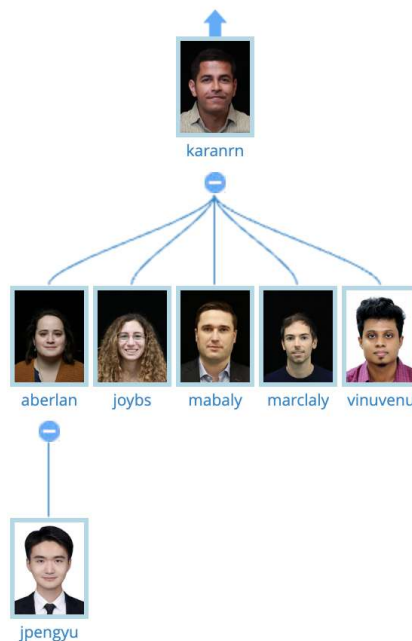


FIGURE 5 – Organigramme de notre sous-équipe *Inventory*

1.2.2 Rôle des membres

Il est essentiel de reconnaître le rôle central de Noah BERLANDI parmi les membres de l'équipe. En tant que mon manager, il assume la responsabilité du périmètre du projet et des livrables que je suis chargé de fournir d'ici la fin de mon stage. Au-delà de cette dimension de supervision, Noah s'est avéré être un conseiller précieux, me guidant sur la meilleure manière de progresser tout au long du déroulement du projet. Son rôle a été à la fois directif, en établissant des orientations claires, et de soutien, en fournissant un accompagnement et des conseils adaptés à chaque étape. Cette combinaison de direction et de soutien a grandement contribué à mon apprentissage et à la réussite de ma mission au sein de l'équipe. Sa présence active a indéniablement renforcé mon expérience professionnelle pendant mon stage.

Il est aussi essentiel de souligner le rôle significatif de Francesca CRISTOFORETTI au sein de l'équipe. Occupant le poste de Sr Business Intelligence Engineer, elle a joué un rôle crucial en fournissant un soutien technique précieux pendant la réalisation de missions techniques. La profondeur de sa connaissance des techniques internes a été une ressource inestimable pour moi, facilitant la compréhension des aspects plus complexes du projet. Francesca a démontré une expertise technique exceptionnelle et une disposition à partager ses connaissances, ce qui a grandement contribué à mon apprentissage et à la réussite des tâches techniques. Sa présence dans l'équipe a renforcé la collaboration et l'efficacité globale du projet.

Il est aussi important de souligner le rôle de soutien joué par Karan NIMRANI en tant que mon skip manager tout au long de mon stage. Respectant la hiérarchie établie, il m'a accordé la liberté d'agir tout en laissant à Noah l'espace nécessaire pour organiser mon travail. Cependant, Karan n'a pas hésité à me fournir de l'aide et des conseils lorsque cela était nécessaire. J'apprécie grandement sa direction au sein de l'équipe, car elle a contribué à créer un environnement de travail favorable à la collaboration et à l'apprentissage.

2 Présentation du sujet

Au cours de mon stage chez Amazon, j'ai été impliqué dans un projet passionnant axé sur la redéfinition des régions (Region Definition Services). L'objectif principal de cette initiative était de remplacer les régions géographiques actuelles, définies par les pays, par des régions de vente mieux adaptées à la capacité opérationnelle d'Amazon. Ce projet revêt une importance stratégique, car il vise à aligner de manière plus efficace la structure régionale avec les exigences opérationnelles et la dynamique du marché.

La mise en œuvre de ce projet vise à mieux accompagner le projet principal en fournissant un soutien essentiel à son développement. En redéfinissant les régions de vente, notre objectif est de créer un environnement plus propice au projet principal, lui permettant de s'adapter de manière plus flexible aux exigences du marché et aux défis opérationnels.

2.1 Connaissances préalables

Avant de présenter le projet, il est nécessaire de présenter quelques connaissances préalables.

- **FC (Fulfillment Center)** : Un FC est un centre de distribution d'Amazon. Il s'agit d'un entrepôt où les produits sont stockés et expédiés aux clients.
- **SC (Sort Center)** : Un SC est un centre de tri d'Amazon. Il s'agit d'un centre où les colis sont triés et expédiés aux clients.
- **DS (Delivery Station)** : Un DS est une station de livraison d'Amazon. Il s'agit d'un centre à partir d'où les colis sont livrés aux clients.
- **Placement system** : Le placement system est un système qui permet de déterminer le FC dans lequel les stocks d'un produit doivent être placés. Il s'agit d'un système qui permet de déterminer le FC le plus adapté pour chaque produit.

