Table des matières

[**Contexte du projet** 2](#_Toc188508222)

[**1.** **Situation initiale** 2](#_Toc188508223)

[**2.** **Situation souhaitée au terme du projet** 2](#_Toc188508224)

[**Description du fichier des écritures** 3](#_Toc188508225)

[**1.** **Frais versus Autres opérations comptables** 3](#_Toc188508226)

[**2.** **Harmonisation des types d’écriture** 4](#_Toc188508227)

[**Les spécifications fonctionnelles et techniques du système attendu** 5](#_Toc188508228)

[**1.** **Définition du bilan comptable** 5](#_Toc188508229)

[**2.** **Les fonctionnalités indispensables** 5](#_Toc188508230)

[**3.** **Les technologies à utiliser impérativement** 6](#_Toc188508231)

[**Travail demandé** 6](#_Toc188508232)

[**1.** **Un modèle de données déjà amorcé** 6](#_Toc188508233)

[**2.** **Ce qu’on attend du data engineer** 7](#_Toc188508234)

[**3.** **Les étapes de la construction de la solution** 7](#_Toc188508235)

[**Exercice 1 : Chargement des fichiers dans les tables (2 pts)** 8](#_Toc188508236)

[**Exercice 2 : nettoyage des colonnes (8 pts)** 8](#_Toc188508237)

[**Exercice 3 : création des contraintes sur les colonnes de tables (4 pts)** 9](#_Toc188508238)

[**Exercice 4 : vers l’établissement du bilan comptable (6pts)** 9](#_Toc188508239)

**Contexte du projet**

## **Situation initiale**

La Société DATA360 est une TPE (Très Petite Entreprise) spécialisée dans le conseil et l’expertise en système informatique. Elle emploie cinq personnes. Les cinq salariés jouent un rôle majeur dans la croissance de l’entreprise. Ils ont chacun un rôle d’opérationnel mais aussi un rôle de commercial.

En tant qu’opérationnel, chaque salarié intervient auprès de clients pour la réalisation des projets informatique. En tant que commercial, chaque salarié est à l’écoute des opportunités que peut offrir le marché de l’informatique. Ils agissent ainsi comme apporteur d’affaires.

Les activités de l’entreprise génèrent un chiffre d’affaires. Ce chiffre d’affaires permet à la fois de rémunérer les salariés mais aussi de dégager les fonds de roulement (fonds de gestion) nécessaires pour le bon fonctionnement de l’entreprise, en particulier le financement des diverses activités liées à la prospection commerciale.

Jusqu’en décembre 2023, la TPE ne dispose pas d’un système performant pour gérer les frais et autres opérations comptables liés à son fonctionnement. Toutes les écritures sont renseignées dans un fichier Excel. Le fichier Excel est disponible dans sur un lecteur réseau de type drive partagé et accessible par l’ensemble des salariés. Chaque salarié y a accès en lecture et écriture. Les saisies dans le fichier présentent des coquilles, certaines informations ne sont pas uniformes. Ceci rend compliquée :

* Le contrôle et la validation des frais
* La production du bilan comptable
* Le suivi de l’évolution du chiffrage d’affaires.

## **Situation souhaitée au terme du projet**

DATA360 s’est récemment dotée d’un système informatique qui lui permet aujourd’hui d’automatiser et de sécuriser un certain nombre de tâches qui jusque-là étaient manuelles. C’est dans ce contexte qu’elle a décidé de rendre automatique dans une certaine mesure :

* La saisie et la consultation de tous les frais
* La saisie et la consultation des autres opérations comptables
* La génération du bilan comptable
* Le suivi de l’évolution du chiffre d’affaires
* La prédiction du chiffre d’affaires dans le future

Après avoir murement défini ses besoins, les 5 salariés sont unanimes pour dire qu’une telle charge de travail peut être absorbée en un mois en y travaillant à temps plein. Pour la réalisation de ce projet, le profil d’un informaticien ingénieur de données ou Data Engineer est nécessaire.

Les paragraphes suivants vont décrire dans les moindres détails le fichier Excel actuellement utilisés.

