

Visualisation et analyse

A partir des rapports de criminalité émanant de tous les quartiers de San Francisco sur 12 années (de 1934 à 1963), on vous demande de réaliser des visualisations dans l'objectif de révéler les tendances cachées.

1. Exploitation du fichier des données

- Télécharger le fichier « train.csv ».
- Ouvrez le fichier et identifiez les principales dimensions et les mesures associées
- Importez le fichier avec le logiciel Tableau Public :
 - Que constatez-vous au niveau des champs des dimensions (volet source de données ou feuille de calcul) ?
 - Corrigez ce problème

2. Visualisation et analyse

Réalisez les visualisations nécessaires pour répondre aux questions suivantes :

- Y a-t-il une tendance annuelle / mensuelle / quotidienne / horaire
 - Quels sont les périodes où la criminalité est la plus élevée et plus basse (heures/jours/semaine/mois)
- Existe-t-il des clusters spécifiques avec un taux de criminalité plus élevé ?
 - Ex. District vs Resolution, Category vs Resolution, Category vs Days, Category vs District...
- Identifiez les zones géographiques présentant les plus fortes criminalités à San Francisco
 - Cette visualisation s'appuie sur les coordonnées géographiques X et Y qui nécessitent une opération de conversion pour être exploitable

3. Tableau de bord

Mettre en place un tableau de bord combinant une visualisation géographique et les visualisations impliquant les catégories des crimes, les résolutions des crimes et les districts. Exploiter les filtres pour analyser la distribution géographique des crimes : district par district, distribution des catégories de crimes, distribution de la résolution de crimes, etc.