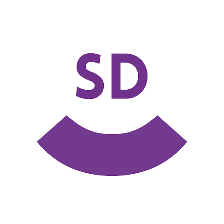
03/06/2024

Clovis Quechon

Conception et implémentation d’une base de données

Saé 2.01



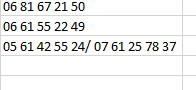
**Présentation du sujet**

Nous avons en notre possession une base de données d’une association en format d’un Excel, notre but est de crée une base de données relationnelle puis de structurer les données pour rendre leur exploitation optimale, être capable intégrer des nouvelles tout en actualisant notre travail via le site “GitHub”.

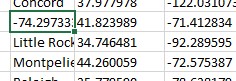
**Traitement des données**

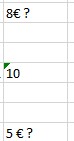
Notre première tâche était de prendre connaissance du Excel, de le modifier pour l’optimiser et de repérer les problèmes possibles dans les données, pour améliorer l'utilisation des données j'ai pris l’initiative d’ajouter une colonne années pour les adhésions. D'après ma première vérification j'ai remarqué les problèmes suivants :

-un numéro de téléphone possédé 2 valeurs différentes j’ai donc supprimé la 2eme valeurs



-Dans le nom de la ville j il y avait une coordonnée j’ai donc supprimé la valeur





-Dans les montants il y avait des « ? » je les ai donc supprimés aussi

**Types des variables**

Numéro adhérent\* (Integer)

Nom (String)

Prénom (String)

Date de naissance (Date)

Rue (String)

Ville (String)

Latitude (String)

Longitude (String)

Etat (String)

Téléphone (String)

Bénévole (String)

Date adhésion\* (Date)

Montant (Integer)

Méthode de payement (String)

Don (Integer)

Année\* (Integer)

Le numéro de téléphone, la latitude et la longitude sont de types string car ils peuvent être mal interpréter par la machine si les variables commencent par un 0, cela permet d’être plus souple avec ses variables

**Normalisation**

R = {Numéro, Nom, Prénom, Date de naissance, Téléphone, Bénévole, Ville, Rue, Latitude, Longitude, Etat, Date adhésion, Année, Montant, Don}

DF1 : Numéro à Nom, Prénom, Date de naissance, Téléphone, Bénévole, Ville, Rue, Latitude, Longitude, Etat

DF2 : Ville, Rue à Latitude, Longitude, Etat

DF3 : Date adhésion, Année, à Montant, Don

Clé candidate {Numéro Adhérent, Date adhésion, Année}

* La table n’était pas 1 NF (Numéro de téléphone) car la tables n’était pas atomique donc j’ai corrigé la table en supprimant les cases contenants plusieurs valeurs

- La table n’était pas 2 NF car une clé primaire définissait des clés non primaires

(DF1) R1{Numéro, Nom, Prénom, Date de naissance, Téléphone, Bénévole, Ville, Rue, Latitude, Longitude, Etat}

R1.2{Numéro, Date adhésion, Année, Montant, Don}

(DF2) Respect de la 2NF

(DF3) R2{Date adhésion, Année, Montant, Don}

R2.2{Numéro, Date adhésion, Année}

Les tables [R1, R2 et R2.2] sont 2NF

- La table est 2NF mais n’était pas 3 NF car une clé non primaire définissait des clés non primaires

La table R1 n’est pas 3 NF selon la DF2 :