**Description du fichier des écritures**

Le fichier Excel qui est utilisé est composé de plusieurs onglets qui chacun contient des données qui peuvent être des frais ou d’autres opérations comptables. Ci-dessous la liste des onglets :

|  |  |
| --- | --- |
| Nom onglet | Description |
| FAC-Banque | Il s’agit du recensement des factures bancaires |
| Ecritures | Contient la liste de ce que la société nomme frais |
| Factures | Contient la liste des factures encaissées par la TPE |
| Salaires versés | Contient la liste des montants versés aux salariés au titre de leur rémunération ou prime |

Comme mentionné plus haut, les saisies dans ce fichier ne sont pas uniformes. Le règlement d’une facture bancaire est une opération comptable, de même l’encaissement d’une facture cliente. Mais au final qu’est-ce qu’un frais et qu’est-ce qu’une opération comptable ? Pourquoi distinguer les deux ?

## **Frais versus Autres opérations comptables**

Un frais fait référence à une note de frais. Il s’agit d’une dépense effectuée par un porteur de la carte bancaire de l’entreprise au nom de l’entreprise. Pour cette dépense un justificatif est obligatoire. Ce dernier peut être une facture ou une toute autre note sans mention du détail. Dans les deux cas les informations qu’on souhaite voir sur le justificatif et à intégrer dans le système sont :

* La date de la dépense
* Le type de la dépense
* Les montants HT, TVA et TTC
* Le moyen de paiement utilisé pour le règlement du frais
* Auprès de quel établissement la dépense a été réalisée
* Le nom qu’on donne à ce frais

On entend par « autres écritures comptables », principalement des encaissements issus des factures client. L’ensemble des encaissements constituent ce qu’on appelle en gestion le Chiffre d’Affaires (CA).

## **Harmonisation des types d’écriture**

Dans le fichier Excel plusieurs écritures sont saisies avec des types totalement différents. Il est temps d’harmoniser les choses en commençant par définir des catégories et sous catégories d’écriture. DATA360 souhaite intégrer dans son système :

| **Catégorie** | **Sous-Catégorie** | **Type de l’opération** |
| --- | --- | --- |
| Trésor Public – Impôts - Taxes | TVA  TVS  IS (Impôt sur la Société)  CFE  RSI  URSSAF  CIPAV  Amende | Décaissement  (Crédit) |
| Auto | Loyer leasing (LLD ou LOA)  Carburant  Réparation auto  Péage  Achat nouvelle clé  Révision auto |
| Restauration | Restauration  Facture Cantine |
| Assurances | HISCOX  MAAF  MACIF |
| Consommable Informatique | Abonnement logiciel  Achat matériel informatique  Achat licence logiciel informatique |
| Formation | Achat cours en ligne  Achat Livres |
| Fonctionnement Entreprise | Création entreprise  Facture Comptable  Achat KBIS  Dépôt de comptes annuels  Facture Juridique  Frais bancaires |
| Confort | Achat vestimentaire  EDF  Forfait mobile  Forfait Internet  Forfait box  Séjour professionnel |
| Rémunération | Salaire  Avance sur salaire  Prime |
| Encaissement | Facture client | Encaissement  (Débit) |

Chaque écriture, dans le nouveau système, devra être rangée dans une sous-catégorie.

**Les spécifications fonctionnelles et techniques du système attendu**

## **Définition du bilan comptable**

Le bilan d’une entreprise est un tableau qui montre sa situation patrimoniale et financière à un instant T, en général la clôture des comptes.​ Il montre ce que l’entreprise possède (actif) et ce qu’elle doit (passif), soit les ressources mobilisées pour financer son actif (emprunts, capital social…).

Le bilan comptable, se compose de deux colonnes : l’actif (à gauche) et le passif (à droite).

L’actif du bilan liste les éléments que possède l’entreprise (l’ensemble de ses biens et de ses créances).

Le passif renseigne sur les créanciers/ayants droit que l’entreprise mobilise pour financer son activité ("à qui appartient quoi et à hauteur de combien" : associés, banques, fournisseurs).

Une image contenant texte

Description générée automatiquement

Ce qui intéresse DATA360 est plus particulière les actifs circulants qui correspondent aux encaissements que l’entreprise fait ainsi que les dettes qui correspondent à toutes les opérations de type décaissement.

