École Polytechnique de Sousse



Année Universitaire: 2025/2026

TP 2: Data Analysis avec Python

Filière: 5 ème année Ingénierie en Informatique

Enseignante : Dr-Ing. Nadia BOUFARES

Nombre de page : 2 Pages

Matière: Data Analytics & Business Intelligence

Cas d'étude - SmartMarket

1. Présentation de l'entreprise

SmartMarket est une entreprise de commerce électronique tunisienne créée en 2018. Elle propose une large gamme de produits électroniques et informatiques : ordinateurs portables, smartphones, tablettes, et accessoires technologiques. L'entreprise opère à la fois via son site web, ses boutiques physiques et un canal téléphonique. Depuis 2020, la direction cherche à mieux comprendre les comportements d'achat de ses clients, optimiser la performance de ses ventes et améliorer la satisfaction client.

2. Problématique

SmartMarket collecte quotidiennement un volume important de données issues de ses ventes, de ses clients, de ses fournisseurs et des enquêtes de satisfaction. Cependant, ces données sont dispersées, souvent incomplètes, incohérentes ou mal structurées. Cette situation empêche l'entreprise de disposer d'indicateurs fiables pour la prise de décision. Le service BI souhaite donc mettre en place un processus d'analyse et d'intégration des données afin de : nettoyer, transformer, modéliser et visualiser les informations clés nécessaires à la stratégie de l'entreprise.

3. Objectifs de l'étude

- 1. Identifier le chiffre d'affaires total par année et par canal de vente.
- 2. Calculer le top 10 des produits les plus vendus.
- 3. Déterminer le taux de remise moyen par catégorie de produit.
- 4. Analyser la performance des ventes par région et par période.
- 5. Détecter les transactions incohérentes ou dupliquées.
- 6. Nettoyer les informations clients (emails invalides, doublons, villes mal écrites).
- 7. Standardiser les catégories de produits et les formats de prix.
- 8. Corriger les remises hors plage (supérieures à 100% ou négatives).
- 9. Calculer le revenu réel de chaque transaction.
- 10. Identifier les clients inactifs depuis plus de 12 mois.
- 11. Mesurer la satisfaction moyenne des clients par ville et par année.
- 12. Identifier les produits avec les notes de satisfaction les plus basses.
- 13. Analyser la fiabilité des fournisseurs et leurs délais moyens de livraison.
- 14. Évaluer le pourcentage de livraisons en retard.
- 15. Comparer les ventes en ligne avec celles en boutique et téléphone.
- 16. Identifier les corrélations entre remises et satisfaction client.
- 17. Déterminer le panier moyen par client.
- 18. Créer une segmentation des clients selon leur comportement d'achat.
- 19. Construire un modèle en étoile pour le Data Warehouse (faits et dimensions).
- 20. Définir les KPI clés : chiffre d'affaires, taux de satisfaction, délai de livraison, fiabilité fournisseur.

4. Travaux attendus

Les apprenants devront :

- Analyser le jeu de données fourni (SmartMarket raw.xlsx) et identifier les anomalies.
- Nettoyer les données à l'aide de Python (pandas, numpy, etc.).
- Proposer un modèle en étoile pour le Data Warehouse.
- Identifier les transformations nécessaires à la cohérence des données.
- Définir et calculer les KPI répondant aux besoins exprimés.