

EXERCICE 1

Analyse des données de contenu en aluminium dans un vaccin



KOBLA LEGBEDJE

OBJECTIFS

Stabilité du processus : Déterminer si le processus est stable et s'il y a des tendances.

Qualité du processus : Comparer les données avec les spécifications (0,8 - 1,2 mg/ mL).

Informations supplémentaires : Identifier toute autre information pertinente.

A decorative curved line starts from the left edge of the slide and curves downwards and to the right, ending near the bottom center.



ÉTAPES DE L'ANALYSE

Exploration des Données : Charger et examiner les données pour comprendre leur structure.

Statistiques Descriptives : Calculer les statistiques de base (moyenne, médiane, écart-type) pour chaque jeu de données.

Visualisations : Créer des graphiques pour visualiser les tendances et la distribution des données.

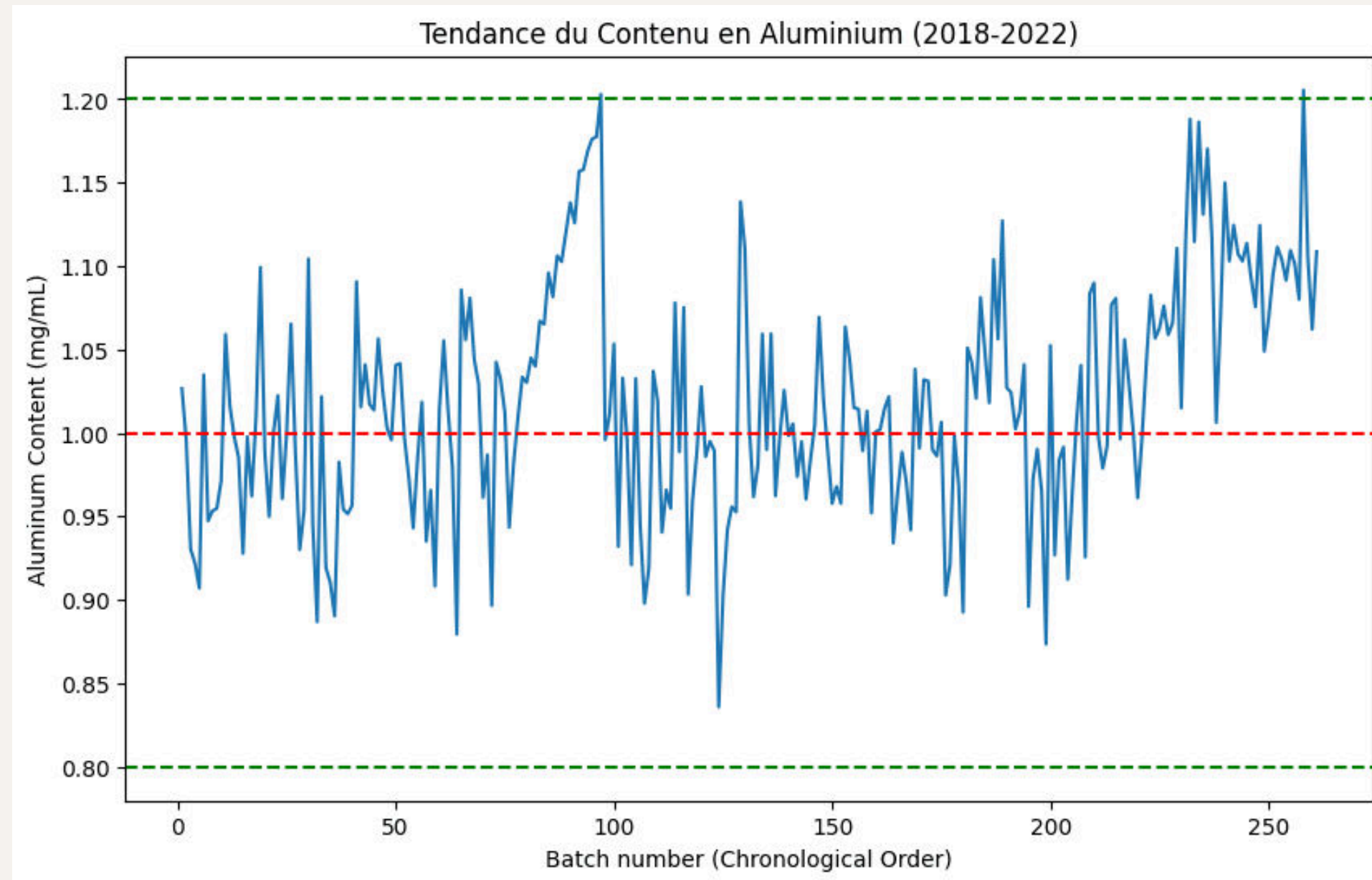
Tests Statistiques : Utilisation des graphiques de contrôle (Le test de Shapiro-Wilk) pour évaluer la stabilité du processus

