

# Détectez des faux billets avec Python

05 juillet 2024

Mohamad ALI

## Zone euro

- Euro
- Euros hors l'UE



Organisation nationale de  
lutte contre le faux-  
monnayage (ONCFM)

Created with mapchart.net

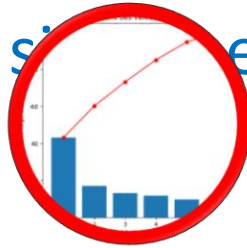


**MISSION** : proposer une analyse des groupements de pays que l'on peut cibler pour exporter nos poulets.

1. Descriptions des données et Analyse exploratoire



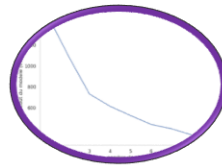
2. régression linéaire simple et multiple



3. Réaliser une **ACP**

4. Régression logistique

5. la méthode des **k-means**

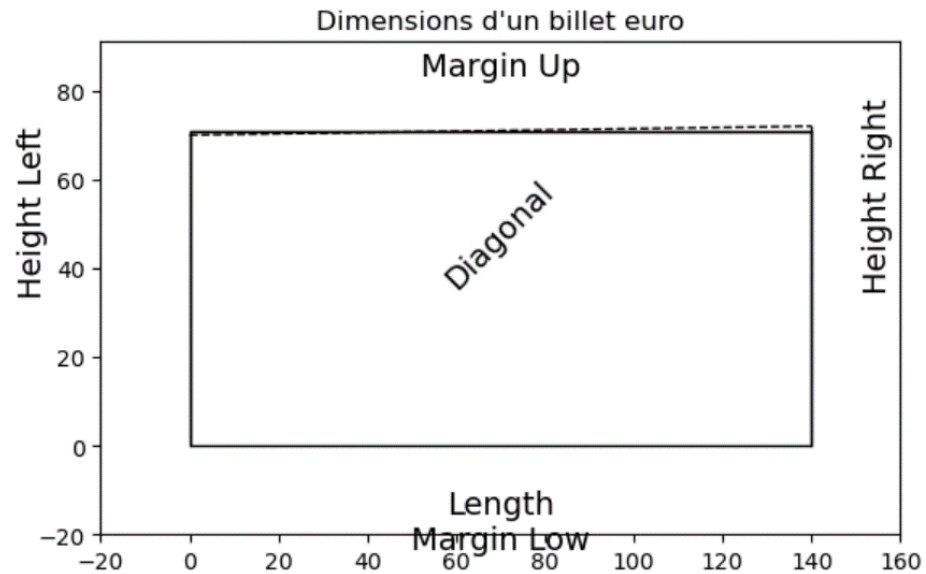


6. Comparaison entre les 2 Models

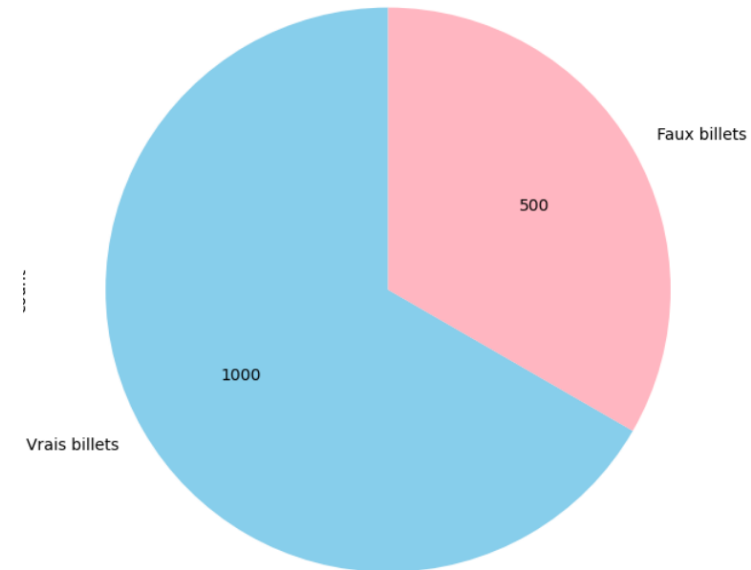
 ONCFM

# DESCRIPTIONS DES DONNÉES

## Dimensions géométriques



## Répartition des vrais et faux billets



NA

```
is_genuine      0
diagonal        0
height_left     0
height_right    0
margin_low      37
margin_up       0
length          0
```

# Matrice des Corrélations Linéaires de Pearson

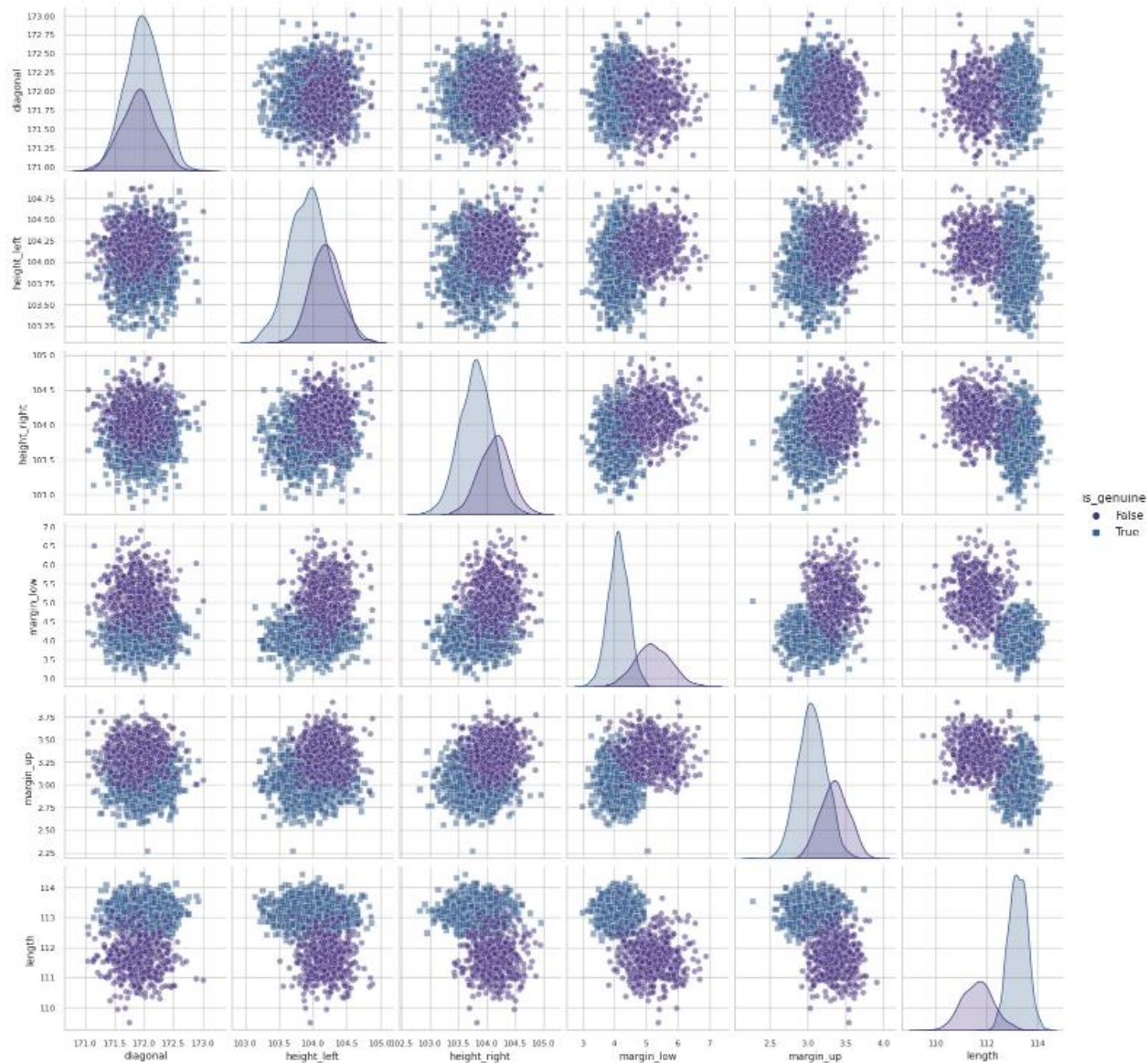


	is_genuine	diagonal	height_left	height_right	margin_low	margin_up	length
is_genuine	1.000000	0.132756	-0.379833	-0.485092	-0.786177	-0.606262	0.849285
diagonal	0.132756	1.000000	0.019472	-0.024492	-0.110639	-0.055649	0.097587
height_left	-0.379833	0.019472	1.000000	0.242279	0.308724	0.246522	-0.320863
height_right	-0.485092	-0.024492	0.242279	1.000000	0.390948	0.307005	-0.401751
margin_low	-0.786177	-0.110639	0.308724	0.390948	1.000000	0.429600	-0.668445
margin_up	-0.606262	-0.055649	0.246522	0.307005	0.429600	1.000000	-0.520575
length	0.849285	0.097587	-0.320863	-0.401751	-0.668445	-0.520575	1.000000

