

# RAPPORT



## CONCEPTION ET IMPLÉMENTATION D'UNE BDDR



S2 ; 2023-2024  
BUT SD Dole : SAE 2.0 I

Hamza KABILOVIC  
[dorine.tabary@univ-fcomte.fr](mailto:dorine.tabary@univ-fcomte.fr)

RAPPORT .....	1
ANALYSE DES DONNÉES .....	3
Analyse.....	3
VALIDATIONS DES DIFFÉRENTES DF.....	4
2 NF .....	4
3 NF .....	5
LES LIENS :.....	5
CREATION DES REQUÊTES / VISUALISATION .....	7
Quel est le nombre d'adhérents par année ?.....	7
Quel est le nombre d'adhérents par Ville ?.....	7
Quel est le nombre d'adhésions par adhérents ?.....	8
Quel est le montant total des adhésions par type de paiement ? .....	8
Quelle est la moyenne d'âge des adhérents ? .....	8
Quel est le montant moyen des adhésions par années ?.....	9
Quel est le montant total des dons par année ? .....	9



## ANALYSE DES DONNÉES

Tout d'abord, le travail demandé est celui d'analyser les données afin de se rendre compte de ce que l'on a devant soi. Ici nous avons une association AVEC qui est partiellement subventionnée par les sommes versées par les adhérents sous forme d'adhésion ou de don. AVEC enregistre chaque année ces informations dans un fichier Excel.

Le but du travail demandé ici est de créer une base de données qui permettra le suivi des adhésions et des dons au fil des années, avec un niveau de granularité plus fin que celui qui existe aujourd'hui, et de présenter un tableau de bord basé sur les résultats des requêtes formulées par l'association

### Analyse

Après une analyse préliminaire, nous avons identifié que certaines données de ce rapport ne respectaient pas la première forme normale (1NF) et compromettaient la cohérence des données.

**Téléphones multiples** : Une personne possédait deux numéros de téléphone différents. Nous avons donc créé un nouvel attribut nommé "Téléphone Bis" pour résoudre ce problème de double téléphone.

**Enregistrement invalide** : Une personne avait un enregistrement contenant seulement le caractère "/". Cet enregistrement a été supprimé.

**Dons mal formatés** : Certaines personnes avaient des enregistrements de dons suivis d'un signe "€" ou d'un "?", ce qui compromettait la cohérence des données. Nous avons supprimé ces signes pour ne laisser que des valeurs numériques cohérentes.

**Paiement** : certaines choses voulant dire la même chose ne sont pas écrites pareil, on uniformise..

**Attribut Bénévole** : Certaines personnes avaient "OUI, Ponctuelle" avec "Ponctuelle" écrit de manière aléatoire. Nous avons uniformisé ces enregistrements pour garantir la cohérence des données.

**Attribut sans enregistrement** : L'attribut "Expr1034" ne contenait aucun enregistrement et a donc été retiré.

**Année** : une année était 3017 donc on l'a changée en 2017

Ces modifications ont permis de valider la conformité de notre base de données à la première forme normale (1NF).

Nous avons donc une relation Universelle :  
 $R = \{\text{Numero\_adherent}^*, \text{Annee}^*, \text{Date\_adhesion}, \text{Montant}, \text{Don}, \text{Paiement}, \text{Nom}, \text{Prenom}, \text{Naissance}, \text{Rue}, \text{Ville}, \text{Latitude\_Ville}, \text{Longitude\_Ville}, \text{Etat}, \text{Telephone}, \text{Telephone\_Bis}, \text{Benevole}\}$

## VALIDATIONS DES DIFFÉRENTES DF

Les différentes DF :

DF1 : { Numero\_adherent\* }  $\rightarrow$  {Nom, Prenom, Naissance, Telephone, Telephone\_Bis, Benevole }

DF2 : {Ville}  $\rightarrow$  {Latitude\_Ville, Longitude\_Ville, Etat}

DF3 : { Numero\_adherent\* }  $\rightarrow$  {Ville, Rue}

DF4 : { Numero\_adherent\*, Annee\* }  $\rightarrow$  { Date\_adhesion, Montant, Don, Paiement }

### 2 NF

La relation initiale est en 1NF (atomicité).

La DF1 ne respecte pas la 2NF, nous allons ainsi procéder à la décomposition de la relation universelle R.

R1 = { Numero\_adherent\*, Nom, Prenom, Naissance, Telephone, Telephone\_Bis, Benevole }

R2 = { Numero\_adherent \*, Annee\*, Date\_adhesion, Montant, Don, Paiement, Rue, Ville, Latitude\_Ville, Longitude\_Ville, Etat}

La DF3 ne respecte pas la 2NF, nous allons ainsi procéder une nouvelle fois à la décomposition de la relation R2.

