Rapport d’analyse - TP Business Analytics & IA

# Phase 1 : Préparation des Données

**Aperçu des données**:

|  |  |  |  |
| --- | --- | --- | --- |
| Période | Nombre de jours | Ventes moyennes | Ecart-type |
| 2023-01-01 00:00:00 à 2024-06-30 00:00:00 | 547 | 1897€ | 872€ |

**Anomalies injectées des données**:

|  |  |
| --- | --- |
| Types d’anomalies | Valeur |
| normal | 538 |
| Stock-out | 3 |
| promotion | 3 |
| Tech-issue | 2 |
| Black-friday | 1 |

***Name: count, dtype: int64***

## Questions de Réflexion :

1. 1. Quels patterns observez-vous dans les données ?

- Tendance croissante des ventes dans le temps (base\_sales augmente).  
- Saisonnalité hebdomadaire : ventes plus fortes le week-end.  
- Saisonnalité mensuelle : pics durant certains mois comme décembre.  
- Présence de bruit aléatoire rendant les données réalistes.

1. 2. Comment identifieriez-vous manuellement les anomalies ?

- En visualisant les courbes des ventes (points anormaux).  
- En comparant les ventes au score Z (>3 ou <-3).  
- En utilisant la méthode IQR (hors de Q1-1.5×IQR ou Q3+1.5×IQR).  
- En repérant les pics ou les chutes soudaines.

# Phase 2 : Développement IA

**Performance des algorithmes:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Algorithmes** | **Précision** | **Rappel** | **F1** |
| anomaly\_iso | 0.333 | 1.000 | 0.500 |
| anomaly\_zscore | 0.333 | 0.222 | 0.267 |
| anomaly\_iqr | 0.273 | 0.333 | 0.300 |
| anomaly\_combined | 0.300 | 1.000 | 0.462 |

## Questions d’analyse:

1. 1. Quel algorithme performe le mieux ?

L’algorithme Isolation Forest a obtenu le meilleur F1-score (0.500), suivi de la méthode combinée (F1 = 0.462).

1. 2. Quels sont les avantages/inconvénients de chaque méthode ?

- Isolation Forest : très bon rappel, détecte bien les anomalies, mais faible précision (faux positifs).  
- Z-score : simple mais peu efficace dans des séries bruitées.  
- IQR : robuste mais moins sensible.  
- Méthode combinée : bon compromis entre rappel et précision.

# Phase 3 : Analytics des Anomalies

## Questions business :

1. 1. Quelles sont les causes les plus fréquentes d’anomalies ?

- 43.3% des anomalies ont lieu pendant le week-end → effet week-end probable.  
- 76.7% des anomalies surviennent durant les mois saisonniers (novembre, décembre, janvier) → forte influence saisonnière.  
- Les anomalies de type spike représentent 40%, et les drops 16.7%.

1. 2. Comment adapter la stratégie selon le type d’anomalie ?

- Pour les spikes : renforcer les campagnes marketing, prévoir du stock supplémentaire, profiter de l'élan.  
- Pour les drops : vérifier les bugs techniques, surveiller les ruptures de stock, lancer des offres pour compenser.  
- Pour les anomalies saisonnières : anticiper avec des prévisions mieux ajustées.  
- Pour les week-ends : adapter les équipes et les promotions à ces périodes.