

Importation, analyse et visualisation de données d’enquête dans un modèle multidimensionnel de trajectoires de vie

Rapport de projet

**Etudiants :**  
BENKAOUR Salwa

FLEURY Pierre

JORDAN Célia

**Enseignant :**

PELLIER Damien

**Référent :**

GENSEL Jérôme

Table des matières

[Introduction 3](#_Toc106268060)

[Contexte 3](#_Toc106268061)

[L’objectif du projet 3](#_Toc106268062)

[Réalisation du projet 3](#_Toc106268063)

[Base de données – Importation 3](#_Toc106268064)

[Base de données – Vérification 5](#_Toc106268065)

[Requêtes – Analyse 5](#_Toc106268066)

[Visualisation – Interface 5](#_Toc106268067)

[Visualisation – Prototype 5](#_Toc106268068)

[Idées abandonnées 6](#_Toc106268069)

[Web sémantique 6](#_Toc106268070)

[Une interface complète 6](#_Toc106268071)

[Gestion de projet 6](#_Toc106268072)

[Conclusion 7](#_Toc106268073)

[Remerciements 7](#_Toc106268074)

Introduction

Contexte

Notre stage a lieu dans un contexte universitaire. En effet, nous l’effectuons sous la supervision de l’équipe STEAMER et plus précisément de Jérôme Gensel.

Un premier travail concernant les trajectoires de vie a été fait par David Noël qui a effectué une thèse sur «Une approche basée sur le Web Sémantique pour l’étude de trajectoires de vie ».

A partir de cette thèse nous avons récupérer un modèle multi point de vue permettant de mettre en ordre les données fournies par l’INED. En effet, l’INED a fourni un document csv contenant plusieurs informations sur 5000 individus.

L’objectif du projet

Notre travail consiste dans un premier temps à mettre ces données sous forme de base de données relationnelle en suivant le modèle multi point de vue. Pour cela nous utilisons PostGreSQL. Dans un deuxième temps nous utilisons Express, VueJS et NodeJS ce qui nous permet de développer un prototype d’application permettant de donner des pistes sur la logique à utiliser pour l’application finale qui sera développée pour exploiter les données. L’objectif de ce prototype est de visualiser les données selon les dimensions spatiales et temporelles. Notre objectif, dans le contexte décrit ci-desus, est d’apprendre de nouvelles technologies et d’apprendre à gérer un projet en équipe.

Réalisation du projet

Base de données – Importation

Pour construire la Base de données nous avons suivi le modèle de David Noël. Nous avons donc 17 tables, les voici avec ce qu’elles contiennent :