R2a = { Numero\_adherent\*, Ville, Rue}

R2b = { Numero\_adherent \*, Annee\*, Date\_adhesion, Montant, Don, Paiement, Latitude\_Ville, Longitude\_Ville, Etat}

La DF4 ne respecte pas non plus la 2NF, une nouvelle décomposition est requise.



$R2b1 = \{\text{Numero\_adherent}^*, \text{Annee}^*, \text{Date\_adhesion}, \text{Montant}, \text{Don}, \text{Paiement}\}$

$R2b2 = \{\text{Numero\_adherent}^*, \text{Annee}^*, \text{Latitude\_Ville}, \text{Longitude\_Ville}, \text{Etat}\}$

Nous avons ainsi 4 relations majeures à « garder » : R1, R2a, R2b1, R2b2. Mais il manque une DF non utilisée, donc on va réaffiner.

### 3 NF

La relation est en 2NF.

Il existe une DF2 non utilisée qui peut être utilisée. Nous allons procéder à la décomposition des tables. Nous avons ainsi ;

$R3 = \{\text{Ville}^*, \text{Latitude\_Ville}, \text{Longitude\_Ville}, \text{Etat}\}$

$R3a = \{\text{Numero\_adherent}^*, \text{Année}^*\}$

Nous avons ainsi décomposé notre relation Universelle et nous n'avons plus de DF non utilisées, nous nous retrouvons avec :

### LES LIENS :

Un lien 1 à N existe entre VILLE.Ville et ADRESSE.Ville car plusieurs personnes peuvent vivre dans une seule et même ville.

Un lien 1 à N existe entre ADHERENT\_BASE.Numero\_adherent et ADHESION.Numero\_adherent car une personne peut être inscrite sur plusieurs années.

Un lien 1 à 1 existe entre ADHERENT\_BASE.Numero\_adherent et ADRESSE.Numero\_adherent car une personne possède une seule adresse.

Ce qui nous donne comme décomposition des tables :



**Ainsi, la décomposition a abouti à des relations normalisées suivantes :**

**ADHERENT\_BASE** pour les informations de base des adhérents.

**VILLE** pour les coordonnées géographiques et l'état des villes.

**ADRESSE** pour associer les adhérents à leurs adresses.

**ADHESION** pour enregistrer les détails des adhésions annuelles des adhérents.



Nous devons ensuite créer une base de données relationnelle à partir de ces données, J'ai décidé de commencer sur phpMyAdmin afin de pouvoir créer facilement les tables et les requêtes car étant plus user-friendly de mon opinion. J'ai ainsi créer ce code :

<https://drive.google.com/file/d/1AzMkmtGkEBCJQds7eX00iuUI1aHTUKSt/view?usp=sharing>

Malheureusement, sa création fuit en vain, en effet, aucun site que j'ai trouvé ne réussissait à me l'importer en entier en raison de sa taille, malgré le fait d'essayer de le compresser, rien y fait, je n'ai pas réussi, j'aurais bien voulu faire toute la SAE 201 sur MariaDB mais ce ne serait pas le cas, pour les requêtes je vais désormais passer à ACCESS qui d'une façon plus simple, importe les données et fait les liaisons entre les tables.

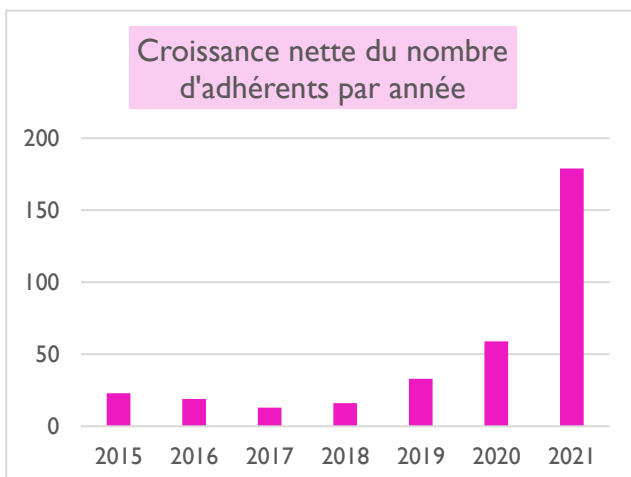
Nous allons passer à la création de requêtes permettant de répondre à des questions existentielles afin de bien comprendre et analyser les données que nous avons, et nous allons aussi créer la visualisation de celles-ci.

## CREATION DES REQUETES / VISUALISATION

Quel est le nombre d'adhérents par année ?