## **Les fonctionnalités indispensables**

Le système attendu doit permettre impérativement :

* d’enregistrer et catégoriser **toutes** les opérations
* d’interroger le système pour obtenir les détails de chaque opération
* d’interroger le système pour générer à la volée le bilan comptable de l’année en cours
* d’interroger le système pour générer à la volée le bilan comptable des années précédentes
* d’interroger le système pour calculer à la volée le CA de l’année en cours ainsi que ceux des années précédentes
* Estimer au vue de la connaissance des données le CA des années à venir

## **Les technologies à utiliser impérativement**

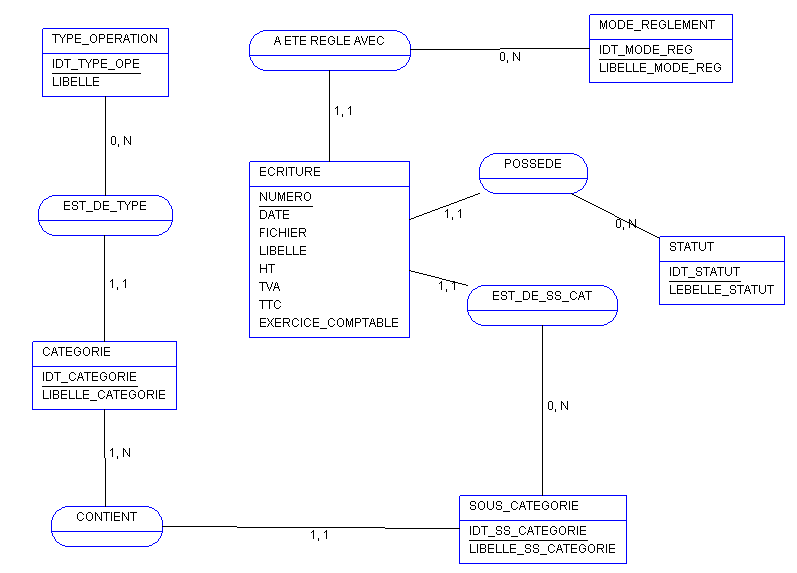
La TPE a acquis récemment dans son système informatique un certain nombre de logiciels qu’elle souhaite utiliser pour construire le nouveau système comptable. Ce sont :

* PostgreSQL version 16
* PgAdmin

**Travail demandé**

## **Un modèle de données déjà amorcé**

Un salarié de la société a pris le sujet au sérieux et a proposé une modélisation pour répondre à ce besoin. Ci-dessous le MCD produit, il est cohérent avec le besoin.



Le MLD correspond est le suivant :

TYPE\_OPERATION(IDT\_TYPE\_OPE, LIBELLE)

CATEGORIE(IDT\_CATEGORIE, LIBELLE\_CATEGORIE, #IDT\_TYPE\_OPE)

SOUS\_CATEGORIE(IDT\_SS\_CATEGORIE, LIBELLE\_SS\_CATEGORIE, #IDT\_CATEGORIE)

MODE\_REGLEMENT(IDT\_MODE\_REG, LIBELLE\_MODE\_REG)

STATUT(IDT\_STATUT, LIBELLE\_STATUT)

ECRITURE(NUMERO, DT\_ECRITURE, FICHIER, LIBELLE, HT, TVA, TTC,EXERCICE\_COMPTABLE, #IDT\_MODE\_REG, #IDT\_STATUT, #IDT\_SS\_CATEGORIE)

## **Ce qu’on attend du data engineer**

On attend du data engineer une implication totale dans le projet et une force de proposition. Il doit :

* Comprendre le besoin exprimé dans le présent document
* Comprendre le fichier Excel existant
* Créer une base de données
* Créer les tables qui correspondent au MLD
* Proposer une méthode optimale pour transférer les données de la feuille Excel vers le nouveau système
* Produit les informations sur la base des requêtes SQL

## **Les étapes de la construction de la solution**

Comme dans tous les projets informatiques, le chef de projet va construire un planning avec le data engineer afin de suivre les différents jalons du projet. Les étapes macro de ce planning pourraient être :

