



Amazon – Analyse des avis produits

Projet Data Engineering & Analyse avancée



Ismaël SYLLA

Soutenance – Certification
Décembre 2025



PLAN DE LA PRÉSENTATION

1. Introduction & contexte du projet
2. Analyse stratégique et identification du besoin
3. Architecture data globale
4. Data Engineering & pipelines
5. Analyse avancée des avis & NLP
6. Gouvernance, déploiement & pilotage
7. Conclusion & perspective



PARTIE 1



Introduction & contexte du projet

Objectifs de la partie

- I. Présenter le contexte global du projet
- II. Exposer la problématique métier
- III. Définir les objectifs et la vision d'ensemble

Compétences RNCP mobilisées

C1 – Analyse de l'écosystème et du contexte
C8 – Cadrage et formalisation du besoin
C26 – Pilotage et vision globale du projet



Introduction générale du projet



- ❑ Amazon : acteur majeur du e-commerce et du cloud
- ❑ Les avis clients : levier clé de confiance et de performance
- ❑ Problématique : volume massif + données non structurées
- ❑ Besoin d'une exploitation data-driven des avis



Présentation d'Amazon et écosystème



- 3^e marque mondiale (Interbrand 2024)
- E-commerce, AWS, logistique, services numériques
- Culture data & IA au cœur des décisions
- Environnement international et réglementé



Pourquoi les avis clients sont stratégiques



- Impact sur :
 - Confiance des acheteurs
 - Classement des produits
 - Recommandations algorithmiques
- Source directe de feedback métier
- Forte valeur mais faible structuration



Problématique globale



- Trop d'avis → difficulté à identifier les plus pertinents
- Risque de faux avis ou avis peu utiles
- Manque de priorisation intelligente
- Besoin d'automatisation et de transparence



Vision du projet



- Construire une solution data-driven
- Automatiser l'analyse et la priorisation des avis
- Intégrer NLP & scoring de pertinence
- Respecter les contraintes réglementaires



Périmètre du projet



- Analyse des avis textuels
- Enrichissement par métadonnées (rating, image, achat)
- Pipeline data & NLP
- Restitution analytique



Positionnement du projet



- ❑ Projet orienté **Data Engineering**
- ❑ Forte dimension **analyse et valeur métier**
- ❑ Prise en compte :
 - des contraintes techniques
 - des contraintes réglementaires
 - des contraintes organisationnelles



Transition vers l'analyse stratégique



- Comprendre l'écosystème Amazon
- Identifier forces, faiblesses, opportunités
- Définir les besoins métier réels
- Prioriser les cas d'usage



PARTIE 2



Analyse stratégique et identification du besoin

Objectifs de la partie

- I. Analyser le contexte et les enjeux du projet
- II. Identifier les besoins métier prioritaires
- III. Définir le cas d'usage à plus forte valeur

Compétences RNCP mobilisées

C1 – C6
C8



ANALYSE STRATÉGIQUE (SWOT)



FORCES	FAIBLESSES
<ul style="list-style-type: none">• Marque mondiale forte et reconnue• Réseau logistique très performant• Leadership e-commerce & cloud (AWS)• Capacités avancées en IA, NLP et Big Data• Large base de clients fidèles	<ul style="list-style-type: none">• Dépendance aux vendeurs tiers• Manque de transparence perçu des algorithmes• Complexité organisationnelle• Risques juridiques et réglementaires• Défis de sécurité à grande échelle
OPPORTUNITÉS	MENACES
<ul style="list-style-type: none">• Croissance des marchés émergents• Exploitation stratégique de l'IA et des LLM• Valorisation des retours clients• Innovation logistique et durable• Fidélisation et hausse du panier moyen	<ul style="list-style-type: none">• Concurrence intense (Alibaba, Walmart, Shein...) • Pression réglementaire (RGPD, CCPA, AI Act)• Cyberattaques et fuites de données• Perte de confiance dans les avis produits• Marché en évolution rapide



Approche méthodologique (Approche Design Thinking)



- ❑ Comprendre → Observer → Définir
- ❑ Idéation → Prototype → Test
- ❑ Approche centrée utilisateur
- ❑ Itération continue



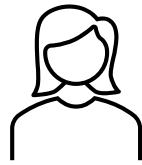
IDENTIFICATION DES PARTIES PRENANTES



- Clients finaux
- Équipes produit
- Équipes data & analytics
- Équipes support et qualité



Personas



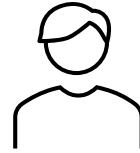
- Emma
- 21 ans
- Etudiante en psychologie

Besoin identifié :

- Bénéficier d'un service de livraison rapide.

