

Détection de faux billets



Les étapes



- Description des données
- Analyse en composantes principales
- Classification par knn
- Test

170 billets,
6 mesures,
1 catégorie.

6 mesures exprimées en mm

- la longueur,
- la hauteur côté gauche,
- la hauteur côté droit,
- la marge entre le bord supérieur du billet et l'image de celui-ci,
- la marge entre le bord inférieur du billet et l'image de celui-ci,
- la diagonale.

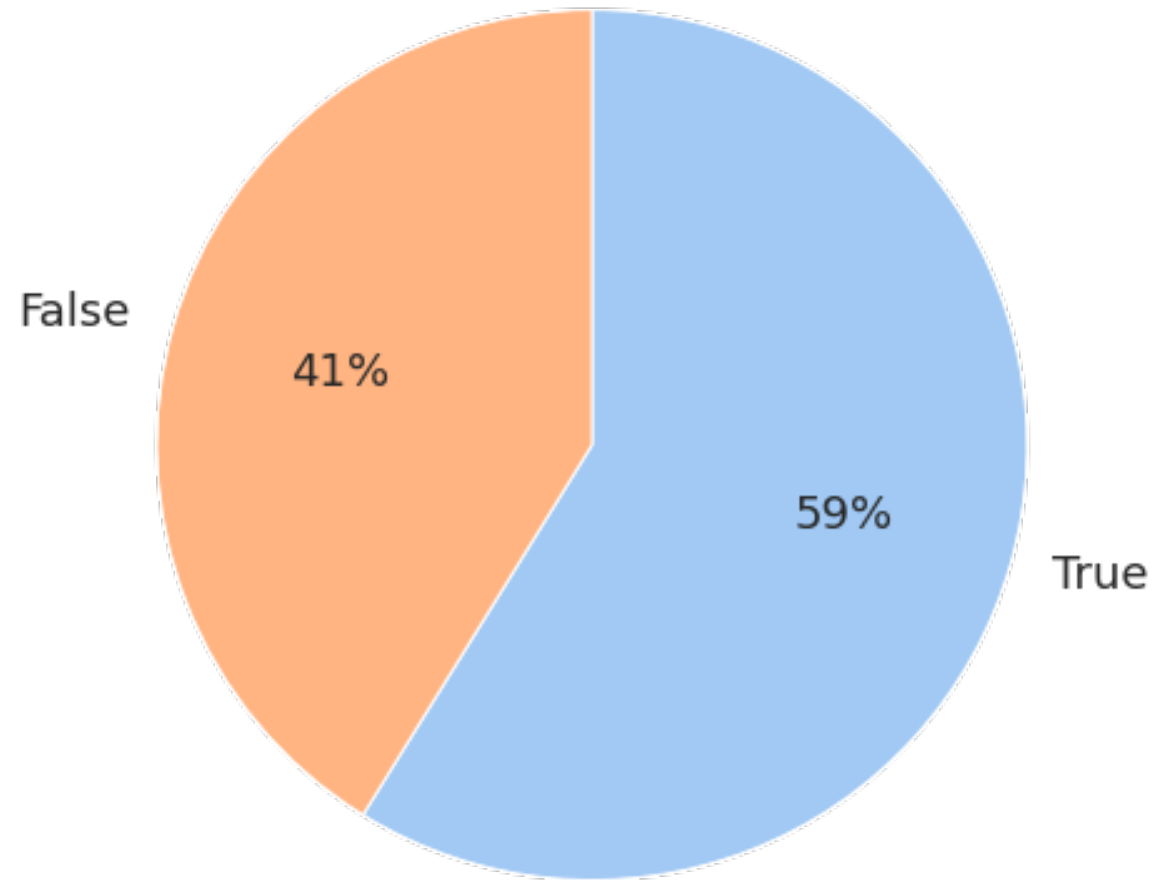
1 catégorie

- la qualité du billet (vrai ou faux).