Le voyage d'un produit peut être résumé comme suit :

- **Achat et stockage** : Amazon achète des produits auprès des fournisseurs et les stocke dans les FC. Au cours de cette étape, le **Placement system** détermine le FC pour stocker les produits.
- **Commande et livraison** : Les clients passent des commandes sur le site web d'Amazon. Les produits sont expédiés aux clients à partir des FC.

Selon le passage entre ces centres, nous pouvons diviser le processus de livraison en trois parties :

- **First mile** : Le premier kilomètre du processus de livraison. Il s'agit du passage entre le FC et le SC.
- **Middle mile** : Le kilomètre du milieu du processus de livraison. Il s'agit du passage entre le SC et le DS.

- **Last mile** : Le dernier kilomètre du processus de livraison. Il s'agit du passage entre le DS et le client.

Il faut tenir en compte que tous les trois passages vont engendrer des coûts. Par exemple, le passage entre le FC et le SC va engendrer des coûts de transport, et le passage entre le SC et le DS va engendrer des coûts de tri. Dans un cas spécifique, le colis peut directement être expédié depuis le FC à le DS, sans passer par un SC. Nous l'appelons **DI**, acronyme de **Direct Injection**.

2.2 Projet principal : Projet Möbius

Le **Placement system** utilisent les coûts de transport réels et la vitesse historique à partir de la période des trois dernières semaines la plus récente. En tant que tels, les systèmes de placement ne sont pas conscients des éventuels changements dans la conception du réseau de transport. Par conséquent, le **Placement system** n'est pas en mesure de réagir rapidement aux changements de conception du réseau de transport. Il a tendance de placer les stocks d'un même produit dans le même FC, à la place de les placer dans des FC différents.

En examinant les stations DIG1 et DRM2 dans la zone statistique métropolitaine de Londres (MSA), les deux sont directement connectées à LCY2 et LCY3 (Figure suivante) avec une heure limite de commande à minuit et des coûts d'expédition similaires. Le **Placement system** ne sépare pas les stock des deux type de produits dans les deux FC.

Cette décision, entraînant une diminution des stocks partagés entre LCY2 et LCY3, nous oblige à maintenir des voies directes (DIs) à partir des deux centres de traitement (FCs) vers les deux sites de distribution (DS). Cependant, si nos systèmes recommandaient de diviser les colis et de les répartir pour augmenter la disponibilité à LCY2 et LCY3, il serait probablement possible de retirer deux DIs en passant le colis par un SC(Figure de droite).

Après la répartition, nous avons des stocks des deux produits dans tous les deux FC. Selon la distance, les coûts **b** et **c** deviendraient plus faibles. Alors que les coûts **a** et **d** deviendraient plus élevés. Les avantages liés aux coûts de transport longue distance compenseraient alors l'impact négatif sur les coûts et la vitesse de la chaîne logistique.

2.3 Redéfinition des régions

2.3.1 Contexte

Les régions géographiques existantes sont définies par chaque pays et l'Union Européenne. Cette nomenclature s'appelle **NUTS**(Nomenclature des unités territoriales statistiques). Il s'agit d'un découpage territorial destiné à faciliter les comparaisons entre pays, ou entre régions, d'un même ensemble.

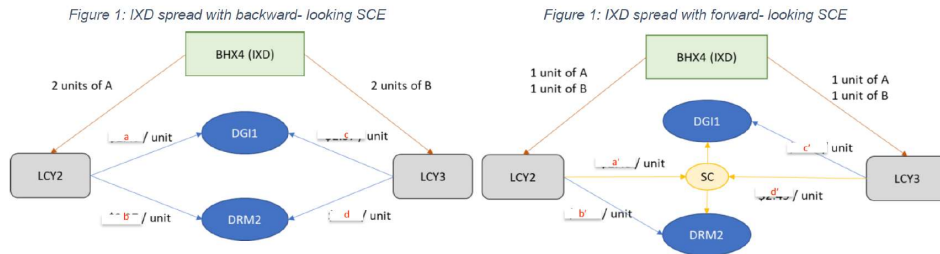


FIGURE 6 – Un exemple des stocks distribués [7]

Ces unités territoriales sont définies uniquement pour les besoins statistiques et ne constituent pas forcément des unités administratives officielles, mais souvent des groupements de ces unités administratives, en fonction de leur population résidente moyenne dans le pays correspondant.