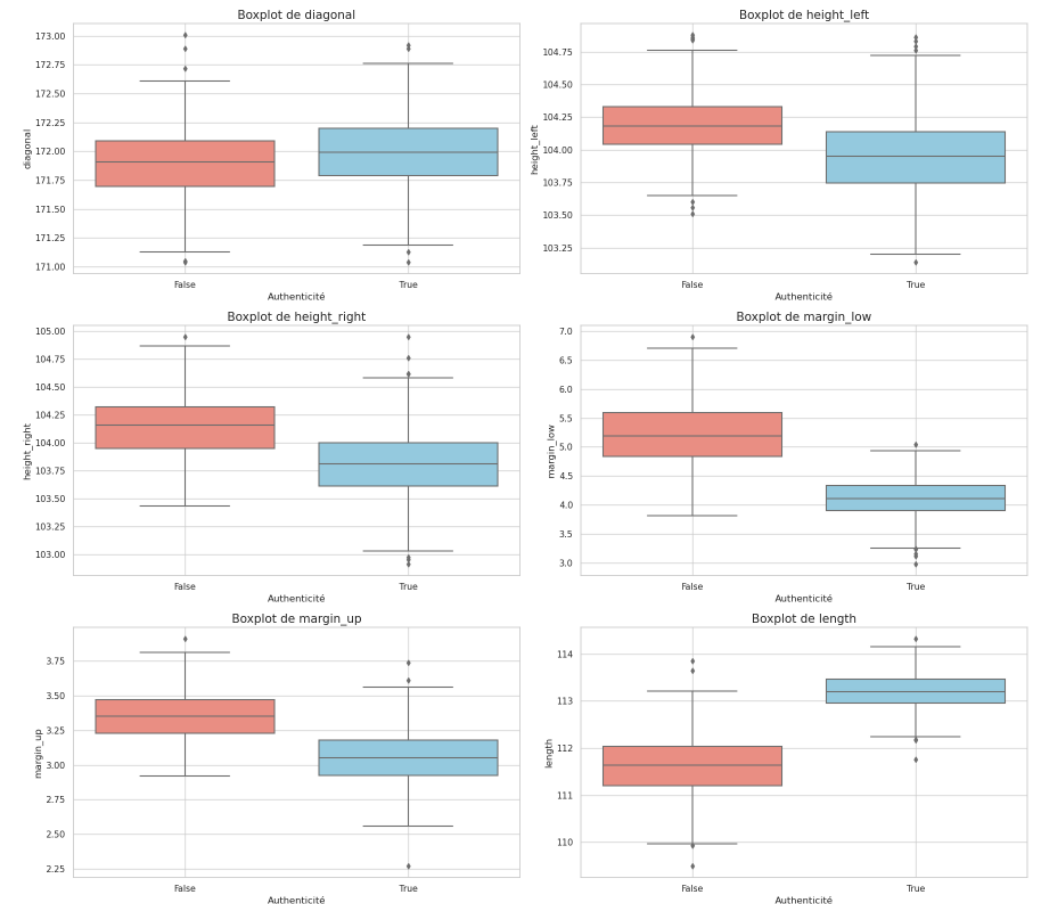


# Visualisation des données numériques

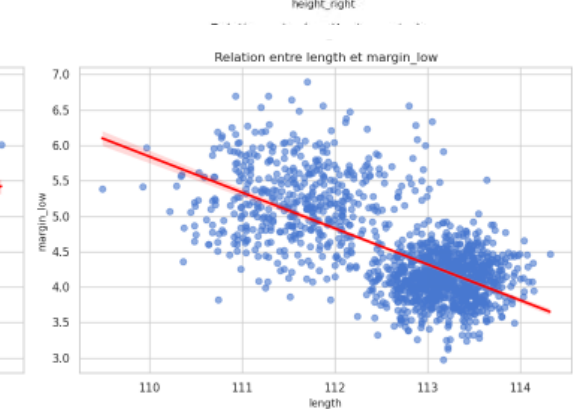
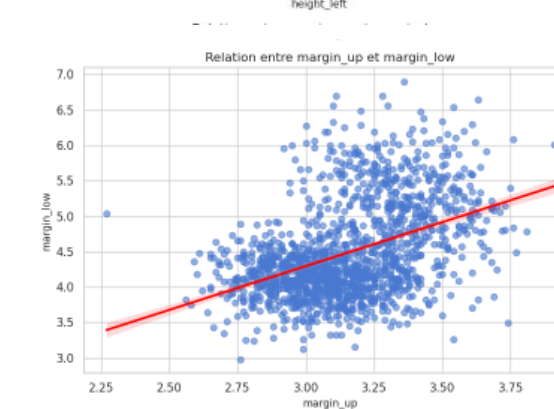
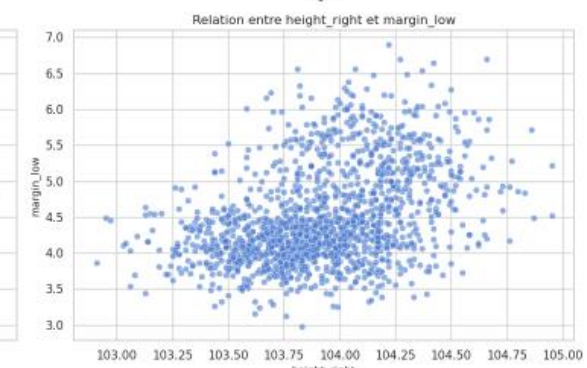
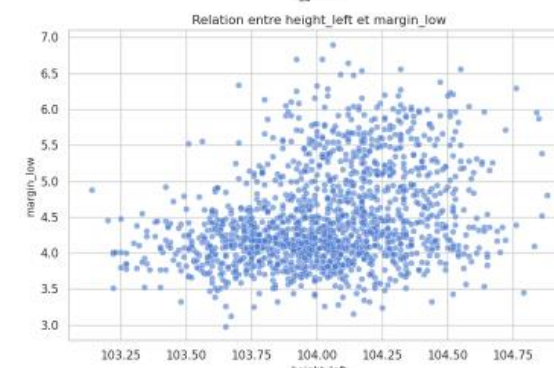
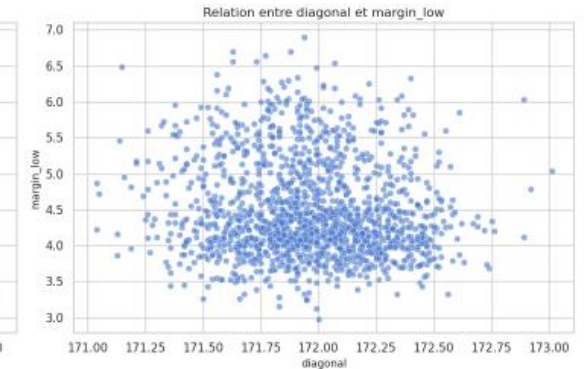
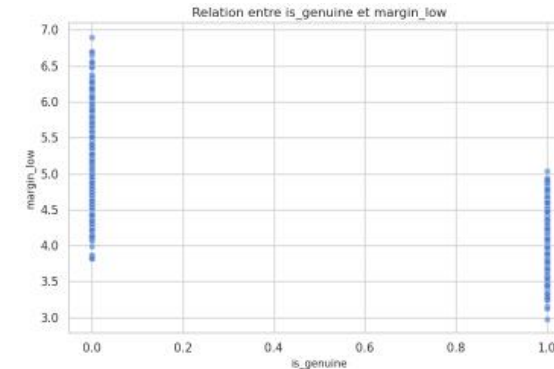
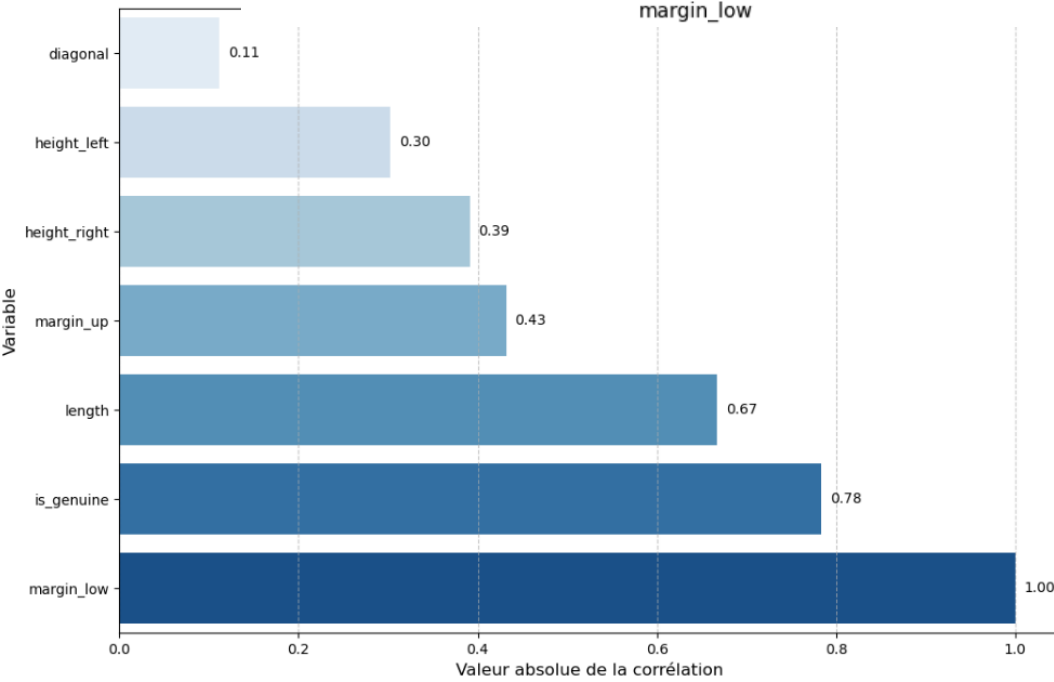
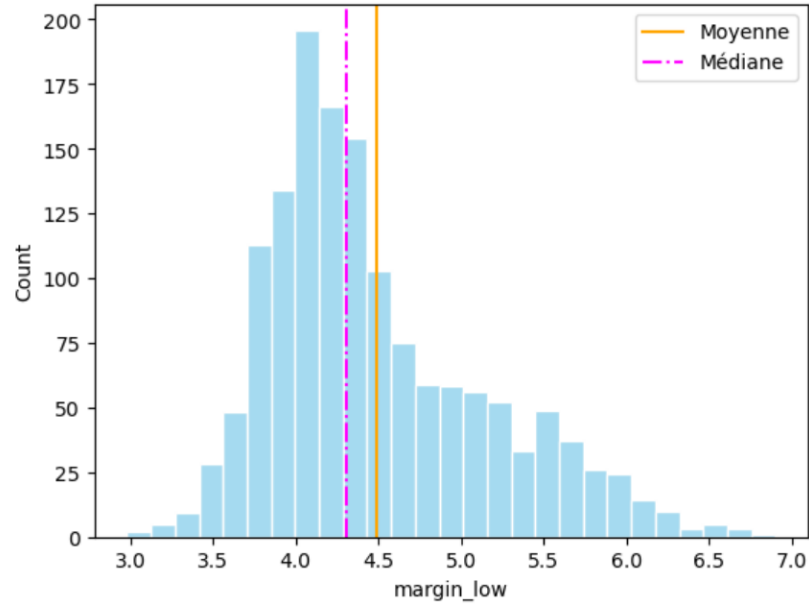
Pairplot des données numériques avec indication de l'authenticité