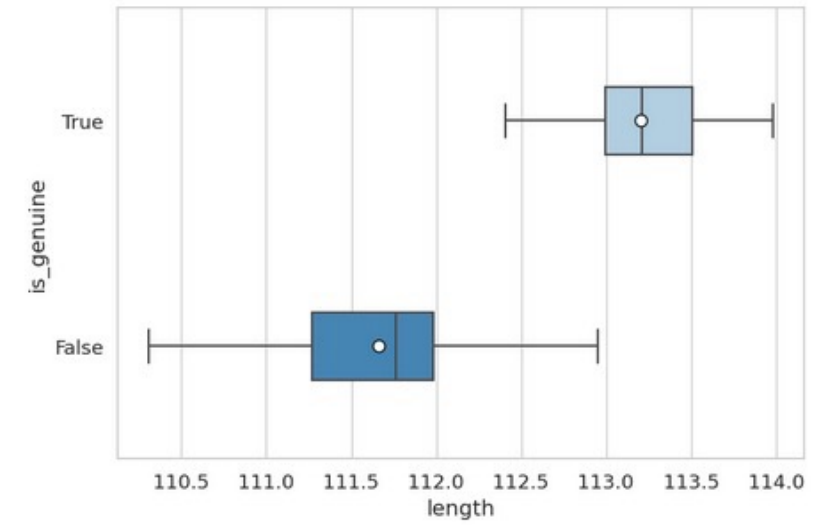
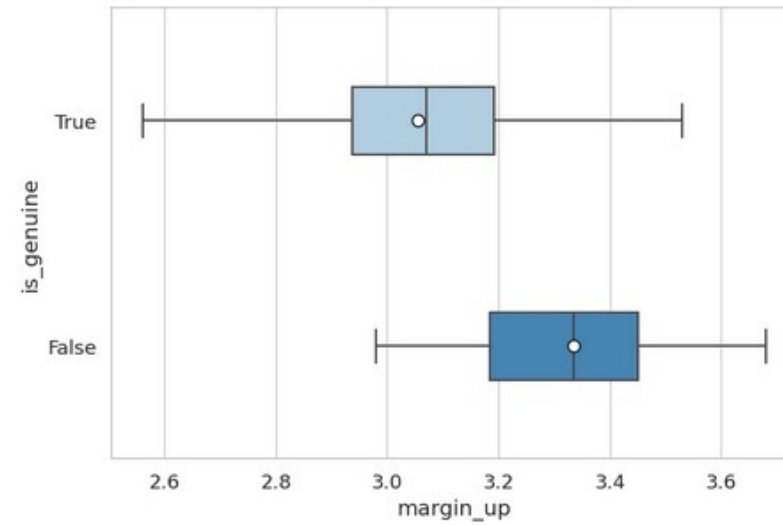
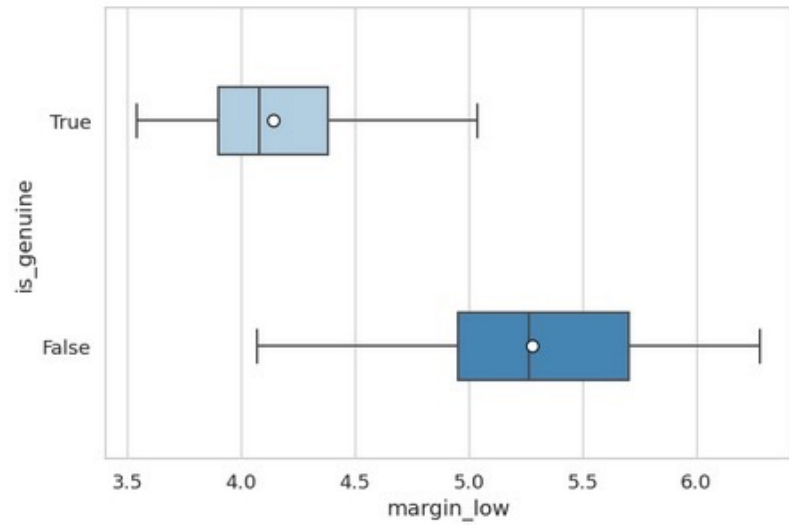
