

Projet : quels impacts, enjeux et caractéristiques du lien entre la santé et l'économie ?

Auteur : Grégoire fuchs

Consignes :

Projet Final CS50 SQL

Le projet final du cours CS50 SQL consiste à développer votre propre base de données en utilisant SQL. Vous êtes **encouragé à créer quelque chose** qui vous intéresse, qui résout un problème substantiel, qui **impacte positivement** les autres ou qui change le monde. Vous pouvez **collaborer avec un ou deux camarades, mais chaque membre du groupe doit contribuer** équitablement.

Quand le réaliser

Le projet doit être soumis avant **le 10 mai 2025 à 23h59**.

Idées de projets

Voici quelques idées pour vous inspirer :

- Une base de données pour trouver vos chansons préférées, représentant des artistes, des playlists, des albums, etc. (comme Spotify ou Apple Music).
- Une base de données pour gérer vos finances personnelles, stockant les soldes des comptes bancaires, les transactions, les budgets, etc. (comme Mint, Quicken, ou votre application bancaire préférée).
- Une base de données pour aider les gens à trouver des amis lorsqu'ils déménagent dans une nouvelle ville, représentant des personnes, des villes, des événements, des connexions, etc. (comme Meetup ou Bumble BFF).

Si vous vous inspirez d'informations trouvées sur internet, <u>veuillez citer d'où viennent vos informations</u>.

Spécifications

Votre projet final doit être composé de trois fichiers :

- ****Document de conception** (en PDF ou Markdown) : Un document rigoureux décrivant le but, la portée, les entités, les relations, les optimisations et les limitations de votre base de données. Il doit inclure un diagramme de relation d'entités. Ce document doit expliquer vos choix de conception et comment vous avez équilibré la création et la manipulation de la base de données. Vous ajouterez un diagramme d'entité-relation décrivant vos tables. Vous pouvez l'écrire directement dans le Readme à l'aide de [mermaid](#).**

Document de conception : **=faire d'abord le schéma entité relation !**

Sommaire :

I. But et utilité de la base de donnée : Pourquoi avoir créer cette base de donnée ?	4
2. But et enjeu:	5
3. Contexte de création et tables créées.	7
Etape 1 :	7
2. Acteurs :	7
1. Entreprises :	7
2. Individu	7
3. Pourquoi les utiliser ?	8
4. Application 1 /2/3	8
Etape 2 : 1. Sport et effet économiques ?	8
1. Table 1 : sport_sante : Activité physique et sédentarité	8
2. Table des effets du sports :	9
3. Tabagisme	11
1. Table 1 : Tabac :	11
Effets du sommeil sur la productivité :	12
4. Alimentation et hydratation	13
1. Table 1 : table de constats d'une santé fragilisée	13
2. Table des effets de l'alimentation.	13
Sur la deuxième application, cette partie n'est que partiellement automatisée avec la deuxième application qui gère les effets du côté de l'employé ; et il doit déclarer son niveau de fatigue.	13
4. Evaluer les impacts économiques des troubles du sommeil	15
4. Risquophilie :	16
Table 2 : tracking des effets :	17
6. Santé mentale et stress	18
3. Hypothèses de création	19
3. Niveaux	20
4. Choix de simuler ces données.	20
Variables	21
Limite	22
1. Etablir un lien causal est impossible	22
1. Justifier des clés primaires.	22
2. Justifier les relations de manière générale .	23
1. Un employeur emploie un ou plusieurs employés (mais 1 employé=1 employeur)	23
2. Un employé utilise plusieurs applications 1,2:création d'une table applications ½ (décision de scission)	23
3. Employés vers profils de santé.(1,N)	23
4. Application 3 : mesure par le manager	24
5. Un profil de santé peut avoir un ou plusieurs effets/degrés d'effets dans le temps	24
SCHEMA Mermaid	25

III.Code	26
- **Script de schéma**	31
2.Résultat positif en faisant “Run query”	33
En texte ici :	33
3.	39
Partie 4: script de requêtes :	39

I.But et utilité de la base de donnée :Pourquoi avoir créer cette base de donnée ?

*Cette base de donnée a été **créée dans le but un peu atypique, de tenter de créer des liens entre la santé et l'économie.***

Cette base de données a été conçue dans un objectif **novateur et ambitieux** : explorer les liens complexes et souvent subtils entre **la santé publique** et l'économie, **sur des agrégats tels que le chômage ou encore l'inflation.**
Cette démarche est

2. But et enjeu:

L'intuition initiale **de ce projet** est qu'il existe une **interdépendance profonde** entre ces deux domaines, où l'état de santé d'une population peut influencer **significativement**

1. sa productivité

2. puis celle de l'entreprise: euros/heure perdu.

3. et donc la performance de l'économie entière : perte de %variable PIB/heure par rapport à pib/heure/habitant;

Pourquoi ?

Cette expérience a pour but de mettre en évidence **des liens de régression** un peu inhabituels en économie.

En effet, la relation classique économie => économie ; devient remplacée par **la relation santé=>économie .Par conséquent**, les mesures vont consister, **dans le cadre d'une expérience économique**, en des mesures mesurées par 3 applications :

1.Appli 1 : mesure les causes de santé sur l'employé/les choix

2.Appli2 : mesure les effets perçus par l'employé

3.Appli 3 : mesure les effets perçus par l'employeur

1.Un choix(pas de “maladie”) correspond à des choix du quotidien comme le fait de *boire un verre d’alcool, de tabacs...*qui est mesurée par des variables **autodéclarées**.

2.A un impact_sur_sante dont la gravité à déterminée ; mais mesurée soit grâce aux capteurs du téléphone (microphone pour le sommeil ,accéléromètre , pour le sport) ; soit l’auto-déclaration grâce à la première application téléchargée par l’individu par l’appli1

3. A un impact sur la productivité de l’entreprise :

1.Renseignée du côté de l’employé par l’appli 2 (disponible sur ordinateur/smartphone)

-time tracking : temps de travail

-productivité ;

2.Du côté du manager : l’appli 3 ; qui permet également d’avoir une mesure subjective de la qualité du travail

4. Macroéconomie : ces choix ont donc un impact sur le bonheur des individus autant que sur l’économie par la croissance économique plus élevée !

