Rapport

Alice Thomassin - 12113182

Table des matières

Je dois recommencer à écrire pour reprendre le contrôle sur mon travail. Dans les dernières semaines, je n’ai pas avancé. Je me demande pourquoi. La journée où j’ai pris le train avec ma mère pour aller à Londres, j’avais l’intention d’avancer. J’ai fait quelques fonctions, mais je n’ai pas vraiment travaillé. J’ai pris beaucoup de temps dans les dernières semaines à me renseigner sur divers calculs statistiques. Je regardais des tutos sur youtube et je prenais des notes. Je me faisais un peu l’illusion d’avancer sans vraiment le faire.

Je crois que j’ai commencé à décrocher quand j’ai perdu mon travail et que j’ai eu à apprendre à utiliser Git. J’ai ensuite réussi à produire ma base unique, mais je n’avais toujours pas ajouté le module sur la fraternité. J’ai ensuite eu à documenter mes variables et là, j’ai perdu beaucoup de temps à me renseigner pour savoir comment l’automatiser. Je crois que ça a affecté mon assurance, parce que je n’arrivais pas à trouver de solutions.

Il y a ensuite eu la quesiton de pondérer les variables. Là, ça aussi ça m’a annéanti. Ça faisait beaucoup et je tournais en rond, sans savoir quoi prioriser.

Quand j’ai commencé à trouver une solution pour la documentation des variables, j’ai réalisé que je n’avais même pas porté attention aux variables manquantes. Là, je me suis lancée dans les démarches de construire les modalités. Je devais donc rédiger des fonctions, ce que j’ai longtemps évité. Le problème, c’est que je ne savais pas si je devais codifier les valeurs manquantes sur la base finale ou si je devais le faire au fur et à mesure. Éventuellement, j’ai réalisé que sur ma base finale, il me manque 6 prénoms. J’ignore s’il existe d’autres variables pour lesquelles j’ai des valeurs manquantes. J’ai donc décidé de me relancer dans le projet de produire une base unique, tout en modifiant au fur et à mesure les modalités manquantes.

Cette suite m’intimide et me décourage. Je crois que c’est la raison pour laquelle je suis inactive. Il y a aussi le rapport de stage qui doit être rendu dans quelques jours, ce dernier que je n’ai toujours pas entamé. Bref, c’est un peu la panique tranquille et profondément inconfortable. Non seulement ça, mais en plus, il y a toutes sortes de catastrophes qui se déroulent aux États-Unis et dans la politique mondiale en générale et j’ai le malheur d’avoir développé un intérêt pour ces questions. Je passe donc mes journées à regarder des vidéos youtube sur le sujet plutôt que d’avancer.

Je dois aussi ajouter que depuis les deux dernières semaines, je suis toujours seule dans ma salle de travail. Ça m’empêche terriblement de ressentir suffisament de pression pour avancer.

Je ne sais pas si je me trouve des excuses ou si ce sont les réelles causes de mon décrochage. Je crois aussi que le contenu que je consomme évoque de fortes émotions chez moi alors que R me fait basculer au neutre. Ce neutre ne m’attire pas. On dirait que j’ai trouvé une nouvelle source d’intérêt et que j’ai du mal à m’adoner à quoi que ce soit qui soit moins stimulant, ou pire, qui me fasse douter en moi.

Pour le moment, j’ignore un peu où je me situe dans mon script. J’ai le vertige et le coût d’entrée me semble élevé. Pourtant, je n’ai vraiment pas le choix. Plus j’attend, pire ce sera. Je pensais qu’écrire allait me faire du bien, mais je n’en suis plus trop certaine. Je crois que c’est une autre manière pour moi de procrastiner.

J’ai modifié les fonctions pour que les variables soient en character plutôt qu’en integer ou en numérique. Comme ça, dans le dfSummary() on voit mieux la répartition des choix de réponse et on évite d’avoir de l’information inutile comme les moyennes. D’autant plus qu’en ajoutant les modalités des valeurs manquantes, on obtient des stats qui n’ont plus de sens (comme une année moyenne de 2039).

NOPE ÇA FONCTIONNE PAS parce que je ne peux plus ordonner mes variables. Elles doivent demeurer en valeurs numériques.

Je crois avoir terminé le redoublement scolaire. Il me reste à voir la liaison avec la base générale. Ensuite je peux passer aux interruptions.

Je dois faire le codebook une fois que la base est jointe avec ceux qui n’ont pas participé au module pour savoir la part de ceux non concernés et ceux qui n’ont juste pas répondu au module

Okay alors, je ne vais pas modifier sc\_redoubl ni sc\_rdb\_redoubl, mais je vais créer une nouvelle variable corrigé. l’enjeu c’est de préserver les 0 dans les modules généraux, mais de conservers le 1 de ceux qui n’ont pas répondu aux sous-modules. Comme ça je peux recoder les valeurs manquantes de ces personnes avec 555 et non juste de NAs.

La nouvelle variable peut être sc\_rdb\_cor et les modalités seraient

0: pas redoublé  
1: redoublé et renseigné  
33: redoublé mais pas renseigné  
55: pas répondu  
NA: pas participé

Ordre à suivre pour grands modules

1. GM: renommer les variables
2. GM: Unir les doublons de manière locale
3. GM: Pause
4. PM: réduire les variables pour ne garder que celles du sous-module et quelques variables de liaison comme id\_anonymat, id\_date\_creation et l’ajout de id\_link
5. PM: Renommer les quelques variables avec lesquelles je vais travailler
6. PM: Corriger les doublons en modifiant que l’id\_anonymat qui se situe dans la base M1\_identity
7. PM: Réduire les variables pour n’avoir que celles du module
8. Unir les doublons, mais deux situations
   * 1. de manière locale au sein de la même base de données
     2. de manière exter

# 1. Présentation de la structure d’accueil

Créé en 1945, l’Institut national des études démographiques rassemble, produit et publie les savoirs pluridisciplinaires nécessaires à la diffusion des connaissances démographiques aussi bien destinées aux communautés scientifiques, qu’à la population générale ou à la vie politique. Le 9 juillet dernier, le directeur de l’Ined, François Clanché et deux directeurs de recherche, Anne Solaz et Laurent Toulemon, se sont rendus à l’Assemblée nationale à l’occasion d’une commission sur la baisse de la natalité en France. Mon tuteur de stage, Gil Bellis, a participé au 30e Congrès international de la population, la plus grande conférence scientifique internationale sur les questions et la recherche en matière de population.(« 30e Congrès international de la population » 2025) Le soucis de la vulgarisation du

# 2. Place du ou de la stagiaire dans l’organigramme

# 3. Description des missions de stage

# 4. Bilan critique

## 4.1 Difficultés rencontrées et solutions trouvées

## 4.2 Points positifs / négatifs

## 4.3 Compétences mobilisées / acquises / à améliorer

# 5. Annexe facultative avec extrait de programmes, codes, résultats, documents produits (avec l’autorisation de l’organisme d’accueil)

« 30e Congrès international de la population ». 2025. <https://www.ined.fr/fr/actualites/rencontres-scientifiques/autres-rendez-vous/30e-congres-international-de-la-population-ipc2025/>.