(DF1) Ok

(DF2) R3{Ville\*, Rue\*, Latitude, Longitude, Etat}

R3.2{Numéro\*, Nom, Prénom, Date de naissance, Téléphone, Bénévole, Ville, Rue}

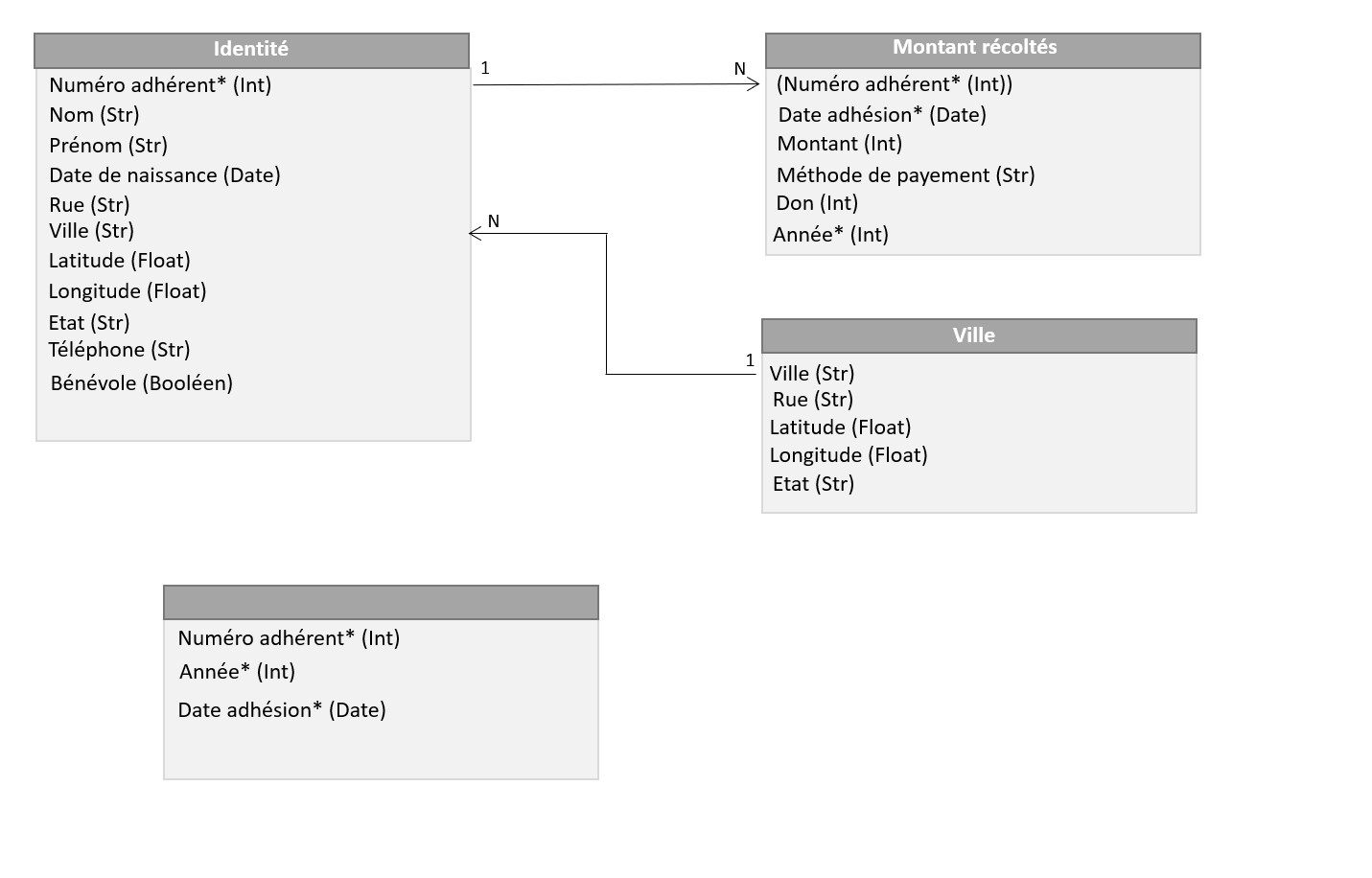
(DF3) Ok

Les tables [R2, R2.2, R3 et R3.2] sont 3NF

- La table est 3NF et aucun non primaire défini des éléments de clés primaire donc elle est BCNF

**Schéma relationnelle**

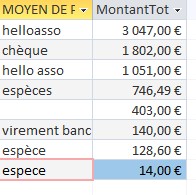
En suivant ma normalisation j’obtiens se schéma relationnelle découpé en 3 parties :



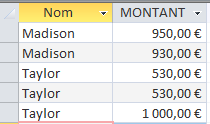
**Les premiers résultats des requêtes**

Pour pouvoir mieux utiliser les données j'ai adapté mon Excel pour pouvoir l'importer dans Access selon mes modifications