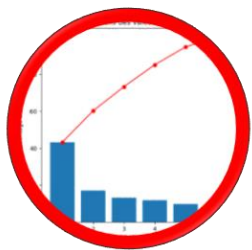


## Analyse is\_genuine en fonction des différentes variables

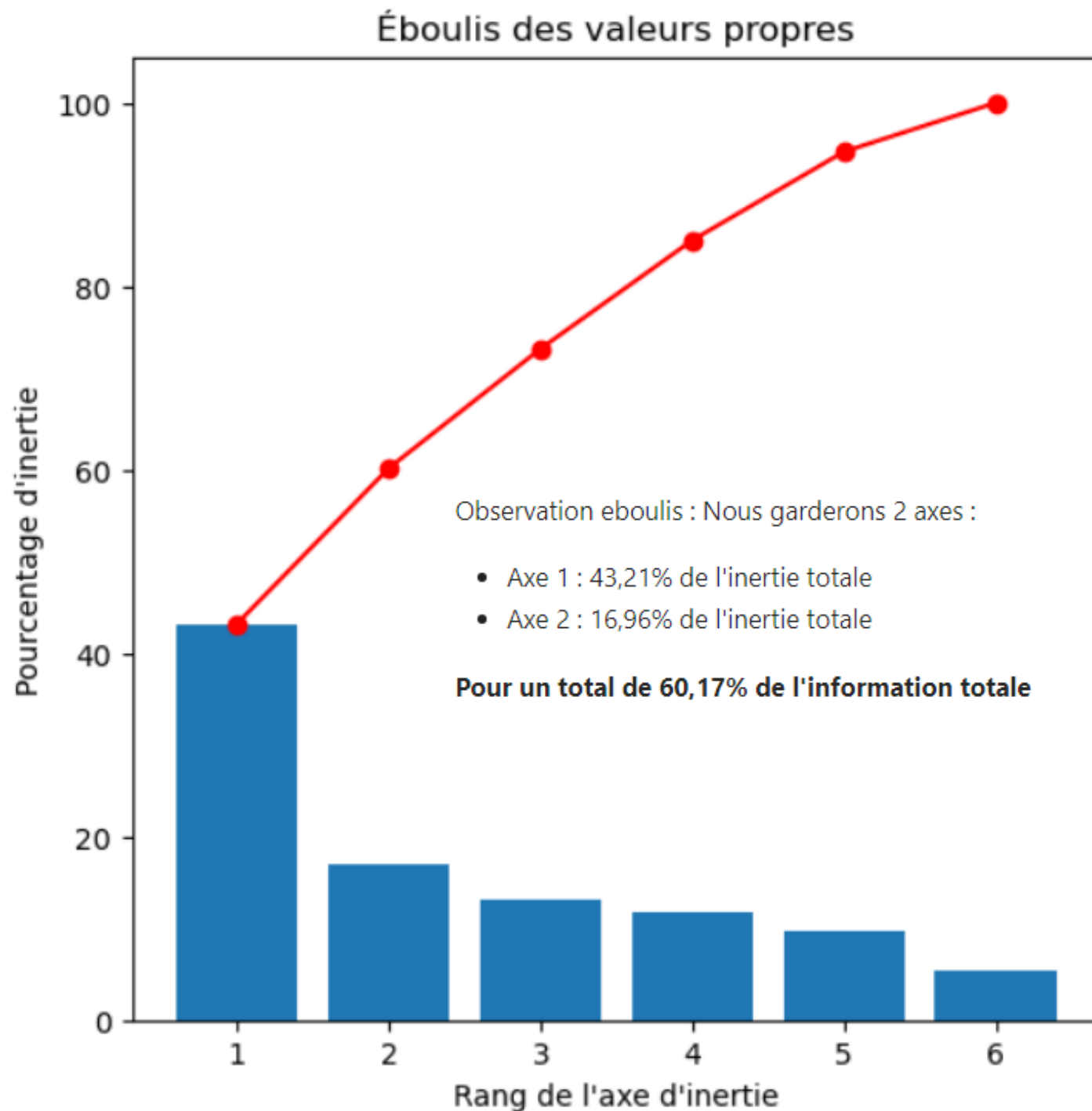


# Liens entre is\_genuine et les autres Variables



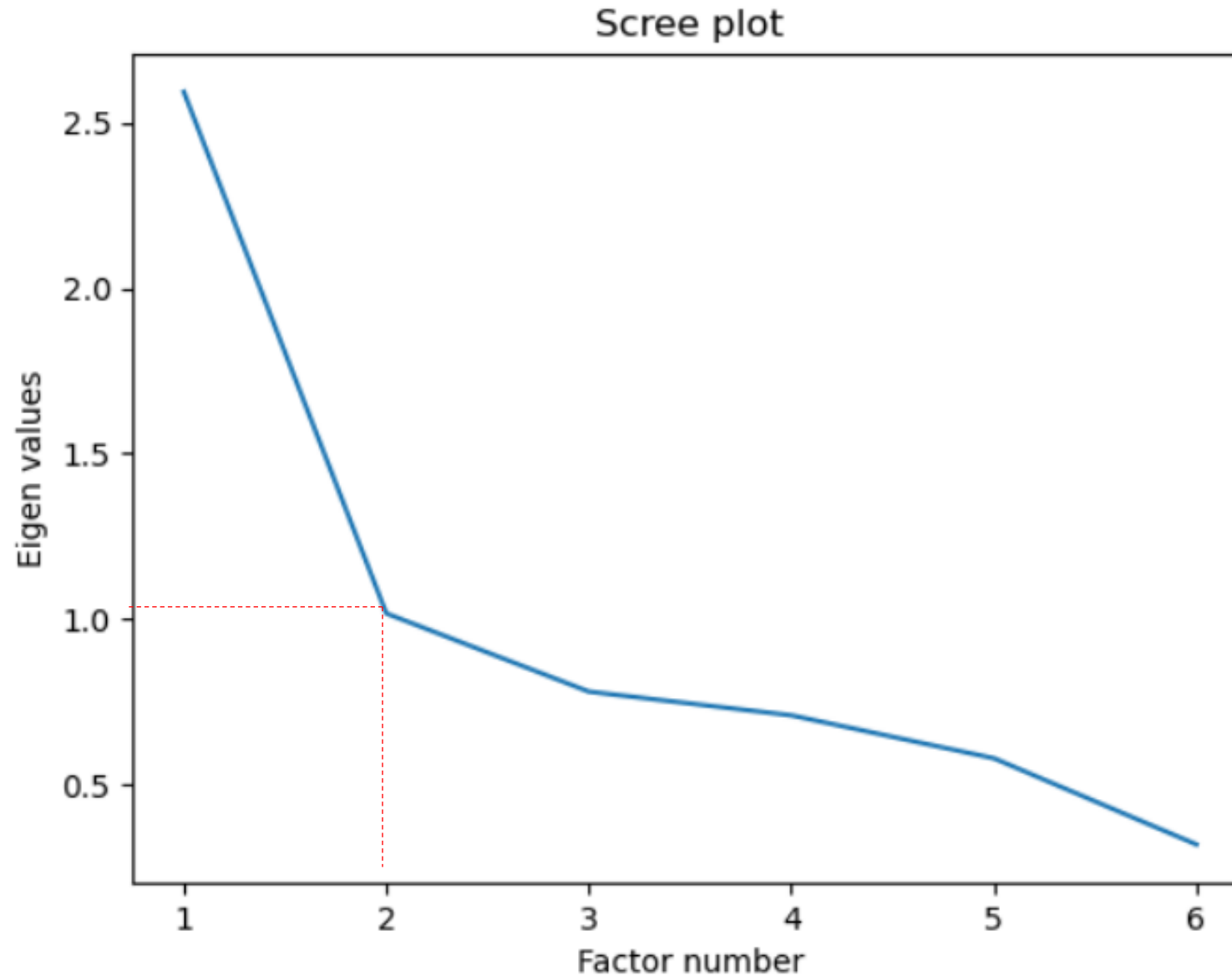
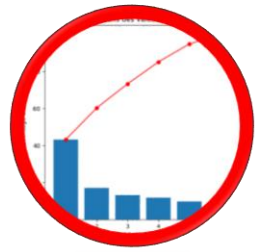