* ***Personne***
  + Son numéro d’enquêté
  + La date de naissance de l’enquêté
  + La référence du lieu où l’évènement s’est produit
* ***Conjoint***
  + Le numéro d’enquêté dont il est le conjoint
  + L’année ou l’évènement s’est produit
  + L’évènement
  + La référence du lieu où l’évènement s’est produit
* ***Enfant***
  + Le numéro d’enquêté dont il est l’enfant
  + Rang des naissances
  + L’année ou l’évènement s’est produit
  + L’évènement
  + La référence du lieu où l’évènement s’est produit
* ***Parent***
  + Le numéro d’enquêté dont il est le parent
  + Le parent 1 ou 2.
  + L’année ou l’évènement s’est produit
  + L’évènement
  + La référence du lieu où l’évènement s’est produit
* ***Localisation***
  + La référence du lieu où l’évènement s’est produit
  + La commune
  + Le département
  + Le code géographique
  + Le pays
  + L’unité territoriale
  + La longitude
  + La latitude
  + Le point géométrique avec posGis
* ***Familial\_trajectory***
  + Le numéro d’évènement
  + Le numéro de l’épisode
  + Le numéro d’enquêté
* ***Familial\_episode***
  + Le numéro de l’épisode
  + Le numéro d’évènement associé
  + Le numéro d’enquêté
  + Le rang
  + L’événement
  + L’année
  + Le lieu
* ***Familial\_event***
  + Le numéro de l’événement
  + Le numéro de l’épisode associé
  + Le numéro d’enquêté
  + Le rang
  + L’événement
  + L’année
  + Le lieu
* ***Résidential\_trajectory***
  + Le numéro d’évènement
  + Le numéro de l’épisode
  + Le numéro d’enquêté
* ***Résidential\_episode***
  + Le numéro de l’épisode
  + Le numéro d’évènement associé
  + Le numéro d’enquêté
  + Le rang
  + L’événement
  + L’année
  + Le lieu
* ***Résidential\_event***
  + Le numéro de l’événement
  + Le numéro de l’épisode associé
  + Le numéro d’enquêté
  + Le rang
  + L’événement
  + L’année
  + Le lieu
* ***Professional\_trajectory*** 
  + Le numéro d’évènement
  + Le numéro de l’épisode
  + Le numéro d’enquêté
* ***Professional\_episode***
  + Le numéro de l’épisode
  + Le numéro d’évènement associé
  + Le numéro d’enquêté
  + Le rang
  + L’événement
  + L’année
  + Le lieu
* ***Professional\_event***
  + Le numéro de l’événement
  + Le numéro de l’épisode associé
  + Le numéro d’enquêté
  + Le rang
  + L’événement
  + L’année
  + Le lieu
* ***Leisure\_trajectory***
  + Le numéro d’évènement
  + Le numéro de l’épisode
  + Le numéro d’enquêté
* ***Leisure\_episode***
  + Le numéro de l’épisode
  + Le numéro de l’événement associé
  + Le numéro d’enquêté
  + Le rang
  + L’événement
  + L’année
  + Le lieu
* ***Leisure\_event***
  + Le numéro de l’événement
  + Le numéro de l’épisode associé
  + Le numéro d’enquêté
  + Le rang
  + L’événement
  + L’année
  + Le lieu

La mise en place de la Base de données nous a pris du temps car nous avons effectué des ajustements tout au long du projet pour répondre aux attentes du client. En effet, plusieurs changements ont été effectués notamment pour la trajectoire familiale qui ne fonctionnait pas comme les autres trajectoires. Le document *schema\_3B* illustre les relations des tables entre elles.

Concernant la répartition des tâches, Pierre et Célia se sont occupés de la base de données. Pierre a principalement travaillé sur la création des tables et les liens des différentes tables entre-elles et sur les relations - clés primaires et étrangères. Célia a principalement travaillé sur le remplissage des tables. Salwa à travailler sur les différentes requêtes qui peuvent être posées à la base de données.

Base de données – Vérification

Le plan de test reprend la méthode utilisée pour vérifier la cohérence de la base de données par rapport au fichier.

Requêtes – Analyse

En deuxième partie, nous devons faire une interface permettant d’exploiter les données, pour cela nous devons savoir ce qu’il est possible de faire avec le prototype.   
Nous répartissons les taches, Salwa trouve des requêtes qu’il est possible de faire par rapport au fichier que nous a fourni l’INED.

Pierre et Célia écrivent les requêtes avec PostgreSQL afin de vérifier que cela est interrogeable sur la base de données.

De plus, un travail commun est effectué afin de trier les données selon leur « type » nous avons ainsi des données spatial ou temporel et des données de point de vue qui sont familiales, professionnelles, résidentielles ou de voyage.

Visualisation – Interface

Pour le développement de l’application Maxime Sage se joint au projet dans le cadre également d’un stage mais qui pour lui se termine fin juillet.

Dans un premier temps, Maxime et Salwa se concentre sur une interface à proposer pour représenter les données.

Pierre et Célia continuent sur les requêtes et l’analyse des données.

Visualisation – Prototype

L’application se fait avec *Vue.js* pour le Frontend et *Node.js* pour le Backend. L’idée est de récupérer les données dans la base de données et de les stocker dans un fichier JSON qui sera ensuite utilisé par le frontend pour récupérer les différentes données selon les objets sélectionnés.

Pour le développement du prototype, nous remarquons qu’il y a plusieurs rubriques qui le compose. Le manuel d’utilisation décrit ces rubriques. Notre stage ne nous permet pas de finir l’application par manque de temps. Nous nous concentrons donc sur la partie *Individual* qui définit l’interrogation des données concernant un seul individu à la fois.

Pour commencer, une rapide formation individuelle s’impose car aucun de nous n’avait utiliser Vue.js et Node.js.

