

Projet de Statistique

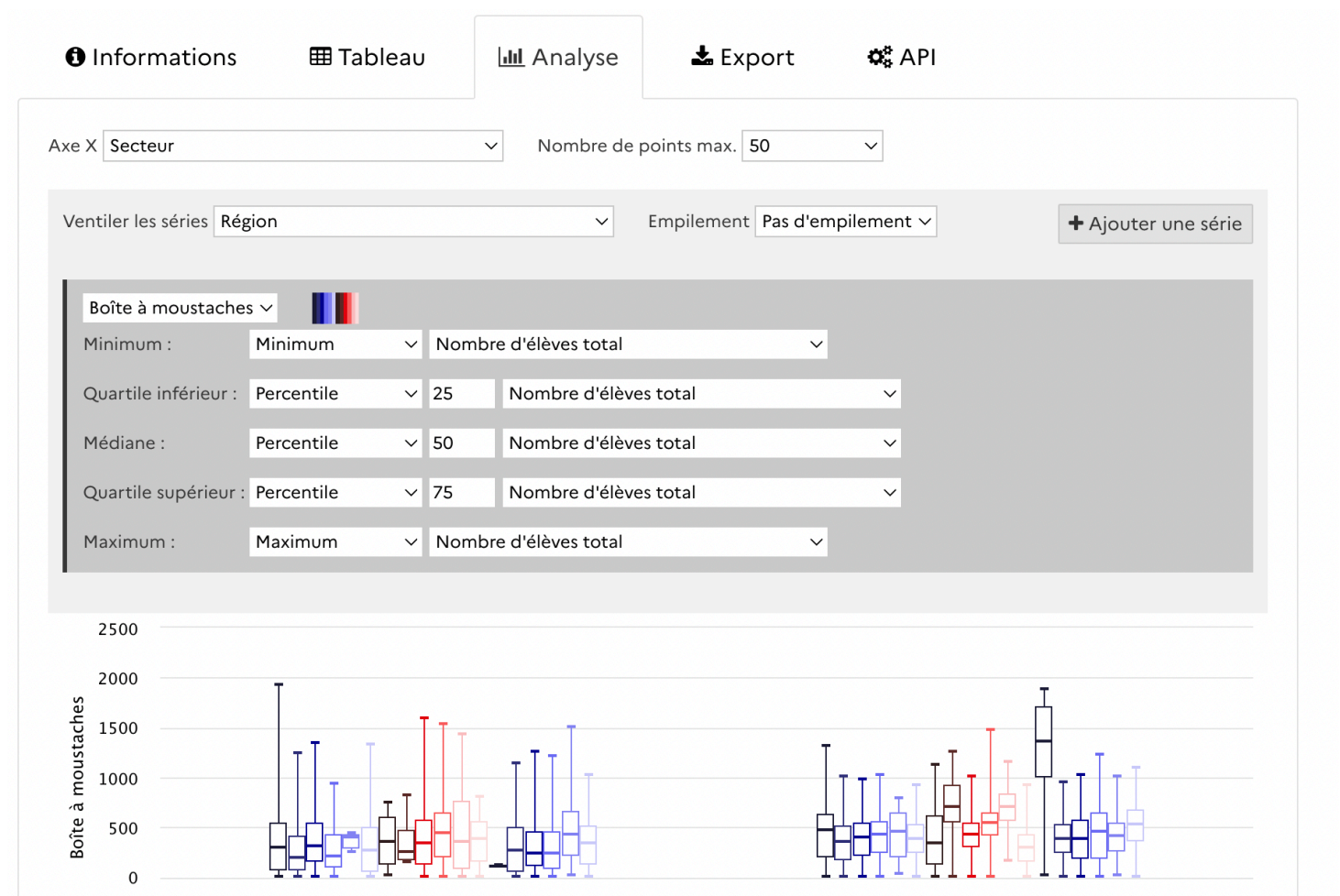
Cadre

Ce projet vous permet d'exercer vos compétences en **python**, **analyse de données**, **nettoyage**, et en **statistiques**. Il est attendu une qualité maximale concernant les **éléments statistiques** (précaution sur les conclusions, formulation d'hypothèses et leur vérification, modèles appropriés).

Présentation

Vous allez travailler sur des données de l'enseignement secondaire (collège). Les données sont disponible sur le site du gouvernement : <https://data.education.gouv.fr/explore/dataset/fr-en-college-effectifs-niveau-sexe-lv/>.

Vous pouvez explorer les données dans un premier temps directement sur leur site.



Questions

Question 1 : Certains collèges ont t'ils davantage de décrochages scolaire que d'autres ? (réduction du nombre d'élèves entre la 6e et la 3e)

Question 2 : Cette tendance a-t'elle un lien avec la taille du collège ? (volume d'étudiants)

Question 3 : Est-ce la même conclusion pour les garçons que pour les filles ?

Question 4 : Est-ce que le constat change en fonction de la région ?

Question 5 : Y-a-t'il un lien avec le fait d'être en collège public ou privé ?

Recommandations

- Evaluer les corrélations via des **tests statistiques**.
- Bien vérifier les **hypothèses** des tests et des modèles.