Laurent NOWIK / Aurélie GODARD / Joseph LARMARANGE

La collecte des données en démographie

La collecte de données est à la base de la discipline démographique qui, par sa nature inductive, doit d'abord collecter des informations démographiques, les regrouper et les réduire, pour ensuite en tirer une synthèse permettant d'identifier de régularités et de proposer des théories.

Les démographes attachent par conséquent une grande importance à cette phase initiale de collecte des données : sans collecte fiable, il serait impossible de bâtir de réelles connaissances sur la population étudiée d'un point de vue numérique.

Plusieurs formes de collectes existent.

A) Les populations se déplacent pour se faire « enregistrer ».
Exemple : les dispositifs permanents et institutionnalisés tels que les enregistrements d'événements à l'état civil (autrefois les registres paroissiaux) ou les registres de populations dans certains pays européens, registres qui présentent l'intérêt d'enregistrer les mouvements de population (et de dénombrer les habitants d'une commune par exemple).

B) Des enquêteurs vont à la rencontre de la population.
Exemples : cas des grandes enquêtes nationales exhaustives (Recensement Général de la Population - RGP) ou de celles bâties sur des échantillons représentatifs importants (enquête emploi ou enquête famille de l'INSEE) ; cas des enquêtes sur échantillons plus réduits (moins de 10.000 personnes), afin de collecter des données moins précisément, mais avec plus de souplesse et à coût moindre (le problème essentiel étant pour ces dernières leur degré de fiabilité qu'il convient de bien estimer).

Si les personnes remplissent elles-mêmes le questionnaire, on dit qu'il est auto-administré. La fiabilité des résultats est toujours moins bonne dans ce cas, car il se peut que la personne interrogée ne comprenne pas les questions posées. Il faut donc des questions simples (test préalable).

C) Un mixte des deux premiers cas existent : cas d'une enquête en pays en développement où les enquêteurs se rapprochent du « terrain », mais où les populations isolées sont également encouragées à se « rapprocher » des enquêteurs.

D) *Les données sont recueillies à partir de certains fichiers ou encore des données sont extraites d'informations administratives*. Les données sont susceptibles de renseigner indirectement sur des phénomènes démographiques (ex : fichier des cartes de séjour pour apprécier le nombre d'immigrés légaux vivant en France).

Quelle exhaustivité ?

On voit ainsi que certaines enquêtes sont exhaustives (elles portent sur l'ensemble d'une population délimitée par son territoire ou une variable précise : ethnie, sexe, âge…), alors que d'autres ne portent que sur une partie de la population.

Certaines enquêtes sont réalisées tantôt auprès d'échantillons représentatifs, tantôt auprès d'échantillons tirés aléatoirement (mais dont on vérifiera ensuite leurs caractéristiques avant de faire des généralisations).

Plusieurs types de méthodes

Il existe également plusieurs types de méthodes de collecte (liées pour partie aux formes précédentes) :

- Enquête à un seul passage : RGP et enquête famille par exemple (permet une connaissance de la population selon une perspective transversale, mais on peut aussi, par questionnement rétrospectif, tenter de reconstituer des biographies et reconstruire du « longitudinal »)
- Enquête à passages répétés pour le suivi précis de certains événements démographiques : enquête emploi, enquête sur la santé des enfants dans certains pays en développement, sur la connaissance de l'état de santé des populations âgées, observatoires de population… Ces approches permettent un suivi longitudinal de la population (cas des Systèmes de Surveillance Démographique : SSD - http://web.idrc.ca/fr/ev-42993-201-1-DO_TOPIC.html). On mesure par exemple la fréquence des maladies et des décès afin de saisir les conséquences numériques de ces phénomènes sur le rythme de croissance et sur la composition des populations.
- Enquêtes répétitives sur des « panels » : à certaines dates, on interroge un panel représentatif de la population totale (les personnes qui constituent le panel sont parfois les mêmes pour plusieurs enquêtes). Cette méthode convient bien au suivi temporel des opinions d'une population : on peut ainsi constituer des « baromètres » : par exemple sur les pratiques contraceptives (pour lutter contre le VIH en Afrique), sur les souhaits en matière de fécondité, sur les relations familiales et au sein du couple (…), et sur aussi beaucoup d'autres sujets pas spécifiquement démographiques, à caractère social et économique (méthode de certains instituts de sondage).

Un exemple de grande enquête réalisée en France : l'enquête famille de 1999 (associée au RGP de 1999)

Source : l'article de Courrier des statistiques n° 93, mars 2000 (Francine Cassan, François Héran, Laurent Toulemon).
http://www.insee.fr/fr/ffc/docs_ffc/cs93e.pdf

Introduction

À l'occasion du recensement de la population de mars 1999, 380 000 hommes et femmes vivant en domicile ordinaire ont rempli un bulletin complémentaire sur le thème de leur « histoire familiale », incluant des questions sur leurs origines, leurs enfants, leurs périodes de vie en couple et leur parcours social, ainsi que sur les langues d'usage au sein de leur famille, nationales et régionales.
Ces mêmes informations ont été recueillies auprès de 6 600 personnes vivant en communauté, dont 1 700 détenus de sexe masculin.
Profondément rénové, le questionnaire démographique de cette édition 1999 de l'enquête Famille permet de nombreuses études, couvrant les histoires familiales dans toute leur complexité.

Principe de l'enquête

Rappel du cours magistral : l'information démographique en France provient de deux sources principales. **Le recensement** qui fournit des données sur l'état de la population à une date donnée et **l'état civil** qui renseigne sur les mouvements de population (naissances, mariages, décès).
Mais ces deux sources sont trop sommaires pour éclairer le comportement des acteurs sur certains aspects : nombre d'enfants, calendrier des naissances, types d'union, remariage, etc. : l'objectif principal du recensement est de fournir une information localisée, à une échelle géographique très fine, notamment pour proposer l'effectif de la population légale de chaque commune, authentifié par un décret ; l'état civil a comme première fonction de définir l'identité des personnes.

Associée au recensement de la population depuis 1954 (ce qui en fait l'une des plus anciennes enquêtes par sondage de l'Insee), **l'enquête Famille** vient donc compléter le dispositif. Son objectif premier est le suivi du développement des *nouvelles formes familiales*, à travers un questionnaire *rétrospectif et à caractère biographique* reconstituant l'histoire démographique des générations.

En effet :

- le recensement ne fait pas de différence entre personnes mariées et personnes remariées (situation légale et situation de fait), il ne repère précisément ni les couples non mariés ni les relations entre adultes et enfants d'un même ménage.
- De son côté, l'état civil distingue entre enfants « légitimes » et « naturels », mais n'informe pas sur le type de relation qui unit les parents des enfants nés hors mariage (co-résidence éventuelle, durée de la cohabitation).

De plus, chaque édition de l'enquête Famille consacre une batterie de questions à l'étude d'un thème particulier lié aux préoccupations du moment, comme l'activité féminine en 1982 ou la garde des jeunes enfants en 1990. Pour 1999, on a choisi la transmission des langues et parlers d'une génération à l'autre (l'exploitation des questions sur les langues et parlers apportera une information très attendue sur la diversité linguistique de notre pays et la dynamique de l'intégration nationale tout au long de ce siècle).

Principe de réalisation de l'enquête famille : dans une fraction des secteurs d'agents recenseurs constitués pour la collecte du recensement (un sur cinquante depuis 1962), les personnes entrant dans le champ de l'enquête Famille reçoivent, en plus de leur bulletin individuel (BI), un questionnaire spécifique qu'elles sont invitées à remplir elles-mêmes et à remettre à l'agent recenseur avec les autres documents du recensement. Entre autres avantages, cette organisation permet de décharger le BI d'un certain nombre de questions démographiques qui figurent dans les recensements étrangers (et qui sont recommandées par l'ONU), mais qu'il serait coûteux de poser à l'ensemble de la population (nombre d'enfants que l'on a déjà eus, état matrimonial détaillé, dates de décès du conjoint ou des enfants...). Réciproquement, l'enquête Famille utilise les informations collectées dans le recensement. L'appariement s'effectue à partir des identifiants utilisés pour le recensement, de la date de naissance et du sexe.