Il existe trois niveaux de NUTS, qui sont définis comme suit :

- NUTS 1 : Les régions de niveau 1 sont les régions les plus grandes. Elles sont utilisées pour les comparaisons entre pays. Il y a 98 régions de niveau 1 en Europe.
- NUTS 2 : Les régions de niveau 2 sont les régions intermédiaires. Elles sont utilisées pour les comparaisons entre régions. Il y a 276 régions de niveau 2 en Europe.
- NUTS 3 : Les régions de niveau 3 sont les régions les plus petites. Elles sont utilisées pour les comparaisons entre régions. Il y a 1342 régions de niveau 3 en Europe.



FIGURE 7 – Plan de la NUTS1[8]

Lors du déroulement de projet Möbius, il a été observé que les régions existantes ne sont pas adaptées aux exigences opérationnelles et à la dynamique du marché. En effet, les régions actuelles sont définies par les pays, ce qui ne permet pas de répondre aux besoins opérationnels.

Nous pouvons relever les problèmes suivants :

- Les régions eux-mêmes ne sont pas logiques. Par exemple, la région NUTS1 ES30 en Espagne est bordée par une autre région, et à première vue, il semble plus logique que ces deux régions soient fusionnées.
- Parmi les NUTS1, nous observons des régions assez longues. Dans ce cas-là, il se pourrait que les communes aux extrémités d'une région soient plus proches d'un autre FC situé dans une autre région que de la leur. Par exemple, la région ES52 en Espagne est très longue, et les communes aux extrémités de cette région peut être couvert par un autre dans une autre région. Par contre, le système actuel ne permet pas de prendre en compte ce genre de situation. Ce dernier attribue encore ces communes au FC de ES52.

En plus, le **Placement system** agrègent les coûts de transport et la vitesse au niveau du code géographique NUTS1. En général, il y a généralement 10 codes NUTS1 dans chaque pays de l'Union européenne, tandis que la granularité pertinente pour le réseau de transport est le niveau de la station de livraison.



FIGURE 8 – Plan de la NUTS1/2 en Espagne [9]

En ne considérant que la connectivité moyenne agrégée au niveau NUTS1, les systèmes ratent l'opportunité d'obtenir une exécution moins chère/plus rapide pour les sous-régions.

Par exemple, en France, il y a 13 codes NUTS1. Cependant, il y a 25 stations de livraison. En agrégeant les coûts de transport et la vitesse au niveau NUTS1, le système ne peut pas prendre en compte les différences entre les stations de livraison.

Il est donc nécessaire de redéfinir les régions à un niveau plus fin, pour que les coûts de transport et la vitesse puissent être agrégés au niveau de la station de livraison.

2.3.2 Objectifs

Lorsqu'un FC est choisi d'être le principal FC d'une région pour satisfaire les commandes de cette dernière, le FC devient le **Mothership** de cette région. L'objectif, dans ce cas-là, comprend deux aspects principaux :

- Assurer que chaque FC est utilisé en tant que **Mothership**.
- Assurer que chaque code postal dispose d'au moins un **Mothership**.

En fin de compte, notre objectif est de trouver un centre de distribution principal (**Mothership**) pour l'ensemble des codes postaux des pays de l'EU5 (la France, l'Allemagne, l'Espagne, l'Italie et le Royaume-Uni). Dans chaque pays, les codes postaux partageant le même mothership deviennent ainsi une nouvelle région, avec chaque mothership optimisant la distance et le coût de livraison.

2.4 Planning

La première phase implique de regrouper les données d'entrée des modèles, d'exécuter ces modèles, puis de comparer les résultats afin de sélectionner le meilleur modèle pour nos besoins.

Une fois que nous avons choisi le modèle, il a été nécessaire de modifier et affiner les paramètres d'entrée, analyser les résultats générés, afin d'ajuster le modèle de manière à obtenir les résultats les plus satisfaisants pour nos besoins.

Au cours du déroulement, certaines tâches pouvaient être réalisées simultanément. Par exemple, le développement de l'outil et l'exécution du modèle, car nous disposions déjà des résultats des modèles avec des données d'entrée différentes de celles que nous avions prévues.