Au final, **ce projet permettra aussi de mieux mesurer les risques liées à certaines actions du quotidien ; et fréquentes ; mais qui peuvent donc avoir des impacts importants sur la vie professionnelle de l’individu à terme ; voire sur l’économie d’un pays .**

Ainsi, prendre ces mesures ; c'est réaliser de la prévention économique !

3.Contexte de création et tables créées.

Ainsi, *une colonne avec un individu sera mise en place et qui aura donc*

-un identifiant primaire liée au nom prénom afin de garantir l'unicité

-des caractéristiques comportementales liées à la santé et à la prise de risque : prise d'alcool ; nombre de verres d'alcools ;

Etape 1 :

2.Acteurs :

1.Entreprises :

Clé primaire : id_entreprise

; en effet, cela permet d'éviter d'évincer des entreprises elles ont le même nom prénom !

Nom_entreprises

Benefices

Heures

Benefices/heures

2.Individu

Clé primaire : id_individu

Nom/prénom

Benefices

Heures

3. Pourquoi les utiliser ?

Elles **seront par la suite dans toutes les tables afin de bien pouvoir identifier chaque individu précisément ;et garantir l'absence de doublon , notamment en ajoutant des restrictions sur les doublons .**

Clé primaire : id_application

4. Application 1 /2/3

Cléprimaire propre

Disponibilité : tous types/que smartphone (1, 3).

But : ces applications vont permettre de prendre les mesures ; il est donc intéressant de créer une table au départ , que l'on pourra mettre en **clé étrangère dès qu'elle a été utilisée, permettant donc de clairement montrer la source des données prises.**

Etape 2 : 1.Sport et effet économiques ?

1.Table 1 : sport_sante : Activité physique et sédentarité

- sport_sante_appli 1 : date et heure pour l'appli 1 de la fin de l'analyse
- minutesactivite_physique_par_semaine (entier)
permet de connaître l'impact de la durée de sport .
Estimation par l'appli 1 grâce à l'accelerometre En général, l'OMS souligne qu'il faut a minima 20 minutes pour réussir à
- nbjours_sport/semaine: **impact de la régularité du sport** (hyp : + effet sur la santé , notamment sur les *hormones du cerveau comme la dopamine* par la régularité....

- nbtypes_activites_par_semaine (entier) :
impact de la diversité du sport (hyp : +adhésion au sport ...)
- nombre_descalier_monté par jour :minutes :
autodéclarée
- nombre_minutes_debout : **variable autodéclarée**
- nombre_minutes_assis : **pareil**

2.Table des effets du sports :

Littérature : Les coûts individuels moyens de **perte de productivité** sont significativement plus élevés chez les inactifs que chez les actifs selon **Kari, Nerg et All (NIH,**. doi: [10.1249/MSS.00000000000003037](https://doi.org/10.1249/MSS.00000000000003037))

L'inaction face à l'inactivité physique pourrait coûter à la santé publique mondiale environ 47,6 milliards USD par an

- date_heure_declaration_appli2
- gain_productivite_heures
Estime les heures gagnées grâce à une meilleure forme (moins de fatigue, meilleure concentration), mesurable via **un time tracker**, permettant **d'isoler le moment du sport ; et un time tracker** permettant **d'isoler le temps du travail.**
- **sport_absenteisme_jours : par géolocalisation**

comptabilise le nombre d'absence les semaines sans sport de l'individu.

- **pas_sport_absentimsme**
comptabilise le nombre d'absence les semaines avec sport de l'individu.

2. consommation d'alcool correspondra à cette nomenclature :

- *consomme_alcool_ au moins une fois (booléen : Oui/Non*
- *frequence_alcool_jour/**semaine** : 1 à 7*
- *nb_verres_moyen/jouravec**alcool**(nombre moyen de verres par séance)*

Table effets de l'alcool mesurés les jours d'alcoolisme :

maladie_hepatique_alcoolique &

stade_maladie_hepatique : l'alcool est la première cause de cirrhose et de stéatose ; la gravité (stade) est directement corrélée aux coûts de soins et à la mortalité.

hypertension_artérielle : la consommation excessive d'alcool augmente la pression artérielle, facteur de risque cardiovasculaire et source d'arrêts de travail.

battements cardiaques : l'alcool peut endommager le muscle cardiaque, réduisant l'endurance et la capacité de travail.

troubles_mentaux_alcool : les troubles psychiques de dépression ou encore d'**addiction** réduisent la concentration, augmentent l'absentéisme et le présentéisme

nombre_remarquesavertissements_mauvaisequalitedetravail_ : sur un an .

nombre_remarquesmauvaiscomportement :

application_time_tracker_tempsdetravail : à l'aide d'une application de time tracker, vérifier si la personne a bien travaillé ; pendant combien de temps

3. Tabagisme

1.Table 1 :Tabac :

-clé primaire : nom prénom

- **consomme_tabac** (booléen 012: oui:classique/électronique/non)

- **cigarettes_par_jour** (nombre moyen de cigarette par jour)
- **consommation_drogues** (liste/catégorie : Cannabis, Cocaïne, etc.)
- **frequence_drogues_jour** (par catégorie)

Enquêtes soulignant un lien direct sur la santé puis l'économie :

Cost of Cigarette Smoking–Attributable Productivity Losses, U.S., 2018

En 2018, le coût total des pertes de productivité liées à la morbidité attribuable au tabagisme aux États-Unis a été estimé à 184,9 milliards USD

En 2018, les pertes de productivité liées à la morbidité attribuable au tabac ont été estimées à 184,9 milliards USD aux États-Unis, dont 9,4 milliards pour l'absentéisme et 46,8 milliards pour le présentéisme .

Table 2 :

Effets du sommeil sur la productivité :

clé primaire : automatisée :

clé étrangère : **-entreprise** NOT NULL UNIQUE: chaque employé doit avoir un unique employeur ! , et doit pouvoir être relié à une seule entreprise.