L'édition 1999 de l'enquête Famille est la première de la série à s'adresser également à des hommes. L'extension aux hommes permet d'apporter des informations fiables sur la paternité, sujet mal connu.
Pour permettre que les jeunes hommes célibataires, qui ne voient pas toujours l'intérêt de répondre à un questionnaire sur la famille, répondent comme les autres, on a rebaptisé le questionnaire « **Étude de l'histoire familiale** » (au lieu de « Étude sur les familles ») : même si l'on n'a pas fondé soi-même une famille, chacun a une histoire familiale.

La taille totale de l'échantillon a été fixée à 445 000 personnes, le partage hommes/femmes ne se faisant pas à l'équilibre mais dans un rapport 3/5, soit 167 000 hommes et 278 000 femmes. Cet échantillon devait en outre être conçu de sorte à être représentatif au niveau régional. Le taux de participation ayant été un peu plus faible que prévu (enquête non obligatoire contrairement au recensement : taux de refus de l'ordre de 14 %), on a donc recueilli in fine 380 000 bulletins, auprès de 145 000 hommes et 235 000 femmes.
Avec les données du recensement, on peut connaître précisément les caractéristiques des personnes qui ont refusé de répondre : sexe, âge, état matrimonial, catégorie socioprofessionnelle, taille du ménage… On élabore alors des coefficients de redressement pour corriger les résultats de l'enquête (afin qu'ils soient pleinement représentatifs).

Annexe

Concernant l'absence de question sur le PACS dans le Recensement de 1999 :
Si l'INSEE ne prend pas le Pacs en compte dans recensement, cela tient aux exigences de la loi. Le Pacs n'a pas à être révélé à un tiers si les individus ne le souhaitent pas, car il ne s'agit pas d'un statut matrimonial légal d'un point de vue juridique puisqu'il ne s'agit pas d'un mariage. Or, il est obligatoire, sous peine de sanction, de répondre à l'ensemble des questions posées dans le cadre du recensement. Il n'est donc pas possible légalement de demander aux personnes recensées si elles sont pacsées.
Pour plus d'informations concernant les statistiques sur le PACS, qui sont peu nombreuses, voir Patrick FESTY, « Pacs : l'impossible bilan » *in Population & Sociétés*, 369, INED, juin 2001.
http://www.ined.fr/publications/pop_et_soc/pes369/PES369.pdf

Analyse du questionnaire de l'enquête Famille

http://joseph.larmarange.free.fr/docs/demo_tours2004/bull33h.pdf

Enfants, beaux-enfants, petits-enfants, arrière-petits-enfants.

Comme dans les versions précédentes, le questionnaire commence par des questions sur les enfants de la personne enquêtée, *y compris les enfants adoptés*. Pour chaque enfant, on demande le sexe, la date et le lieu de naissance ainsi que, pour les enfants adoptés, la date de l'accueil de l'enfant. Le cas échéant, les âges au départ du domicile parental et éventuellement au décès (avec une précision sur les enfants mort-nés, que les personnes ont à cœur de mentionner) permettent de retracer les périodes pendant lesquelles la personne a élevé ses enfants. Les questions sur les enfants adoptés permettent de savoir - c'est une première - qui sont les adoptants, d'un point de vue social et démographique, et comment se sont diffusées les diverses pratiques de l'adoption depuis la Grande Guerre (le lieu de naissance de l'enfant permettant de séparer adoption nationale et adoption internationale).

Viennent ensuite des questions sur les beaux-enfants, entendus ici comme les enfants d'un conjoint (ou d'un ex-conjoint) que la personne a élevés. Pour les beaux-enfants, les questions sont identiques à celles des enfants, permettant ainsi, pour certaines études, d'assimiler enfants et beaux-enfants élevés.

Puis trois questions viennent compléter l'histoire parentale : l'accueil d'autres enfants, la présence de petits-enfants et arrière-petits-enfants avec leur nombre et l'âge de l'aîné d'entre eux, ce qui permettra enfin de savoir à quel âge en moyenne l'on devient grand-parent ou arrière-grand-parent. Cette dernière information n'est pas directement demandée dans le questionnaire, mais facilement déduite à partir de la connaissance relative à l'âge des parents et des enfants : principe général que l'on retrouvera à de nombreuses reprises dans un questionnaire.

Périodes de vie en couple, avec ou sans mariage.

Les questions sur l'histoire conjugale se concentrent sur les unions de fait, avec ou sans mariage. Après une question sur la situation actuelle, on demande les dates de début et de fin des première et dernière périodes de vie en couple, ainsi que les dates éventuelles de mariage, de rupture, de divorce et de décès du conjoint.

Sur la recommandation de la Commission nationale de l'informatique et des libertés, on a limité l'interrogation à deux unions au maximum, la première et la dernière, la notion d'union étant restreinte pour le passé aux unions ayant donné lieu à une co-résidence d'au moins six mois.

Des questions sur les caractéristiques du dernier conjoint (date et lieu de naissance, mariage éventuel avant l'union, profession) permettent de situer socialement les couples rompus.
D'un point de vue d'étude démographique, il aurait été préférable de poser des questions sur le premier conjoint, pour les personnes ayant vécu plusieurs unions, mais les tests ont montré la réticence des personnes à évoquer trop précisément un conjoint qui a été « remplacé » par un autre dans l'histoire conjugale.

Les questions sur les périodes de vie en couple et sur les enfants et beaux-enfants livrent une description précise des diverses situations familiales (familles « monoparentales » et « recomposées »). Aussi mettent-elles sur le même plan les cohabitations sans mariage et les couples mariés, en incluant éventuellement les enfants n'habitant pas avec le parent ou le beau-parent. Leur dimension rétrospective permet, en outre, de retracer la succession des types de famille dans lesquels les personnes ont pu vivre.

Au total, l'enquête Famille aborde trois calendriers, celui des naissances (complet), celui des unions (simplifiée) et celui de l'insertion sociale (sommaire).

Questions aux étudiants :

Le questionnaire sur l'histoire familiale permet-il de dire combien de personnes n'ont pas d'enfants ?
-> oui, grâce à la question 2, mais évidemment cette information n'a pas le même sens selon que les répondants ont 25 ans ou 75 ans, car on interroge alors des personnes qui ont fini leur vie féconde (après 50 ans) et d'autres qui peuvent ne pas l'avoir finie, voire ne pas l'avoir commencée.

Le questionnaire sur l'histoire familiale permet-il de donner le nombre d'enfant qu'ont eu les femmes au cours de leur vie féconde ?
-> oui, grâce à la question 3, mais les femmes de moins de 50 ans n'ont pas encore terminée leur vie féconde. On considère que celles qui ont plus de 50 ans l'ont terminée, mais seules celles qui sont encore en vie sont interrogeables. On a donc un effet de sélection et il faut alors faire l'hypothèse que les femmes qui sont décédées avant 50 n'ont pas eu une fécondité différentes de celles des femmes qui sont toujours en vie.

Le questionnaire permet-il de dire à quel âge les personnes interrogées ont eu en moyenne leur premier enfant ?
-> oui, grâce à la question 3, pour les femmes comme pour les hommes (âge de la paternité) et on connaît aussi la durée intergénésique.

Le questionnaire permet-il de dire à quel âge la mère du répondant a eu son premier enfant ?
-> non, on ne connaît pas l'âge de la mère ?