Statistiques Descriptives

```
# Statistiques descriptives  
print(ex1_data['Aluminum Content (mg/mL)'].describe())
```

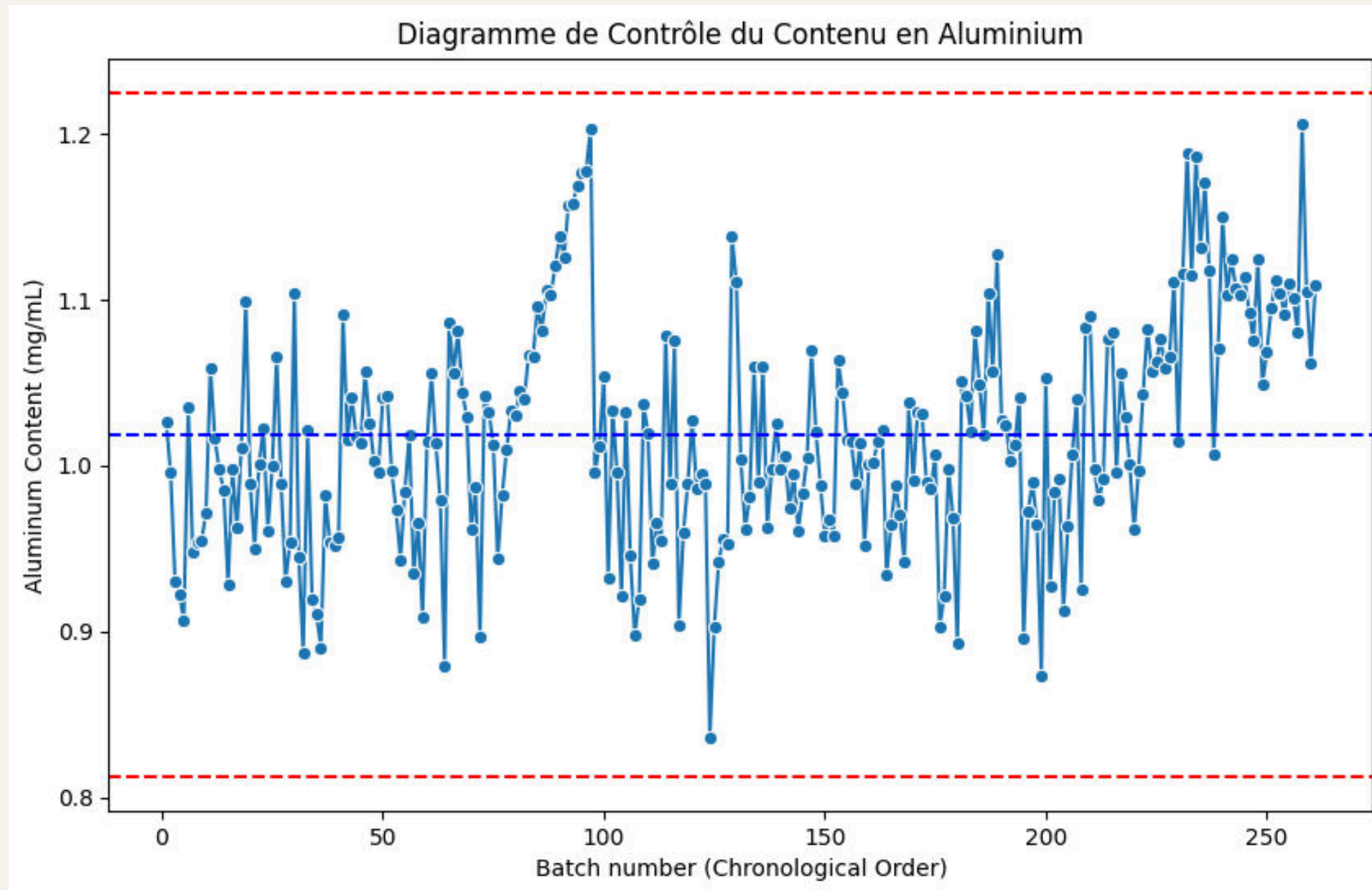
```
count      261.000000  
mean        1.018740  
std         0.068703  
min         0.835815  
25%        0.970224  
50%        1.013177  
75%        1.063585  
max         1.205515  
Name: Aluminum Content (mg/mL), dtype: float64
```