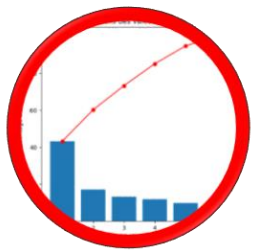
# ANALYSE EN COMPOSANTES PRINCIPALES (ACP)



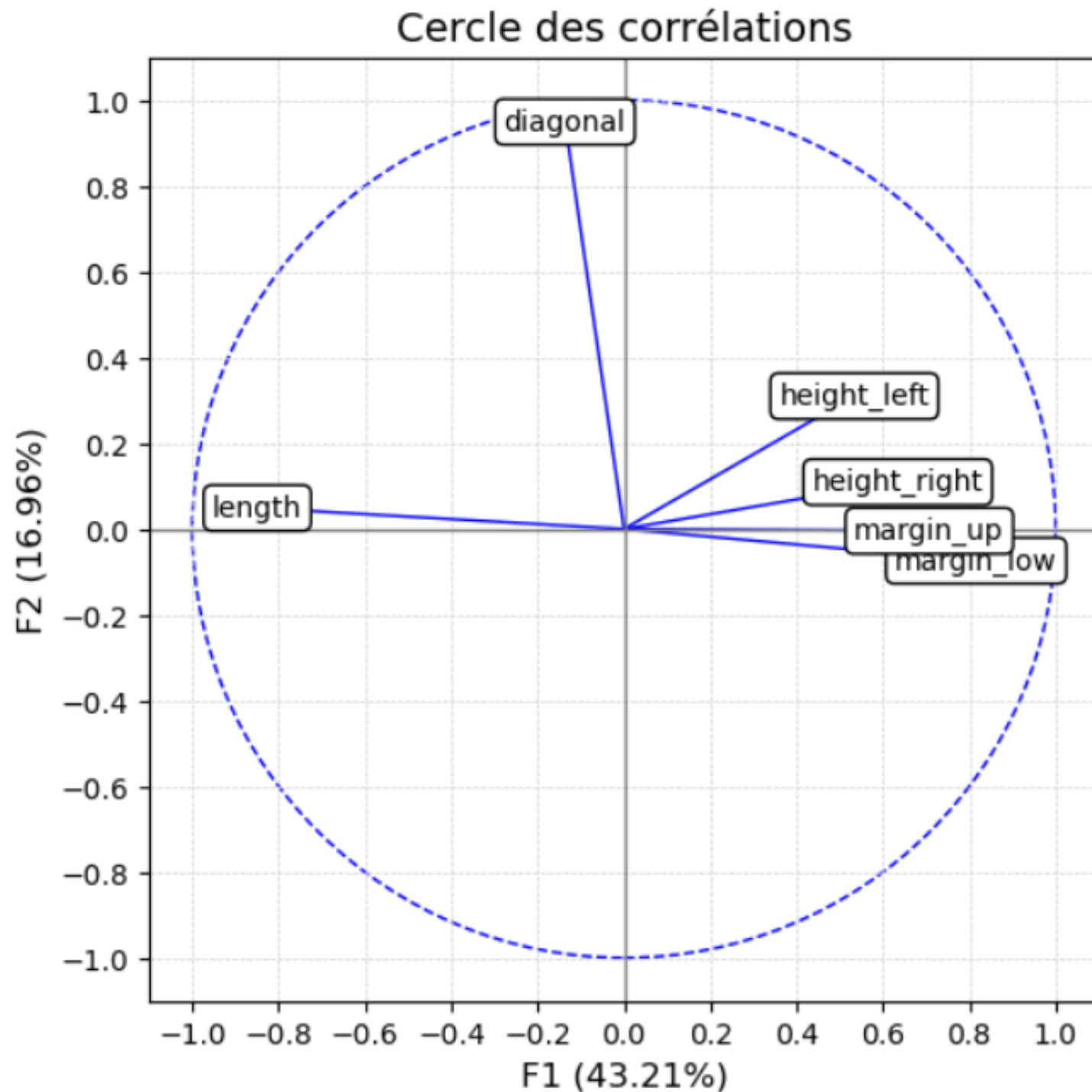


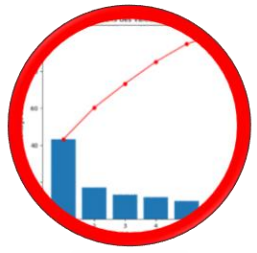
# Affichage de la Variance Cumulée Expliquée