Que peut-on dire concernant la fécondité de la mère du répondant ?
-> voir cadre 17

Que peut-on dire concernant la fécondité du père du répondant ?
-> rien

Le questionnaire permet-il de recueillir une information sur la situation actuelle et l'évolution des comportements démographiques dans les différents groupes sociaux ?
-> oui, car le bulletin individuel du recensement croisé avec le bulletin de l'enquête famille donne les caractéristiques socioprofessionnelles du répondant

Le questionnaire permet-il de recueillir une information sur des comportements démographiques au sein de différentes régions ?
-> oui, car le bulletin individuel du recensement tout comme ce (cartouche du) questionnaire sur l'histoire familiale est localisée (commune et district) et parce que l'échantillon a été choisie de telle sorte qu'il soit représentatif au niveau régional.

Le questionnaire permet-il d'estimer, par exemple, la mortalité infantile (de moins d'un an) des enfants nés en France en 1980 ?
-> oui, grâce à la question 3 puisqu'on demande si l'enfant est né en France (des femmes qui vivent en France aujourd'hui peuvent avoir accoucher à l'étranger).

Le questionnaire permet-il de calculer le nombre de grands-parents et d'arrière grands-parents vivant en France ?
-> oui, questions 7 et 8

Le questionnaire renseigne-t-il sur la promotion sociale intergénérationnelle ?
Oui avec bulletin du RGP : grâce à la question 18 et grâce aux questions sur l'activité professionnelle dans le bulletin individuel du RGP.

- Le questionnaire permet-il de calculer la durée des unions achevées (légales et de fait) ?
-> oui, question 14

Présentation du questionnaire sur la « taille des familles »

http://joseph.larmarange.free.fr/docs/demo_tours2004/quest.pdf

Nous avons imaginé une situation d'enquête. Il s'agit d'un exercice puisque l'enquête ne répond à aucune demande (il n'y a pas de client). Cependant, nous l'avons conçue avec rigueur et une enquête réalisée pour une institution serait menée de façon comparable.

Les objectifs pédagogiques

- Comprendre la spécificité d'un questionnaire démographique
  Le questionnaire est centré sur des variables typiquement démographiques ; il n'y a pas de recueil d'opinions (aspects beaucoup plus fréquents dans des questionnaires sociologiques -> voir cours et TD en deug 2)
- Identifier les étapes d'une procédure de construction d'un questionnaire démographique (associant rigueur et préparation technique)
  - Une réflexion est nécessaire en amont sur le choix des variables et leur exploitation.
  - Le choix des variables conduit à l'énoncé de questions qui seront formulées pour recueillir le mieux possible l'information souhaitée.
  - Des variables collectées, d'autres peuvent être déduites.
- Effectuer la passation du questionnaire (20 questionnaires par étudiant), sa codification puis sa saisie (sur une feuille de calcul Excel).
- Faire une analyse du fichier de collecte : analyse des spécificités des questionnaires collectés (par groupe de 5) et analyse d'une variable (au choix de l'étudiant) sur l'ensemble des questionnaires fusionnés. Cette analyse, qui comparera « nos » résultats à des résultats officiels, permettra d'indiquer la fiabilité de « notre » questionnaire

Ce travail sera évalué et constituera une des deux notes du contrôle continu de TD

Le questionnaire

Il s'agit d'un questionnaire permettant de repérer les générations ascendantes et descendantes d'un individu
Le questionnaire permet alors de recueillir des informations sur 4 « générations ».

- celle interrogée âgée de 50 à 70 ans : la génération notée « G »
- celle de leurs parents, vivants ou décédés : la génération notée « G+1 »
- celles de leurs enfants : « G-1 »,
- et des petits enfants : « G-2 »

Les informations collectées permettront de produire certaines connaissances relatives à la mortalité, la fécondité, la taille des familles, le statut matrimonial. Un des buts étant d'apprécier les évolutions entre générations.

Errata : à la question v08, il s'agit de « *y compris votre père* » et non « *y compris vous* » et à la question v12, il s'agit de « *y compris votre mère* » et non « *y compris vous* ».

Repérage des informations démographiques

A partir de ce questionnaire, il est possible de calculer de nombreux indicateurs démographiques. Certains peuvent également être calculés dans l'enquête Famille, certains existent dans le RGP. Des comparaisons sont alors possibles. D'autres sont spécifiques à ce questionnaire et apportent donc une information supplémentaire :

Exemples :

- Âge et statut professionnel (évolution des actifs, des retraités, en fonction de l'âge)
- Âge moyen à la première naissance en G+1, G et G-1 pour des femmes n'ayant pas toutes terminées leur vie féconde (différence en fonction du sexe du répondant).
- Écart d'âge maximal entre enfants de premier et dernier rang en G et G-1
- Proportion d'orphelins en G et âge au décès des parents (à analyser par génération).
- Taille des fratries dans les 3 générations les plus âgées (G+1, G et G-1)
- Descendance finale femmes et hommes de G+1 et G
- Descendance atteinte à différents âges pour G+1 et G
- Âge au décès en G+1, G-1 et G-2 (surestimation attendue de la longévité, car effet de sélection lié à la survie de G)
- Mortalité infantile (avant 1 an) et infanto-juvénile (avant 5 ans) dans la génération G-1.
- Statut matrimonial légal et situation de couple réelle, génération G
- Intervalles intergénésiques (entre deux naissances), générations G et G+1. Comparaison Hommes / Femmes.
- Etc.

La passation du questionnaire : un exemple

Exemple d'entretien à coder :

« L'étudiant fait partie du TD 2, son groupe est le 5, il s'agit du 18ème questionnaire. Il interroge sur le marché de sa commune une personne qui semble faire partie de la cible. Elle paraît un peu jeune, mais l'étudiant veut éviter d'avoir tous « ses » individus au centre du groupe d'âge. En fait, après premières questions, on apprend que la personne vit à La Riche et qu'elle est née en 1941 (elle aura donc 63 ans dans l'année, mais fait beaucoup plus jeune que son âge : comme quoi, les impressions sont trompeuses !).

L'homme (car c'est un individu de sexe masculin) déclare d'abord être marié. Quand je lui demande s'il est toujours en première union ou s'il s'est remarié, il m'indique alors qu'il n'est pas vraiment marié, mais qu'il vit en concubinage, mais qu'il ne fait pas la différence.

Que faut-il donc lui demander ?

Il se trouve que c'est un homme qui a divorcé de sa première femme en 1975 et qui a vécu maritalement avec son actuelle conjointe à partir de 1974 (c'est lui qui le dit).

Que fait-on de ces dates ? (rien)

L'homme se dit « rentier », mais il indique cependant qu'il touche une retraite d'un emploi qu'il a exercé pendant une vingtaine d'années entre 1960 et 1982. (on coche donc « retraité »).

Il ne connaît pas la date de naissance de son père, mais il sait que ce dernier est décédé un an après son divorce à l'âge de 56 ans (donc année de naissance en 1920) d'un accident brutal de la route (évidemment les conditions du décès ne sont pas relevées ici, mais « notre » homme est loquace). Notre répondant avait 2 oncles et a toujours une tante du côté de son père. Il sait aussi que la mère de son père avait eu un enfant décédé en bas âge. Ses parents avaient aussi adopté un enfant français handicapé âgé de 3 ans au moment de l'adoption.

Sa mère, fille unique, est née en 1922 et est toujours en vie. Elle s'est remariée.

Notre homme a une sœur aînée et un frère cadet. L'aînée est de la génération 1939 et le frère va avoir 59 ans cette année.

Le répondant a eu 2 garçons avec sa première femme. Le premier est né en 1963 et le second a 4 ans de moins. Avec sa compagne actuelle, il a eu 2 jumelles dont une est décédée à la naissance en 1980. Son fils aîné a deux enfants mais il ne sait pas précisément leur âge car depuis le divorce il ne voit plus ses garçons. Sa fille a une petite fille qui vient d'avoir 1 an il y a tout juste deux semaines. ».

