**Master 2e année Démographie**

**Dossier enquête génération**

**Cours de Mr Cordazzo**

Vous disposez des bases de l’enquête génération 2010 du Cereq, interrogation à 5 ans. Base des caractéristiques et informations complètes (ind2010\_champcomplet au format SAS et txt) et d’une base xls avec l’identifiant et la situation résidentielle pour chaque mois de l’enquête (HMOISX) ( 01 = Vit chez ses parents ; 02 = Vit en couple ; 03 = Vit seul (y compris foyer, colocation)).

Nous nous intéressons à la thématique suivante : **Les trajectoires résidentielles et professionnelles des jeunes**

**Partie 1**

**Préparation du fichier *Gen2010\_HMOIS***

* Préparer les données sur les trajectoires résidentielles pour une analyse de séquence (choix de la durée d’observation, ajustement des données dans la base au niveau des colonnes, correctifs cellules vides etc.). Les données sont en date, il les faut en durée depuis la sortie d’études (utilisation d’un programme *boucle*).

|  |  |  |
| --- | --- | --- |
| **HMOISX** | Calendrier Habitat mensuel | Alph4 |
| - |

* *Le* ***calendrier Habitat*** *décrit, mois après mois, la situation de l’enquêté selon qu’il habitait chez ses parents, en couple ou seul (y compris foyer, colocation).*
* ***HMOIS1*** *correspond à novembre 2009. Le dernier mois est* ***HMOIS69*** *: il correspond à juillet 2015. La description se fait à partir du mois suivant la date de fin de formation (Q20) et est tirée à l’identique du mois d’interrogation au mois* ***HMOIS69****.*
* *Chaque mois du calendrier habitat est codé sur 4 positions. Les deux premières positions renseignent sur le mode d’habitat du jeune mois par mois :*
* *01 = Vit chez ses parents*
* *02 = Vit en couple*
* *03 = Vit seul (y compris foyer, colocation)*
* *Les situations sont exclusives et hiérarchisées. Par exemple, un jeune déclarant vivre en couple chez ses parents sera classé en ‘01’. Les deux dernières positions de la variable donnent le numéro de la séquence habitat.*
* *Remarque (s) : A partir de Génération 2010, on ne pose plus la question pour le dernier mois de formation (Q20)*

**Préparation du fichier *indiv10\_champcomp***

* Sélection des données à minima à utiliser pour l’analyse (trajectoires professionnelles, sexe, durée d’obtention du premier emploi, âge à la fin des études, CSP des parents, niveau d’études des parents, niveau de diplomation à la fin des études) et Identifiant et pondération.
* Traitement des données (correctifs) : doublons, données incomplètes, recodage, imputations ou non, etc.