

# PROJET 16 : Prédiction du Box-Office de Cinéma

## OBJECTIF DU PROJET

Les films sont des investissements à haut risque. Prédire les revenus avant la production aide les studios à budgétiser et à commercialiser efficacement.

## FICHIER DE DONNÉES

- box\_office.csv
- 800 lignes
- Inclut des valeurs manquantes, valeurs aberrantes et déséquilibres pour un apprentissage réaliste

## TÂCHES À RÉALISER

### ■ TÂCHE PRINCIPALE

Prédire Revenus

### ■ TÂCHES SECONDAIRES (OPTIONNELLES)

1. Classer comme "Flop", "Succès", "Blockbuster"
2. Recommander le "Mois de Sortie Optimal"
3. Analyser l'efficacité "Budget vs Revenu"
4. Prédire la Note IMDb (Proxy de qualité)

## LIVRABLES ATTENDUS

1. \*\*Exploration des Données (EDA)\*\*
  - Analyse statistique descriptive
  - Visualisations des distributions
  - Analyse des corrélations
  - Identification des valeurs manquantes et aberrantes
2. \*\*Ingénierie des Fonctionnalités\*\*
  - Création de nouvelles variables pertinentes

- Transformation des variables existantes
- Encodage des variables catégorielles
- Normalisation/Standardisation si nécessaire

### 3. \*\*Modélisation\*\*

- Choix et justification du/des modèle(s)
- Entraînement et validation
- Optimisation des hyperparamètres
- Évaluation avec métriques appropriées

### 4. \*\*Résultats et Interprétation\*\*

- Performance du modèle principal
- Résultats des tâches secondaires (si réalisées)
- Insights et recommandations
- Visualisations des résultats

## CONSEILS

- Commencez par une exploration approfondie des données
- L'ingénierie des fonctionnalités est CRUCIALE pour ce projet
- Documentez votre démarche et vos choix
- Comparez plusieurs approches si possible
- Pensez à l'aspect pratique et utilisable de vos résultats

Bon travail ! ■