Réponse avec la codification ci-dessous :

Corrigé de l'exemple :

UNIVERSITE FRANCOIS RABELAIS

Enquête démographique sur la taille des familles

« Madame, Monsieur, nous réalisons actuellement une enquête sur la taille des familles en Indre-et-Loire. Cette enquête est totalement anonyme et les données individuelles resteront confidentielles conformément aux principes de la CNIL (Commission Nationale Informatique et Liberté). Nous vous remercions de nous consacrer quelques minutes afin de répondre à nos questions ».

Consignes : Utiliser un stylo bleu ou noir pour remplir le questionnaire, un stylo rouge pour la codification. Les absences d'information (questions non posées) seront codées en 98. NSP signifie *« Ne sait pas »* et NC *« Non Concerné »*.

Identifiant n° TD : 2 n° Groupe : 5 n° Questionnaire : 0 1 8 — id1 2 ; id2 5 ; id3 0 1 8 ; ident 2 5 0 1 8

Caractéristiques individuelles

Dans quel département se situe votre résidence principale ? 3 7
Si différent de 37 ou Refus ➔ ARRET DU QUESTIONNAIRE

En quelle année êtes-vous né(e) ? 1 9 4 1 — v01 1 9 4 1
Si ? 1933 ou ? 1954 ou Refus ➔ ARRET DU QUESTIONNAIRE

Sexe de la personne interrogée ? ☑1 Homme ☐2 Femme — v02 1

Quelle est situation matrimoniale ? — v03 0 5
☐1 Célibataire ☑ Vit maritalement ☐ Marié ☐8 En instance de divorce ☐9 Veuf/Veuve ☐10 Divorcé
☐99 Refus ☐6 Premier mariage ☐7 Remariage
☐2 Jamais marié ☐3 En inst. de div. avec conjoint précédent ☐4 Veuf(ve) ☑5 Divorcé(e)

Êtes-vous ? ☐1 en activité temps plein ☐2 en activité temps partiel ☐3 en préretraite progressive — v04 0 5
☐4 en préretraite totale ☑5 en retraite ☐6 chômeur ☐7 autres (au foyer, pension de réversion ou d'invalidité)
☐99 Refus

Calculé à partir de l'âge au décès. 1920=1976-56

Père

En quelle année est né votre père ? 1 9 2 0 ☐9996 NSP/Père inconnu ☐9999 Refus — v05 1 9 2 0
Si NSP, demandez l'âge actuel ___ du père ou bien l'année du décès ___ en plus de l'âge au décès.
Est-il toujours en vie ? ☑0 Non ☐1 Oui (➔ v07=997) ☐96 NSP ☐99 Refus — v06 0
À quel âge est-il décédé ? 5 6 ans ☐996 NSP ☐999 Refus — v07 5 6
Si NSP, demandez éventuellement l'année du décès 1976 ou bien l'âge du répondant au décès de son père ___.
Combien de frères et sœurs votre père avait-il ? 4 (en comptant les personnes décédés. Si demi-frères ou demi-sœurs, ne compter que ceux de même mère. Ne pas compter les enfants adoptés).
Donc votre grand-mère paternelle (la mère de votre père) a eu 5 *enfants, y compris votre père ?* — v08 5
☐96 NSP ☐99 Refus

Mère

En quelle année est née votre mère ? 1 9 2 2 ☐9996 NSP/Mère inconnue ☐9999 Refus — v09 1 9 2 2
Si NSP, demandez l'âge actuel ___ de la mère ou bien l'année du décès ___ en plus de l'âge au décès.
Est-elle toujours en vie ? ☐0 Non ☑1 Oui (➔ v11=997) ☐96 NSP ☐99 Refus — v10 1
À quel âge est-elle décédée ? ___ ans ☐996 NSP ☐999 Refus — v11 9 9 7
Si NSP, demandez éventuellement l'année du décès ___ ou bien l'âge du répondant au décès de son père ___.
Combien de frères et sœurs votre mère avait-elle ? 0 (en comptant les personnes décédés. Si demi-frères ou demi-sœurs, ne compter que ceux de même mère. Ne pas compter les enfants adoptés).
Donc votre grand-mère maternelle (la mère de votre mère) a eu 1 *enfants, y compris votre mère ?* — v12 1
☐96 NSP ☐99Refus

Fratrie

Combien votre mère a-t-elle eu d'enfants, y compris vous et ceux qui sont maintenant décédés ?
(ne pas tenir compte des enfants adoptés) 3 enfants ☐96 NSP ☐99 Refus — v13 3
Parmi ces enfants, combien sont nés avant vous ? 1 (si l'aîné mettre 0) ☐96 NSP ☐99 Refus — v14 1
Quelle est l'année de naissance du premier enfant ? 1 9 3 9 ☐9996 NSP(âge actuel ___) ☐9999 Refus — v15 1 9 3 9
Quelle est l'année de naissance du dernier enfant ? 1 9 4 6 ☐9996 NSP(âge actuel 58) ☐9999 Refus — v16 1 9 4 6
NB : 1946 = 2003-58.

Descendance **ATTENTION : On considère que l'âge actuel est l'âge au 31/12/2003**

Combien avez-vous eu d'enfants (y compris ceux qui sont décédés aujourd'hui) ? 4 enfants — v17 4
(ne pas compter les enfants adoptés) *Si aucun ou Refus ➔ Fin questionnaire* ☐96 NSP ☐99 Refus

Pour chacun de vos enfants en allant du plus âgé au plus jeune ?

$	Sexe 1 M 2 F	Année de naissance Si NSP, mettre âge entre ( )	Toujours en vie ? 1 Oui 0 Non	Si décédé, âge au décès ? Si NSP, mettre année entre ( )	A-t-il (eu) des enfants ? 1 Oui 0 Non	Si oui, combien (y compris décès) ? NC = 97	Année de naissance de l'aîné Si NSP, mettre âge entre ( ) NC = 9997
01	1	1963 ( )	1	97 ( )	1	2	9996 (NSP)
02	1	1967 ( )	1	97 ( )	0	97	9997 ( )
03	2	1980 ( )	0	00 ( )	0	97	9997 ( )
04	2	1980 ( )	1	97 ( )	1	1	2003 ( )
05		( )		( )			( )
06		( )		( )			( )
	enf1	enf2	enf3	enf4	enf5	enf6	enf7

Si plus de 6 enfants, prendre un second questionnaire et numérotez les lignes. Pensez à inscrire l'identifiant.

Nous vous remercions d'avoir répondu à nos questions.

La variable correspondant à chaque cellule du tableau est obtenue en prenant la racine de la variable (dernière ligne du tableau) et en y ajoutant $ et le rang de l'enfant (première colonne).

Exemple :
La variable de la seconde colonne et troisième ligne est : ***enf2$03***.

Lignes vides :
Coder en 97 sauf enf2 et enf7 en 9997.

Quelques éléments méthodologiques de codification

Sur la majorité des logiciels de statistiques, les données sont saisies sous la forme d'un immense tableau où chaque ligne correspond à un questionnaire, et où les différentes variables sont représentées en colonne. Le nombre de variables correspond au nombre d'informations connues pour chaque individu, et non au nombre de questions posées dans le questionnaire. Certaines questions peuvent en effet devoir être codées à l'aide de plusieurs variables. C'est notamment le cas avec les questions à choix multiples.

Cas 1 : Vous demandez la CSP. Les personnes interrogées ne peuvent répondre qu'une seule des modalités proposées. ➔ La question sera codée en une seule variable où à chaque modalité on attribuera un code.

