# Chapitre 3 : Méthodologie de l’étude

3.1 Type de recherche  
  
Cette étude adopte une approche quantitative et analytique visant à modéliser le phénomène d’attrition client dans une banque commerciale. L’objectif est de comprendre les déterminants de ce phénomène et de proposer un modèle prédictif permettant d’identifier les clients à risque de départ.

3.2 Source et collecte des données  
  
Les données utilisées proviennent du système d'information de la banque. Elles ont été collectées à partir de différentes bases de données internes, notamment celles liées aux transactions, aux caractéristiques des clients, aux interactions avec le service client, et aux enquêtes de satisfaction. Un échantillon représentatif a été extrait afin de garantir la qualité et la fiabilité de l’analyse.  
  
Un processus de nettoyage des données a été effectué pour traiter les valeurs manquantes, les doublons et les incohérences éventuelles.

3.3 Variables de l’étude  
  
Deux types de variables ont été considérés :  
- Les variables explicatives : âge, sexe, ancienneté, solde moyen, nombre d’opérations, fréquence des visites en agence, réclamations, etc.  
- La variable cible : statut du client (actif ou clôturé)  
  
Ces variables permettront de construire un modèle de prédiction de l’attrition.

3.4 Techniques d’analyse  
  
L’analyse des données sera réalisée en plusieurs étapes :  
- Statistiques descriptives pour caractériser la population étudiée  
- Analyse bivariée pour explorer les liens entre les variables explicatives et la variable cible  
- Modélisation prédictive à l’aide d’algorithmes de machine learning (régression logistique, arbre de décision, random forest, etc.)  
  
Les performances des modèles seront évaluées à l’aide de métriques telles que l’accuracy, le recall, la précision et l’AUC (Area Under the Curve).

3.5 Outils et environnements de travail  
  
L’analyse sera menée à l’aide des outils suivants :  
- Python (avec les bibliothèques pandas, scikit-learn, matplotlib, seaborn)  
- Jupyter Notebook pour l'exploration et la modélisation  
- Microsoft Excel pour la visualisation initiale et les tableaux de synthèse  
- Power BI ou Streamlit pour une restitution interactive des résultats  
  
Ces outils offrent une grande souplesse et une capacité d’analyse poussée des données.