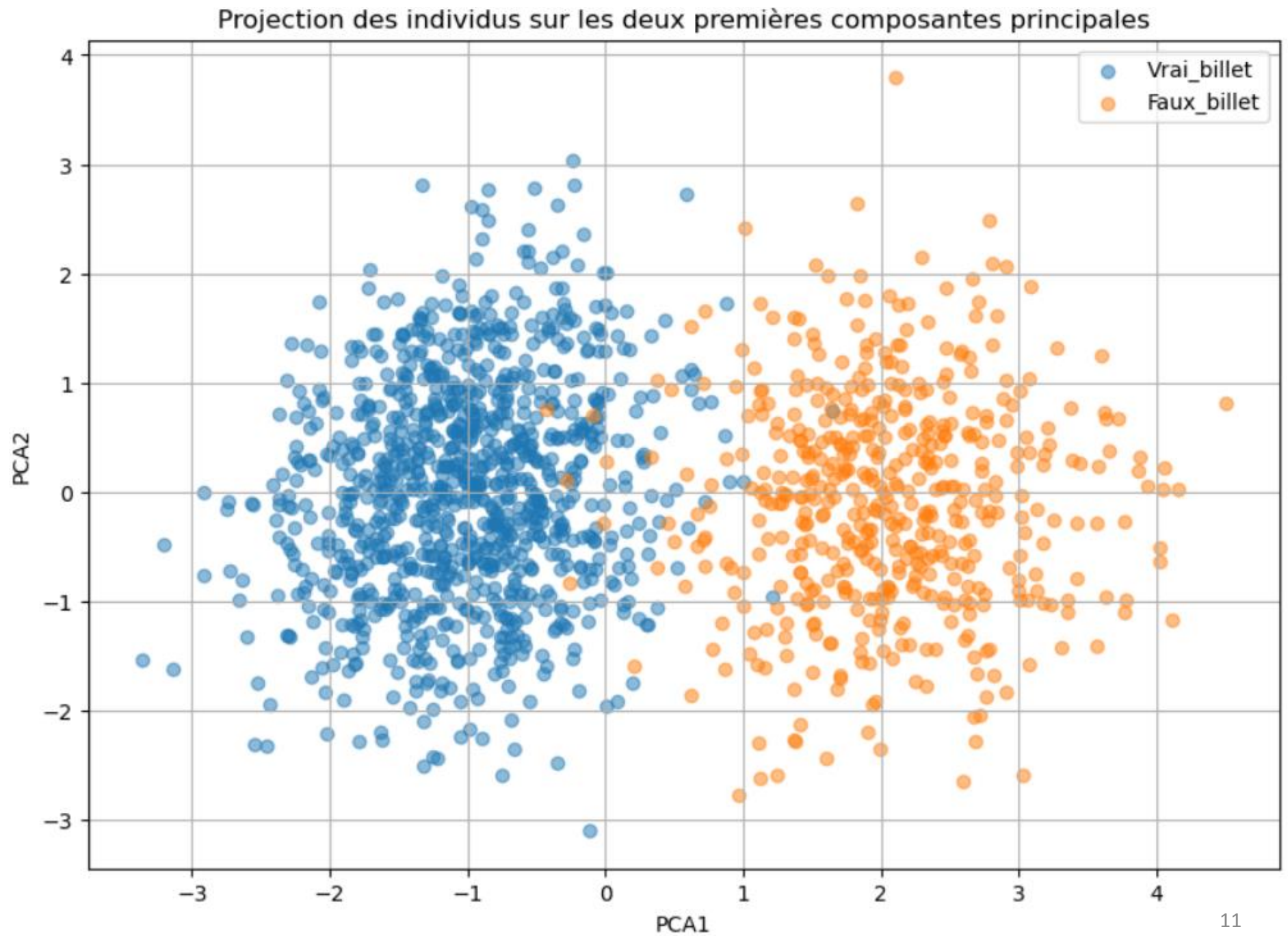


# CERCLE DES CORRÉLATIONS

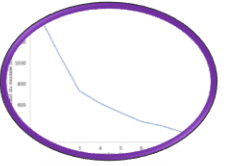




## Projection d'individus

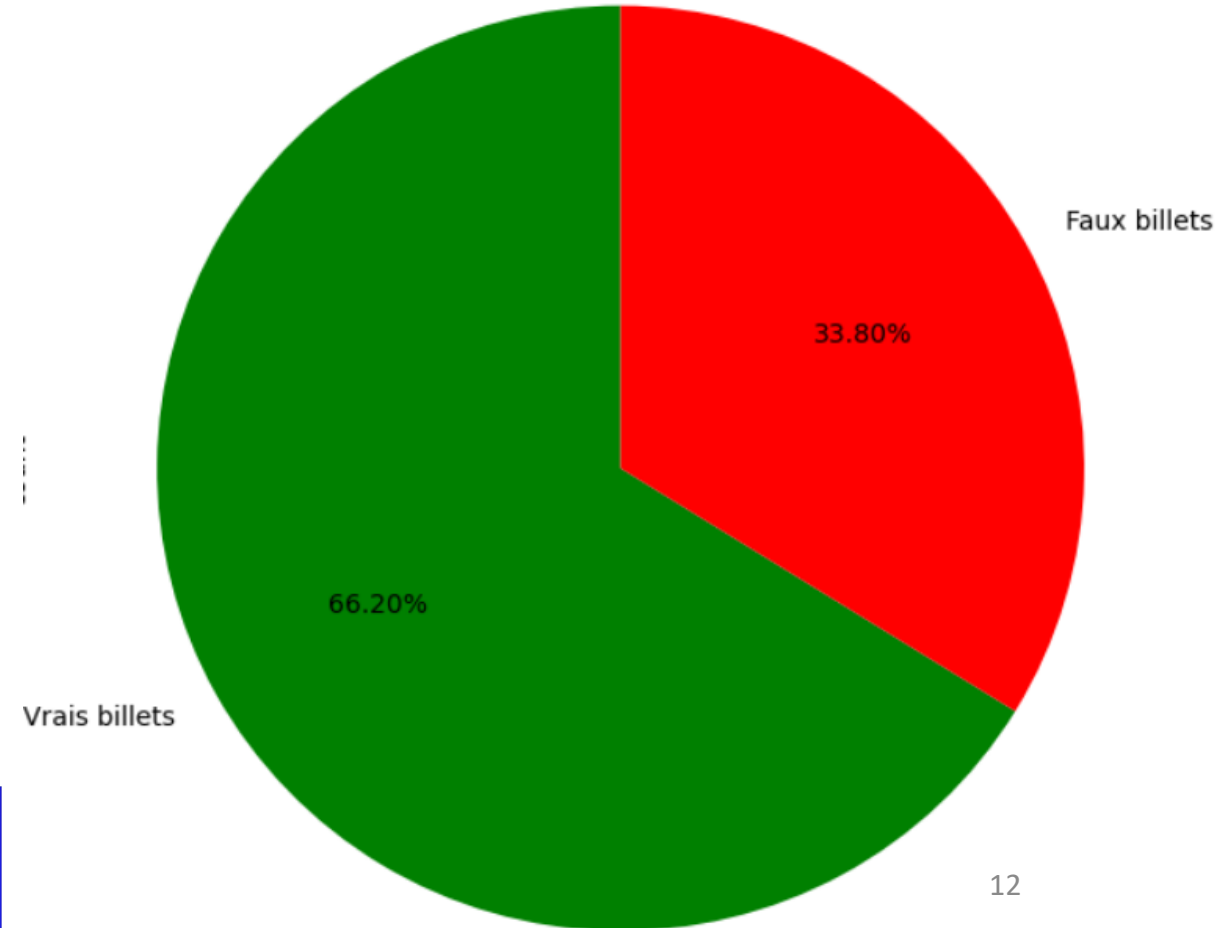
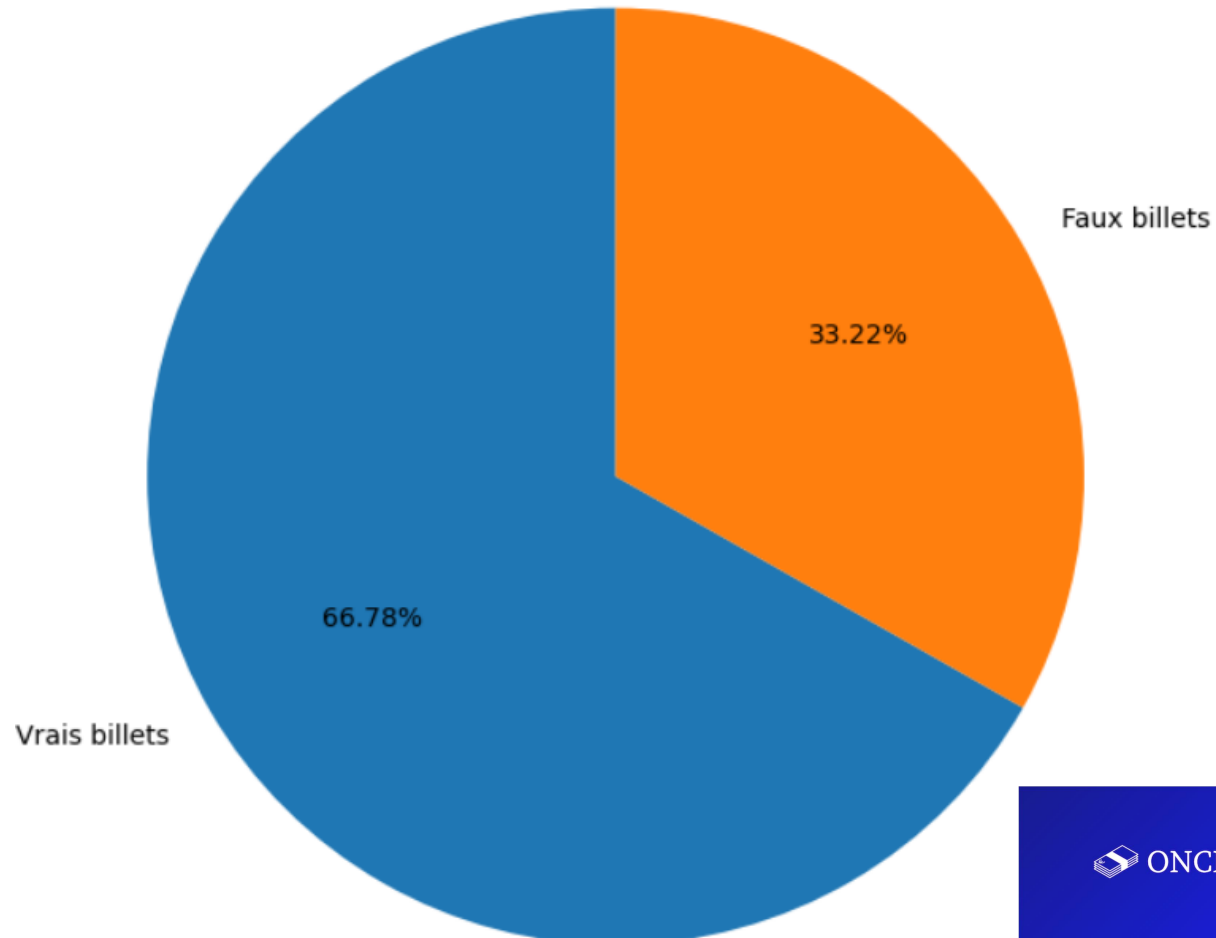


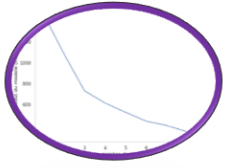
# Diviser les données d'entraînement et de test