Cas d'utilisation :

- Optimisation logistique et amélioration de la livraison dans les zones rurales.



- Rayan
- 25 ans
- Business Analyst chez Amazon

Besoin identifié :

- Pouvoir automatiser l'analyse et la classification des avis clients sur les produits.

Cas d'utilisation :

- Classer les avis et identifier leur pertinence.



- Mathis
- 42 ans
- entrepreneur

Besoin identifié :

- Améliorer la visibilité et la performance de ses produits.
- Comprendre les critères influençant le classement des produits.

Cas d'utilisation :

- Analyse des avis clients et tendances de satisfaction.
- Transparence sur les algorithmes de classement produits.



Design Thinking – Personas



- Rayan
- 25 ans
- Business Analyst chez Amazon

Besoin retenu :

✓ Classer les avis et identifier leur pertinence.

Valeurs ajoutées :

Analyse rapide et plus fiable.

Elimination d'un travail manuel et gain de temps.

Levier et première brique pour d'autres besoins identifiés.



Priorisation des cas d'usage



□ Priorité : classification & pertinence des avis

- Fort impact métier
- Faisabilité technique
- Base pour d'autres cas d'usage



- Stockage & Data Lake : Amazon S3
- Data Warehouse : Snowflake
- Analyse & ML : NLP, Zero-Shot, VADER
- Plateformes avancées : Streamlit,
SageMaker



- RGPD : protection et anonymisation des données
- CCPA : droits des utilisateurs
- AI Act : transparence des algorithmes
- RSE : IA responsable et durable



Veille réglementaire statique



Texte	Sujet	Points clés	Date	Mots-clés utilisés
RGPD (Regulation (EU) 2016/679)	Protection des données personnelles	- Cadre clé de conformité - Impact direct sur le NLP	Texte initial 27 avr. 2016 (applicable depuis 2018)	RGPD, data privacy, anonymisation, consentement, EU compliance
CCPA (California Consumer Privacy Act)	Réglementation américaine sur la protection des données personnelles	- Cadre californien de protection des données - Impact opérationnel	Adopté en 2018 – renforcé en 2023 (CPRA)	CCPA, CPRA, data privacy US, consentement, droit à la suppression
AI Act (Règlement UE)	Régulation IA (transparence, obligations)	- Transparence et contrôle - Impact technique	2024–2025 (mesures d'application 2025).	IA Acte, réglementation IA, transparence, conformité UE
RSE Numérique (Responsabilité Sociétale & Environnementale)	Réglementation / bonnes pratiques	- IA responsable et sobre - Dimension éthique et sociale	2024- 2025	RSE numérique, CSRD, Green AI, IA responsable, durabilité



PARTAGE DE LA VEILLE

- Documentation centralisée
- Partage collaboratif (outils internes)
- Mise à jour continue





Conclusion de la phase stratégique

- Besoin clairement défini
- Cas d'usage priorisé
- Nécessité d'une architecture scalable
- Passage à la conception technique



PARTIE 3



Architecture data globale

Objectifs de la partie

- I. Présenter l'architecture data cible du projet
- II. Justifier les choix technologiques retenus
- III. Garantir la sécurité, la traçabilité et la conformité des données

Compétences RNCP mobilisées

C9 – C11
C14
C15



OBJECTIFS DE L'ARCHITECTURE



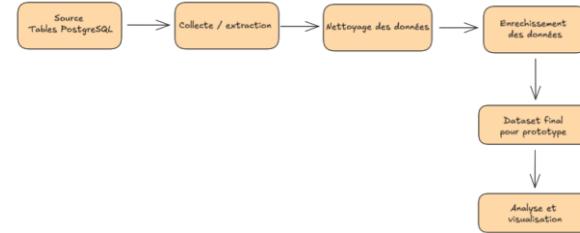
- Centraliser et structurer les données
- Garantir la qualité et la traçabilité
- Permettre l'analyse et la valorisation métier
- Assurer la scalabilité et la gouvernance



VISION GLOBALE DE L'ARCHITECTURE



- ❑ Architecture orientée **Data Plateforme**
- ❑ Séparation claire des responsabilités
- ❑ Flux de données maîtrisé de bout en bout