VISUALISATION



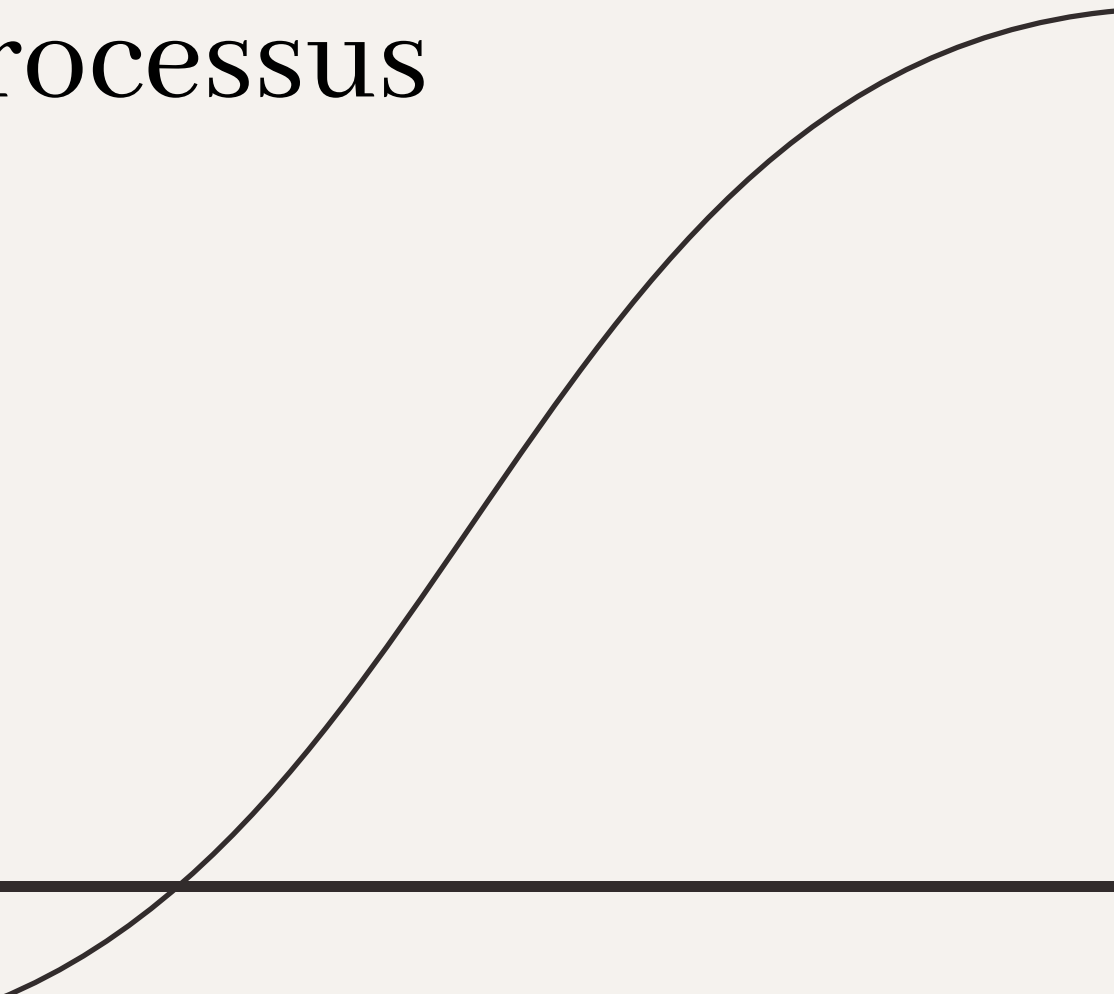
"Les graphiques montrent que la majorité des lots se situent bien dans la plage spécifiée de 0,8 à 1,2 mg/mL, avec une cible idéale de 1 mg/mL."