-sante_tabac NOT NULL unique: permet notamment de relier à une personne ayant un problème de santé de façon unique et obligatoire.

date_heure

4.Alimentation et hydratation

1. Table 1 : table de constats d'une santé fragilisée

*Cette partie est remplie via une application **qui demande à l'utilisateur, en fin de journée, de faire un récapitulatif des actions effectuées et mises en place !***

- clé primaire automatique : **nom_prenom_personne** :
en effet, **on souhaite une seule mesure par personne !**
- date_heures_declaration_alim
- nb_portions_fruits_legumes_par_jour (entier)
- nb_consommation_fast_food (fréquence)
- nb_eau_litres_par_jour (réel): autodéclaré
- Type de **repas un jour J (appli)**: lourd, sucré, équilibré : avec une photo
- **Degré_de_fatigue_**: autodéclaré

2.Table des effets de l'alimentation.

Sur la deuxième application, cette partie n'est que partiellement automatisée **avec la deuxième application qui gère les effets du côté de l'employé ; et il doit déclarer son niveau de fatigue.**

Sur une troisième application **sa hierarchie peut indiquer le nombre d'erreurs /vitesses d'exécution...**

**Table 2 : effets_sommeil Effets du sommeil sur la
productivité :**

clé primaire : automatisée :

clé étrangère : -entreprise NOT NULL UNIQUE: chaque employé doit avoir un unique employeur ! , et doit pouvoir être relié à une seule entreprise **afin d'estimer l'impact sur son bénéfice/heures..**

-sante_sommeil NOT NULL unique: permet notamment de relier **chaque effet** à une personne ayant un problème de santé de façon unique et obligatoire.

- date_heure

1.niveau_fatigue_apres:échelle subjective

Cette mesure subjective, standardisée sur une échelle simple, permet de lier directement chaque type de repas (lourds, sucrés, équilibrés) à la sensation de **somnolence**

Un tiers de la population souffre de malnutrition, conduisant à une baisse de productivité au travail et à des coûts accrus pour l'employeur .

2.impact_vitesse_exécution_tâche_time_tracking :
la personne est-elle ralentie par la
personne_timetracking

3.Impact_heures_absenteisme

4..impact_erreurs_tâches_appli_entreprises.

5.Evaluation_finale_perte_argent/heure_entreprise_benefice :

4. Evaluer les impacts économiques des troubles du sommeil

1. Constat : **Table des troubles de santé sante_sommeil**
Grâce à l'application de détection de problème de santé, **tant que le smartphone est placé à côté du lit**, les données du sommeil sont **automatiquement estimées par l'application**.

Clé primaire :

- **nom_prénom** : car l'identifiant doit rester unique, ce qui permet la clé primaire

Variables :

- nom_prenom
- date_heure_enregistrement
- heure_debut_sommeil
- heure_fin_sommeil
- heures_sommeil_moyennes (réel)
- qualite_sommeil (échelle de 1 à 20): **déclarée**
- **durée du réveil**

2. Estimer les effets

Effets du sommeil sur la productivité :

clé primaire : automatisée :

clé étrangère : -entreprise NOT NULL UNIQUE: chaque employé doit avoir un unique employeur ! , et doit pouvoir être relié à une seule entreprise.

-sante_sommeil NOT NULL unique: permet notamment de relier à une personne ayant un problème de santé de façon unique et obligatoire.

- date_heure
- nb_erreurs_travail_par_mois : rapporté par un manager qui est mis au courant **par sms par l'appli de "tracking"** de les insomnie, elie directement la fatigue à la performance métier (retards, erreurs) avec preuves/exemples
- .jours_absence_fatigue : mesure l'absentéisme spécifique au sommeil, traduisant un coût direct pour l'employeur.
- reduction_productivite_pourcentage : **en utilisant une variable de la table entreprises;** permet de calculer l'ampleur de la **perte de performance**, et d'estimer son impact économique en heures et en coûts.

4.Risquophilie :

Table 1 : Comportements à risque routier ou domestique, *détecté par l'application d'autodiagnostic obligatoire.*

Clé ID : risques_id

Clé entreprises : entreprises : **permet de faire le lien avec les bénéfices ! sur la productivité (en euros/heures)**

- date_heure_detection_action_risquophilie
- ceinture_**securite_portee** (pourcentage de trajets):déclaration sur l'application

- **vitesse_excessives** (nombre d'infractions ou auto-déclaré) : **détectée par l'application 1 automatiquement**
- **nombre_pauses_travail** : détecté automatiquement
- **probleme_comportements_sante** : détecté par la hiérarchie (appli3); subjectif !

Table 2 : tracking des effets :

clé primaire : automatisée :

clé étrangère : **-entreprise NOT NULL UNIQUE**: chaque employé doit avoir un unique employeur ! , et doit pouvoir être relié à une seule entreprise.

-risque_tabac NOT NULL unique: permet notamment de relier à une personne ayant un problème de santé de façon unique et obligatoire.

date_heure

- **temps_horaire_moyenne: avec time tracking,**

Mesure concrète du nombre de tâches ou d'unités de travail réalisées par heure, **automatisée par application**

- **nombre_erreur_**
- **perte_au_niveau_entreprise**
- **estimation_avec_representativité échelle**

6. Santé mentale et stress

1. "Troubles" ou pas ?

- **Clé primaire: trouble1_id**
- **nom_diagnostic_hôpital**
- **nb_migraines** : Les migraines entraînent souvent des journées complètes d'invalidité, impactant fortement la disponibilité et la performance.
- **depression_lutte**

2. Prise en charge et donc trouble mental vérifiable: **nb_psychologues_consulte**

3. **nb_jours_medication** : **évalue la régularité des effets sur la santé du trouble mental**. En effet, *il est surtout intéressant de voir lorsque la maladie se déclare, et de préférence temporairement (non génétique)*.

TABLE 2 : Effets_tracking(suivi) automatisé:

clé primaire : automatisée :

clé étrangère : **-entreprise NOT NULL UNIQUE**: chaque employé doit avoir un unique employeur ! , et doit pouvoir être relié à une seule entreprise.