* Nettoyer le fichier Excel actuel qui est fourni
* Créer les tables du MPD
* Choix de la méthode optimale de transfert des données existantes
* Transférer les données dans une table de travail
* Split des données dans les tables du MPD
* Choix de la bonne stratégie de création des rapports de bilans (requêtes SQL)

Dans les détails, on procéderait de la manière suivante, sous forme d’exercice :

### **Exercice 1 : Chargement des fichiers dans les tables (2 pts)**

* Créer les tables du MPD
  + Consignes à respecter :
    - toutes les colonnes sont de type TEXT
    - aucune contrainte PK et FK ne doit être créées à ce stade
* Alimenter chacune des tables avec le contenu des fichiers csv fournis
  + 6 fichiers CSV pour les 6 tables du MPD

### **Exercice 2 : nettoyage des colonnes (8 pts)**

1. Après avoir exécuté les requêtes ci-dessous, dites quelles actions ont-elles sur les données et la structure de la table ECRITURE :

update ecriture set ht = replace(ht, ' ', '');

update ecriture set ht = replace(ht, ',', '.');

update ecriture set tva = replace(tva, ' ', '');

update ecriture set tva = replace(tva, ',', '.');

update ecriture set ttc = replace(ttc, ' ', '');

update ecriture set ttc = replace(ttc, ',', '.');

1. Proposez une ou plusieurs requêtes qui change le type de données des colonnes **HT, TVA, TTC** de **TEXT** à **DECIMAL**.
2. Proposez une ou plusieurs requêtes qui change le type de données des colonnes **IDT\_MODE\_REG, IDT\_STATUT** et **IDT\_SS\_CATEGORIE** de **TEXT** à **INTEGER**.
3. Proposez une requête SQL qui supprime toutes les écritures qui n’ont pas de date d’écriture.
4. Proposez une requête qui renomme la colonne **DT\_ECRITURE** en **DATE\_ECRITURE**
5. Proposez une requête qui change le type de **DATE\_ECRITURE** de **TEXT** à **DATE**.
6. Proposez une requête SQL qui supprime toutes les écritures qui n’ont pas de **NUMERO**
7. Changez le type de données de la colonne **EXERCICE\_COMPTABLE** de **TEXT** à **INTEGER**.

### **Exercice 3 : création des contraintes sur les colonnes de tables (4 pts)**

Proposez des requêtes sql pour répondre aux besoins de contraintes suivants :

* Chaque table doit disposer d’une contrainte de clé de primaire conformément au MLD
* Créer les contraintes de clés étrangères si besoin conformément au MLD
* Les colonnes suivantes ne doivent pas admettre du null :
  + Au niveau de la table ECRITURE : DATE\_ECRITURE, EXERCICE\_COMPTABLE, TTC
  + Au niveau de la table MODE\_REGLEMENT : LIBELLE\_MODE\_REG
  + Au niveau de la table STATUT : LIBELLE\_STATUT
  + Au niveau de la table CATEGORIE : LIBELLE\_CATEGORIE
  + Au niveau de la table SOUS\_CATEGORIE : LIBELLE\_SS\_CATEGORIE

A ce stade toutes les tables sont correctes en structure et en données. Les contraintes de clés primaire ainsi que les clés étrangères ont été créées.

### **Exercice 4 : vers l’établissement du bilan comptable (6pts)**

Il vous est demandé à présent de :

1. Exécutez et commentez la requête suivante :

select distinct exercice\_comptable from ecriture order by 1;

1. Proposer une requête SQL qui liste le nombre d’écritures ainsi que la somme des montants HT, TVA, TTC des écritures par année comptable. Cette liste doit être triée sur l’année de l’exercice comptable, de la plus récente à la plus ancienne.
2. Proposer une requête SQL qui affiche le montant moyen TTC des écritures de catégorie « Fonctionnement Entreprise »
3. Proposer une requête SQL qui affiche le nombre, la somme HT, la somme TTC par type d’opération
4. Proposez une requête SQL qui permet d’afficher le bilan comptable de l’année 2022 sous la forme (ne pas tenir compte des montants affichés, seule la forme compte ) :

Une image contenant texte

Description générée automatiquement