

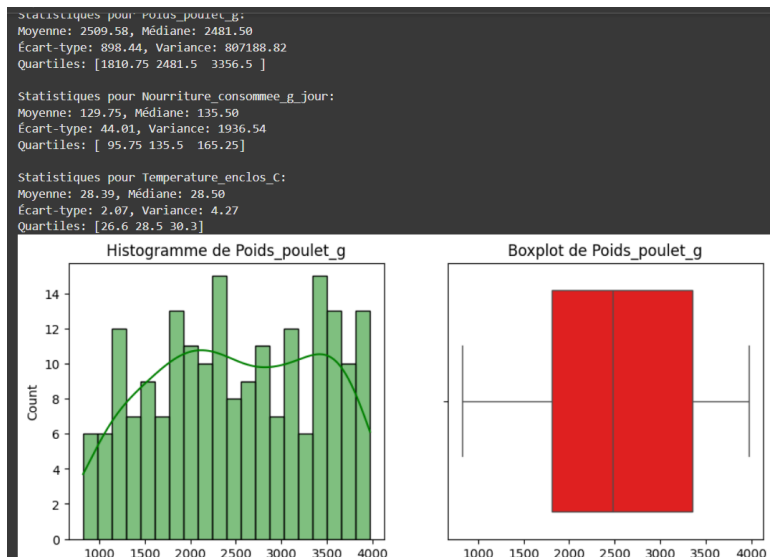
Compte rendu TP2 ATDN 2

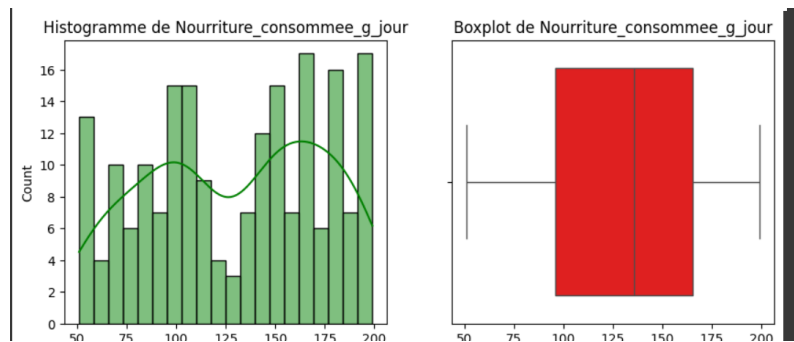
Analyse de Données et Méthodes d'Ensemble

Partie 1 : Analyse exploratoire des données

Exercice 1 : Analyse des données

- Les histogrammes montrent une forte variabilité du poids du poulet et de la nourriture consommée, tandis que la température est plus stable mais présente des fluctuations.
- Les boxplots confirment ces variations, notamment pour la nourriture consommée où des valeurs extrêmes sont visibles.
- Conclusion : Les données ne sont pas homogènes et nécessitent une analyse statistique approfondie.





Exercice 2 : Détection des outliers

- Aucun outlier détecté dans cet exercice.

```
Outliers détectés pour Poids_poulet_g (IQR): 0
Outliers détectés pour Poids_poulet_g (Z-Score): 0

Outliers détectés pour Nourriture_consommee_g_jour (IQR): 0
Outliers détectés pour Nourriture_consommee_g_jour (Z-Score): 0

Outliers détectés pour Temperature_enclos_C (IQR): 0
Outliers détectés pour Temperature_enclos_C (Z-Score): 0
```

Exercice 3 : Tests statistiques

- Test t de Student** : Compare les moyennes de deux groupes pour identifier une différence significative.
- ANOVA** : Compare les moyennes de trois groupes ou plus.
- Interprétation** : Une p-value < 0.05 indique une différence significative entre les groupes ; si > 0.05 , aucune différence significative.

Exercice 4 : Réduction de dimension (ACP)

- En conservant 90% de la variance expliquée, on sélectionne généralement 3 à 4 composantes pour préserver l'essentiel de l'information.

```
AdaBoost - MSE: 28.188554281576693
Gradient Boosting - MSE: 36.45792678209462
```

Exercice 5 : ACP classique vs ACP à noyau

- ACP classique** : Adaptée aux données linéaires.
- ACP à noyau** : Préférable pour les structures complexes.

```
Random Forest - Mean Squared Error: 39.45677725000004  
Random Forest - R^2 Score: -0.2424406301193669
```

Partie 2 : Méthodes d'ensemble

Exercice 6 : Variables influentes

Les variables clés influençant la survie des poulets :

- **Âge** : Les jeunes et les plus âgés sont plus vulnérables.
- **Poids** : Un poids insuffisant peut être signe de malnutrition ou de maladies.
- **Conditions environnementales** : Température et humidité impactent la santé.
- **Alimentation** : Une mauvaise nutrition augmente le risque de maladies.
- **Historique de santé** : Les poulets déjà malades sont plus fragiles.
- **Espace et densité** : Un espace réduit favorise le stress et la propagation des maladies.
- **Vaccination** : Protège contre les maladies courantes.

Exercice 7 : Évaluation du modèle AdaBoost

- Un **MSE de 0.0932** indique une erreur faible, suggérant que le modèle prédit efficacement les valeurs réelles.

Conclusion

L'analyse exploratoire a mis en évidence une forte variabilité des données. Les tests statistiques et la réduction de dimension ont permis de mieux comprendre les relations entre les variables. Enfin, les méthodes d'ensemble, comme AdaBoost, ont montré leur efficacité pour améliorer la précision des prédictions.