De plus, les sous-tâches de l'exécution des modèles s'étendent sur toute la période, car ce sont des tâches récurrentes. Chaque fois que nous obtenions une nouvelle version des résultats, il était nécessaire de les évaluer et de les analyser. Grâce à ces analyses, nous pouvions identifier ce qui devait être modifié. Nous avons répété ces tâches jusqu'à ce que les résultats nous satisfassent.



FIGURE 9 – Le planning du projet

En raison de mon expérience antérieure, j'ai anticipé des imprévus dès le début du développement du projet, et la réalité a confirmé mes prévisions. Par exemple, j'ai rencontré des difficultés lors de la mise en place de notre application sur la plateforme, car nous devons répondre à des exigences internes pour déclarer tous les packages utilisés ainsi que la version, en plus de respecter les formats d'entrée/sortie spécifiques.

D'ailleurs, des retards ont été observés lorsque nous attendions les données de la part de nos parties prenantes, entraînant ainsi des retards de notre côté. Nous avons dû nous adapter afin d'ajuster le planning pour le rendre plus conforme à la réalité.

Il est essentiel de reconnaître que le planning, bien qu'il serve de guide précieux, ne doit pas être considéré comme un ensemble de règles immuables. Notre expérience dans le développement du projet a clairement démontré que des imprévus peuvent survenir, que ce soit en raison de délais dans l'obtention de données des parties prenantes ou de difficultés techniques inattendues. La flexibilité et l'adaptabilité sont des qualités cruciales dans la gestion de projets, permettant d'ajuster le planning en fonction des réalités du terrain.

2.5 Technologies et outils utilisés

Durant mon stage, j'ai utilisé plusieurs technologies et outils.

Développement :

- **Langage de programmation** : Python 3.10, PostgreSQL 13.4
- **IDE**¹ : Visual Studio Code, DataGrip
- **Service de cloud** : Amazon Web Services

Bureautique :

- **La suite Microsoft Office** : pour rédiger les rapports, sortir des fichiers Excel et faire le planning : Microsoft Word,
- **L^AT_EX** : Pour rédiger ce rapport de stage
- **Suite de productivité Salesforce** : *Quip* pour la prise de notes partagées et *Slack* pour la communication interne
- **Présentation** : Microsoft PowerPoint

Système de contrôle de version : Amazon Code²

1. Integrated Development Environment, une suite de logiciel qui fournit des outils essentiels pour le développement des logiciels

2. Un service interne de contrôle de version développé par Amazon

Gestion des tâches : Trello³, Notion⁴

2.6 Contributions

Au cours de ce projet, mes contributions ont été axées sur le développement, avec la réalisation des différentes étapes par mes propres moyens. En tant que membre actif de l'équipe, j'ai assumé la responsabilité de la mise en œuvre des développements.

Par ailleurs, il est important de souligner que la collaboration avec mon manager a été cruciale. Son soutien m'a été particulièrement précieux dans la collecte des données nécessaires auprès de nos parties prenantes. Grâce à cette collaboration efficace, nous avons pu surmonter les défis liés à l'obtention des informations essentielles pour le projet, consolidant ainsi notre démarche collaborative au sein de l'équipe.

À la fin de mon stage, j'ai déjà construit un flux stable de mon travail quotidien, et il existe des processus constants à suivre. La personne qui me remplace pourrait directement reprendre mon travail. Je vais aussi l'intégrer pour qu'elle puisse continuer doucement mon travail sans confusion de sa part.

Un sous-paragraphe

Lorem ipsum dolor sit amet

Une section non numérotée

3. Un outil de gestion de projet en ligne, inspiré par la méthode Kanban de Toyota

4. Une application de prise de notes, de bases de données, des tableaux Kanban, de wikis, de calendriers et de rappels

3 Réalisation

INTRODUCTION DES MODELE!

3.1 Rassemblement des données et exécution des modèles

3.1.1 Contexte

Une liste non ordonnée :

- Un élément de niveau 1 ;
 - Un élément de niveau deux
 - * Un élément de niveau trois ;
 - * Un second élément de niveau trois.
 - Retour au niveau deux.
- Un autre élément de niveau 1.

3.2 Application du *Mapping*

3.2.1 Contexte

Dans le cadre de la définition des régions, nous cherchons à trouver le centre de distribution principal (**Mothership**) pour chaque code postal à cinq chiffres (zip5) dans un pays, et les codes postaux partageant le même mothership formeraient ainsi une région. Dans ce cas, disposer d'une carte pour visualiser l'apparence des régions nous aiderait à évaluer les décisions prises par les modèles.