-troublemental_tabac NOT NULL unique: permet notamment de relier à une personne ayant un problème de santé de façon unique et obligatoire.

date_heure

- `date_heures_tracking_detection_probleme_santé`
- `jours_absence_stress` et `erreurs_par_mois` : Mesurent l'absentéisme et le présentéisme inefficace liés au stress et aux maux
- **`productivite_horaire_base`** vs. **`productivite_horaire_stress`** : Comparaison directe de la performance pour chiffrer l'effet du stress **grâce au suivi de la gestion du temps sur ordinateur.**
- **`reduction_productivite_perc`** : Synthétise l'impact sur l'efficacité au travail en pourcentage, facilitant l'estimation des coûts économiques associés.
- **`evolution_productivite_prise_en_charge`** : mesure l'évolution de la productivité après prise en charge

De surcroît, **il m'apparaît important de souligner que ces choix restent assez cohérents par rapport à la littérature.**

3.Hypothèses de création

En effet, une population en bonne santé est généralement **plus productive, créative**, et capable de soutenir activement l'économie locale ou nationale. À l'inverse, des problématiques sanitaires importantes peuvent freiner le **développement économique** en entraînant une

- augmentation **des dépenses publiques** de santé de l'Etat
- **indirectement sur le marché du travail** : une diminution de **la productivité du travail des travailleurs**
- en impactant directement **le marché du travail** à travers des **absences fréquentes ou prolongées**.

Ainsi, il apparaît évident **qu'une table "employeur/emploi"** devra être créée afin de prendre en compte les **impacts indirects** d'une santé affaiblie de l'individu !

3.Niveaux

Toutefois, il est important de **souligner** qu'il s'agit d'un point de vue très macroéconomique ; et non issu d'une analyse microéconomique. Ce travail cherchera alors analyser ces deux niveaux.

4.Choix de simuler ces données.

D'emblée, il paraît important de souligner **que ces informations sont des informations personnelles, voire sensibles**. En conséquence, en raison de la RGPD, et afin de préserver la vie privée des individus, ces données n'ont jamais été collectées publiquement.

En effet, **dans la littérature économétrique, seule l'espérance de vie est utilisée comme proxy de santé ;** ou des variables issues d'études **spécialisée sur une thématique** , sont mises en place !
En conséquence,

Variables

La création de cette base de données vise ainsi à fournir un outil concret permettant de réaliser des analyses approfondies et éclairées. Elle permet, par exemple, de :

- **Identifier les corrélations statistiques, mais pas causales**, entre certains indicateurs économiques (taux de chômage, niveau de revenu moyen, inflation, etc.) et des indicateurs sanitaires assez classiques (**espérance de vie, taux de maladies chroniques, accès aux soins, etc.**).
- Explorer **des politiques précises** comment des interventions économiques **d'un Etat** (*comme les subventions publiques pour la recherche des entreprises économiques ou les politiques fiscales: des taxes par exemple sur le sucre comme en France , ou sur l'alcool*) peuvent avoir des impacts directs ou **indirects sur la santé publique**.
- Fournir aux décideurs politiques, chercheurs et analystes économiques des données solides pour évaluer **les coûts économiques des problèmes de santé publique** notamment en proposant une

colonne qui chiffre le coût par année pour la Sécurité sociale autant que le coût sur la productivité pour un travailleur donné.

Limite

1. Etablir un lien causal est impossible

Ces analyses ne pourront refléter toutefois la réalité ; même si elle s'étaient appuyées sur des données réelles.

Afin de favoriser l'originalité de ce travail ; les relations entre les tables ; il m'est apparu intéressant d'ajouter des variables qui sont rarement mesurées, sur un même panel d'individus, voire jamais.

Par ailleurs, l'économie elle-même influence considérablement les conditions de santé d'une population.

Ainsi , il est clair qu'il existe un problème de causalité inverse entre ces deux mécanismes, comme le souligne Aghion dans le Bénéfice de la santé , il est compliqué d'isoler l'effet de la santé , car ce dernier est déjà en partie capté par la variable “Education” sur le PIB.

Ce problème peut alors créer une situation où l'effet attribué dans la santé, peut être sous-estimé.

II. Relations.

1. Justifier des clés primaires.

Ainsi, j'ai décidé de ne pas mettre de clé primaire comme **le nom_prenom**. En effet, il me reste important de souligner **notamment**, que cela empêcherait que plusieurs mesures pour un même individu puissent être prises durant l'expérience !

2. Justifier les relations de manière générale .

1. Un employeur emploie un ou plusieurs employés (mais 1 employé=1 employeur)

C'est une relation **plusieurs vers un** : en effet un employé ne peut qu'appartenir à une entreprise ; mais une entreprise peut avoir plusieurs employés.

De plus, un unique EMPLOYEUR peut utiliser ||--II APPLICATION : utilise pour refléter que les employeurs peuvent aussi utiliser des applications.

2. Un employé utilise plusieurs applications 1,2:création d'une table applications 1/2 (décision de scission)

Comme annoncé avant : un **employé va utiliser plusieurs applications (1,2)**. Une application a forcément un ou plusieurs employés. Mais une application peut être utilisée par 0 ou 1 employé

3. Employés vers profils de santé.(1,N)

Un profil de santé va correspondre ici à la table d'identification **du problème de santé**.

Un acteur correspond aux variables **employeurs/employés, qui correspondent à entreprise/individu d'auparavant**, mais renommer afin de gagner en clarté !.

On peut souligner une relation de type 1 acteur donc 1 individu ou 1 entreprise, peut donc avoir plusieurs enregistrements de profils de santé /lignes de causes de santé car une maladie ou un même problème peut resurgir ou évoluer plusieurs fois , et que les mesures doivent être répétées sur les applications .

A contrario, un profil de santé =1 ligne d'enregistrement, ne peut correspondre qu'à un seul acteur, car les problèmes de santé sont personnels.