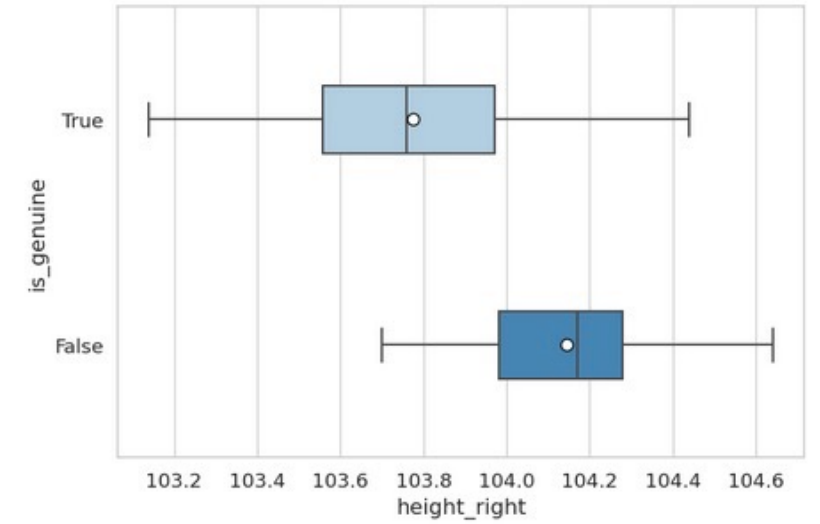
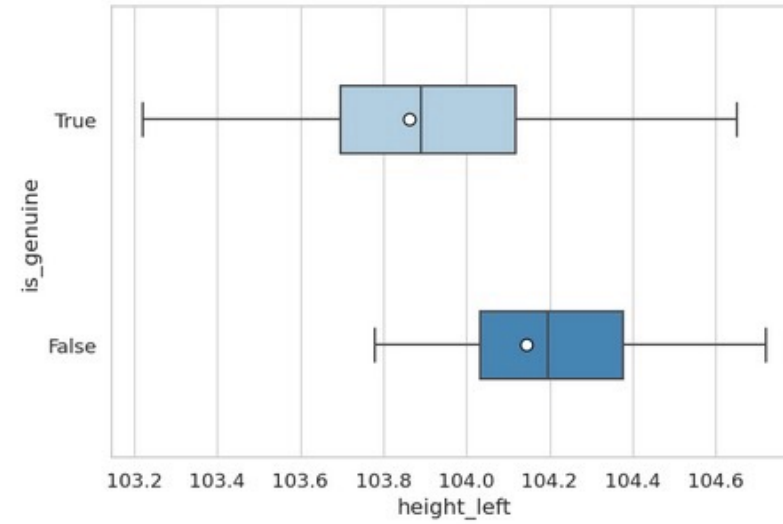
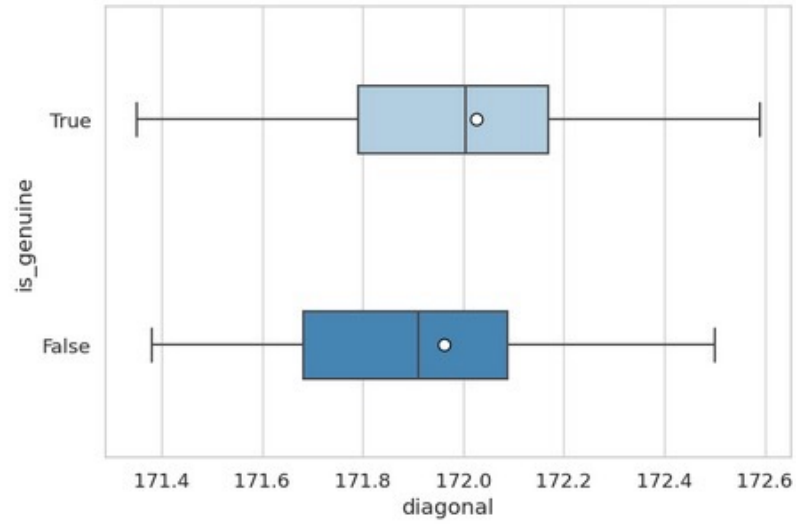
Observations

