



Rapport SAE 2.01

Conception et implémentation d'une base de données relationnelle

Quelle est l'évolution du nombre d'adhérents et des dons au sein de l'association?

Est-ce que les membres fideles impacte-t-il reelment plus l'association du point de vu financier ?

Encadré par :

MME DORINE TABARY

Réalisé par :

EL BADAoui MARWAN
ALAHYAN SALMA
TARAZA YASMINE

TABLE DES MATIERES

Identification du besoin	3
Analyse des sources de données	4
Informations personnelles des membres	4
Informations sur les adhésions et les contributions	4
Informations relatives à la réunion du 21 décembre 2024	5
État initial des données	5
Préparation des données	5
Nettoyage manuel	6
Réorganisation des données avec Python	6
Résultat du traitement initial	7
Formatage des dates	7
...	...



1. Identification du besoin

La SAÉ 2.01 vise à concevoir et mettre en œuvre une base de données relationnelle destinée à optimiser la gestion des informations d'une association. Celle-ci répertorie les données concernant ses membres, les adhésions annuelles, les contributions, ainsi que la préparation d'un événement de fin d'année.

Actuellement, ces informations sont saisies dans un fichier Excel, ce qui limite la structuration, la qualité et l'analyse des données. Le projet aura alors pour mission :

- D'élaborer une base de données relationnelle solide, intégrant les diverses entités et connexions identifiées ;
- De formuler des requêtes SQL efficaces pour extraire les informations utiles ;
- De créer des visualisations synthétiques à partir des résultats obtenus, pour les besoins de suivi ou de communication.

Ce travail suit une approche complète qui comprend l'analyse des besoins, la documentation technique, et passe par la modélisation, l'implémentation et l'alimentation de la base.

2. Analyse des sources de données

Le document Excel fourni, intitulé dataAsso.xlsx, regroupe toutes les informations historiques relatives aux membres, à leurs adhésions et dons annuels, ainsi qu'à une réunion du 21 décembre 2024. Ce fichier ne comporte qu'une seule feuille, où chaque ligne représente un membre, et chaque colonne présente une information spécifique. Cette disposition des données en largeur peut rendre leur analyse difficile si elles ne sont pas réorganisées.

2.1 Informations personnelles des membres

Les premières colonnes du fichier concernent les informations d'identité et de localisation des membres :

- Nom, Prénom
- Date de naissance
- Rue, Ville, Latitude Ville, Longitude Ville, État
- Téléphone

Ces informations permettent d'identifier chaque membre de manière unique et fournissent une perspective géographique de la répartition des membres.

2.2 Informations sur les adhésions et les contributions

Le fichier présente des groupes de colonnes répétitifs pour chaque année, de 2015 à 2024, qui concernent :

- DATE D'ADHÉSION XXXX
- MONTANT XXXX
- DON XXXX
- MÉTHODE DE PAIEMENT XXXX
- BUREAU XXXX

Ainsi, les informations des membres peuvent varier d'une année à l'autre, selon qu'ils ont adhéré ou non, le montant versé, la méthode de paiement choisie, ou le rôle occupé au sein du bureau.

Remarque essentielle : Une telle structure, avec une colonne pour chaque année, ne convient pas à une base de données relationnelle. Il est nécessaire de transformer ces groupes en format long (voir la section "Prétraitement").

2.3 Informations relatives à la réunion du 21 décembre 2024

Le fichier inclut, enfin, des colonnes spécifiques à la réunion :

- Heure d'arrivée (Réunion du 21 décembre 2024 – arrivée prévue à 10h, 12h, 14h, 16h)
- Heure de départ (Réunion du 21 décembre 2024 – départ à 12h, 14h, 16h, 18h)
- Participation au repas
- Préférences alimentaires (végétalien, végétarien, etc.)
- Commentaires / Allergies
- Présence ou non

Ces informations sont essentielles pour comprendre la participation. Cependant, les données étant réparties sur plusieurs colonnes contenant la mention "oui", cela complique leur exploitation sans un traitement préalable.