1. Cela va donc se traduire par des clés **étrangères dans les tables de cause** par une relation où un profil de **santé (Clé primaire)** peut avoir une ou

plusieurs mesures de cause dans le temps . Cela **entraîne** donc des **doublons possibles** des id_individus et d'id_entreprises !

id_individu INT UNIQUE NOT NULL

id_entreprises INT UNIQUE NOT NULL

2. Cela peut donc se traduire notamment par **une variable de cause qui serait prise par une des trois applications** . Ainsi, l'ID d'application *devrait donc être présent* avec doublons **dans le but notamment** de correspondre ainsi à **l'application 1, puis 2 ou encore 3**. Son ID **correspond donc au numéro de l'application utilisé**, qui peut être égal à **1,2,3** au final.

Pourquoi cette clé n'est pas égale à zéro ?

*En effet, dans ce cadre, il n'est **pas possible qu'une variable** indiquée ne provienne pas d'une application dans le cadre de l'expérience.*

id_application INT NOT NULL

4.Application 3 : mesure par le manager

Idem application 1/2

5.Un profil de santé peut avoir un ou plusieurs effets/degrés d'effets dans le temps

Ensuite, la relation (1,1) entre **ces mêmes tables de santé et les tables d'effets** .

Un même profil de santé, donc **une seule ligne** peut avoir plusieurs mesures des effets. **En effet, on pourra tenter de les agréger par la suite**, et une hausse forte de cigarette/droge peut avoir un effet **sur plusieurs jours**.

A contrario, un effet ne peut avoir **qu'un seul et unique enregistrement de symptômes de santé lié à chaque effet**. En effet, un effet est **forcément relié à une cause...**

Donc la relation finale **est de type :TROUBLE_MENTAL ||--o{ EFFETS_TROUBLE**

SChéma Mermaid

Ainsi, **après de multiples transformations afin de rendre le diagramme le plus lisible possible, et cohérent** ; voici le diagramme final !

III.Code

```
`mermaid
%%{init: {"theme": "base", "themeVariables": {"primaryColor": "#ffa",
"edgeLabelBackground": "#fff8"}}}%
erDiagram
    direction TB

    %% Employers and Employees
    EMPLOYEUR {
        int id_entreprise PK
        varchar nom
        varchar nom_manager
        decimal benefices
        decimal heures
        decimal benefices_par_heure
    }
    EMPLOYE {
        int id_individu PK
        varchar nom
        varchar prenom
        int id_entreprise FK
        decimal salaire
    }

    %% Application for Cause Measurements
    APPLICATION_CAUSE {
        int id_application_cause PK
        int id_individu FK
        varchar nom
        varchar disponibilite
        datetime date_heure
        text but
        text acteur
        varchar smartphone
    }

    %% Cause Tables
    SPORT_SANTE {
        int id_sport PK
        int id_individu FK
        int id_entreprise FK
        int id_application_cause FK NOT NULL
```

```

    int minutes_activite
    int nb_jours_par_semaine
    int nb_types_activite
    int nb_escalier
    int minutes_debout
    int minutes_assis
}
ALCOOL {
    int id_alcool PK
    int id_individu FK
    int id_entreprise FK
    int id_application_cause FK NOT NULL
    boolean consomme
    int freq_par_semaine
    decimal nb_verres_moyen
}
TABAC {
    int id_tabac PK
    int id_individu FK
    int id_entreprise FK
    int id_application_cause FK NOT NULL
    boolean consomme_tabac
    int cigarettes_par_jour
    text consommation_drogues
    int freq_drogues_jour
}
SOMMEIL {
    int id_sommeil PK
    int id_individu FK
    int id_entreprise FK
    int id_application_cause FK NOT NULL
    time heure_debut
    time heure_fin
    decimal heures_moyennes
    int qualite_echelle
    decimal duree_reveil
}
ALIMENTATION {
    int id_alim PK
    int id_individu FK
    int id_entreprise FK
    int id_application_cause FK NOT NULL
    int nb_portions_fruits
    int nb_fast_food
    decimal litres_eau
    varchar type_repas
    int degre_fatigue
}

```

```

RISQUE {
    int id_risque PK
    int id_individu FK
    int id_entreprise FK
    int id_application_cause FK NOT NULL
    decimal ceinture_pourcentage
    int infractions_vitesse
    int nb_pauses
    text comportement_sante
}

```

```

TROUBLE_MENTAL {
    int id_trouble PK
    int id_individu FK
    int id_entreprise FK
    int id_application_cause FK NOT NULL
    varchar diagnostic
    int nb_migraines
    boolean depression
    int nb_psychologues
    int nb_jours_medication
}

```

%% Application for Effect Measurements

```

APPLICATION_EFFECT {
    int id_application_effect PK
    int id_individu FK
    varchar nom
    varchar disponibilite
    datetime date_heure
    text but
    text acteur
    varchar smartphone
}

```

%% Effect Tables

```

EFFETS_SPORT {
    int id_effet_sport PK
    int id_sport FK
    int id_application_effect FK NOT NULL
    decimal gain_productivite
    int absenteisme_sans_sport
    int absenteisme_avec_sport
    decimal impact_PIB_par_habitant
}

```

```

EFFETS_ALCOOL {
    int id_effet_alcool PK
    int id_alcool FK
    int id_application_effect FK NOT NULL
}

```

```

    boolean maladie_hepatique
    int stade_hepatique
    boolean hypertension
    int battements_cardiaques
    boolean troubles_mentaux
    int nb_remarques_travail
    int nb_mauvaises_remarques
    decimal temps_travail_heure
    decimal impact_PIB_par_habitant
}
EFFETS_TABAC {
    int id_effet_tabac PK
    int id_tabac FK
    int id_application_effect FK NOT NULL
    int respirations_par_minute
    varchar etat_peau
    varchar type_test
    decimal valeur_co_ppm
    decimal capacite_vitale_forcee
    decimal volume_expiratoire_max
    int frequence_respiratoire
    int score_coherence
    int duree_test
    text commentaire
    decimal impact_PIB_par_habitant
}
EFFETS_SOMMEIL {
    int id_effet_sommeil PK
    int id_sommeil FK
    int id_application_effect FK NOT NULL
    int nb_erreurs_par_mois
    int jours_absence_fatigue
    decimal reduction_productivite_perc
    decimal impact_PIB_par_habitant
}
EFFETS_ALIMENTATION {
    int id_effet_alim PK
    int id_alim FK
    int id_application_effect FK NOT NULL
    decimal impact_vitesse_execution
    int impact_absenteisme
    int impact_erreurs
    decimal perte_argent_heure
    decimal impact_PIB_par_habitant
}
EFFETS_RISQUE {
    int id_effet_risque PK
    int id_risque FK