- La quantité de données est faible, 170 billets.
- Le jeu de données est légèrement déséquilibré, 100 vrais billets pour 70 faux.

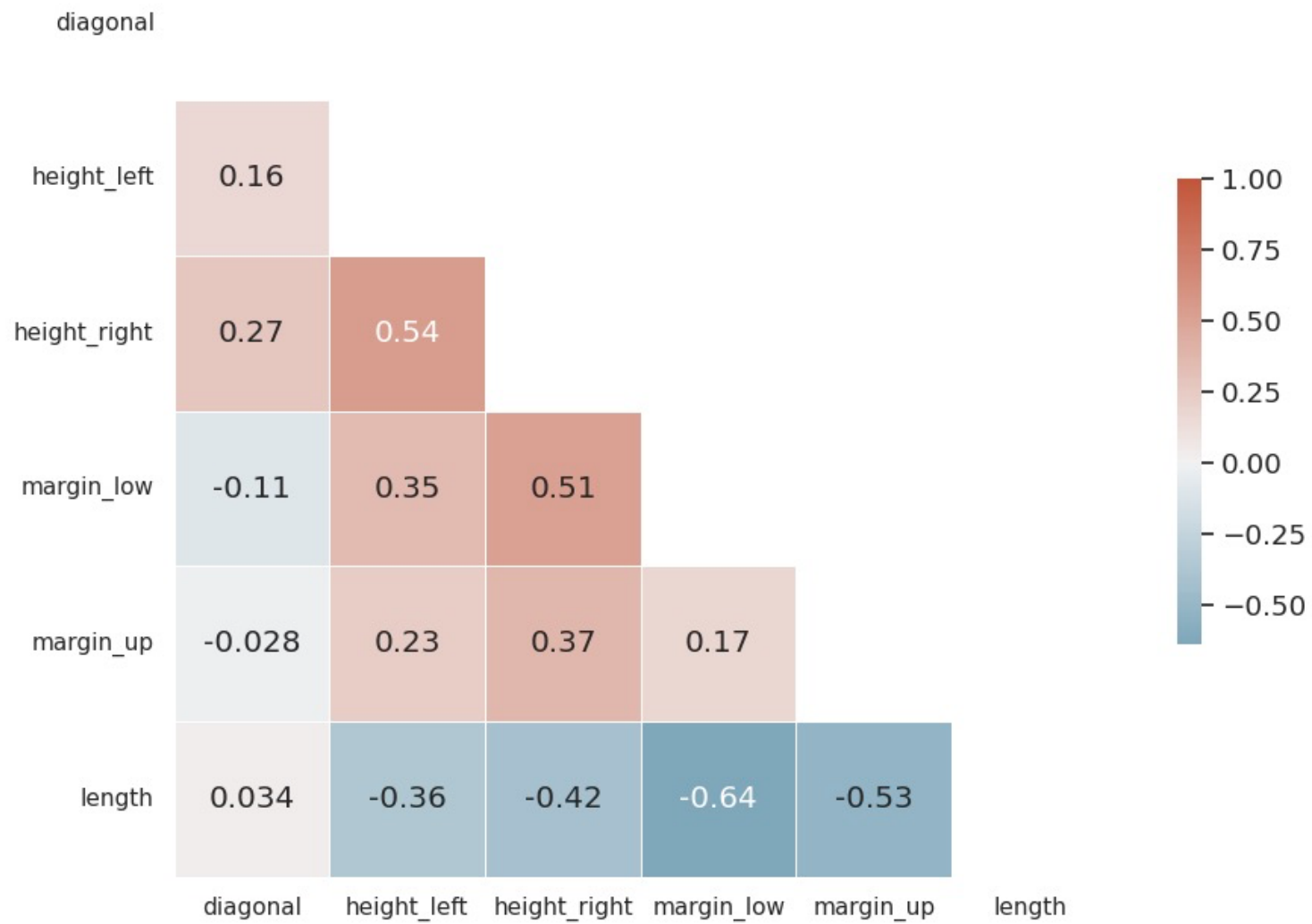
Classification des 170 billets