Pour répondre à cette question nous avons la requête ci-contre ;

```
SELECT Annee, COUNT(DISTINCT
Numero_adherent) AS AdherentsDifférents
FROM ADHESION
GROUP BY Annee;
```



Ce graphique nous permet de voir la nette évolution du nombre d'adhérents qui ont été dans l'association. une petite baisse dans les années 2017-2018 qui contraste avec une explosion du nombre d'adhérents plus tard

Quel est le nombre d'adhérents par Ville ?

Nous avons donc la requête :

```
SELECT VILLE.Ville, COUNT(DISTINCT
ADRESSE.Numero_adherent) AS
AdherentsDifférents
FROM ADRESSE AND VILLE WHERE
VILLE.Ville = ADRESSE.Ville
GROUP BY VILLE.Ville;
```

Cela n'a pas de sens de faire une visualisation sur cette étant donné le fait que il existe une cinquantaine de villes différentes dans le fichier. Voici les résultats les plus extrêmes, Jackson = 10 personnes différentes ! Frankfort = 9 personnes. Raleigh Lansign et Salt Lake City à égalité, 7.

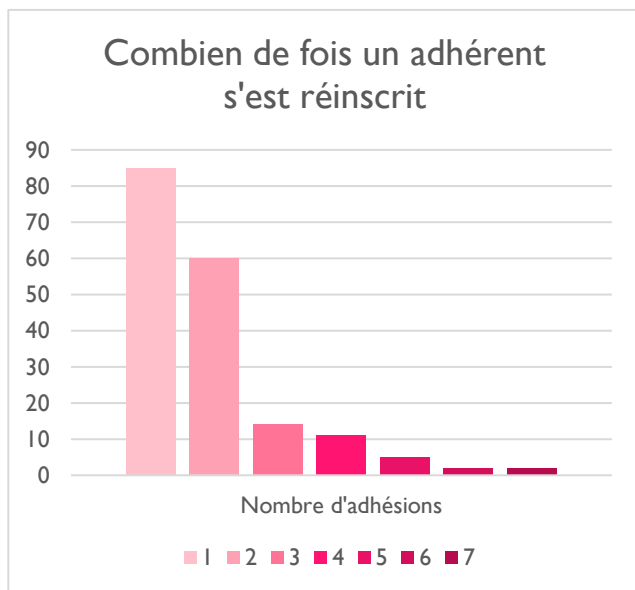
Six villes se partagent la couronne du « 1 seule personne habite ici », il s'agit de Boston, Des Moines, Austin, Oklahoma City, Springfield et Pierre.

## Quel est le nombre d'adhésions par adhérents ?

Cette question relève du fait que certains adhérents se réinscrivent plusieurs années de suite, ainsi, nous aimerions savoir la disparité qui existe entre les adhérents « One Time » et les autres.

La requête qui nous permet de savoir ces choses, la voilà, la visualisation sera faite d'une différente façon, la requête et la visualisation sont complémentaires :

```
SELECT Numero_adherent, COUNT(*) AS
NombreAdhesions
FROM ADHESION
GROUP BY Numero_adherent;
```

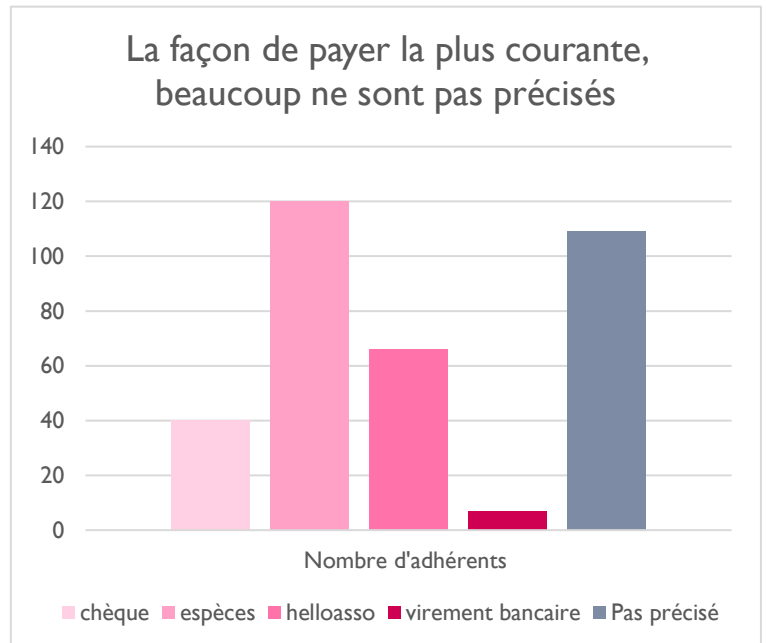


Nous pouvons voir un bon nombre d'adhérents qui sont venus 2 années différentes, sinon, les chiffres baissent plus le nombre d'années monte, ce qui était attendu.