Au début, nous avons utilisé l'outil analytique interne *FA* pour réaliser le mappage des motherships au niveau des codes postaux (zip5). Cet outil interne offre la possibilité de créer des cartes visuelles enrichies, permettant l'utilisation de couleurs distinctes et de légendes pour représenter chaque catégorie de manière claire et intuitive.

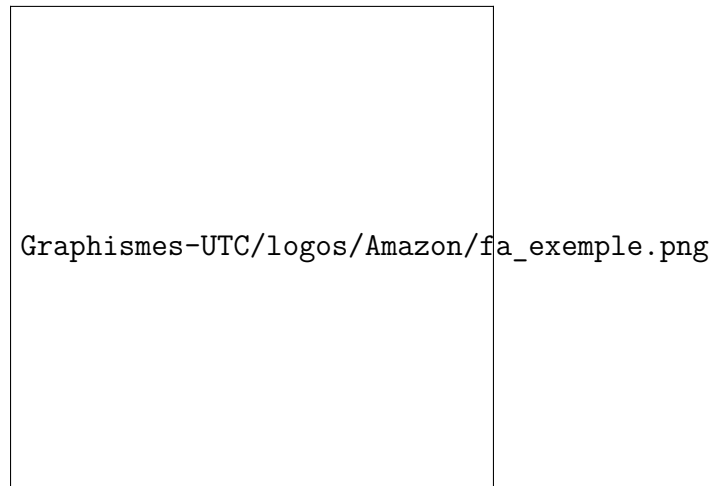


FIGURE 10 – Exemple d’une carte par *FA*

Cependant, nous avons rapidement identifié deux problèmes majeurs :

- Il existe un format strict pour le fichier d’entrée, et le traitement des données en lui-même prend du temps



FIGURE 11 – Exemple d’un fichier d’entrée

- Comme nous réalisons le mappage pour l’ensemble du pays, l’outil *FA* a tendance à être instable pour des cas comme l’Allemagne (DE) où nous avons un grand nombre de codes postaux.

3.2.2 Mission

La mission était particulièrement axée sur la création d’un outil automatisé visant à simplifier et à rendre efficace le processus de visualisation des données. L’objectif était de

concevoir une solution qui élimine le besoin d'une intervention manuelle, permettant ainsi une génération automatique de visualisations.

3.2.3 Action

Outils et techniques utilisés

Pour automatiser complètement le processus, la mise en place d'une plateforme avec un serveur dédié était nécessaire. Cette infrastructure a été cruciale pour héberger l'application et assurer sa disponibilité continue.

Heureusement, Amazon a mis en place une plateforme interne *DasBoard* qui nous permet de créer des applications et de gérer les versions.

DasBoard est une plateforme d'orchestration de travaux distribuée conçue pour les applications de modélisation développées par les chercheurs en recherche. Elle automatise la configuration de l'infrastructure logicielle (ressources AWS, paquet de code, ensemble de versions et pipelines de déploiement continu) pour administrer leurs applications avec une interface utilisateur unifiée. Elle favorise le partage d'applications avec d'autres utilisateurs, tout en facilitant l'intégration d'applications via les flux de travail DasBoard. Cette plateforme offre une interface utilisateur pour n'importe quel outil en ligne de commande. Ces outils en ligne de commande sont appelés "Applications" dans le contexte de DasBoard.

Pour la construction de l'application, le choix du langage Python pour ce projet s'est révélé judicieux en raison de ses nombreuses bibliothèques pratiques et de sa réputation bien établie pour ses performances en matière de traitement de données. Python offre un écosystème riche de bibliothèques spécialisées qui facilitent le développement rapide et efficace d'applications axées sur la manipulation et l'analyse de données.

Première version

Dans la première phase, notre vision consiste à créer une carte au format image pour visualiser le mappage dans chaque pays. Cette approche permettra une représentation graphique claire et intuitive du maillage géographique, offrant ainsi une vue d'ensemble de la répartition des régions définies.

La première version de l'application vise à établir une structure générale, un flux de travail complet, et une application fonctionnelle sans erreurs. Pour atteindre cet objectif, la première étape consiste à traiter les données à partir de la sortie des modèles, en extrayant les informations nécessaires. Ensuite, il est essentiel de préparer les données finales en ajoutant des informations géographiques pertinentes. Enfin, la phase de réalisation du mappage intervient, utilisant des bibliothèques spécialisées pour créer une représentation visuelle du maillage géographique.