Cas 2 : Vous demandez à des étudiants les deux raisons principales pour lesquelles ils se sont inscrits à l'université. Vous avez donc deux informations que vous devrez coder dans deux variables. Chaque modalité aura un code et le code de la première modalité sera mis dans la première variable, celui de la seconde modalité dans la seconde. Il est possible ainsi de coder des réponses hiérarchisées (on demande la raison principale, puis la seconde, puis la troisième, etc.). Pensez à prévoir dans la seconde variable un code indiquant que la personne n'a donné qu'une seule réponse au lieu de deux. (Adaptez suivant la circonstance).

Cas 3 : Vous demandez si les personnes possèdent certains équipements et vous proposez une liste. Dans le cas présent il est donc possible de cocher toutes les cases ou aucunes. On cherche donc à savoir, pour chaque modalité, si la case a été cochée. On créera donc autant de variables que d'équipements, chaque variable comportant deux réponses : Oui ou Non.

Le dictionnaire de codification

Ce document est nécessaire dans toute enquête et doit être réalisé avec soin. C'est lui qui indique comment l'information recueillie va être saisie informatiquement afin de pouvoir procéder aux analyses. Un dictionnaire de codification mal réalisé pourra entraîner une perte massive d'informations et rendre certaines analyses irréalisables.

Le dictionnaire de codification indique également les variables filtres. Par exemple, si le père de l'individu est toujours en vie (v06=1), alors la question v07 ne concerne pas cet individu. Le dictionnaire de codification précise donc que la question v07 ne concerne que les individus ayant un père décédé (base : v06=0).
Dans cette enquête, un code est attribué lorsque les personnes sont NC (non concernées). Il s'agit du code 97. Pour les variables codées sur 3 ou quatre chiffres, on utilise alors les codes 997 et 9997. Le dictionnaire précise à chaque fois dans quel cas le code 97 s'applique (NB : le symbole <> signifie *différent de*).

D'autre part, nous avons décidé de distinguer trois situations :

- **Refus :** La personne ne veut pas répondre à la question.
  Code 99 (999 et 9999 si variable à trois ou quatre chiffres).
- **NSP (Ne Sait Pas) :** Il ne s'agit pas d'un manque de volonté de la personne enquêtée mais elle ne connaît pas ou ne se souviens pas de l'information qui lui est demandée.
  Code 96 (996 et 9996).
- **AI (Absence d'Information) :** idéalement, ce code ne devrait pas ce rencontrer dans la base. Il est utilisé en cas d'oubli de l'enquêteur, c'est-à-dire quand une information n'a pas été reportée sur le questionnaire.
  Code 98 -998 et 9998).

Lorsque un de ces trois codes est appliqué à une question filtre, le code 97 s'applique alors aux questions situées sous ce filtre.

Le masque de saisie

C'est un fichier Excel conçu pour la saisie des questionnaires.
Ce fichier n'autorise à la saisie que les valeurs prévues pour chaque variable. De plus, des contrôles ont été mis par rapport aux variables filtres ainsi que des contrôles de cohérence (par exemple, une femme ne peut pas avoir d'enfant l'année de ces cinq ans).
L'ensemble de ces contrôles implique qu'en cas d'erreur de saisie à la question filtre, il faut d'abord corriger celle-ci avant de pouvoir saisir les variables suivantes.
L'ensemble des contrôles sont fournis dans le dictionnaire de codification et s'affichent à l'écran pendant la saisie. Toutes les variables doivent être impérativement renseignées. Seule exception, dans le cas d'une femme ayant **n** enfants, les variables correspondant aux enfants de rang **n+1** et plus pourront rester vide à la saisie. Elles seront traitées par la suite informatiquement pour être remplacées automatiquement par 97, 997 ou 9997 selon le cas.

Consignes aux étudiants

- Au sein de chaque TD, constitution de groupes de 5 étudiants avec une personne coordinatrice nommée.

- Chaque étudiant soumettra 20 fois le questionnaire (une partie peut se faire par téléphone en sollicitant le bouche à oreille : penser à sa famille, aux amis de ses parents, aux amis des amis ; une autre partie devra se faire dans des lieux publics : sortie d'une boulangerie, jour de marché, lors de l'attente à la gare routière, etc.). Les personnes ayant refusé de répondre seront comptées par les étudiants (dénombrement en fonction du sexe) sur la « fiche récapitulative de terrain ».

- Critères à retenir pour chaque étudiant : sélectionner 12 femmes et 8 hommes âgés de 50 à 70 ans. Attention : les étudiants devront avoir des personnes <u>dans tous les groupes d'âges quinquennaux</u>. Il faudrait privilégier ce type de distribution :

Age	femmes	Hommes	total
50-54	4	3	7
55-59	3	2	5
60-64	3	2	5
65-69	2	1	3
	12	8	20

…plutôt que celle-ci (type courbe de Gauss) : c'est ce qui se produira pourtant si les étudiants n'y prennent pas garde, en se fiant seulement à la « tête des gens » :

Age	femmes	hommes	total
50-54	2	1	3
55-59	5	3	8
60-64	5	3	8
65-69	2	1	3
	12	8	20

- Les étudiants doivent effectuer la codification manuelle de chaque questionnaire à l'aide du dictionnaire de codification et faire la saisie des données sur le fichier Excel proposé. Le masque de saisie est téléchargeable sur le site de Joseph Larmarange (http://joseph.larmarange.free.fr). En cas de problème de codification ou problème pour saisir un questionnaire en particulier, contactez un enseignant. Un questionnaire où il manque plus des deux tiers des réponses (refus, AI) sera écarté. Il sera néanmoins rendu à l'enseignant avec la mention « NON SAISI ».

- Dans chaque groupe, un étudiant est chargé de faire une présentation de 3 pages maximum des questionnaires de son groupe (répartition de la population en fonction des variables, calculs de certaines moyennes). Avis sur les caractéristiques du sous-échantillon.

- Les autres étudiants du groupe choisissent un indicateur (ex : l'âge moyen à la première naissance) et font une comparaison des données de l'enquête (2000 à 3000 questionnaires) avec des données nationales ou régionales. Ces dernières pourront être données aux étudiants ou seront à rechercher dans des revues de l'Insee, diverses publications démographiques… Des précisions et certaines données seront fournies ultérieurement.

- Echéances : les groupes d'étudiants devront rendre pour le 17 et 19 mars leurs questionnaires codifiés et saisis, soit sur disquette, soit de préférence par e-mail (si par e-mail au plus tard les 16 et 18 mars) : **un seul fichier par groupe (100 questionnaires saisis par fichier si groupe de 5, 120 si groupe de 6)**. Les enseignants se chargeront de la concaténation des différents fichiers et mettront ensuite à disposition des étudiants la totalité du fichier. Le fichier saisi devra se nommer sous la forme td2gr3.xls (format Excel) pour le groupe 3 du td 2. La fiche bilan de collecte de chaque groupe devra être rendue en même temps que le fichier saisi.

- Les questionnaires papier seront également à rendre. La codification sera vérifiée. Au dos du questionnaire, l'étudiant indiquera où/comment il a réalisé chaque passation (en 2 lignes maximum). En cas de problème avec le répondant, l'étudiant devra aussi l'indiquer au dos du questionnaire.

- En TP d'informatique, les étudiants vont découvrir Excel lors des prochaines séances. Mais la saisie des questionnaires devra se faire en dehors des séances.

UNIVERSITE FRANCOIS RABELAIS

Enquête démographique sur la taille des familles

« Madame, Monsieur, nous réalisons actuellement une enquête sur la taille des familles en Indre-et-Loire. Cette enquête est totalement anonyme et les données individuelles resteront confidentielles conformément aux principes de la CNIL (Commission Nationale Informatique et Liberté). Nous vous remercions de nous consacrer quelques minutes afin de répondre à nos questions ».

Consignes : Utiliser un stylo bleu ou noir pour remplir le questionnaire, un stylo rouge pour la codification. Les absences d'information (questions non posées) seront codées en 98. NSP signifie *« Ne sait pas »* et NC *« Non Concerné »*.