Répartition des vrais et faux billets dans train

Répartition des vrais et faux billets dans test

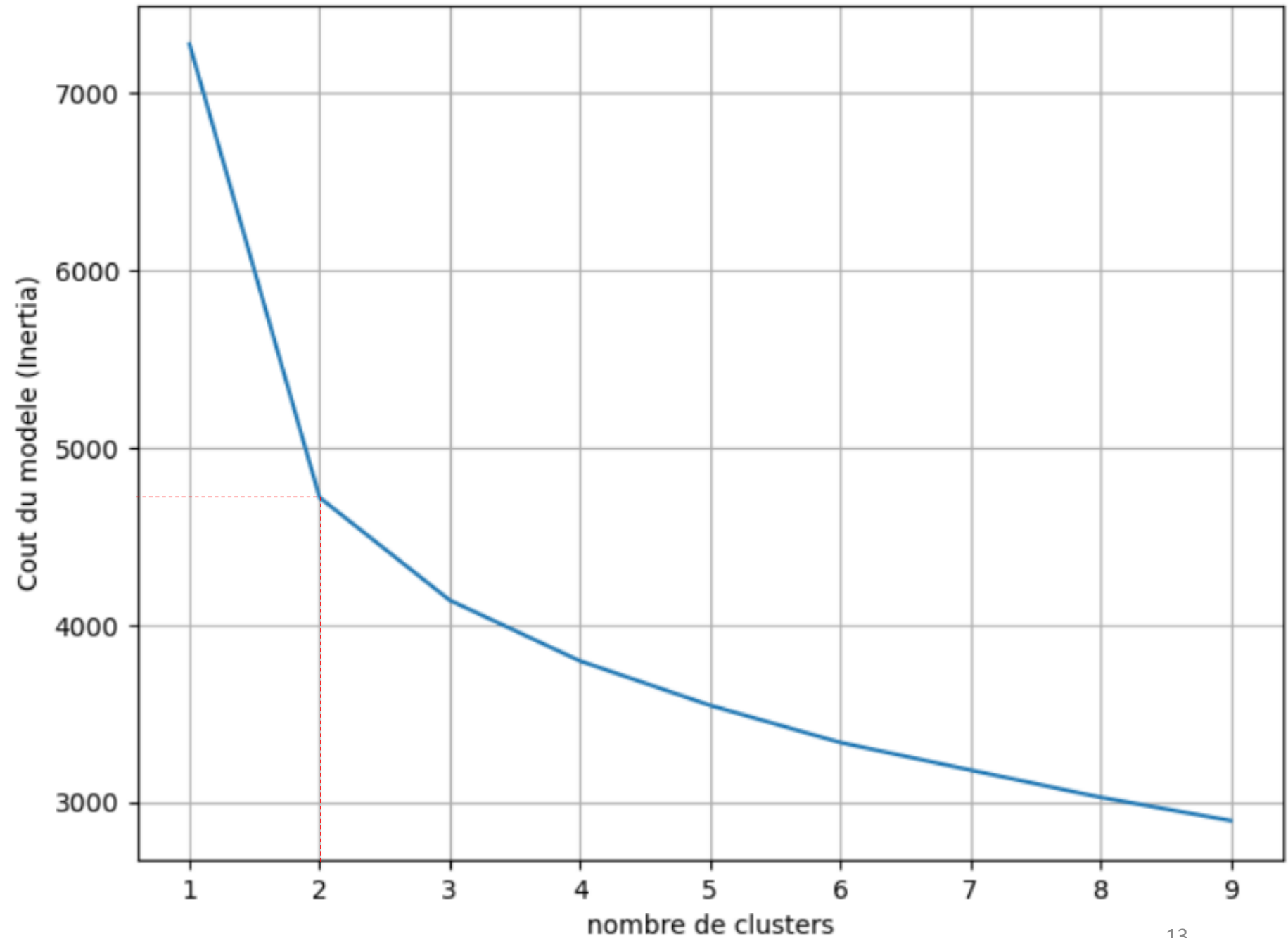




# K-MEANS

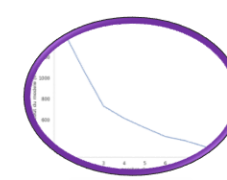
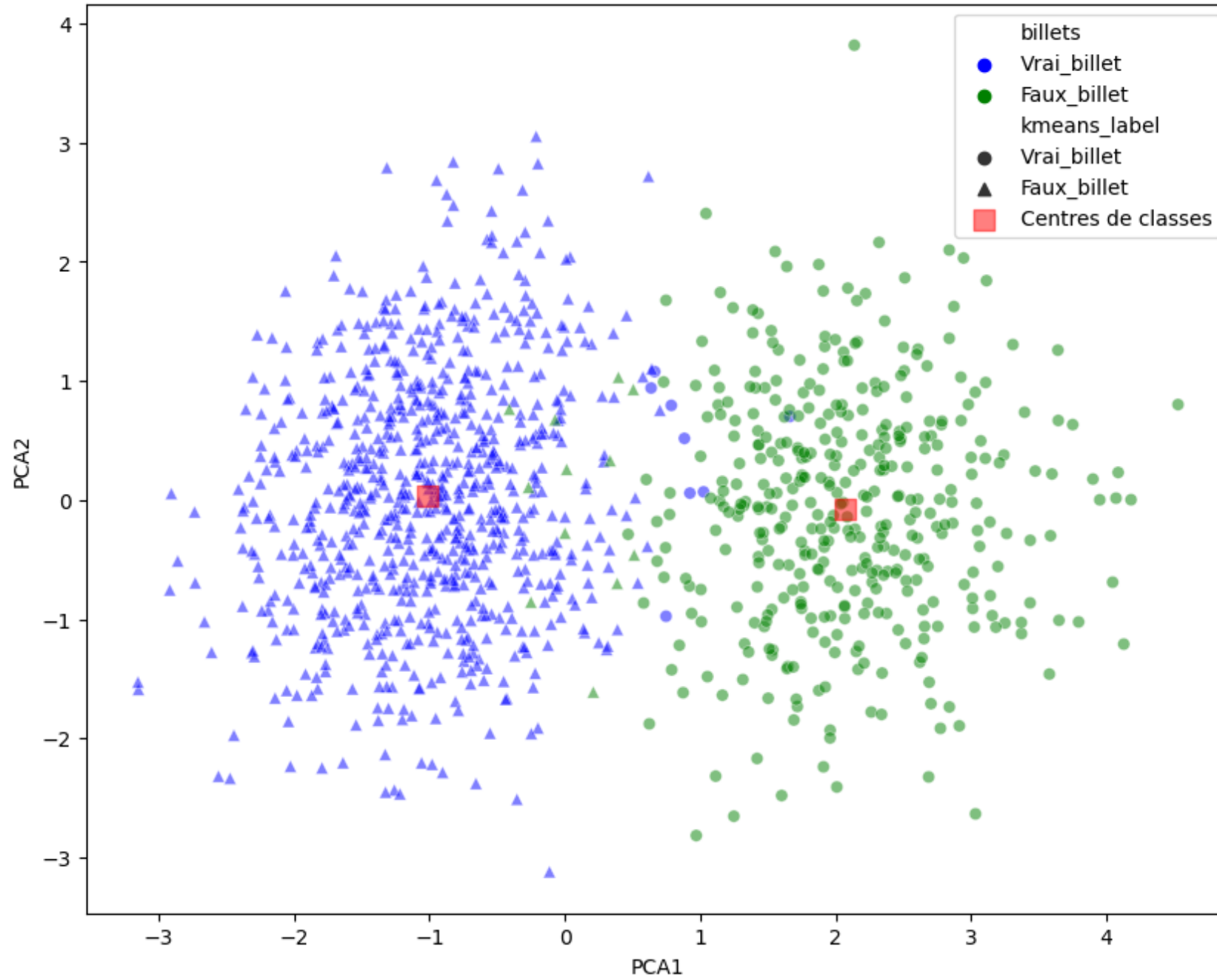
## MÉTHODE