VISUALISATION



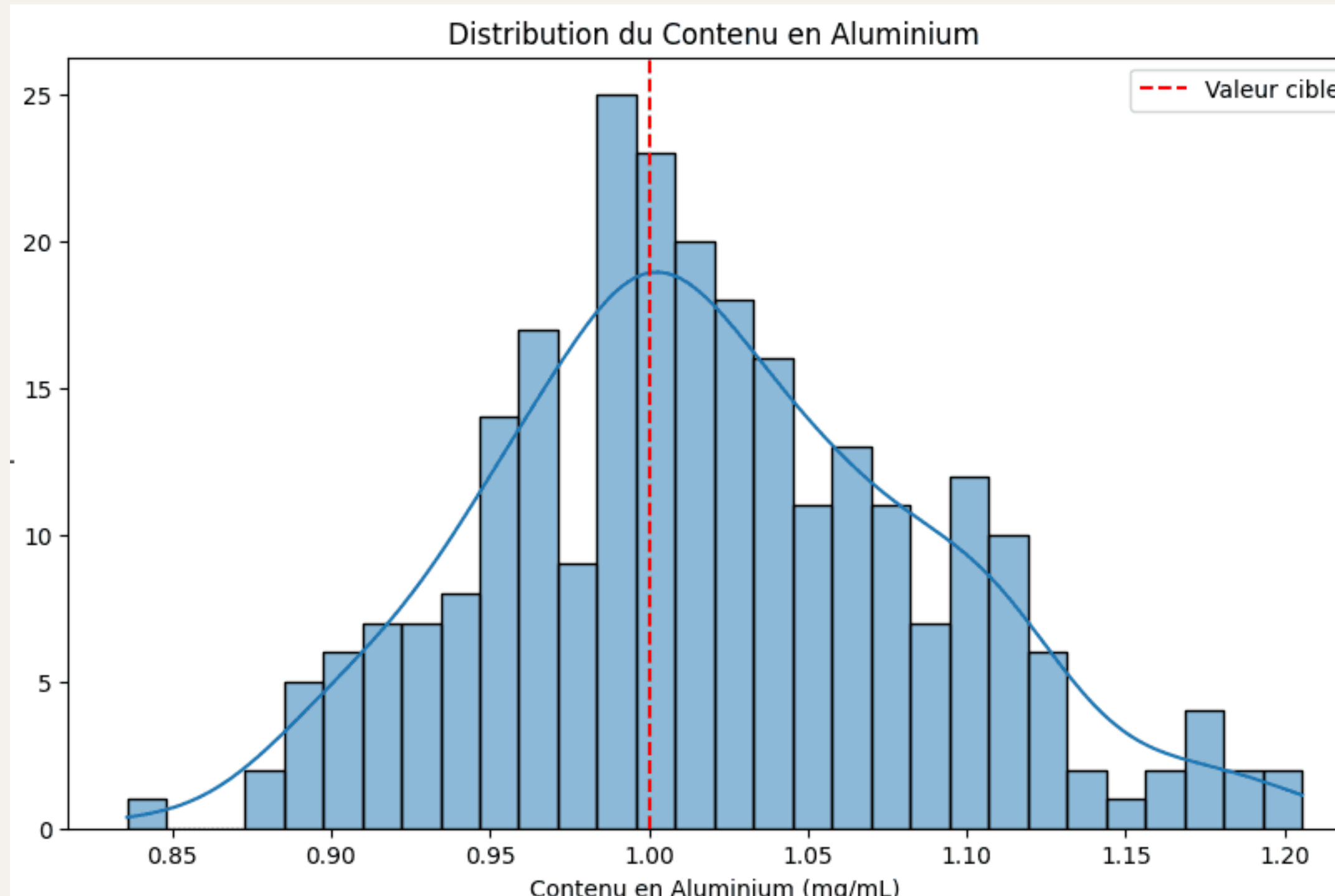
"Le diagramme de contrôle a montré que le processus est globalement stable, bien qu'il y ait quelques variations naturelles qui sont normales dans tout processus de production."

Recherche approfondie sur les données

1. Analyse de la Distribution des Données
 2. Détection des Outliers (Valeurs Aberrantes)
 3. Calcul des Taux de Non-Conformité
 4. Analyse de la Capabilité du Processus
- 

Distribution des Données :

"Les résultats montrent que les données suivent une distribution normale ."