Identifiant

n° TD : ☐ n° Groupe : ☐ n° Questionnaire : ☐☐☐

id1 ☐
id2 ☐
id3 ☐☐☐
ident ☐☐☐☐☐

Caractéristiques individuelles

Dans quel département se situe votre résidence principale ? ☐☐

Si différent de 37 ou Refus ➔ ARRET DU QUESTIONNAIRE

En quelle année êtes-vous né(e) ? 1☐☐☐ — v01 ☐☐☐☐

Si ≤1933 ou ≥1954 ou Refus ➔ ARRET DU QUESTIONNAIRE

Sexe de la personne interrogée ? ☐1 Homme ☐2 Femme — v02 ☐

Quelle est votre situation matrimoniale ? — v03 ☐☐

☐1 Célibataire ☐ Vit maritalement ☐ Marié ☐8 En instance de divorce ☐9 Veuf/Veuve ☐10 Divorcé
☐99 Refus
→ ☐6 Premier mariage ☐7 Remariage
→ ☐2 Jamais marié ☐3 En inst. de div. avec conjoint précédent ☐4 Veuf(ve) ☐5 Divorcé(e)

Êtes-vous ? ☐1 en activité temps plein ☐2 en activité temps partiel ☐3 en préretraite progressive — v04 ☐☐
☐4 en préretraite totale ☐5 en retraite ☐6 chômeur ☐7 autres (au foyer, pension de réversion ou d'invalidité)
☐99 Refus

Père

En quelle année est né votre père ? 1☐☐☐ ☐9996 NSP/Père inconnu ☐9999 Refus — v05 ☐☐☐☐
Si NSP, demandez l'âge actuel ☐ du père ou bien l'année du décès ☐ en plus de l'âge au décès.
Est-il toujours en vie ? ☐0 Non ☐1 Oui (➔ v07=997) ☐96 NSP ☐99 Refus — v06 ☐☐
→ ***À quel âge est-il décédé ?*** ☐☐☐ ans ☐996 NSP ☐999 Refus — v07 ☐☐☐
Si NSP, demandez éventuellement l'année du décès ☐ ou bien l'âge du répondant au décès de son père ☐.
Combien de frères et sœurs votre père avait-il ? ☐☐ (en comptant les personnes décédés. Si demi-frères ou demi-sœurs, ne compter que ceux de même mère. Ne pas compter les enfants adoptés).
Donc votre grand-mère paternelle** (la mère de votre père) **a eu** ☐☐ **enfants, y compris votre père ? — v08 ☐☐
☐96 NSP ☐99 Refus

Mère

En quelle année est née votre mère ? 1☐☐☐ ☐9996 NSP/Mère inconnue ☐9999 Refus — v09 ☐☐☐☐
Si NSP, demandez l'âge actuel ☐ de la mère ou bien l'année du décès ☐ en plus de l'âge au décès.
Est-elle toujours en vie ? ☐0 Non ☐1 Oui (➔ v11=997) ☐96 NSP ☐99 Refus — v10 ☐☐
→ ***À quel âge est-elle décédée ?*** ☐☐☐ ans ☐996 NSP ☐999 Refus — v11 ☐☐☐
Si NSP, demandez éventuellement l'année du décès ☐ ou bien l'âge du répondant au décès de son père ☐.
Combien de frères et sœurs votre mère avait-elle ? ☐☐ (en comptant les personnes décédés. Si demi-frères ou demi-sœurs, ne compter que ceux de même mère. Ne pas compter les enfants adoptés).
Donc votre grand-mère maternelle** (la mère de votre mère) **a eu** ☐☐ **enfants, y compris votre mère ? — v12 ☐☐
☐96 NSP ☐99Refus

Fratrie

Combien votre mère a-t-elle eu d'enfants, y compris vous et ceux qui sont maintenant décédés ?
(ne pas tenir compte des enfants adoptés) ☐☐ enfants ☐96 NSP ☐99 Refus — v13 ☐☐
Parmi ces enfants, combien sont nés avant vous ? ☐☐ (si l'aîné mettre 0) ☐96 NSP ☐99 Refus — v14 ☐☐
Quelle est l'année de naissance du premier enfant ? 1☐☐☐ ☐9996 NSP(âge actuel ☐) ☐9999 Refus — v15 ☐☐☐☐
Quelle est l'année de naissance du dernier enfant ? ☐☐☐☐ ☐9996 NSP(âge actuel ☐) ☐9999 Refus — v16 ☐☐☐☐

Descendance

Combien avez-vous eu d'enfants (y compris ceux qui sont décédés aujourd'hui) ? ☐☐ enfants — v17 ☐☐
(ne pas compter les enfants adoptés) *Si aucun ou Refus ➔ Fin questionnaire* ☐96 NSP ☐99 Refus

Pour chacun de vos enfants en allant du plus âgé au plus jeune ?

$	*Sexe* 1 M 2 F	*Année de naissance* Si NSP, mettre âge entre ( )	*Toujours en vie ?* 1 Oui 0 Non	*Si décédé, âge au décès ?* Si NSP, mettre année entre ( )	*A-t-il (eu) des enfants ?* 1 Oui 0 Non	*Si oui, combien (y compris décès) ?* NC = 97	*Année de naissance de l'aîné* Si NSP, mettre âge entre ( ) NC = 9997
01	☐	☐☐☐☐ (☐)	☐	☐☐ (☐)	☐	☐☐	☐☐☐☐ (☐)
02	☐	☐☐☐☐ (☐)	☐	☐☐ (☐)	☐	☐☐	☐☐☐☐ (☐)
03	☐	☐☐☐☐ (☐)	☐	☐☐ (☐)	☐	☐☐	☐☐☐☐ (☐)
04	☐	☐☐☐☐ (☐)	☐	☐☐ (☐)	☐	☐☐	☐☐☐☐ (☐)
05	☐	☐☐☐☐ (☐)	☐	☐☐ (☐)	☐	☐☐	☐☐☐☐ (☐)
06	☐	☐☐☐☐ (☐)	☐	☐☐ (☐)	☐	☐☐	☐☐☐☐ (☐)
	enf1	enf2	enf3	enf4	enf5	enf6	enf7

Si plus de 6 enfants, prendre un second questionnaire et numérotez les lignes. Pensez à inscrire l'identifiant.

Nous vous remercions d'avoir répondu à nos questions.

La variable correspondant à chaque cellule du tableau est obtenue en prenant la racine de la variable (dernière ligne du tableau) et en y ajoutant $ et le rang de l'enfant (première colonne).

Exemple :
La variable de la seconde colonne et troisième ligne est : ***enf2$03***.

Lignes vides :
Coder en 97 sauf enf2 et enf7 en 9997.

Dictionnaire de codification

AI : Absence d'Information
NC : Non Concerné
NSP : Ne Sais Pas

Les variables **c01** à **c09** sont des variables calculées à partir des variables recueillies.
Les variables **cenf01** à **cenf06** Sont des variables calculées spécifiques au fichier *enfants*.