SOURCES DE DONNÉES



- ❑ Base transactionnelle PostgreSQL
- ❑ Données structurées (produits, acheteurs, avis)
- ❑ Données enrichies (images, métadonnées)

DBeaver 25.2.4 - <amazon_db> Script-1

Fichiers Edition Navigation Rechercher Editeur SQL Base de données Fenêtres Aide

Navigateur de bases de données Projets

Filter connections by name

amazon_db localhost:5433

Bases de données amazon_db

Schémas public

Tables

- buyer
- carrier
- cart
- cart_items
- category
- customer
- customer_payment
- customer_shipping
- daily_deals
- discount
- orders
- payment_details
- product
- product_images
- product_reviews
- return
- review
- review_images
- seller
- seller_products
- seller_reviews
- shipment
- shipping_details
- subscription
- wishlist_item

Tables étrangères

Files - General

10M
8K
8K
8K
24K
35M
12M
27M
16K
16K
21M
12M
17M
40M
9,1M
8K
32M
29M
4,5M
7,3M
16M
8K
28M
4,8M
3,5M
...



DATA LAKE : ZONE D'ATTERRISSAGE



- ❑ Stockage des données brutes
- ❑ Historisation des données
- ❑ Découplage source / analytique

The screenshot shows the AWS S3 console interface. On the left, a sidebar menu includes 'Compartiments', 'Gestion des accès et sécurité', and 'Gestion du stockage et informations'. The main area displays a list of objects under the 'raw/' folder. The table has columns for 'Nom' (Name), 'Type' (Type), 'Dernière modification' (Last modified), 'Taille' (Size), and 'Classe de stockage' (Storage class). Objects listed include 'http://', 'same/', 'category/', 'order/', 'product_detail/', 'product/', 'master_product/', and 'missed'. Action buttons like 'Copier l'URI S3', 'Copier l'URL', 'Télécharger', 'Ouvrir', 'Supprimer', 'Actions', 'Créer un dossier', and 'Charger' are visible at the top of the object list.



STRUCTURATION BRONZE / SILVER / GOLD



- ❑ **Bronze** : données brutes
- ❑ **Silver** : données nettoyées et enrichies
- ❑ **Gold** : données prêtes pour l'analyse

amazon-industry-insights-raw-data infos

Objets Métadonnées Propriétés Autorisations Métriques Gestion Points d'accès

Objets (4)

Copier l'URI S3 Copier l'URL Télécharger Ouvrir Supprimer Actions Crée un dossier Charger

Les objets sont les entités fondamentales stockées dans Amazon S3. Vous pouvez utiliser l'inventaire Amazon S3 pour obtenir une liste de tous les objets de votre compartiment. Pour que d'autres personnes puissent accéder à vos objets, vous devez leur accorder explicitement des autorisations. En savoir plus

Rechercher des objets en fonction du préfixe

Nom	Type	Dernière modification	Taille	Classe de stockage
bronze/	Dossier	-	-	-
gold/	Dossier	-	-	-
raw/	Dossier	-	-	-
silver/	Dossier	-	-	-



- ❑ Centralisation des données analytiques
 - ❑ Optimisation des requêtes
 - ❑ Séparation stockage / calcul