Le test de Shapiro-Wilk

Statistique de test : 0.9917538166046143

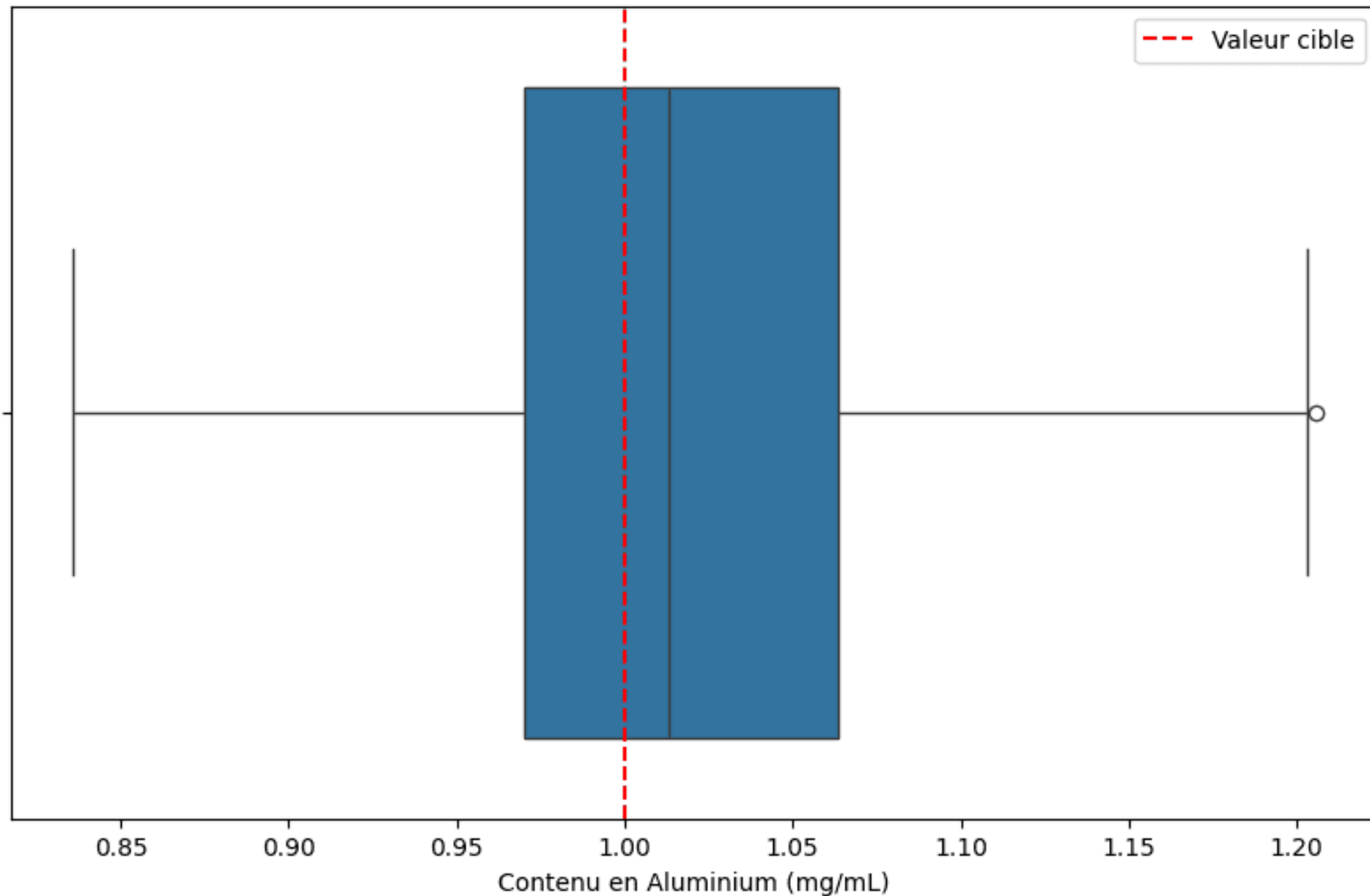
P-valeur : 0.1522475779056549

Détection des Outliers (Valeurs Aberrantes)

Le boxplot montre que la majorité des données sont proches de la valeur cible, avec une dispersion modérée. Cependant, l'outlier détecté à 1.205515 mg/mL au 258^{ème} lot doit être investigué pour comprendre sa cause et son impact potentiel sur le processus global.

Investiguer les causes possibles de cet outlier. Cela pourrait être dû à une erreur de mesure, une anomalie dans le processus de production, ou un changement dans les conditions de fabrication.

