

**Sujet :** Octroi de prêts hypothécaires

**Dataset :** (Site officiel gouvernement USA)

<https://ffiec.cfpb.gov/data-browser/data/2023?category=states> ( data )

<https://ffiec.cfpb.gov/documentation/publications/loan-level-datasets/lar-data-fields> ( explication des variables )

Les variables clés de la dataset HMDA (à peu près 900 000 instances et 99 colonnes chaque année de 2018 à 2023) incluent **démographiques** (race, ethnie, sexe), **financières** (revenu, scores de crédit, ratios d'endettement), et **caractéristiques du prêt** (montant, taux d'intérêt, objectif). De plus, la **localisation géographique** et les **caractéristiques de la propriété** (type et occupation) révèlent les disparités régionales dans l'octroi de crédits. La variable à prédire est **action\_taken** , elle représente la réponse à la demande de prêt .

**Contexte :** En Californie, où le marché immobilier est très compétitif et les prix élevés, l'accès au crédit immobilier est crucial pour permettre aux habitants de devenir propriétaires. Les institutions financières utilisent divers critères pour évaluer les demandes de prêts, comme le revenu ou la dette, mais des problèmes subsistent, avec des demandes méritantes parfois refusées et d'autres acceptées à tort. Ces décisions peuvent refléter des biais socio-économiques, influencés par des facteurs comme la race, l'ethnicité ou la localisation. Analyser ces inégalités à travers le dataset HMDA permettrait d'optimiser le processus d'octroi de prêts, en garantissant plus d'équité tout en améliorant la rentabilité pour les banques.

**Utilisateurs/Personnes/Organisation :** Les **institutions financières** et les **prêteurs** cherchent à améliorer le processus d'approbation des prêts hypothécaires en réduisant les biais et en optimisant l'évaluation des risques. Ils ont besoin d'outils simples et rapides pour analyser les demandes de prêts et identifier les disparités dans les pratiques de prêt.

**Objectif : ( valeur ajoutée )**

L'objectif est de développer un modèle de classification efficace pour l'approbation ou le rejet des demandes de prêts immobiliers en Californie, tout en identifiant les principaux facteurs influençant ces décisions et en détectant d'éventuels biais discriminatoires liés à des caractéristiques démographiques telles que l'ethnicité et le revenu.