

# Rapport d'avancement

---

Travail réalisé pour l'analyse des données afin d'optimiser notre prédiction :

1. **Calcul de la fréquence des Catégories de crimes à San Francisco.** Cela a permis de sélectionner les Catégories de crimes les plus représentés: *Larceny/Theft, Other Offenses, Non-Criminal, Assault*, représentant 55% des crimes commis. Nous avons pu également sélectionner les crimes les moins représentés, 18 catégories de crimes ayant chacun un pourcentage inférieur à 1%
2. **Calcul de la fréquence des Sous Catégories de crimes à San Francisco.** Ces résultats sont utiles uniquement à titre informatif.
3. **Calcul de la fréquence des crimes pour chaque district.** Cela a permis de révéler les Districts où étaient réalisés le plus grand nombre de crimes. Southern, Northern et Mission sont les Districts les plus importants.
4. **Calcul de la Catégorie de crime avec la plus haute fréquence pour chaque district.** Il ressort que les Catégories Larceny/Theft et Other Offenses sont les catégories de crimes les plus présentent dans chaque District de San Francisco ce qui est en concordance avec le calcul des fréquences de Catégories de crimes à San Francisco. Une seule Exception, le District Tenderloin, où la Catégorie Drug/Narcotic est le premier facteur de crime.
5. **Calcul total du nombre de crimes par jour de la Semaine.** Ces résultats sont uniquement à titre informatif, il ressort que le nombre de crimes commis à San Francisco est plus important le Vendredi (133 734 crimes) et le Mercredi (129 211 crimes) et inversement le nombre de crimes est moins important le Dimanche (116 707 crimes).
6. **Calcul de la Fréquence des Crimes par Heure et par District.** Cela a permis de constater 2 pics d'augmentation du nombre de crimes dans chaque District à 12h et 17h-18h. Une seule Exception, il s'agit encore du District Tenderloin où le nombre de crimes augmente progressivement de 4h jusqu'à 17h.  
Grâce au graphique des fréquences nous avons également remarqué que les crimes avaient majoritairement lieu entre 08h et 00h et beaucoup moins entre 00h et 8h.
7. **Calcul de la Fréquence des Crimes par Heure et par Catégorie.** J'ai sélectionné pour cette analyse les 4 catégories de crimes ayant la plus haute fréquence à San Francisco. On remarque pour cette analyse un pic du nombre de crimes à 12h et 17h-18h pour les 2 premières catégories (Larceny/Theft, Other Offenses) en relation avec la fréquence des crimes par heure et par district, et un seul pic de crime à 12h pour les 2 catégories restantes (Non-Criminal, Assault).

→ Autres Analyses à réaliser :

- Quelles sont les adresses dans San Francisco où il se produit le plus de crimes.
- Pour chaque adresse quelle est la catégorie de crime la plus représentée.