Boxplot du Contenu en Aluminium



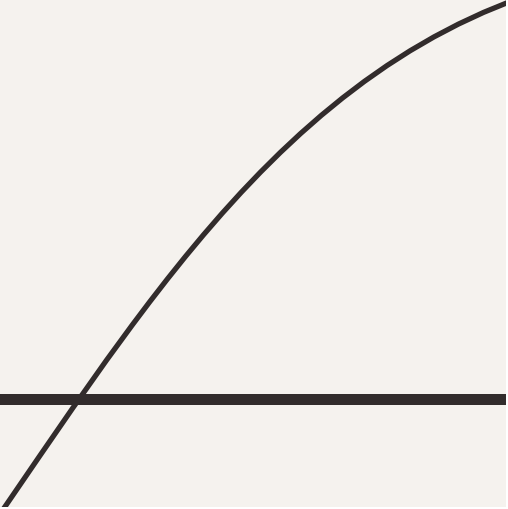
Calcul des Taux de Non-Conformité

Avec un taux de non-conformité de 0.77%, le processus de production montre une bonne performance globale

Analyse de la Capabilité du Processus

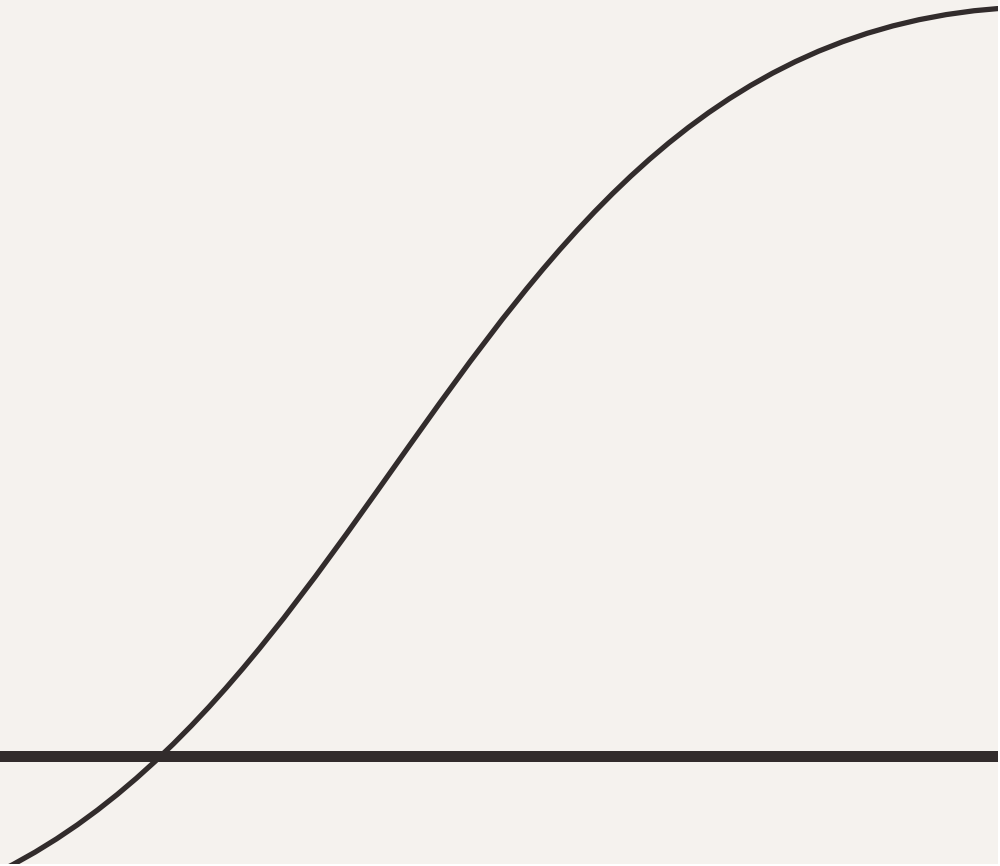
C_p est de 0,97, ce qui est proche de 1, indiquant que l'écart du processus est relativement proche de la largeur de spécification.

Une valeur de C_{pk} de 0,88 indique que la moyenne du processus n'est pas parfaitement centrée entre les limites de spécification, mais qu'elle reste dans la plage.