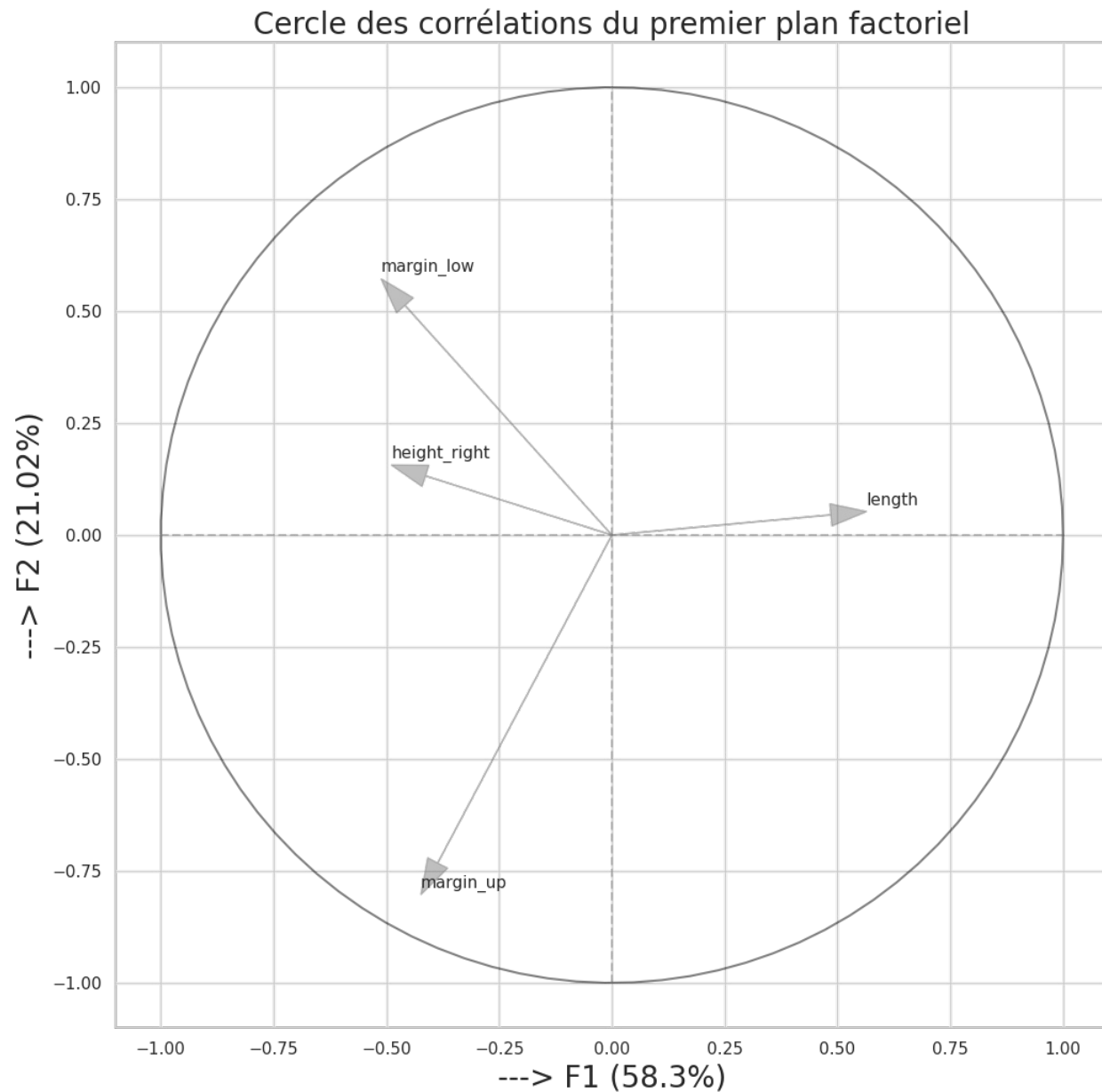


Boxplots des 6 variables explicatives



Matrice des corrélations

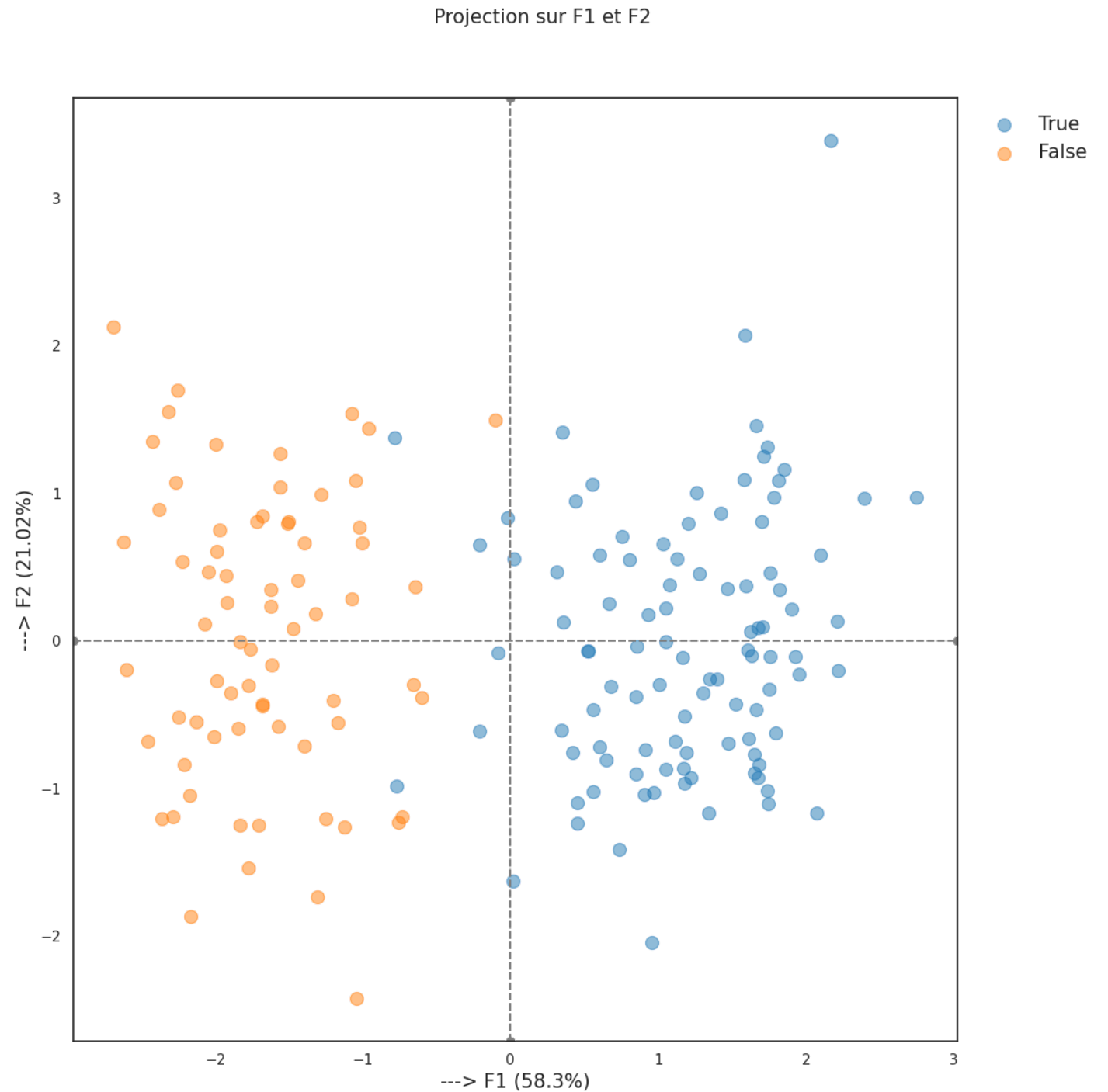




- Axe F1 : corrélé positivement à la longueur et négativement aux marges.
- Axe F2 : corrélé positivement à la diagonale.

Projection

On peut séparer les billets par une droite d'abscisse -0.75 , selon les valeurs prises par la variable synthétique F1.