Pour le traitement des données, vu que dans une première phase, nous avons tenté d'utiliser le modèle Polaris, dont la sortie comprend les unités de différentes catégories, de la source (FC) à la destination (DS), j'ai entrepris de calculer la somme de chaque catégorie, puis réalisé une comparaison entre elles. Le processus débute en lisant le fichier `nd_solution.csv` dans un `DataFrame` Pandas, un outil puissant de manipulation de données en Python. Une fois les données chargées, la tâche consiste à regrouper les informations en fonction de différents champs et à agréger la colonne "units" en faisant la somme de ses valeurs par la fonction `groupby` fournie par `Pandas.DataFrame`.

Après cette étape, nous avons identifié la meilleure FC pour chaque DS (destination). Pour obtenir la meilleure FC pour chaque code postal, il a été nécessaire de fusionner les données précédentes avec les paires zip-ds fournies par notre partie prenante. Ces informations sont stockées dans le `S3 bucket` de notre compte `AWS`, car les exigences de *DasBoard* ne nous autorisent pas à enregistrer des fichiers prédéfinis.

Pour ce faire, j'ai créé une fonction `get_s3`. Cette fonction récupère un fichier depuis un `S3 bucket` en utilisant le SDK `AWS boto3`. Elle nécessite le nom du bucket, un `ARN` de rôle `IAM` avec accès, et le nom du fichier (clé) à récupérer. En assumant le rôle, elle obtient des informations d'identification temporaires. Ensuite, elle utilise un client `S3` pour récupérer le fichier spécifié. La fonction prend également en compte le format du fichier (CSV, JSON, ou GeoJSON) pour lire son contenu dans un `DataFrame Pandas` ou `GeoPandas`.

En effectuant une simple fusion (`merge`) de ces ensembles de données, nous avons obtenu le mapping final entre les codes postaux et les FCs, établissant ainsi une correspondance optimale entre les zones géographiques et les centres de traitement logistique.

Dans l'exemple de la carte précédente, chaque code postal est représenté par un polygone. Sa forme et sa localisation sont généralement définies par un fichier `JSON` ou `GeoJSON`. Pour des raisons de sécurité et de conformité, ces données sont fournies sur notre plateforme interne. En effectuant une simple fusion(`merge`) sur le code postal, j'ai obtenu le `DataFrame` final qui associe chaque code postal à son polygone correspondant.

Notre objectif était de créer une image pour chaque pays, et la bibliothèque `GeoPandas` m'a permis d'atteindre cet objectif de manière efficace. En utilisant `GeoPandas`, j'ai pu manipuler et visualiser les données géographiques, représentant ainsi graphiquement les polygones associés à chaque code postal.

Deuxième version

Bien que la première version ait été bien accueillie par mon manager, j'ai décidé d'aller plus loin, car je me suis rendu compte que la visibilité de chaque pays était limitée en raison de la taille de l'image. Avec l'ajout des couleurs, il était parfois même plus difficile d'identifier chaque catégorie. Dans ce contexte, j'ai pensé à créer une nouvelle version qui permettrait aux utilisateurs de zoomer pour examiner une région spécifique.

Après avoir exploré cette problématique, j'ai identifié une bibliothèque qui me permettrait de réaliser cet objectif : **Folium**. **Folium** est une bibliothèque Python qui facilite la création de cartes interactives basées sur des données géospatiales. Son intégration avec d'autres bibliothèques telles que **Leaflet.js** permet d'ajouter des fonctionnalités de zoom et de panoramique, offrant ainsi une expérience utilisateur améliorée pour l'exploration de données géographiques.

```
*** exemple de l'utilisation ***
import folium
# Create a map centered on a specific location
mymap = folium.Map(location=[37.7749, -122.4194], zoom_start=12)
# Add a marker with a popup message
folium.Marker(
    location=[37.7749, -122.4194],
    popup='San Francisco, CA',
    icon=folium.Icon(color='blue')
).add_to(mymap)
# Save the map as an HTML file
mymap.save('example_map.html')
```

Lors de l'utilisation de cette bibliothèque, il est essentiel de mettre en place une fonction dédiée à la génération d'un ensemble de couleurs ainsi qu'une autre fonction spécifiquement dédiée à la définition de styles. La première fonction sert à créer une palette de couleurs harmonieuse et distincte, tandis que la seconde permet de personnaliser les styles d'éléments graphiques tels que les marqueurs, les lignes ou les polygones sur la carte interactive.

