

Rapport d'Analyse — 20250916_160724

Profilage

- **Lignes** : 30
- **Colonnes** : 5

Prétraitement

- Doublons supprimés : 0
- Outliers remplacés :
 - Salaire : 3 valeurs aberrantes
- Colonnes uniformisées : Ville

Imputation

- **Age** → median (13.33% imputé, asymétrique ($|skew|=0.921 \geq 0.5$))
- **Salaire** → mean (23.33% imputé, symétrique ($|skew|=0.284 < 0.5$))
- **Ville** → most_frequent (36.67% imputé, catégoriel (mode))

Distorsion avant/après imputation

Colonne	Méthode	% valeurs imputées	Distorsion (%)
Age	median	13.33%	6.15%
Salaire	mean	23.33%	17.69%
Ville	most_frequent	36.67%	73.68%

Résultats du modèle

- Accuracy : 100.00%
- F1-score : 100.00%

Versions sauvegardées

- **Données brutes** → data_versions\raw.parquet
- **Données imputées** → data_versions\imputed.parquet