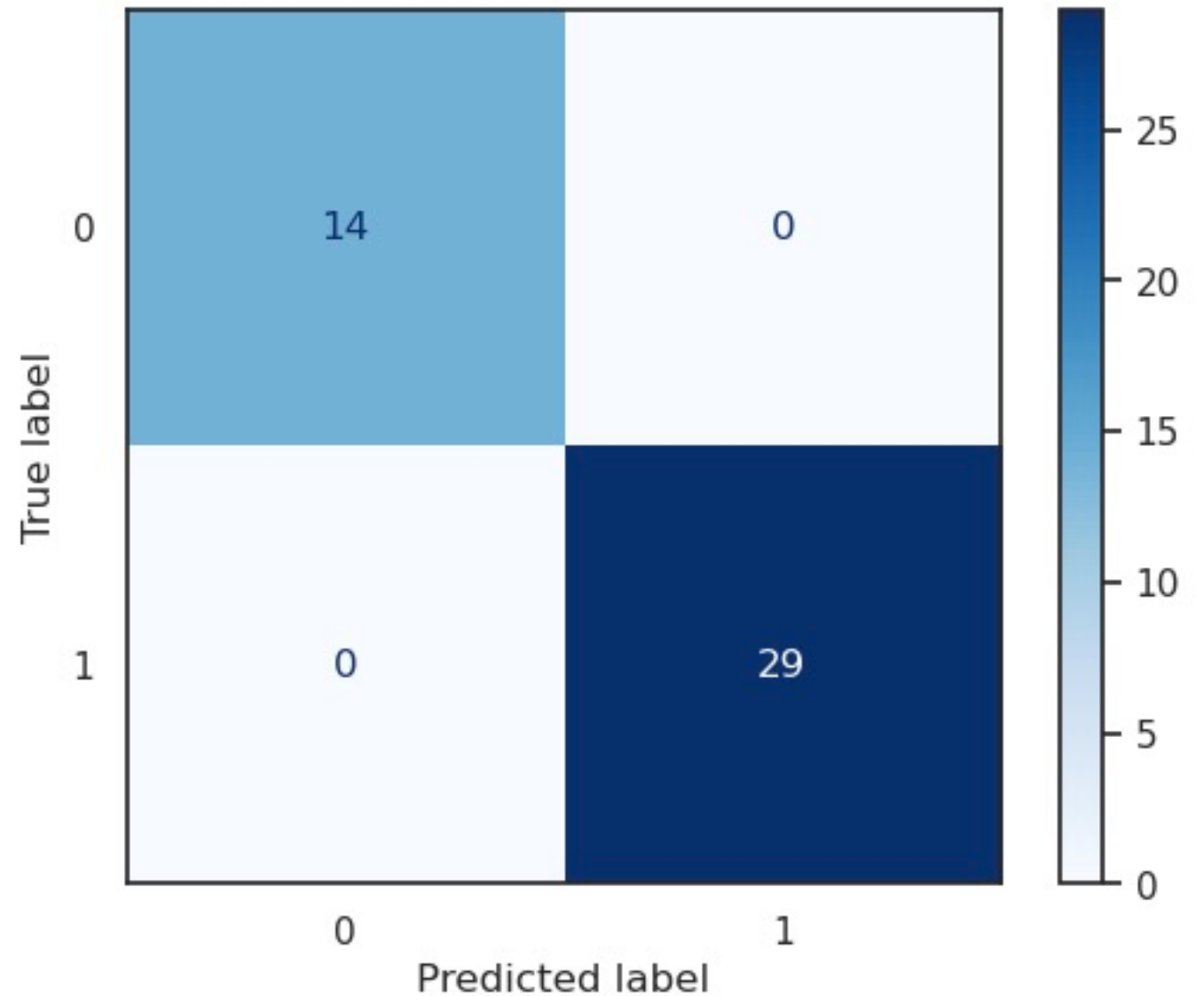
Matrice de confusion

La représentation des résultats sous forme de matrice permet le calcul de différents critères de performance du modèle.

		Classe réelle	
		-	+
Classe prédite	-	True Negatives <i>(vrais négatifs)</i>	False Negatives <i>(faux négatifs)</i>
	+	False Positives <i>(faux positifs)</i>	True Positives <i>(vrais positifs)</i>

Résultats du modèle

- spécificité = $vn / (vn + fp)$
- spécificité = $19 / (0 + 19) = 1$
- La **spécificité**, c'est-à-dire le **taux** de faux billets classés faux **est égal à 1**, le maximum possible.



Test