Après ma création de mon ma base de données dans Access et la création de mes requêtes j'ai rencontré un problème avec l’orthographe du moyen de paiement



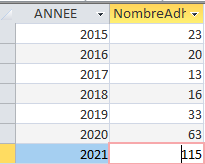
**Personne ayant donné un montant > 500**

SELECT Identité.Nom, adherent.MONTANT

FROM Identité INNER JOIN adherent ON identité.[NUMERO ADHERENT] = adherent.[NUMERO ADHERENT]

WHERE adherent.MONTANT > 500;

Cette requête permet de voir les plus gros montant avec le nom de l’adhèrent, on retrouve Un Nom qui se répète

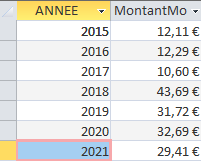
**NBAdherents par année**

SELECT ANNEE, COUNT(\*) AS NombreAdherents

FROM adherent

GROUP BY ANNEE;

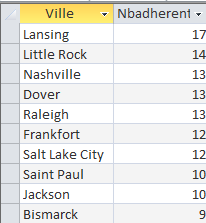
Le nombre d’adhérents a augmenté au fil des années cela montre qui l’association marche bien et prend de l’ampleur

**Le montant moyen par année**

SELECT AVG(MONTANT) AS MontantMoyen

FROM adherent;

On remarque que la moyenne des montant a doublé à partir de 2018 mais malgré le nombre élevé d’adhérents en 2021 le montant lui a même baissé comparé à 2020

**NB adherent par villes**

SELECT Identité.Ville, Count(adherent.[NUMERO ADHERENT]) AS Nbadherent

FROM Identité LEFT JOIN adherent ON Identité.[NUMERO ADHERENT] = adherent.[NUMERO ADHERENT]

GROUP BY Identité.Ville

ORDER BY Nbadherent DESC;

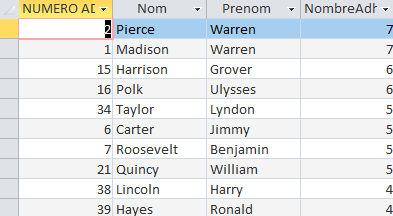
Ayant une longue liste de ville ou habite les adhérents on remarque qu’ils sont repartis dans tout l’Amérique du Nord avec certaines villes qui se différencie des autres (la capture d’écran ne contient que les 10 villes contenantes le plus d’adhérents)

**Nombre d’année des adhérents**

SELECT adherent.[NUMERO ADHERENT],Identité.Nom,Identité.Prenom,COUNT(adherent.[NUMERO ADHERENT]) AS NombreAdhesions

FROM adherent INNER JOIN Identité ON adherent.[NUMERO ADHERENT] = Identité.[NUMERO ADHERET]

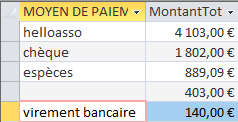
GROUP BY adherent.[NUMERO ADHERENT], Identité.Nom, Identité.Prenom

****ORDER BY NombreAdhesions DESC;

Cette requête permet de voir la fidélité des adhérents et aussi l’arrivée des nouveaux adhérents On remarque que des personnes soutiennent la cause depuis longtemps (la capture d’écran ne contient que les 10 adhérents qui ont le plus de nombre d’adhérassions)

**Totaux des montants des moyens de payements**

SELECT adherent.[MOYEN DE PAIEMENT],SUM(adherent.MONTANT) AS MontantTotal

FROM adherent

GROUP BY adherent.[MOYEN DE PAIEMENT]

ORDER BY SUM(adherent.MONTANT) DESC;

Nous retrouvons des modes de payements qui sont plus utilisés que les autres comme « helloasso » nous avons aussi des montants avec aucun mode de payement qui n’était pas spécifié

**Moyenne d’âge**

SELECT AVG((Date() - [DATE DE NAISSANCE] )/ 365) AS AgeEnAnnees

FROM Identité;

La moyenne d’âge est proche du début de la quarantaine

**Visualisations**

Visualisations