

SAE302 - Intégration de données dans un datawarehouse

Projet

Projet SAÉ 302

Dans ce projet, nous vous demandons de reproduire l'ensemble du processus décisionnel étudié en TP, mais en l'appliquant à un autre contexte métier, que vous choisirez vous-même (par exemple : ventes, logistique, santé, éducation, sport, anime, etc.).

Vous pourrez notamment utiliser des jeux de données publics (par exemple sur Kaggle ou data.gouv.fr). Merci de prendre un jeu de données non nettoyé, afin que vous puissiez appliquer quelques transformations dessus (par exemple : supprimer les doublons, gérer les valeurs manquantes, harmoniser les valeurs, etc.).

Pour ce projet, vous devrez :

- identifier un contexte métier clair et expliquer les objectifs décisionnels (identifier les questions métiers),
- décrire précisément les sources de données utilisées (formats, volumes, origine),
- indiquer la source des données (URL, plateforme, référence),
- mettre en œuvre l'ensemble du pipeline BI :
 - sources → ETL → Data Warehouse → reporting,
- justifier les choix de modélisation (faits, dimensions),
- produire des tableaux de bord répondant aux questions métier identifiées.

L'ensemble de ce travail devra être documenté dans un rapport final de 5-15 pages.