Database Explorer		AMAZON_DB / AMAZON / REVIEWS		Last Update: 1 day ago 21 rows 21 columns	
Amazon		Reviews			
Reviews		Reviews			
REVIEW_ID	BEST_RATING	REVIEWER_ID	ITEM_ID	P.D.	PRODUCT_NAME
1	96005	27ab6e0f-212c-2f95-8d4ef76e800230c329f18710416...	800.MJ822W	Samsung The 1TB Internal Storage Case Box Harder Container PINK Jewelry Box...	AMAZON
2	96013	0307049a-7616-0778-7379-000000000000	96013	Erika Hippopotamus Bath for Adults - Paraben-Free Natural Ylang Ylang...	AMAZON
3	96021	06966773-0557-2047-7200-000000000000	96021	SMILENIN Writing Pen for Girls - Premium Fine Point Spring Safety...	AMAZON
4	96028	07000000-0000-0000-0000-000000000000	96028	Wet Seal Girls' Soft Fleece Hoodie - Solid Colors - For The Youngest Girls...	AMAZON
5	96037	5053-3136-0045-0045	96037	Levi's Women's High-Rise Super Skinny Fit Jean - Indigo - 26x30	AMAZON
6	96045	034730-0000-0000-0000-000000000000	96045	Tom's One Tree Olive Oil Hand Facial Moisturizer 10 fl. oz.	AMAZON
7	96053	000000-0000-0000-0000-000000000000	96053	Dermablend Ageless Skin Care Cream, Flawless Optically ERASE The Appearance...	AMAZON
8	96061	04d4bf-0000-0000-0000-000000000000	96061	Skinner Black Scrubber for Shower Exfoliating, 2.5" x 10"	AMAZON
9	96069	000000-0000-0000-0000-000000000000	96069	Fog Horn Snorkel Mask & Mouthpiece - Diving Snorkeling Mask with Snorkel...	AMAZON
10	96077	39e4-0000-0000-0000-000000000000	96077	Skull Dental Chair Headrest - Stainless Steel Dental Chair Professional O...	AMAZON
11	96086	M19a00-0000-0000-0000-000000000000	96086	Granulomatous Bristle Smoothing Hair Brush for Women and Men - Medium...	AMAZON
12	96094	5e53-0000-0000-0000-000000000000	96094	Father Shaver Styling Professional Barber - Everyday Hair Styling Tool - Combi...	AMAZON
13	96102	03e3ed-0000-0000-0000-000000000000	96102	Parfait Long Straight Black Wig Synthetic Wigs for Women Natural Middle Part...	AMAZON
14	96110	000000-0000-0000-0000-000000000000	96110	The Original 100% Pure Organic 15ml Eye Cream - The Best Eye Cream of GLAQUED Eyebrow...	AMAZON
15	96118	997-0000-0000-0000-000000000000	96118	Bath Scrubber Body Brush Shower Scrubber Back Brush with Long Handle, Ergo...	AMAZON
16	96126	Ma15cf-0000-0000-0000-000000000000	96126	Sweet Baby Shampoo, Sulfate Free, Nutri-Pure, No Parabens, Phthalates, Dyes, Endoc...	AMAZON
17	96134	7995-0000-0000-0000-000000000000	96134	Salonix Body Wash (Scented)	AMAZON
18	96142	000000-0000-0000-0000-000000000000	96142	2	AMAZON



GESTION DES DONNÉES NON CONFORMES

(Gestion des rejets et logs)



- ❑ Données rejetées stockées séparément
- ❑ Journalisation des erreurs
- ❑ Tracabilité des traitements

Mongo Express Database: airflow_logs > Collection: transform_load_logs > Document 693eeef787f4504d8e41b3520

Editing Document: 693eeef787f4504d8e41b3520

```
1 {  
2     "_id": ObjectId('693eeef787f4504d8e41b3520'),  
3     "message": "Loading tables from s3...",  
4     "level": "INFO",  
5     "logger": "utils.review.processor",  
6     "timestamp": ISODate("2025-12-14T17:10:16.071Z"),  
7     "dag_id": null,  
8     "task_id": null,  
9     "run_id": null  
10 }
```

Back Save

Delete

Mongo Express Database: airflow_logs > Collection: extract_to_s3_logs > Document 693eeef6b2d306782f9466bcb

Editing Document: 693eeef6b2d306782f9466bcb

```
1 {  
2     "_id": ObjectId('693eeef6b2d306782f9466bcb'),  
3     "message": "Starting anonymize data for table category...",  
4     "level": "INFO",  
5     "logger": "unusual_prefix_4e201c21eed689f79219de2715e92d8f62d84ac5_extract_to_s3",  
6     "timestamp": ISODate("2025-12-14T17:10:03.131Z"),  
7     "dag_id": null,  
8     "task_id": null,  
9     "run_id": null  
10 }
```

Back Save

Delete



ORCHESTRATION DES FLUX



- Enchaînement automatisé des traitements
- Planification des tâches
- Gestion des dépendances

Mongo Express Database: airflow_logs

Collections

	View	Export	[JSON]	Import	
extract_to_s3_logs					<button>Del</button>
transform_load_logs					<button>Del</button>
transform_logs					<button>Del</button>

Database Stats

Collections (incl. system.namespaces)	3
Data Size	16.1 KB
Storage Size	61.4 KB
Avg Obj Size #	210 Bytes
Objects #	77
Indexes #	3



SÉCURITÉ & TRACABILITÉ (gouvernance)