Var.	Libellé	Modalités
id1	Numéro du TD	*1 à 6*
id2	Numéro du groupe	*1 à 9*
id3	Numéro du questionnaire dans le groupe	*1 à 200*
ident	Identifiant	*=id1*10.000+id2*1000+id3*
v01	Année de naissance	*1934 à 1953*
v02	Sexe	1. Homme 2. Femme
v03	Situation matrimoniale	1. Célibataire 2. Vit maritalement (jamais marié) 3. Vit maritalement (en instance de divorce avec conjoint précédent) 4. Vit maritalement (Veuf(ve)) 5. Vit maritalement (Divorcé(e)) 6. Marié (Premier mariage) 7. Remarié 8. En instance de divorce 9. Veuf/veuve 10. Divorcé(e) 98. AI 99. Refus
v04	Activité	1. en activité temps plein 2. en activité temps partiel 3. en préretraite progressive 4. en préretraite totale 5. en retraite 6. chômeur 7. autres (au foyer, pension de réversion ou d'invalidité,…) 98. AI 99. Refus
v05	Année de naissance du père	*Sur 4 chiffres* 9996. NSP / Père inconnu 9998. AI 9999. Refus *v05<(v01-10)* *v05>(v01-100)*
v06	Père toujours en vie ?	0. Non 1. Oui 96. NSP 98. AI 99. Refus

v07	Âge au décès du père	*Sur trois chiffres* *>10* *<130* 996. NSP 997. NC (*si v06<>0*) 998. AI 999. Refus
v08	Descendance de la grand-mère paternelle	*Sur deux chiffres* *>0* *<25* 96. NSP 98. AI 99. Refus
v09	Année de naissance de la mère	*Sur 4 chiffres* 9996. NSP / Père inconnu 9998. AI 9999. Refus *v05<(v01-10)* *v05>(v01-70)*
v10	Mère toujours en vie ?	0. Non 1. Oui 96. NSP 98. AI 99. Refus
v11	Âge au décès de la mère	*Sur trois chiffres* *>10* *<130* 996. NSP 997. NC (*si v10<>0*) 998. AI 999. Refus
v12	Descendance de la grand-mère maternelle	*Sur deux chiffres* *>0* *<25* 96. NSP 98. AI 99. Refus
v13	Descendance de la mère	*Sur deux chiffres* *>0* *<25* 96. NSP 98. AI 99. Refus
v14	Nombre de naissances avant l'interviewé(e) =Rang de l'interviewé(e)-1	*De 0 à 24* 96. NSP 98. AI 99. Refus *<v13*
v15	Année de naissance du premier	*Sur quatre chiffres* *Si v13=1, alors =v01* *≤ v01* *> (v01-60)* 9996. NSP 9998. AI 9999. refus

v16	Année de naissance du dernier	*Sur quatre chiffres* *Si v13=1, alors v16=v01* *≥ v01* *< (v01+60)* 9996. NSP 9998. AI 9999. refus
v17	Descendance de l'interrogé(e)	*De 0 à 20* 96. NSP 98. AI 99. Refus
Pour i variant de 01 à 12		
enf1$i	Sexe de l'enfant i	1. Homme 2. Femme 96. NSP 97. NC si v17<i ou 96 ou 98 ou 99 98. AI 99. Refus
enf2$i	Année de naissance de l'enfant i	*Sur quatre chiffres* *≥(v01+10)* *≤2004* *Si i≥2 et si enf2$(i-1) connu, ≥enf2$(i-1)* 9996. NSP 9997. NC si enf1$i=97 9998. AI 9999. Refus
enf3$i	Enfant i toujours en vie ?	0. Non 1. Oui 96. NSP 97. NC si enf1$i=97 98. AI 99. Refus
enf4$i	Âge au décès	*Sur deux chiffres* *≤ 2004-enf2$i si connu sinon ≤2004-v01-10* 96. NSP 97. NC si enf3$i<>0 98. AI 99. Refus
enf5$i	A eu des enfants ?	0. Non 1. Oui 96. NSP 97. NC si enf1$i=97 98. AI 99. Refus *Si enf2$i connu, et si 2004-enf2$i≤10 alors interdire 1.*
enf6$i	Nombre d'enfants qu'a eu i	*Sur deux chiffres de 1 à 20.* 96. NSP 97. NC si enf5$i<>1 98. AI 99. Refus

enf7$i	Année de naissance de l'aîné	*Sur quatre chiffres.* *Si enf2$i connu, ≥enf2$i+10* *≤2004* 9996. NSP 9997. NC si enf5$i<>1 9998. AI 9999. Refus
Variables Calculées		
c01	Âge par différence de millésimes à la première maternité de la mère du répondant.	*= v15-v09* 97. NC si v09>9000 ou v15>9000
c02	Âge par différence de millésimes à la première maternité ou paternité du répondant.	*=enf2$01-v01* 97. NC si enf2$01>9000
c03	Année de naissance du dernier enfant de l'enquêté.	*=enf2$i avec i=v17* 9997. NC si v17=0 ou si v17>90
c04	Intervalle entre la première naissance et la dernière naissance de l'enquêté.	*=c03-enf2$01* 97. NC si v17<2 ou v17>90 ou enf2$01>9000 ou c03>9000
c05	Intervalle intergénésique moyen sur l'ensemble des naissances de l'enquêté. (Hypothèse : les jumeaux sont négligeables)	*=c04/(v17-1)* 97. NC si c04=97 or v17>90
c06	Intervalle entre la première naissance et la dernière naissance de la mère de l'enquêté.	*=v16-v15* 97. NC si v13>90 ou v13=1 ou v16>9000 ou v15>9000
c07	Intervalle intergénésique moyen sur l'ensemble des naissances de la mère de l'enquêté. (Hypothèse : les jumeaux sont négligeables)	*=c06/(v13-1)* 97. NC si c06=97
c08	État matrimonial légal	1. Célibataire (v03=1 ou 2) 2. Marié(e) (v03=3 ou 6 ou 7 ou 8) 3. Veuf(ve) (v03=4 ou 9) 4. Divorcé(e) (v03=5 ou 10) 97. Valeur manquante (v03>90)
c09	Situation de couple	1. Vit seul (v03=1, 8, 9 ou 10) 2. Vit maritalement (v03=2, 3, 4 ou 5) 3. Marié(e) (v03=6 ou 7) 97. Valeur manquante (v03>90)
Variables spécifiques au fichier enfant		
rang	Rang de l'enfant	
identenf	Identifiant enfant	*=ident*100+rang*
enf1	Sexe de l'enfant i	1. Homme 2. Femme 96. NSP 98. AI 99. Refus
enf2	Année de naissance de l'enfant	*Sur quatre chiffres* 9996. NSP 9998. AI 9999. Refus
enf3	Enfant toujours en vie ?	0. Non 1. Oui 96. NSP 98. AI 99. Refus

enf4	Âge au décès	*Sur deux chiffres* 96. NSP 97. NC si enf3<>0 98. AI 99. Refus
enf5	A eu des enfants ?	0. Non 1. Oui 96. NSP 98. AI 99. Refus
enf6	Nombre d'enfants qu'il a eu	*Sur deux chiffres de 1 à 20.* 96. NSP 97. NC si enf5<>1 98. AI 99. Refus
enf7	Année de naissance de l'aîné	*Sur quatre chiffres.* 9996. NSP 9997. NC si enf5<>1 9998. AI 9999. Refus
cenf01	Année de naissance de l'enfant précédant	*Sur quatre chiffres* 9997. NC si rang=1
cenf02	Année de naissance de l'enfant suivant	*Sur quatre chiffres* 9997. NC si rang=v17
cenf03	Naissance multiple (supposée multiple si né la même année que l'enfant précédent ou l'enfant suivi d'où une surestimation de la gémellité)	0. Non 1. Oui (enf2=cenf01 OR enf2=cenf02) 97. NC ((cenf01>9000 AND cenf02>9000) OR enf2>9000)
cenf04	Intervalle intergénésique avec naissance précédente (NC si naissance multiple).	*=enf2-cenf01* 97. NC (enf2>9000 OR cenf01>9000 OR cenf01=enf2)
cenf05	Intervalle intergénésique avec naissance suivante (NC si naissance multiple).	*=cenf02-enf2* 97. NC (enf2>9000 OR cenf02>9000 OR cenf02=enf2)
cenf06	Âge à la première paternité ou première maternité.	*=enf7-enf2* 97. NC (enf2>9000 OR enf7>9000)

UNIVERSITE FRANCOIS RABELAIS

Enquête démographique sur la taille des familles

Fiche récapitulative de terrain

N° TD : _____ **N° Groupe :** _____

Liste des étudiants dans ce groupe :

1. ..... 4. .....

