

PROJET de Spécialisation #2

“SUPPLY CHAIN” : COMPLEMENTS

OBJECTIFS :

- ↳ Optimiser la chaîne d'approvisionnement à travers une analyse de données poussées et l'entraînement d'un modèle de prédiction.
- ↳ Monter en compétences en **Business Intelligence** sur un ensemble de données du monde réel, axé sur l'analyse de la **chaîne d'approvisionnement**.
- ↳ Analyser, comprendre, structurer et optimiser le processus d'approvisionnement d'une chaîne logistique

RAPPEL du CONTEXTE BUSINESS :

Aujourd'hui, pour tirer son épingle du jeu, lorsque l'on est une entreprise de vente en ligne, il est nécessaire de mettre en place une organisation optimale.

Le challenge est d'anticiper les besoins du client, pour être capable de :

- ↳ **délivrer le bon produit,**
- ↳ **au bon endroit,**
- ↳ **et au bon moment,**
- ↳ tout en maîtrisant de manière optimale **les coûts et la qualité.**

Il est donc nécessaire d'optimiser le fonctionnement de sa chaîne logistique.

CONTEXTE du PROJET :

En tant qu'analyste de données :

- Gérez efficacement les **enjeux des expéditions et des stocks**.
- **Identifiez et rectifiez** les inefficacités de la chaîne d'approvisionnement.
- Élaborez des **tableaux de bord stratégiques** pour alerter les stakeholders sur les défis émergents.
- **Préconisez** des améliorations opérationnelles structurantes pour l'entreprise.

PRECONISATIONS en termes d'organisation et de travail

Phase 1 :

Compréhension des relations entre les données : Par exemple, faire un schéma entité-relation (ERD) illustrant les liens (les clés primaires, clés étrangères, etc.) entre produits, commandes, stocks, achats, fournisseurs...

Documentation des variables : Avant même de commencer l'analyse, produire un mini-dictionnaire de données (nom de la variable, type, signification, lien avec d'autres tables).

Approche itérative : un premier tour de nettoyage, une première visualisation, puis itérations successives pour affiner...

Phase 2 :

Préprocessing de vos données en Python avec gestion des valeurs manquantes, des doublons, des valeurs aberrantes ; **Normalisation / cohérence des formats**

Tenir un document (ou un notebook) expliquant chaque étape de cleaning, pour retracer votre raisonnement.

Phase 3 : Analyse Exploratoire des Données sous Power BI, Tableau, Looker...

- **Statistiques descriptives :**
 - Distribution des variables (moyennes, médianes, écart-types, histogrammes).
 - Taux de retours, fréquences des commandes, délais moyens d'expédition, etc.
- **Visualisations de premier niveau (exemples concrets) :**
 - Évolution des stocks dans le temps par produit ou par catégorie.
 - Répartition des commandes par produit, par client ou par période.
 - Coût moyen d'achat ou d'expédition par fournisseur.

- **Mise en avant de KPI et d'indicateurs opérationnels :**
 - Taux de rupture de stock par produit.
 - Taux de retours (défectueux vs erreurs de livraison).
 - Fiabilité fournisseurs (délai moyen, taux de retard).
- **Analyse des corrélations :**
 - Corrélation entre délais de livraison et taux de retours, par exemple.
 - Corrélation entre prix d'achat et volume commandé.
- **Identification des points problématiques : Retards fréquents, stocks trop élevés ou trop bas, produits sur ou sous-performants, etc.**

Phase 4 : Machine Learning (modèles de prédictions divers selon...)

Choix du ou des cas d'usage (*il s'agit ici exemples*) :

- **Prévision de la demande :** Time Series Forecasting ou Régression pour prédire les quantités de commandes futures par produit.
- **Estimation du délai fournisseur :** Modèle prédictif du délai de livraison en fonction de la période, du fournisseur, de la localisation, etc.
- **Analyse de probabilité de retour :** Modèle de classification (retour / non-retour) basé sur le type de produit, le transporteur, la saisonnalité, etc.