```

```

    int id_application_effect FK NOT NULL
    decimal temps_horaire_moyenne
    decimal taches_par_heure
    int nb_erreurs
    decimal perte_economique
    decimal representativite_echelle
    decimal impact_PIB_par_habitant
}
EFFETS_TROUBLE {
    int id_effet_trouble PK
    int id_trouble FK
    int id_application_effect FK NOT NULL
    int jours_absence
    int nb_erreurs
    decimal productivite_horaire_base
    decimal productivite_horaire_avec_trouble
    decimal reduction_productivite_perc
    decimal evolution_productivite
    decimal impact_PIB_par_habitant
}

```

%% Relationships %%

```

EMPLOYEUR ||--o{ EMPLOYE      : emploie
EMPLOYE ||--o{ APPLICATION_CAUSE : utilise
EMPLOYE ||--o{ APPLICATION_EFFECT : utilise

```

```

APPLICATION_CAUSE ||--o{ SPORT_SANTE      : mesure
APPLICATION_CAUSE ||--o{ ALCOOL           : mesure
APPLICATION_CAUSE ||--o{ TABAC            : mesure
APPLICATION_CAUSE ||--o{ SOMMEIL          : mesure
APPLICATION_CAUSE ||--o{ ALIMENTATION     : mesure
APPLICATION_CAUSE ||--o{ RISQUE           : mesure
APPLICATION_CAUSE ||--o{ TROUBLE_MENTAL   : mesure

```

```

SPORT_SANTE ||--o{ APPLICATION_EFFECT : "diminue les capacités et est
enregistré"
ALCOOL ||--o{ APPLICATION_EFFECT      : "diminue les capacités et est enregistré"
TABAC ||--o{ APPLICATION_EFFECT       : "diminue les capacités et est enregistré"
SOMMEIL ||--o{ APPLICATION_EFFECT     : "diminue les capacités et est enregistré"
ALIMENTATION ||--o{ APPLICATION_EFFECT : "diminue les capacités et est
enregistré"
RISQUE ||--o{ APPLICATION_EFFECT      : "diminue les capacités et est enregistré"
TROUBLE_MENTAL ||--o{ APPLICATION_EFFECT : "diminue les capacités et est
enregistré"

```

```

APPLICATION_EFFECT ||--o{ EFFETS_SPORT      : affiche
APPLICATION_EFFECT ||--o{ EFFETS_ALCOOL     : affiche
APPLICATION_EFFECT ||--o{ EFFETS_TABAC      : affiche

```

APPLICATION_EFFECT ||--o{ EFFETS_SOMMEIL : affiche
APPLICATION_EFFECT ||--o{ EFFETS_ALIMENTATION : affiche
APPLICATION_EFFECT ||--o{ EFFETS_RISQUE : affiche
APPLICATION_EFFECT ||--o{ EFFETS_TROUBLE : affiche

SPORT_SANTE ||--o{ EFFETS_SPORT : entraine
ALCOOL ||--o{ EFFETS_ALCOOL : entraine
TABAC ||--o{ EFFETS_TABAC : entraine
SOMMEIL ||--o{ EFFETS_SOMMEIL : entraine
ALIMENTATION ||--o{ EFFETS_ALIMENTATION : entraine
RISQUE ||--o{ EFFETS_RISQUE : entraine
TROUBLE_MENTAL ||--o{ EFFETS_TROUBLE : entraine

- ****Script de schéma****

Chapitre 3 :

Un script annoté des déclarations SQL (`CREATE TABLE`, `CREATE INDEX`, `CREATE VIEW`, etc.) qui composent le schéma de votre base de données. Si vous utilisez une base de données existante, expliquez comment vous l'avez adaptée ou étendue pour répondre à vos besoins spécifiques.

2. Résultat positif en faisant “Run query”

```
C:\Users\greg\OneDrive\Documents\VS Code\SQL\exolinkedin.sql
CREATE TABLE EMPLOYEUR (
    id_entreprise INTEGER PRIMARY KEY
    nom TEXT NOT NULL
    nom_manager TEXT
    benefices REAL
    heures REAL
    benefices_par_heure REAL
);
CREATE TABLE EMPLOYE (
    id_individu INTEGER PRIMARY KEY
    nom TEXT NOT NULL
    prenom TEXT NOT NULL
    id_entreprise INTEGER
    salaire REAL
    FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
);
CREATE TABLE APPLICATION_CAUSE (
    id_application_cause INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    nom TEXT NOT NULL
    disponibilite TEXT
    date_heure TEXT NOT NULL
    but TEXT
    acteur TEXT
    smartphone TEXT
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
);
CREATE TABLE APPLICATION_EFFECT (
    id_application_effect INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    nom TEXT NOT NULL
    disponibilite TEXT
    date_heure TEXT NOT NULL
    but TEXT
    acteur TEXT
    smartphone TEXT
```

En texte ici :

```
CREATE TABLE EMPLOYEUR (
    id_entreprise INTEGER PRIMARY KEY
    nom TEXT NOT NULL
    nom_manager TEXT
    benefices REAL
    heures REAL
    benefices_par_heure REAL
);
CREATE TABLE EMPLOYE (
    id_individu INTEGER PRIMARY KEY
    nom TEXT NOT NULL
    prenom TEXT NOT NULL
    id_entreprise INTEGER
```

```

    salaire REAL
    FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
);
CREATE TABLE APPLICATION_CAUSE (
    id_application_cause INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    nom TEXT NOT NULL
    disponibilite TEXT
    date_heure TEXT NOT NULL
    but TEXT
    acteur TEXT
    smartphone TEXT
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
);
CREATE TABLE APPLICATION_EFFECT (
    id_application_effect INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    nom TEXT NOT NULL
    disponibilite TEXT
    date_heure TEXT NOT NULL
    but TEXT
    acteur TEXT
    smartphone TEXT
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
);
CREATE TABLE SPORT_SANTE (
    id_sport INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    id_entreprise INTEGER NOT NULL
    id_application_cause INTEGER NOT NULL
    minutes_activite INTEGER
    nb_jours_par_semaine INTEGER
    nb_types_activite INTEGER
    nb_escalier INTEGER
    minutes_debout INTEGER
    minutes_assis INTEGER
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
    FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
    FOREIGN KEY (id_application_cause) REFERENCES
APPLICATION_CAUSE(id_application_cause)
);
CREATE TABLE ALCOOL (
    id_alcool INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    id_entreprise INTEGER NOT NULL
    id_application_cause INTEGER NOT NULL
    consommation INTEGER CHECK(consommation IN (0 1))
    freq_par_semaine INTEGER
    nb_verres_moyen REAL