- Contrôle des accès
- Traçabilité des transformations
- Conformité réglementaire intégrée



Conclusion de la partie architecture



- Architecture cohérente et scalable
- Séparation claire des rôles
- Base solide pour les pipelines



PARTIE 4



Data Engineering & pipelines

Objectifs de la partie

- I. Mettre en œuvre les pipelines de données automatisés
- II. Garantir la qualité, la fiabilité et la traçabilité des données
- III. Industrialiser les traitements et les flux data

Compétences RNCP mobilisées

C12 – C13
C18 – C19



RÔLE DU DATA ENGINEERING



- Transformer des données brutes en données exploitables
- Automatiser les traitements
- Garantir la fiabilité et la qualité des données
- Préparer l'analyse et la restitution métier



VUE D'ENSEMBLE DU PIPELINE DATA



- Pipeline **ETL orchestré par Airflow**
- Architecture **modulaire et découplée**
- Séparation ingestion / transformation / restitution
- Traçabilité et qualité intégrées dès la conception



ORCHESTRATION AVEC AIRFLOW



- ❑ DAG principal d'orchestration
- ❑ DAG d'extraction PostgreSQL → S3
- ❑ DAG de transformation & chargement
- ❑ Communication inter-DAG via XCom



SOURCES & VOLUMÉTRIE DE DONNEES



- Base PostgreSQL transactionnelle
- Tables cœur métier :
 - avis,
 - produits,
 - acheteurs,
 - catégories,
- Données structurées et semi-structurées



INGESTION : POSTGRESQL → S3 (DATA LAKE)



- Extraction depuis PostgreSQL
- Anonymisation des données sensibles
- Stockage en zone raw sur S3
- Gestion des erreurs et alerting



TRANSFORMATION & ENRICHISSEMENT



- Chargement des données depuis S3
- Jointures multi-tables
- Nettoyage et normalisation
- Séparation données propres / rejetées



- Tests de qualité automatisés
- Validation des règles métier
- Génération de rapports qualité
- Seuils d'acceptation contrôlés



CHARGEMENT ANALYTIQUE DANS SNOWFLAKE



- Chargement des données propres dans Snowflake
- Modèle orienté analytique
- Création de vues métiers
- Séparation stockage / calcul



- Logs centralisés dans MongoDB
- Métadonnées d'exécution
- Suivi des volumes traités
- Historisation des runs



GESTION DES ERREURS & ALERTING



- Détection des anomalies
- Alertes email automatiques
- Blocage contrôlé du pipeline
- Reprise sécurisée



- Conteneurisation Docker
- Variables d'environnement sécurisées
- Déploiement reproductible
- Séparation dev / prod



SCALABILITÉ & ÉVOLUTIVITÉ



- Ajout de nouvelles sources
- Augmentation des volumes
- Nouveaux cas d'usage analytiques
- Passage à des traitements plus avancés



Conclusion Data Engineering



- Données fiables et prêtes à l'analyse
- Pipelines automatisés
- Base solide pour la valorisation métier



PARTIE 5



Analyse avancée des avis & NLP

Objectifs de la partie

- I. Analyser les avis clients à partir des données textuelles
- II. Mettre en place un scoring de pertinence des avis
- III. Exploiter le NLP pour produire de la valeur métier

Compétences RNCP mobilisées

C7
C16
C17



OBJECTIFS DE L'ANALYSE AVANCÉE



- Comprendre le contenu réel des avis clients
- Identifier les avis les plus exploitables
- Compléter la note par des signaux qualitatifs
- Aider la prise de décision métier



VUE D'ENSEMBLE DES DONNÉES ANALYSÉES



- ❑ 2 592 avis analysés
- ❑ Rating moyen : 4,46 / 5
- ❑ 31,6 % des avis contiennent des images
- ❑ Longueur texte très hétérogène



DISTRIBUTION DES RATINGS



- Forte concentration en 5★ (≈ 76 %)
- Peu d'avis négatifs
- Déséquilibre naturel des classes



LONGUEUR DU TEXTE VS RATING



- Avis 2★-4★ → textes plus longs
- Avis 5★ → textes plus courts
- Avis négatifs souvent plus argumentés



PRÉSENCE D'IMAGES DANS LES AVIS



- Corrélation positive entre rating élevé et images
- 35,7 % des avis 5★ avec images
- Signal de preuve utilisateur