Ensuite, une première répartition des rôles se fait, Maxime et Célia s’occupe en premier du backend, avec Node, afin de récupérer les données depuis la base de données et de les mettre sous forme d’un JSON.

Pierre et Salwa eux se concentre sur le frontend, avec Vue, afin d’avoir une première interface fonctionnelle en vue.

Dans un deuxième temps, Pierre se concentre sur la mise en place des *checkbox* ainsi que des différentes options de l’application. Salwa se concentre sur les lignes de vie et surtout sur leur visualisation de celle-ci. Maxime se concentre sur la récupération des données dans un JSON ainsi que la mise en place d’un JSON aux normes. Célia se concentre sur OpenLayers afin de faire de la visualisation sur une carte puis il fait un peu de représentation de ligne de vie afin de proposer plusieurs idées pour leur représentation.

Idées abandonnées

Web sémantique

Au début du projet nous devions faire du web sémantique mais comme nous n’avions aucune base il a été décidé que nous utiliserons une base de données.

Une interface complète

Le projet en temps plein n’étant que de 2 mois nous avons vite compris que ne pourrons pas tout réaliser au niveau de l’interface. Nous nous sommes donc concentrés uniquement sur la partie individuelle en laissant de côté les parties sur l’interrogation d’un group d’individu et des métadonnées.

Gestion de projet

Pour la réalisation de notre projet nous avons dans un premier temps consacré les lundis après-midi à sa réalisation. Nous avons également avancé pendant les semaines, principalement celles où nous avions le moins de cours car une semaine sur deux nous avions les alternants et donc les cours était plus dense cette semaine.

Nous nous donnions alors des objectifs de résultat plus important pour le lundi qui suivait afin de proposer quelque chose de nouveau à M. Gensel.

Nous avons également mis en place un Git mais pour la première partie celui-ci n’a pas été utile.

Ensuite vient le stage, nous devons effectuer 35h/semaine au LIG du 02 mai au 24 juin. Pendant les trois premières semaines de ce stage, nous avons concentré nos efforts sur la logique du prototype, c’est-à-dire comment exploiter les données, quel type de visualisation choisir… Pour centraliser nos avancées nous avons utilisé le GitLab que l’équipe STEAMER a mis à notre disposition (<https://gricad-gitlab.univ-grenoble-alpes.fr/steamer/ined-3b>).

A partir de là nous avons alors créé une branche pour chacun dans laquelle chacun de nous pouvait développer sa partie sans crainte. Au niveau pratique, nous avons mis en évidence un manque de connaissances des bonnes manières sur git car nous nous sommes rapidement mélangés dans les branches, les fonctions, les rôles…

Nous avons le plus souvent deux réunions par semaine avec l’équipe, composée de nous-mêmes ainsi que Jérôme Gensel, Samandar Ibragimov, Marlène Villanova-Oliver et Camille Bernard.

Conclusion

Le premier objectif a été atteint, la base de données relationnelle est complète. Pour ce qui est du développement du prototype, nous n’avons terminé qu’une seule des trois sous-parties, même si nous avons déjà réfléchi à la logique d’utilisation des deux autres sous-parties, *Group* et *Metadata*.

Le projet a été très enrichissant malgré un départ tardif. En effet, la demi-journée de stage durant l’année est un temps trop court pour avancer correctement dans un projet comme celui-ci, ce qui a entrainé ce faux départ. Malgré cela, le début du stage à temps plein a été une bonne phase pour nous

Nous avons amélioré nos compétences de travail en groupe, en étant confrontés à des impasses.

C’est du point de vue pratique que le stage a été le plus bénéfique, nous avons aperçu les tenants et certains aboutissants d’un projet de développement d’une application, même si nous nous sommes arrêtés à la proposition d’un prototype d’application.

Remerciements

Nous tenons à remercier Jérôme Gensel pour avoir proposé ce sujet très intéressant et nous avoir permis de réaliser notre projet dans de bonne conditions en étant disponible tout au long du projet.

Nous tenons également à remercier Camille Bernard, Samandar Ibragimov et Marlène Villanova-Oliver pour leur implication tout au long de notre stage.

Merci également à Maxime Sage de nous avoir rejoint en cours de route et de continuer.