2.4 État initial des données

Le fichier contient :

- 9 colonnes d'informations personnelles
- 50 colonnes relatives à l'adhésion/don (5 colonnes sur 10 ans)
- 12 colonnes en lien avec l'événement
- Soit un total de 71 colonnes

Un travail de nettoyage, de restructuration et de mise en conformité est donc essentiel pour créer une base de données relationnelle claire et performante.

2.5 Préparation des données

Avant de procéder à la conception de la base de données relationnelle, une série d'opérations de nettoyage et de préparation ont été réalisées sur le fichier dataAsso.xlsx. Ces tâches ont été principalement réalisées manuellement, suivies d'un complément avec un script Python pour organiser les données.

2.5.1 Nettoyage manuel

Différentes erreurs et incohérences ont été rectifiées directement dans le document Excel :

- Suppression de 2 lignes incorrectes qui étaient soit vides soit inutilisables.
- Formatage des montants :
 - Deux valeurs "?" ont été enlevées de la colonne MONTANT 2021.
 - Deux cellules contenant "8 € ?" et "5 € ?" ont été ajustées manuellement à "8" et "5" afin de faciliter la conversion en format numérique.
- Erreur de localisation : une valeur de longitude s'était retrouvée par erreur dans la colonne Ville. Elle a été supprimée à la main pour restaurer la cohérence.

Ces modifications ont été indispensables pour assurer la qualité des données et éviter les erreurs lors du traitement avec des outils automatisés.

2.5.2 Réorganisation des données avec Python

Un script Python a ensuite été employé à l'aide de l'intelligence artificielle pour transformer les colonnes annuelles et événementielles en une table utilisable, où chaque ligne représente l'adhésion d'un membre pour une année particulière.

Objectif :

- Concentrer toutes les informations d'adhésion dans un format allongé, comportant une seule colonne pour l'Année, la Date_Adhésion, le Montant_Adhésion, etc.
- Simplifier les colonnes liées aux événements (comme l'arrivée, le départ, le régime alimentaire...) pour en faire des champs clairs et faciles à lire.

Explication simplifiée

Ce script :

- Examine chaque membre (chaque ligne dans le tableau Excel)
- Vérifie, pour chaque année, si le membre a souscrit une adhésion
- Le cas échéant, il génère une nouvelle ligne contenant :
 - Les informations personnelles
 - Les détails de l'adhésion pour cette année
 - Les données concernant la réunion de fin d'année
- Enfin, toutes ces lignes sont réunies dans un tableau unique et ordonné, avec une entrée par membre et par année.

2.5.3 Résultat du traitement initial

Le document final produit, nommé `donnees_transformees.xlsx`, présente des informations propres, réorganisées et utilisables provenant de l'association.

Chaque entrée dans le fichier représente l'inscription annuelle d'un membre, incluant l'ensemble des données nécessaires :

- Information sur le membre : nom, prénom, date de naissance, adresse, numéro de téléphone...
- Détails de l'inscription : année, date d'inscription, montant, don, mode de paiement, fonction éventuelle dans le bureau

Détails de participation à l'événement du 21 décembre 2024 : heure d'arrivée, heure de départ, régime alimentaire, remarques, présence, repas

2.5.4 Formatage des dates

Une fois que le fichier final, `donnees_transformees.xlsx`, a été produit, nous avons pris le soin de vérifier manuellement le format des dates.

Certaines colonnes, comme `Date_Adhesion` ou `Date de naissance`, affichaient les dates sous un format incluant l'heure `00:00:00` qui était automatiquement ajoutée par Excel ou Python (par exemple : `2023-03-15 00:00:00`).

Pour rendre les informations plus claires, toutes les dates ont été converties manuellement vers un format simplifié « Date courte » :

`jj/mm/aaaa` (par exemple : `15/03/2023`).

Cette étape garantit une présentation des données plus claire et les rend plus accessibles.

Agence Borcelle, 123 Anywhere St., Any City