- Nettoyage du texte
- Normalisation
- Filtrage des avis trop courts
- Calcul de features textuelles



FEATURES UTILISÉES POUR L'ANALYSE



- Longueur du texte
- Présence d'image
- Rating
- Score de sentiment
- Indicateurs de pertinence



ANALYSE DE SENTIMENT



- Score positif / neutre / négatif
- Complément du rating
- Détection d'écart note / sentiment



IDENTIFICATION DES AVIS PERTINENTS



- Combinaison longueur + image + sentiment
- Priorisation des avis utiles
- Réduction du bruit informationnel



APPROCHE ZERO-SHOT



- ❑ Classification thématique sans entraînementModèles
- ❑ NLP pré-entraînés (Transformers)
- ❑ Détection dynamique des sujets évoqués
- ❑ Forte flexibilité métier



- Exploration interactive
- Filtres par score / rating
- Lecture directe des avis clés
- Approche orientée usage métier

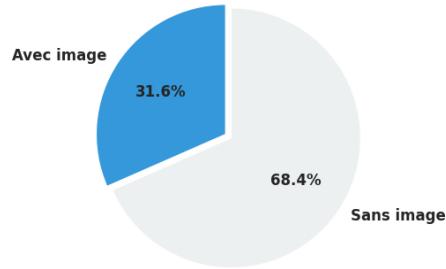


Prototype

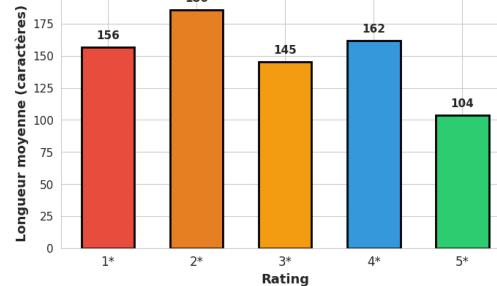


Analyse sur 10 produits (2592 avis)

Présence d'Images



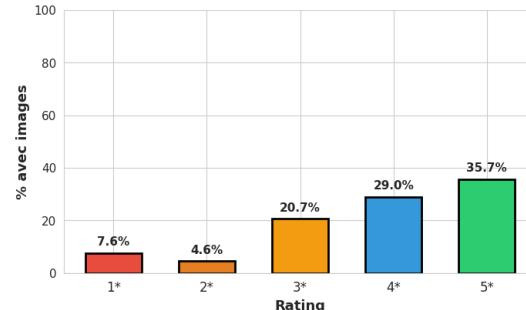
Longueur Moyenne du Texte par Rating



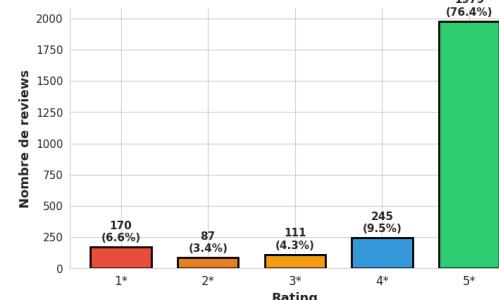
RÉSUMÉ STATISTIQUE

- Reviews totales : 2592
- Avec images : 818 (31.6%)
- Avec commandes : 2592 (100.0%)
- Longueur texte :
 - Moyenne : 117 caractères
 - Médiane : 44 caractères
 - Min : 4 | Max : 1815
- Rating moyen : 4.46/5

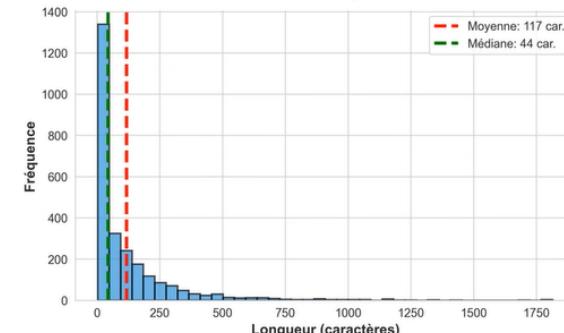
Proportion de Reviews avec Images par Rating



Distribution des Ratings



Distribution de la Longueur du Texte





Justificatif de pondération



30 %

Longueur de texte : plus l'avis est long, plus il y a de détail.

25 %

Sentiment (VADER) : analyse de sentiments pour détecter la satisfaction ou l'insatisfaction des clients.