```

```

        FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
        FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
        FOREIGN KEY (id_application_cause) REFERENCES
APPLICATION_CAUSE(id_application_cause)
    );
CREATE TABLE TABAC (
    id_tabac INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    id_entreprise INTEGER NOT NULL
    id_application_cause INTEGER NOT NULL
    consommation_tabac INTEGER CHECK(consommation_tabac IN (0 1))
    cigarettes_par_jour INTEGER
    consommation_drogues TEXT
    freq_drogues_jour INTEGER
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
    FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
    FOREIGN KEY (id_application_cause) REFERENCES
APPLICATION_CAUSE(id_application_cause)
);
CREATE TABLE SOMMEIL (
    id_sommeil INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    id_entreprise INTEGER NOT NULL
    id_application_cause INTEGER NOT NULL
    heure_debut TEXT
    heure_fin TEXT
    heures_moyennes REAL
    qualite_echelle INTEGER
    duree_reveil REAL
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
    FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
    FOREIGN KEY (id_application_cause) REFERENCES
APPLICATION_CAUSE(id_application_cause)
);
CREATE TABLE ALIMENTATION (
    id_alim INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    id_entreprise INTEGER NOT NULL
    id_application_cause INTEGER NOT NULL
    nb_portions_fruits INTEGER
    nb_fast_food INTEGER
    litres_eau REAL
    type_repas TEXT
    degre_fatigue INTEGER
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
    FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
    FOREIGN KEY (id_application_cause) REFERENCES
APPLICATION_CAUSE(id_application_cause)
);

```

```

CREATE TABLE RISQUE (
    id_risque INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    id_entreprise INTEGER NOT NULL
    id_application_cause INTEGER NOT NULL
    ceinture_pourcentage REAL
    infractions_vitesse INTEGER
    nb_pauses INTEGER
    comportement_sante TEXT
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
    FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
    FOREIGN KEY (id_application_cause) REFERENCES
APPLICATION_CAUSE(id_application_cause)
);
CREATE TABLE TROUBLE_MENTAL (
    id_trouble INTEGER PRIMARY KEY
    id_individu INTEGER NOT NULL
    id_entreprise INTEGER NOT NULL
    id_application_cause INTEGER NOT NULL
    diagnostic TEXT
    nb_migraines INTEGER
    depression INTEGER CHECK(depression IN (0 1))
    nb_psychologues INTEGER
    nb_jours_medication INTEGER
    FOREIGN KEY (id_individu) REFERENCES EMPLOYE(id_individu)
    FOREIGN KEY (id_entreprise) REFERENCES EMPLOYEUR(id_entreprise)
    FOREIGN KEY (id_application_cause) REFERENCES
APPLICATION_CAUSE(id_application_cause)
);
CREATE TABLE EFFETS_SPORT (
    id_effet_sport INTEGER PRIMARY KEY
    id_sport INTEGER NOT NULL
    id_application_effect INTEGER NOT NULL
    gain_productivite REAL
    absenteisme_sans_sport INTEGER
    absenteisme_avec_sport INTEGER
    impact_PIB_par_habitant REAL
    FOREIGN KEY (id_sport) REFERENCES SPORT_SANTE(id_sport)
    FOREIGN KEY (id_application_effect) REFERENCES
APPLICATION_EFFECT(id_application_effect)
);
CREATE TABLE EFFETS_ALCOOL (
    id_effet_alcool INTEGER PRIMARY KEY
    id_alcool INTEGER NOT NULL
    id_application_effect INTEGER NOT NULL
    maladie_hepatique INTEGER CHECK(maladie_hepatique IN (0 1))
    stade_hepatique INTEGER
    hypertension INTEGER CHECK(hypertension IN (0 1))
    battements_cardiaques INTEGER

```

```

troubles_mentaux INTEGER CHECK(troubles_mentaux IN (0 1))
nb_remarques_travail INTEGER
nb_mauvaises_remarques INTEGER
temps_travail_heure REAL
impact_PIB_par_habitant REAL
FOREIGN KEY (id_alcool) REFERENCES ALCOOL(id_alcool)
FOREIGN KEY (id_application_effect) REFERENCES
APPLICATION_EFFECT(id_application_effect)
);
CREATE TABLE EFFETS_TABAC (
    id_effet_tabac INTEGER PRIMARY KEY
    id_tabac INTEGER NOT NULL
    id_application_effect INTEGER NOT NULL
    respirations_par_minute INTEGER
    etat_peau TEXT
    type_test TEXT
    valeur_co_ppm REAL
    capacite_vitale_forcee REAL
    volume_expiratoire_max REAL
    frequence_respiratoire INTEGER
    score_coherence INTEGER
    duree_test INTEGER
    commentaire TEXT
    impact_PIB_par_habitant REAL
    FOREIGN KEY (id_tabac) REFERENCES TABAC(id_tabac)
    FOREIGN KEY (id_application_effect) REFERENCES
APPLICATION_EFFECT(id_application_effect)
);
CREATE TABLE EFFETS_SOMMEIL (
    id_effet_sommeil INTEGER PRIMARY KEY
    id_sommeil INTEGER NOT NULL
    id_application_effect INTEGER NOT NULL
    nb_erreurs_par_mois INTEGER
    jours_absence_fatigue INTEGER
    reduction_productivite_perc REAL
    impact_PIB_par_habitant REAL
    FOREIGN KEY (id_sommeil) REFERENCES SOMMEIL(id_sommeil)
    FOREIGN KEY (id_application_effect) REFERENCES
APPLICATION_EFFECT(id_application_effect)
);
CREATE TABLE EFFETS_ALIMENTATION (
    id_effet_alim INTEGER PRIMARY KEY
    id_alim INTEGER NOT NULL
    id_application_effect INTEGER NOT NULL
    impact_vitesse_execution REAL
    impact_absenteisme INTEGER
    impact_erreurs INTEGER
    perte_argent_heure REAL
    impact_PIB_par_habitant REAL