Feature Engineering :

- Création de variables temps (par exemple, jour de la semaine, mois, saison).
- Indicateurs cumulés (ex. délai moyen précédent, moyenne mobile du stock).
- Catégorisation (ex. regrouper catégories produits, segmentation fournisseur).
- **Séparation entraînement / test :** Veiller à la méthodologie différente selon le type de modèle utilisé (chronologique dans le cas d'une série temporelle, fractionnement 70/30 ou 80/20 sinon).

Évaluation des modèles :

- ↪ Métriques adaptées selon le type de modèle de ML entraîné
- ↪ Courbes de validation (learning curves), etc....

Interprétation des résultats : Lier les constats quantitatifs à des enjeux business (réduction des coûts, amélioration du taux de service, etc.).

Phase 5 : Formulation de recommandations opérationnelles

Synthèse des insights clés

- Mettre l'accent sur les inefficacités repérées.
- Énoncer les produits critiques ou les fournisseurs problématiques (délais trop longs, coûts trop élevés).

Actions correctives et pistes d'optimisation

- Optimiser le stock de sécurité sur certains produits.
- Négocier des SLA (Service Level Agreements) plus stricts avec des fournisseurs.
- Mettre en place un meilleur suivi des retours pour identifier les causes récurrentes.

Plan d'implémentation

- Prioriser les actionsn lliste des améliorations possibles (court terme vs moyen terme).
- Proposer une feuille de route réaliste (ressources, coûts, ordre de mise en œuvre).
- Gains potentiels (efficacité, rentabilité, satisfaction client...)

Conclusion & Perspectives

- Bilan global du projet
- Limites et pistes d'amélioration future

Phase 6 : Livrables attendus

- ↪ **Notebook Détaillé (AED, Stats, Preprocessing, Machine Learning....)**
- ↪ **Dashboard (AED et Stats, KPI,)**
- ↪ **Rapport écrit** : Un rapport, synthétique et stratégique structuré clairement, incluant :
 - Les recommandations opérationnelles précises pour optimiser les stocks et la chaîne d'approvisionnement.
 - Les priorités à mettre en œuvre à court et moyen terme, hiérarchisées selon leur impact potentiel
- ↪ **Slides de présentation orale** :
 - Mise en avant de la problématique et des insights clés.
 - Graphiques les plus parlants.
 - Recommandations et plan d'action.
 - Prévoir une démonstration rapide (si possible) des notebooks et/ou du dashboard.

!! Une présentation orale, démontrant votre capacité à structurer vos idées logiquement et à utiliser efficacement les données pour étayer votre argumentation.

Détails et conseils pratiques supplémentaires

- ↪ **Gestion du temps** : Répartition des tâches au sein du groupe, rétroplanning, Calendrier indicatif (ex. 1 vendredi pour les liaisons, le dictionnaire des variables, le preprocessing, 2 vendredis pour l'exploration et la modélisation, 1 vendredi pour la finalisation, la sythèse, la rédaction du rapport et la préparation de la soutenance).
- ↪ **Bonnes pratiques de versioning** : Utilisation de Git pour suivre l'évolution du projet en équipe (code, notebooks, rapports).
- ↪ **Visualisations riches** : utilisation de cartes (pour localiser les fournisseurs ou entrepôts), de courbes temporelles pour les séries de commandes, de heatmaps pour les corrélations, etc.
- ↪ **Eventuellement, réfléchir à la façon de “raconter” l’histoire des données** : quel est le parcours d’un produit de l’entrepôt au client, quels sont les points de friction ? (Storytelling)

Critères d'évaluation possibles

- ↳ **Qualité de la préparation / nettoyage des données** : clarté de la démarche, justification des choix.
- ↳ **Pertinence de l'analyse exploratoire** : variété et pertinence des indicateurs, bonne communication visuelle.
- ↳ **Bonne compréhension business** : alignement des questions d'analyse et de modélisation avec les besoins de l'entreprise.
- ↳ **Modélisation et évaluation** : pertinence du choix d'algorithmes, respect des bonnes pratiques (split, cross-validation, etc.), qualité des métriques utilisées.
- ↳ **Recommandations** : concrètes, chiffrées, en lien avec les observations et les résultats des modèles.
- ↳ **Présentation et rapport** : structure logique, argumentation claire, synthèse convaincante.