20 %

Présence d'image : 31% des avis contiennent des images (indicateur d'engagement).

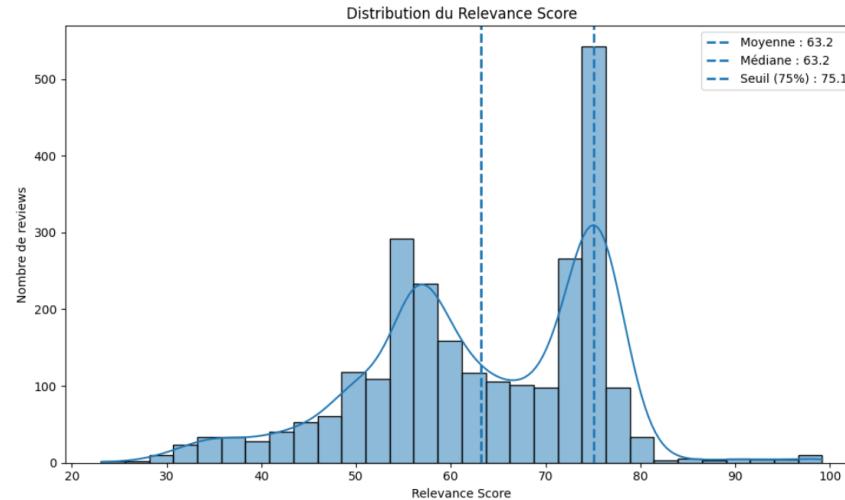
15 %

Note extrême : les avis très positifs ou très négatifs apportent souvent plus d'information.

10 %

Achat vérifié : tous les avis sont vérifiés

Sélection de seuil





Résultats du prototype



Exemple 1 : avis pertinent

- Description : « Love it »
- Présence d'image
- Achat vérifié
- **Sentiment positif fort (VADER)**
- Score de texte (fonction centrée sur 300 caractères)

Exemple 2 : avis non pertinent

- Description : « This product is very good.... (~ 1000 caractères)»
- Présence d'une image
- Achat vérifié
- Sentiment moyen (VADER)
- **Score du texte (fonction centrée sur 300 caractères) qui fait baisser sa valeur**

Limites claires

- Score basé surtout sur la forme et l'émotion
- Survalorisation des avis 5★
- Avis courts parfois trop bien notés
- Avis longs parfois pénalisés
- Pas d'analyse sémantique
- Mesure l'engagement, pas la vraie pertinence



RÉSULTATS OBTENUS



- Hiérarchisation claire des avis
- Identification des avis à forte valeur
- Réduction du bruit informationnel



LIMITES & AXES D'AMÉLIORATION



Limites identifiées	Améliorations possibles
<input type="checkbox"/> Dépendance à la qualité du texte	<input type="checkbox"/> Modèles NLP plus avancés
<input type="checkbox"/> Pondérations perfectibles	<input type="checkbox"/> Apprentissage supervisé
<input type="checkbox"/> Volume de données limité	<input type="checkbox"/> Feedback utilisateur



- Tableaux analytiques
- Indicateurs clés par produit
- Aide à la décision



- ❑ Valorisation des résultats NLP
- ❑ Exploration interactive des avis clients
- ❑ Accès simple pour les équipes métier
- ❑ Approche produit data (MVP)

 **Amazon Product Reviews - Relevance Analysis**

Browse products and view their most relevant reviews based on:

-  Rating score
-  Text quality and length
-  Keyword relevance
-  Presence of images

 How do you want to search?

Browse Products Search by Product ID

 100 products disponibles

 Filter products by name:
Type to search...

 Showing 100 product(s)

 Select a Product:
ISALE! SUPPLY Tallow Shave Soap (Juniper, 4.5 oz) (All Beauty)



- Connexion directe aux tables **Gold** (Snowflake)
- Lecture seule (pas d'impact sur les pipelines)
- Découplage traitement / visualisation
- Sécurité et traçabilité préservées



Apport métier et perspectives



- ❑ Gain de temps pour les équipes
- ❑ Réduction de l'analyse manuelle
- ❑ Meilleure adoption de la solution data
- ❑ MVP évolutif (API, BI, authentification)

★ Most Relevant Reviews

Number of reviews to display: 20

1 review(s) pertinent(s) trouvé(s)

Review #1 | ★★★★ (A/S) - Excellent product | Relevance: 63.84

Description:

An exceptional product that performs as well as or better than advertised. The aroma is pleasing, but the moisturizing effect of the shave soap is very much appreciated. I bought two units and will undoubtedly repeat that order in the future. A very happy customer.