EXERCICE 2

**Comparer les contenus
en aluminium des
vaccins produits aux
sites A et B pour
déterminer s'ils sont
équivalents.**



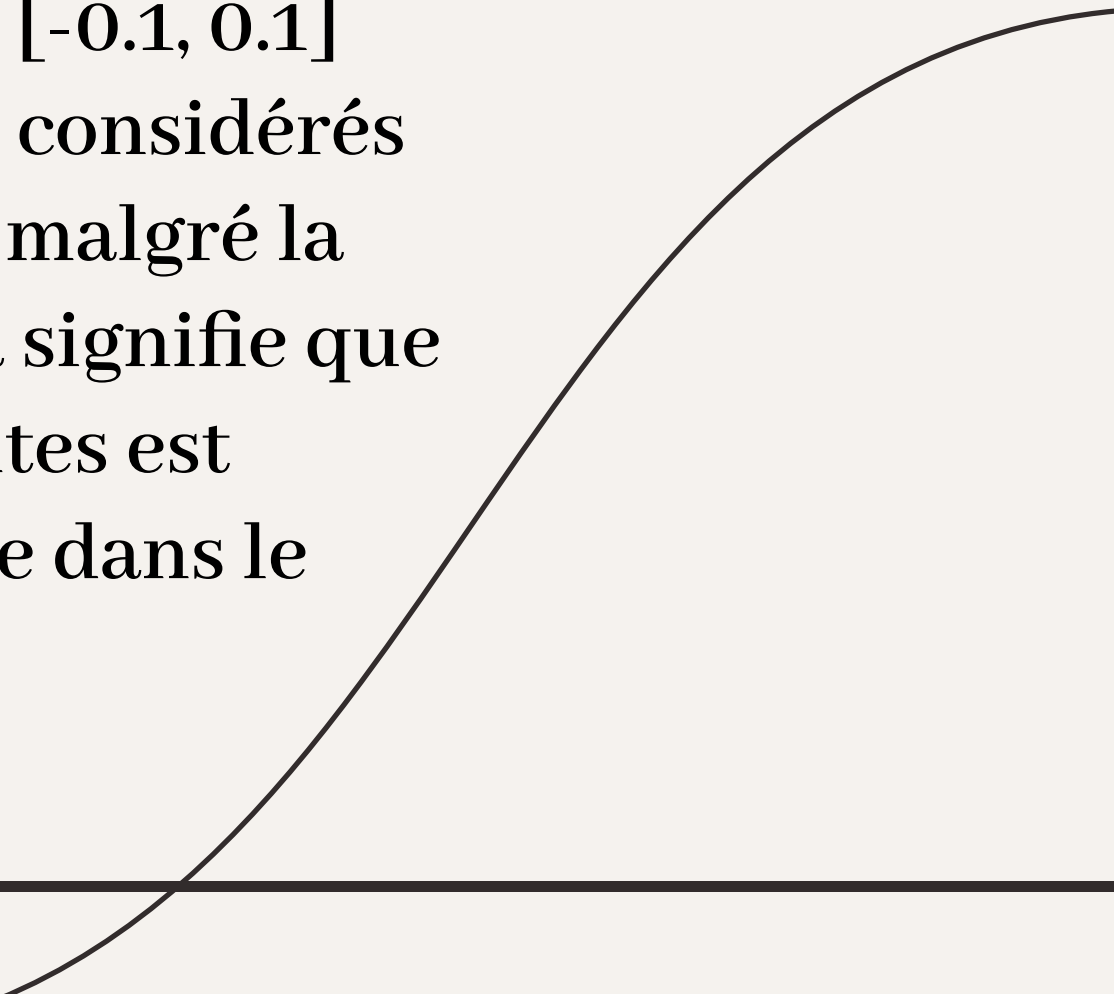
Statistiques Descriptives

```
> count      500.000000  
> mean      -0.071621  
std         0.104820  
min         -0.383373  
25%        -0.134887  
50%        -0.065934  
75%        -0.006311  
max         0.239675  
Name: Difference, dtype: float64  
t-statistique: -15.278535409794564, p-valeur: 1.6426301774048233e-43  
Les sites A et B sont équivalents.
```