## DU COUDE



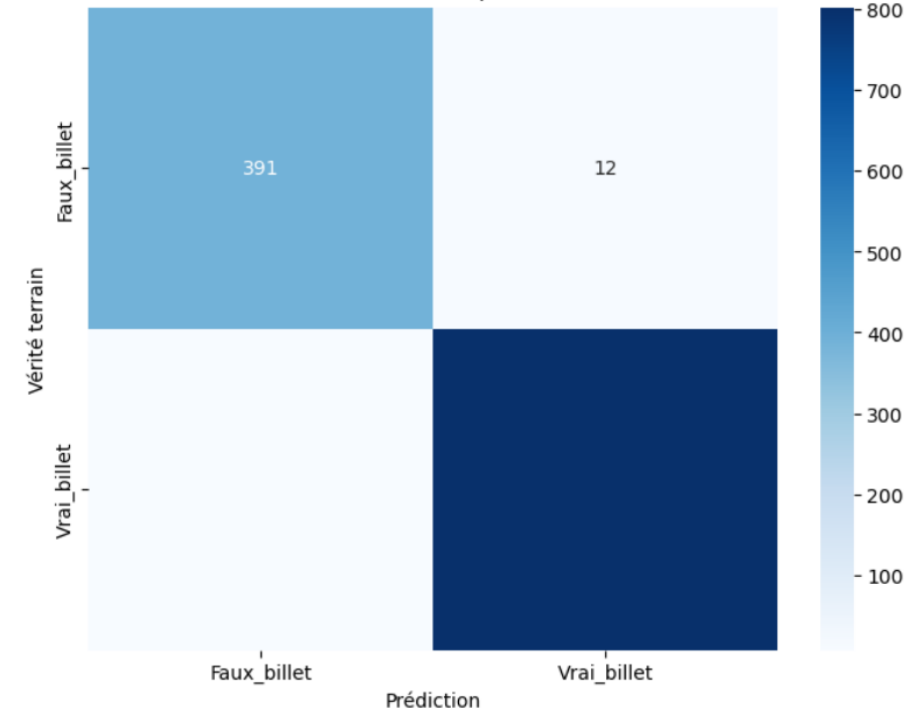


Clusters par composantes PCA pour train

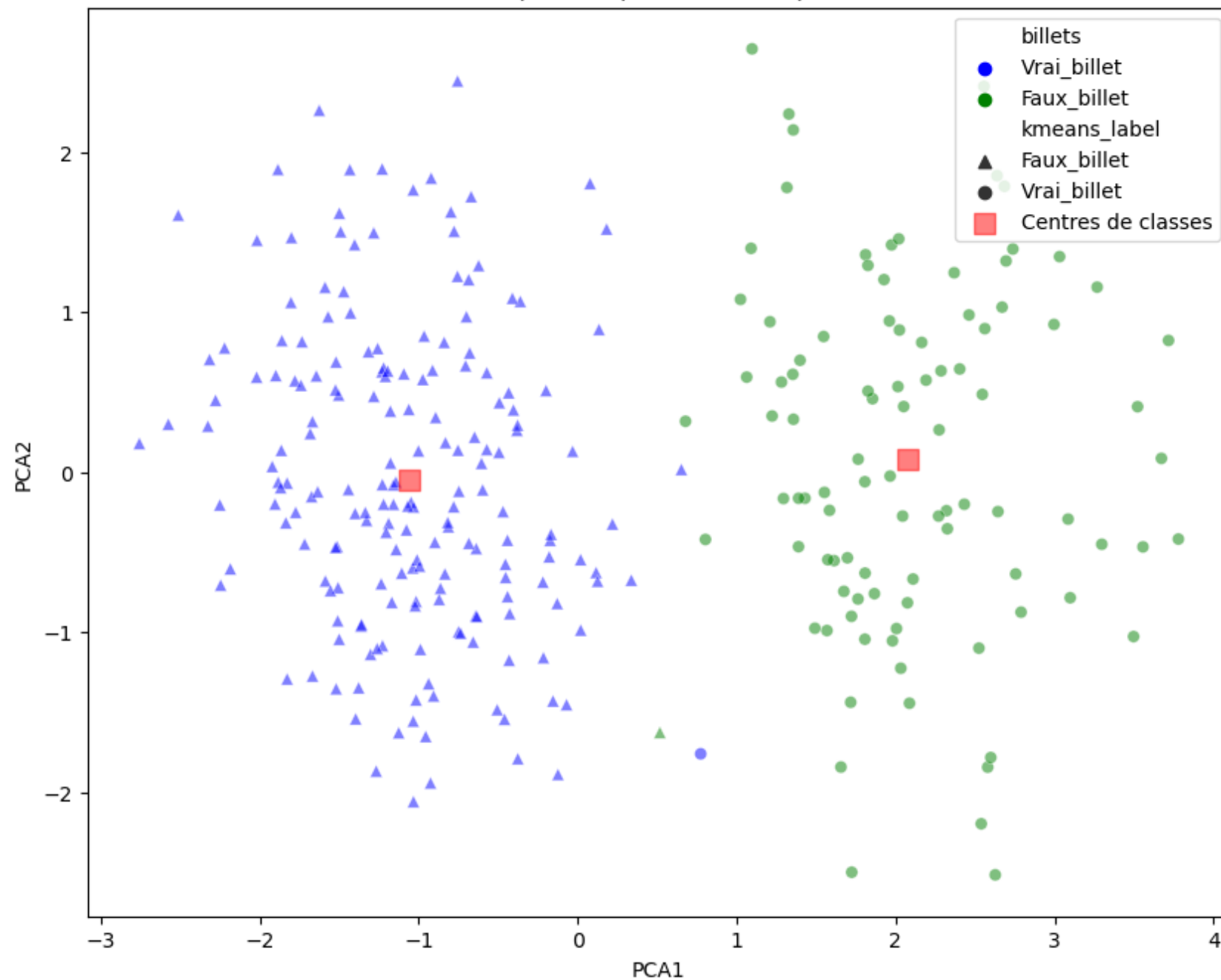


# K-Means sur les données d'entrainement

Matrice de confusion pour K-means

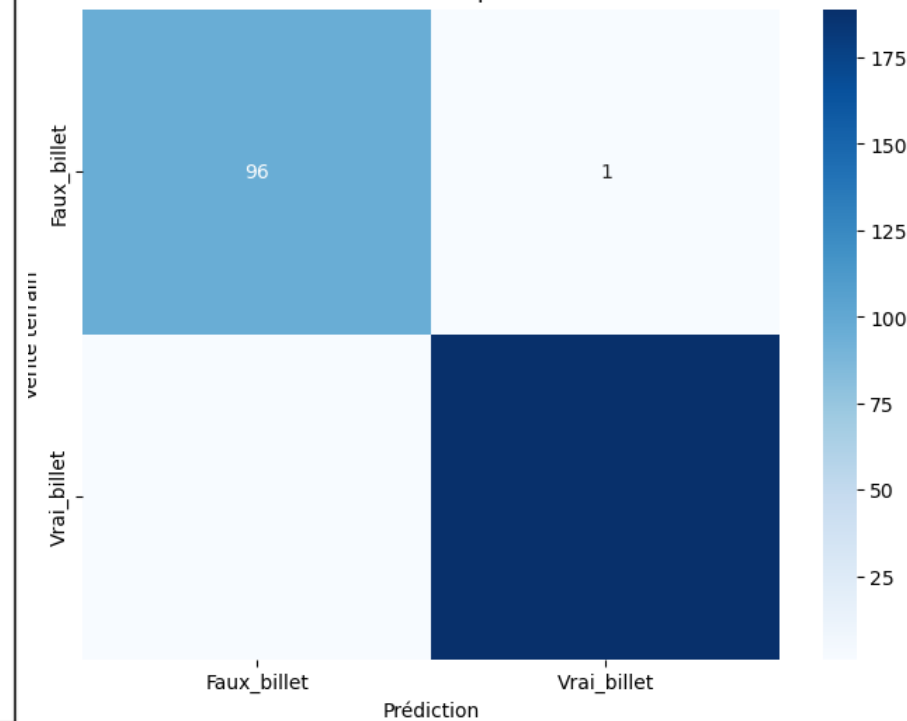


Clusters par composantes PCA pour test

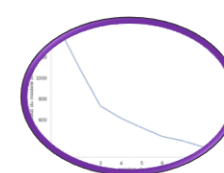
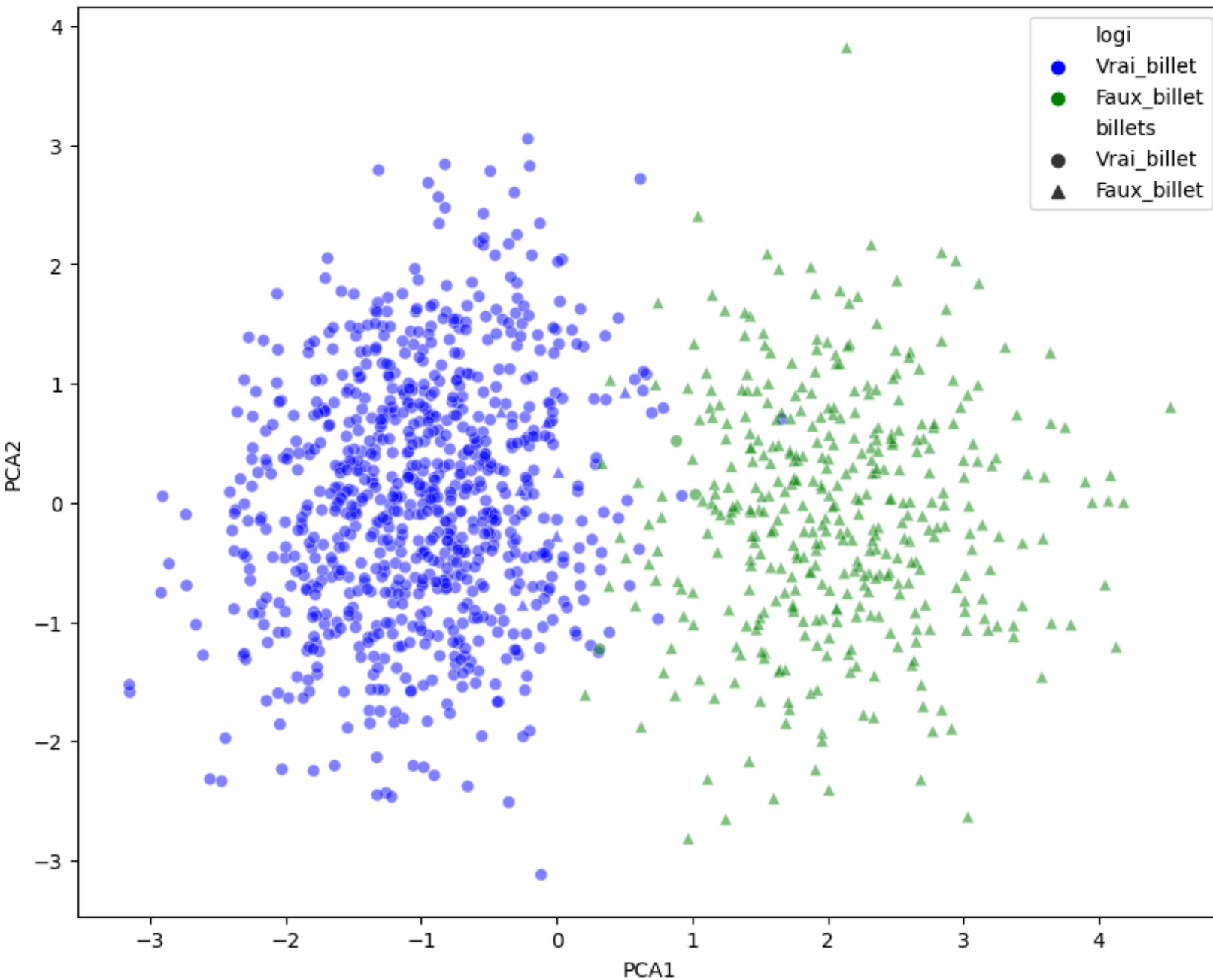