Langue de texte: 200 caractères

Score et Métriques:

Rating: 4/5	Confidence Score: 95.069	Relevance Score: 0.972	Has Images: No
Relevance Score: 63.842	% Test Length Score: 0.985	Extreme Rating: No	Has Orders: Yes

Informations:

Product ID: B0BM4GX6TT

Category: All beauty

Product ID: B0BM4GX6TT

○ Browse Products | Search by Product ID

Enter Product ID: B0BM4GX6TT

Product ID: B0BM4GX6TT

★ Most Relevant Reviews

Number of reviews to display: 20

20 review(s) pertinent(s) trouvé(s)

Review #1 | ★★★★ (A/S) - Excellent product | Relevance: 98.28

Review #2 | ★★★★ (S/S) - Works beyond belief for eyebrows | Relevance: 98.11

Review #3 | ★★★★ (S/S) - Works beyond belief | Relevance: 96.33

Review #4 | ★★★★ (S/S) - Awesome product!! | Relevance: 94.56



- Valeur métier concrète
- Aide à la décision
- Résultats exploitables



PARTIE 6



Gouvernance, déploiement & pilotage

Objectifs de la partie

- I. Assurer la gouvernance et la conformité réglementaire des données
- II. Déployer, superviser et maintenir la solution data
- III. Piloter le projet, les ressources et les risques

Compétences RNCP mobilisées

C20 – C32



GOUVERNANCE GLOBALE DU PROJET



- Gouvernance à 3 niveaux : stratégique, tactique, opérationnel
- Alignement métier / technique
- Décisions structurées et traçables
- Pilotage continu du projet



RÔLES, RESPONSABILITÉS & RACI



- Rôles clairement définis (PO, DE, DS, DevOps, Support...)
- Responsabilités formalisées
- Matrice RACI par phase
- Limitation des zones grises



PILOTAGE AGILE & RYTHME DU PROJET



- Méthodologie Agile / Scrum
- Sprints de 2 semaines
- Rituels : planning, daily, review, rétrospective
- Ajustement continu du backlog



DÉPLOIEMENT & MISE EN PRODUCTION



- Déploiement progressif
- Séparation dev / prod
- Automatisation des traitements
- Sécurisation des accès



- Plan de formation structuré
- Documentation claire et versionnée
- Support à 3 niveaux
- Objectif : autonomie des utilisateurs



- Monitoring des pipelines
- Suivi des incidents
- Indicateurs de qualité
- Boucle d'amélioration continue



MAINTENANCE & ÉVOLUTIVITÉ

Maintien en conditions opérationnelles



- Maintenance corrective et évolutive
- Anticipation de la montée en charge
- Ajout de nouveaux cas d'usage



BUDGET & RESSOURCES

Gestion des ressources



- Choix technologiques maîtrisés
- Priorité aux solutions open source / optimisées
- Arbitrage coûts / valeur



CONDUITE DU CHANGEMENT

Adoption par les utilisateurs



- Documentation claire
- Accompagnement des équipes
- Sensibilisation à l'usage des données



PARTIE 7



Conclusion & perspectives

Objectifs de la partie

- I. Présenter le bilan global du projet
- II. Mettre en évidence les apports et les limites
- III. Ouvrir sur les perspectives d'évolution

Compétences RNCP mobilisées

Synthèse globale du cycle de vie d'un projet
data
Capitalisation des compétences acquises
Vision long terme et amélioration continue



BILAN GLOBAL DU PROJET



- Projet data de bout en bout
- Architecture robuste et scalable
- Valeur métier démontrée
- Gouvernance intégrée



- Enrichissement des modèles NLP
- Intégration de nouveaux jeux de données
- Déploiement à plus grande échelle



POSITIONNEMENT PROFESSIONNEL



- Data Engineer orienté architecture
- Sensibilité gouvernance et conformité
- Capacité à piloter un projet data



CONCLUSION GÉNÉRALE

- Problématique métier adressée
- Solution technique industrialisable
- Vision long terme et responsable





Remerciements



**Merci aux membres du jury
Merci à l'organisme de formation / accompagnement
Merci pour le temps et l'échange**



Merci !