Dans ce cas-là, j'ai créé deux fonctions : **generate_colors** et **style_function**.

La fonction **generate_colors** génère de manière automatisée un ensemble de couleurs distinctes à partir d'une liste de catégories fournie en paramètre. Elle utilise la colormap **'tab20'** de **Matplotlib**, qui contient 20 couleurs prédéfinies, en sélectionnant le nombre de couleurs en fonction du nombre de catégories. En utilisant une compréhension de liste, elle itère à travers la colormap, convertit les valeurs RGB en codes couleur hexadécimaux, puis crée un dictionnaire associant chaque catégorie à sa couleur respective. Ce dictionnaire est ensuite renvoyé par la fonction. Ainsi, elle simplifie le processus d'attribution de couleurs distinctes et cohérentes à chaque catégorie pour des graphiques ou d'autres visualisations.

Cette fonction, nommée **style_function**, stylise les éléments sur une carte en fonction de leur catégorie source. Elle prend un élément en paramètre, extrait la valeur de la propriété "source" des propriétés de cet élément, puis utilise le dictionnaire **category_colors** (généré ailleurs) pour obtenir la couleur attribuée à cette catégorie source. Si la catégorie n'est pas présente dans le dictionnaire, elle utilise la couleur "gris" par défaut. La fonction retourne un dictionnaire avec des propriétés de style telles que **fillColor** (provenant de **category_colors**), **fillOpacity**, **color**, et **weight**.

Finalisation

Jusqu'à présent, l'application a pris une forme claire et a été déployée en production. Lors de l'examen complet de l'application, une nouvelle fonctionnalité a été proposée. Il était nécessaire d'ajouter des repères pour indiquer les emplacements de nos centres de traitement logistique (FC), ainsi que de générer des données sur le pourcentage de volume de chaque FC dans son pays et la distance moyenne entre cet FC et tous les codes postaux qu'il couvre.

Pour intégrer des repères, une approche directe a été adoptée en utilisant la fonction `folium.Marker` fournie par la bibliothèque `folium`. Cette fonction prend en entrée les coordonnées de latitude et de longitude de chaque repère à ajouter. Les repères fournissent également des informations contextuelles, telles que le nom du repère, lorsqu'on survole le repère avec la souris.

Il reste donc

Un sous-paragraphe

Lorem ipsum dolor sit amet

3.3 Analyses approfondies sur le business

Références

- [1] CONTRIBUTEUR. Amazon Wikipedia. URL : <https://fr.wikipedia.org/wiki/Amazon>. (Consulté le 16 novembre 2023).
- [2] VENNGAGE. Amazon History Timeline. URL : <https://fr.venngage.com/templates/social-media/amazon-history-timeline-5e9942ed-1474-44d3-ba8a-3d283eabe34a>. (Consulté le 25 novembre 2023).
- [3] MAHESHONE. All-things-amazon. URL : <https://maheshone.com/amazon-products-and-services/>. (Consulté le 27 novembre 2023).
- [4] Lélia de MATHAREL. Amazon History Timeline. URL : <https://fr.venngage.com/templates/social-media/amazon-history-timeline-5e9942ed-1474-44d3-ba8a-3d283eabe34a>. (Consulté le 20 novembre 2023).
- [5] STATISTA. Annual net sales revenue of Amazon from 2004 to 2022. URL : <https://www.statista.com/statistics/266282/annual-net-revenue-of-amazoncom/>. (Consulté le 27 novembre 2023).
- [6] AMAZON. About Amazon EU. URL : <https://aboutamazon.eu>. (Consulté le 16 novembre 2023).
- [7] Noah BERLANDI. « SCE-DPE pilot SCOT alignment on inputs ». In : Amazon Confidential (2022), p. 1.
- [8] EUROSTAT. NUTS and Statistical regions - level 1. URL : <https://ec.europa.eu/eurostat/documents/345175/7451602/2021-NUTS-1-map.pdf>. (Consulté le 22 novembre 2023).
- [9] EUROSTAT. ESPANA NUTS level 2. URL : <https://ec.europa.eu/eurostat/documents/345175/7451602/2016-NUTS-2-map-ES.pdf>. (Consulté le 22 novembre 2023).

★ ★ ★