# K-Means sur les données de test

Matrice de confusion pour K-means

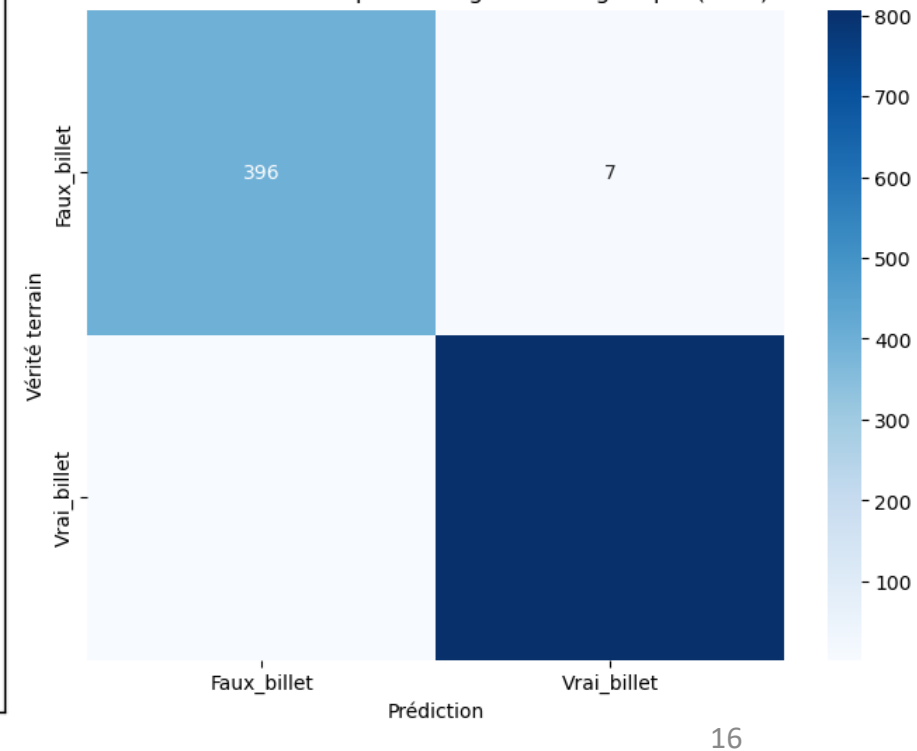


Clusters par composantes PCA pour train (régression logistique)

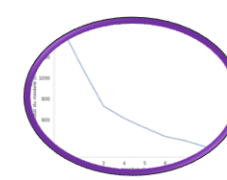
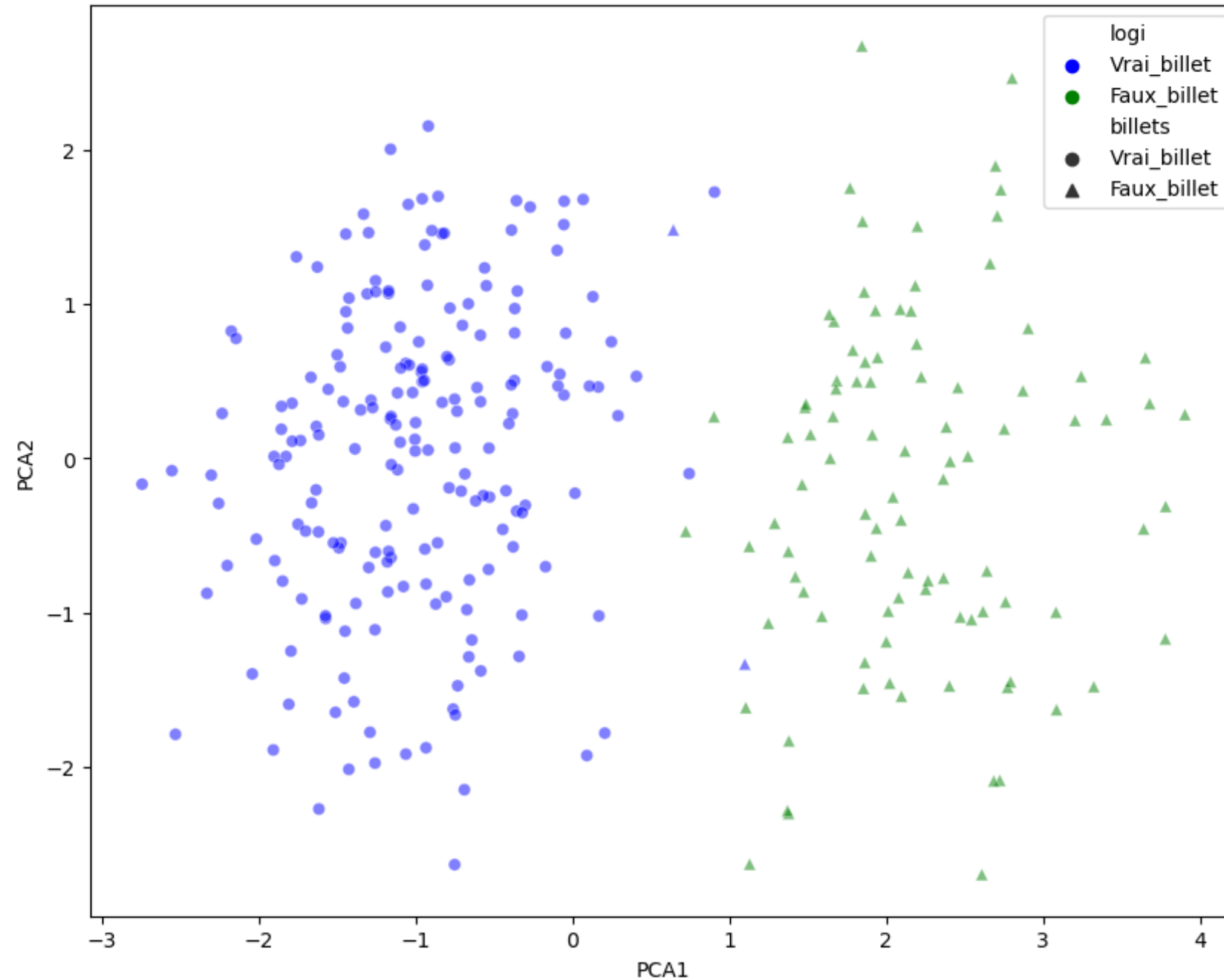


# Régression logistique sur les données d'entrainement

Matrice de confusion pour la régression logistique (train)

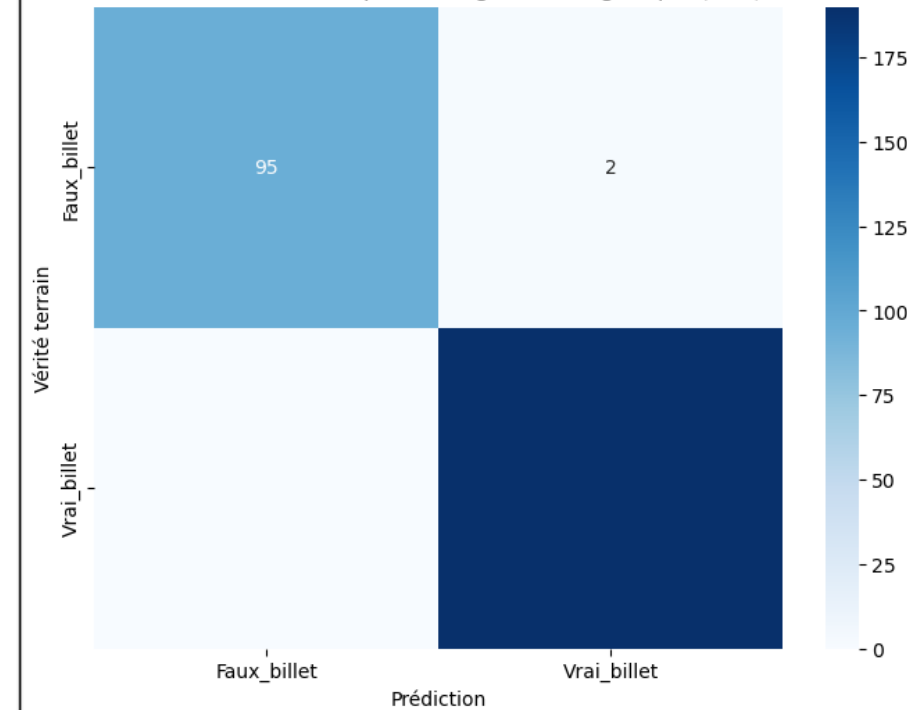


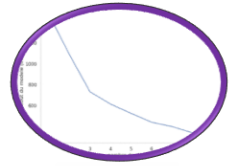
Clusters par composantes PCA pour test (régression logistique)



# Régression logistique sur les données de test

Matrice de confusion pour la régression logistique (test)





# Comparaison des modèles

	Kmeans_Train	Kmeans_Test	Reglog_Train	Reglog_Test
<b>Accuracy</b>	98.35	99.30	99.18	99.30
<b>Precision</b>	98.53	99.47	99.14	98.96
<b>Recall</b>	99.01	99.47	99.63	100.00
<b>F1_Score</b>	98.77	99.47	99.38	99.48
<b>Auc_Roc</b>	98.02	99.22	98.95	98.97



# CONCLUSION GÉNÉRALE

- **Conclusion**
- **Résumé :**
  - Utilisation efficace des techniques de Machine Learning pour répondre à un problème concret.
  - Choix de la régression logistique pour sa performance supérieure dans ce contexte.
- **Perspectives :**
  - Continuer l'amélioration des modèles.
  - Appliquer cette méthodologie à d'autres domaines de détection de fraude.