```

```

        FOREIGN KEY (id_alim) REFERENCES ALIMENTATION(id_alim)
        FOREIGN KEY (id_application_effect) REFERENCES
APPLICATION_EFFECT(id_application_effect)
    );
CREATE TABLE EFFETS_RISQUE (
    id_effet_risque INTEGER PRIMARY KEY
    id_risque INTEGER NOT NULL
    id_application_effect INTEGER NOT NULL
    temps_horaire_moyenne REAL
    taches_par_heure REAL
    nb_erreurs INTEGER
    perte_economique REAL
    representativite_echelle REAL
    impact_PIB_par_habitant REAL
    FOREIGN KEY (id_risque) REFERENCES RISQUE(id_risque)
    FOREIGN KEY (id_application_effect) REFERENCES
APPLICATION_EFFECT(id_application_effect)
);
CREATE TABLE EFFETS_TROUBLE (
    id_effet_trouble INTEGER PRIMARY KEY
    id_trouble INTEGER NOT NULL
    id_application_effect INTEGER NOT NULL
    jours_absence INTEGER
    nb_erreurs INTEGER
    productivite_horaire_base REAL
    productivite_horaire_avec_trouble REAL
    reduction_productivite_perc REAL
    evolution_productivite REAL
    impact_PIB_par_habitant REAL
    FOREIGN KEY (id_trouble) REFERENCES TROUBLE_MENTAL(id_trouble)
    FOREIGN KEY (id_application_effect) REFERENCES
APPLICATION_EFFECT(id_application_effect)
);
CREATE INDEX idx_sport_individu ON SPORT_SANTE(id_individu);
CREATE INDEX idx_alcool_individu ON ALCOOL(id_individu);
CREATE INDEX idx_tabac_individu ON TABAC(id_individu);
CREATE INDEX idx_sommeil_individu ON SOMMEIL(id_individu);
CREATE INDEX idx_alimentation_individu ON ALIMENTATION(id_individu);
CREATE INDEX idx_risque_individu ON RISQUE(id_individu);
CREATE INDEX idx_trouble_individu ON TROUBLE_MENTAL(id_individu);
CREATE VIEW vue_effets_complets AS
SELECT
    s.id_individu
    s.id_application_cause
    f.id_application_effect
    f.gain_productivite
    f.impact_PIB_par_habitant
FROM SPORT_SANTE s
JOIN EFFETS_SPORT f ON s.id_sport = f.id_sport

```

```
/* vue_effets_complets(id_individu id_application_cause id_application
```

3.

Partie 4: script de requêtes :

- ****Script de requêtes**** (en SQL ou Python) : Un script annoté de requêtes SQL (`SELECT`, `INSERT`, `UPDATE`, `DELETE`, etc.) qui vous permet de manipuler la base de données, que ce soit pour modifier des informations ou pour en extraire. Ce script doit démontrer comment utiliser la base de données pour résoudre des problèmes spécifiques.

- ****Sources des données**** : Fournissez les fichiers de données utilisés pour peupler votre base de données (csv, txt, xlsx etc...). Ces fichiers doivent être créés manuellement, générés avec un LLM, ou téléchargés sur internet. Indiquez clairement la provenance de chaque fichier.

Considérations supplémentaires

- ****Équilibre entre création et manipulation**** : Votre projet doit démontrer un équilibre entre la création de la base de données et sa **manipulation**. Si la **création est complexe**, concentrez-vous sur l'explication détaillée de la structure et des choix de conception. Si la base de données est déjà presque faite, concentrez-vous sur l'explication de son utilisation et des requêtes avancées.

- Jeu d'essai NHIS synthétique : schéma relationnel et 60 exemples
- Ce fichier SQL unique est prêt à être collé dans toute session
- SQLite (≥ 3.31). Il supprime d'abord les tables existantes de même nom, crée un schéma en 3^e forme normale (cinq tables), puis insère
- un mini-panel fictif de 2024 (100 ménages, 500 personnes). Enfin,
- soixante requêtes illustrent différentes analyses. Tous les

- commentaires sont rédigés en français clair pour paraître rédigés
- manuellement.

- ****Projets en temps réel vs analytiques**** : Votre base de données peut être conçue pour suivre des informations en temps réel (par exemple, suivi des stocks, gestion des événements) ou pour des analyses approfondies (par exemple, analyse des données de vente, génération de statistiques). Assurez-vous que votre projet est clair sur son objectif principal.

- ****Utilisation de bases de données existantes**** : Si vous utilisez une base de données existante, avec par exemple des fichiers csv, expliquez comment vous l'avez adaptée pour répondre à vos besoins spécifiques.

- ****Format de la base de données**** : La base de données doit être fournie sous forme de fichier SQLite ou hébergée sur un serveur (MySQL, PostgreSQL, MongoDB). Si elle est hébergée, fournissez les informations d'accès nécessaires.

Outils de travail

- ****Libre choix des outils**** : Vous êtes libre d'utiliser les outils de votre choix pour ce projet. Cela inclut, mais n'est pas limité à, Python, SQLite Studio, DataGrip, et CodeSpace.

Diffusion du travail

- ****Dépôt GitHub**** : Créez un dépôt GitHub pour votre projet. Ajoutez tous les membres de votre groupe ainsi que moi-même en tant que collaborateurs. Assurez-vous que tous les fichiers nécessaires (document de conception, scripts SQL, fichiers CSV) sont inclus dans le dépôt.