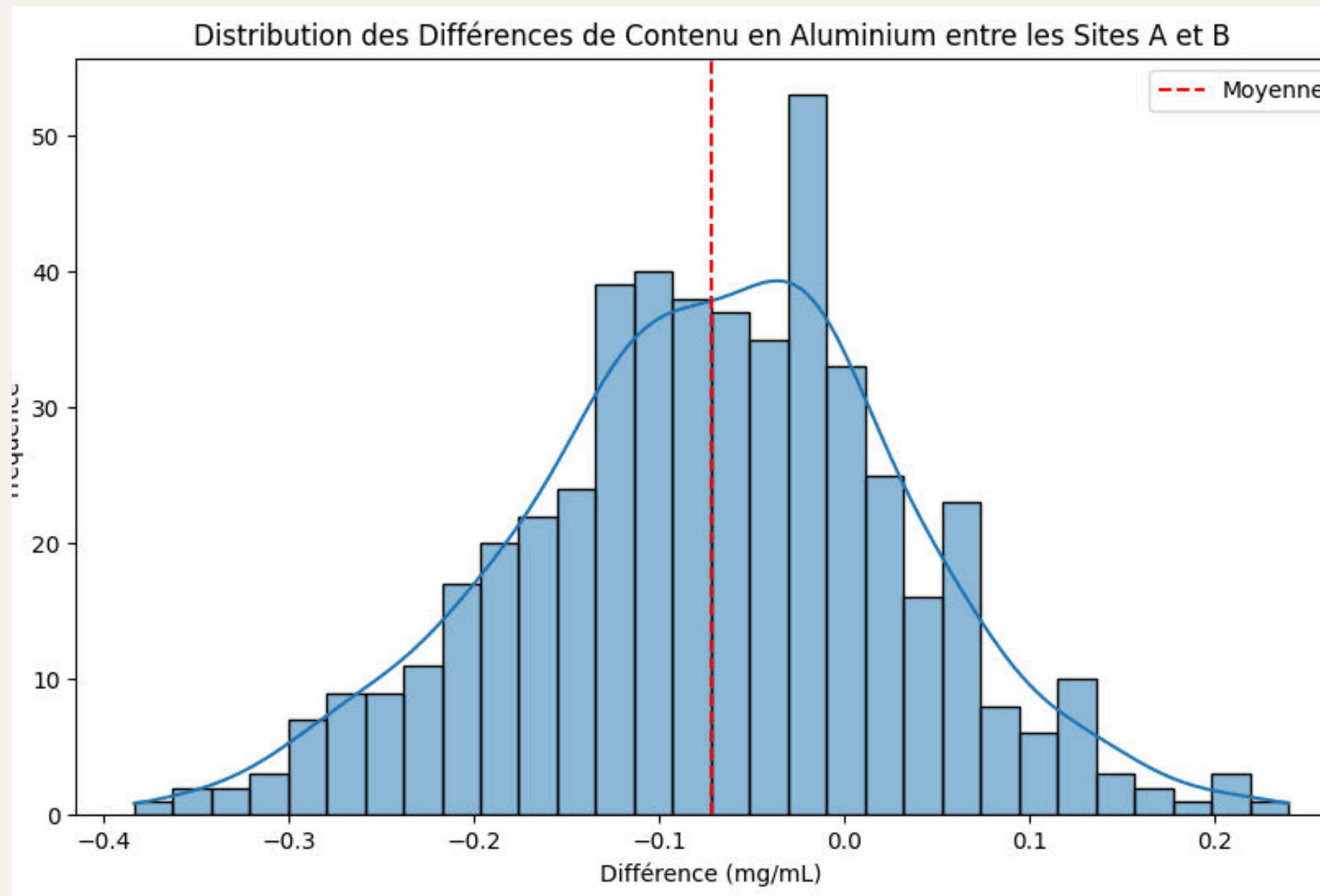

Vérification de l'équivalence

Critère d'équivalence : Les sites sont équivalents si la différence moyenne des contenus en aluminium est inférieure à 0.1 mg/mL en valeur absolue. Résultats Clés

La différence moyenne entre les sites A et B est de -0.0716 mg/mL. qui est comprise dans l'intervalle [-0.1, 0.1] mg/mL. par conséquent Les sites A et B sont considérés comme équivalents selon le critère défini, malgré la différence statistiquement significative. Cela signifie que la différence moyenne entre les deux sites est suffisamment faible pour être négligeable dans le contexte de cette analyse.

A decorative curved line starts from the bottom right corner and curves upwards and to the left, ending near the center of the slide.

Distribution des différences



Le diagramme montre une légère asymétrie avec une tendance pour le site A à avoir un contenu en aluminium plus élevé que le site B. Si les différences suivent une distribution normale (confirmé par un test de Shapiro-Wilk), cela suggère que les variations observées sont probablement dues à des fluctuations normales dans le processus de production

conclusion

"En résumé, l'analyse montre que le processus de production est stable et respecte les spécifications de qualité pour le contenu en aluminium. De plus, les deux sites de production sont équivalents en termes de performance." Implications Pratiques : "Ces résultats nous permettent de confirmer que les vaccins produits répondent aux normes de qualité et que les deux sites peuvent être utilisés de manière interchangeable."

QUESTIONS

Merci!