## Quel est le montant total des adhésions par type de paiement ?

Cette question nous demande quel type de paiement est le plus courant, et le plus généreux, voici la requête qui nous permet d'avoir les informations nécessaires.

```
SELECT Paiement, SUM(Montant) AS MontantTotal
FROM ADHESION
GROUP BY Paiement;
```



Nous apercevons grâce à ce graphique la prédominance des espèces dans les modes de paiements préférés des adhérents, beaucoup ne sont pas précisés tout de même

## Quelle est la moyenne d'âge des adhérents ?

Voici la requête nous permettant d'avoir cette information

```
SELECT AVG(TIMESTAMPDIFF(YEAR, Naissance,
CURDATE())) AS AgeMoyen
FROM ADHERENT_BASE;
```

Cette requête calcule l'âge de chaque adhérent en années en soustrayant (TIMESTAMPDIFF) leur date de naissance de la date actuelle (CURDATE()), puis prend la moyenne de ces âges pour obtenir l'âge moyen des adhérents.

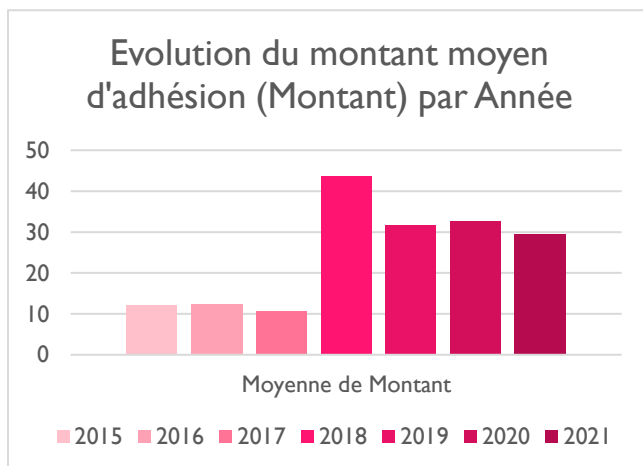


## Quel est le montant moyen des adhésions par années ?

Cette question est intéressante du fait que nous aimerions bien savoir si le fait d'avoir plus de gens au fil des années a une influence sur le montant donné lors des adhésions ou non.

La requête qui nous permet de savoir ces choses, la voilà, la visualisation sera faite d'une différente façon, la requête et la visualisation sont complémentaires :

```
SELECT Annee, AVG(Montant) AS
MontantMoyen
FROM ADHESION
GROUP BY Annee;
```



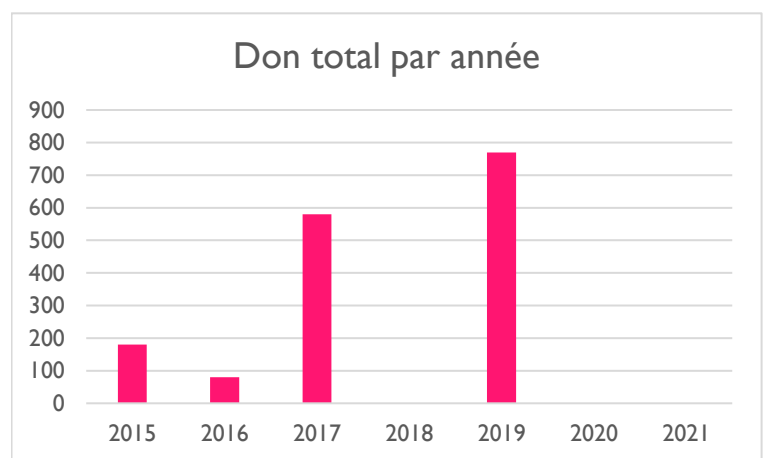
Nous pouvons voir qu'effectivement, en allant de l'avant, la plupart

## Quel est le montant total des dons par année ?

Intéressant de savoir si les gens sont plus généreux au fil des années et que peut-on espérer durant les prochaines années, faire des prévisions selon le nombre etc serait une bonne piste d'études.

La requête qui nous permet de voir la voici :

```
SELECT Annee, SUM(Don) AS MontantTotalDons
FROM ADHESION
GROUP BY Annee;
```



Nous pouvons voir que certaines années, il n'y a pas eu de dons mais que quand il y'en a, c'est significativement plus que la dernière année où il y en a eu, toujours en constante augmentation.

En conclusion, ce projet a permis de transformer une base de données Excel utilisée par l'association AVEC en une base de données relationnelle plus structurée et efficace. En suivant les principes de normalisation, nous avons créé des tables interconnectées pour mieux gérer les informations sur les adhérents, leurs adresses, villes et adhésions. Les requêtes SQL élaborées fournissent des insights précieux sur les tendances d'adhésion et de dons, aidant l'association à mieux comprendre et servir ses membres.