2. ..... 5. .....

3. ..... 6. .....

Refus de répondre au questionnaire : ____ Hommes ____ Femmes

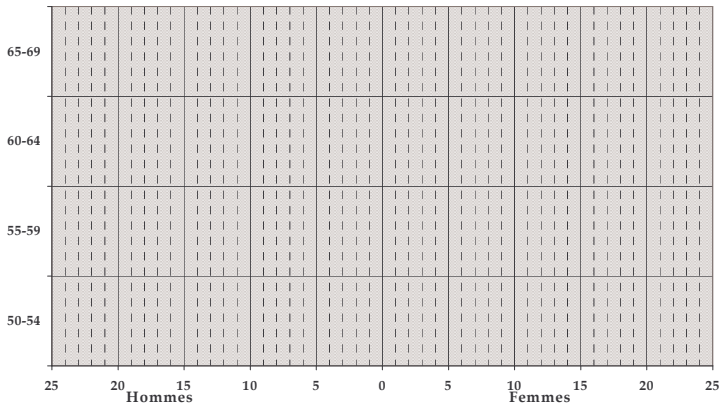
Nombre de questionnaires collectés (a) : _____

Nombre de questionnaires écartés (b) : _____

Nombre de questionnaires saisis (a-b) : _____

Récapitulatif de la collecte (effectifs) :

Géné-rations	Âge au 31/12/03	Hommes	Femmes	Ensemble
1949-53	50-54			
1944-48	55-59			
1939-43	60-54			
1934-38	65-69			
Ensemble				



Faire une pyramide des âges.

Lieux de passation des questionnaires :

.....

.....

.....

.....

.....

Commentaires / Observations :

.....

.....

.....

.....

.....

.....

.....

.....

UNIVERSITE FRANCOIS RABELAIS
DEUG 1 Sociologie – Enseignement de DEMOGRAPHIE
Laurent NOWIK / Aurélie GODARD / Joseph LARMARANGE

TD
2003/2004

Consignes pour l'analyse de l'enquête sur la taille des familles.

Rappel : le travail proposé a 4 objectifs pédagogiques :

1) sensibilisation aux modes de collecte des données démographiques ;
2) mise en pratique d'un dispositif d'enquête et réalisation d'une collecte jusqu'à la numérisation des données codifiées ;
3) analyse d'une petite partie des données collectées (avec appui d'un tableur) ;
4) appréciation de la qualité de la méthode en comparant les résultats de l'enquête à des données nationales, régionales ou départementales.

Le fichier complet est à télécharger avant le 25 avril 2004 à l'adresse :
http://joseph.larmarange.free.fr/enseignements/enseignements.php#demo_tours
(rappel : cette page internet contient aussi des corrections de TD utiles pour vos révisions)

Notation : les étudiants qui ont participé au travail de collecte seront évalués sur 8 points (deux premiers objectifs). Les 12 autres points porteront donc sur la deuxième partie du travail (les deux derniers objectifs).

Deux types d'analyse sont proposés :

- au sein des groupes un étudiant effectuera une présentation des 80 à 120 questionnaires collectés par le groupe. Ce travail dispense de l'objectif 4. *Voir annexe A pour les préconisations détaillées.*
- Les autres étudiants choisiront un sujet de leur choix parmi une liste. L'analyse sera effectuée à partir des fichiers complets résultant de l'enquête (3123 questionnaires collectés utilisables). Chaque étudiant du groupe devra choisir un sujet différent. Les étudiants qui le souhaiteraient pourront travailler par deux, mais dans ce cas le travail à rendre devra être deux fois plus volumineux. *Voir annexe B pour les préconisations détaillées.*

Délai : Le travail est à rendre pour le 4 mai 2004. Aucun dossier ne sera accepté après cette date. Les travaux seront récupérés à l'occasion du contrôle terminal de démographie. Le numéro du groupe de TD de l'étudiant et le numéro du groupe d'enquête au sein du TD doivent figurer en 1ère page. Les dossiers prêts avant cette date pourront être déposés dans le casier de M. Nowik au secrétariat du département de sociologie.

Annexe A – Liste des indicateurs à fournir sur les mini fichiers de 80 à 120 questionnaires.

1) Structure par âge, sexe et état matrimonial (V01xV02xV03)
2) Activité par âge et sexe (V04xV01xV02)
3) Survie des parents en fonction de l'âge du répondant (2 parents vivants, orphelins de père, de mère, de père et mère)
4) Taille moyenne de la fratrie de la mère des répondants (nombre incluant cette mère)
5) Taille moyenne de la fratrie des femmes répondantes (nombre incluant la répondante)
6) Nombre (moyen) d'enfants des femmes répondantes
7) Age moyen à la première naissance pour les mères des répondants (frère aîné ou sœur aînée du répondant, pouvant être le répondant lui-même)
8) Age moyen à la première paternité ou à la première maternité pour les répondants ayant eu des enfants (distinguer le sexe et l'âge du répondant)
9) Age moyen à la première paternité ou à la première maternité pour les enfants des répondants ayant eu eux-mêmes des enfants (distinguer le sexe et l'âge de l'enfant)

NB : Pour le calcul des âges moyens, vous devez d'abord calculer pour chaque individu l'âge par différence de millésime à la première maternité ou paternité. Pour cela, sous Excel, dans une colonne vide, vous écrivez la formule pour le premier individu (=*adresse de la cellule contenant l'année de naissance de l'enfant – adresse de la cellule contenant l'année de naissance du parent*) puis vous faites glisser la formule dans toutes les autres cellules de la colonne (à partir du petit carré en bas à droite du rectangle de sélection).

Chaque dossier doit comporter :

- *Les tableaux associés aux thèmes d'étude demandés ;*
- *une pyramide des âges (pour le premier thème) ;*
- *deux graphiques au choix tirés des tableaux ;*
- *un commentaire succinct (2000 à 3000 caractères espaces compris) des résultats tirés des tableaux.*

Annexe B – Liste des sujets possibles (1 sujet au choix de l'étudiant)

1) La mortalité des ascendants des répondants
2) La taille de la fratrie des parents des répondants (distinguer fratrie du père et fratrie de la mère)
3) L'âge moyen à la première naissance pour les mères des répondants (frère aîné ou sœur aînée du répondant, pouvant être le répondant lui-même)

4) Le statut matrimonial des répondants
5) La taille de la fratrie des répondants et les écarts d'âge entre l'aîné et le cadet (le plus jeune)
6) L'âge moyen à la paternité ou à la maternité pour les répondants ayant eu des enfants (distinguer âge moyen à la première naissance et toute naissance confondue, et faire les calculs en fonction du nombre d'enfants)
7) Les enfants des répondants étudiés du point de vue de leur nombre (descendance finale brute) et des écarts intergénésiques
8) La mortalité des enfants des répondants

9) L'âge moyen à la première paternité ou à la première maternité pour les enfants des répondants ayant eu eux-mêmes des enfants

Important : quel que soit le thème choisi, les étudiants devront prendre en compte l'âge et le sexe du répondant, mais aussi d'autres variables qu'ils jugeront pertinentes. En outre, dans certains cas, l'analyse ne peut se faire qu'en tenant compte de l'âge. Par exemple, pour le sujet n°1, la mortalité des ascendants est nécessairement fonction de l'âge du répondant (la proportion d'orphelins n'est pas la même à 50 et à 70 ans).

Le travail à rendre doit représenter 2 à 3 pages par étudiant. Chaque dossier devra préciser le sujet choisi (avec le numéro) et devra comporter :

- *au moins deux tableaux de données produit par l'étudiant*
- *un graphique (qui n'est pas nécessairement la simple transcription d'un tableau)*
- *un commentaire de 2000 à 3000 caractères (espaces compris) relatif aux données produites*
- *la recherche d'une donnée nationale, régionale ou départementale permettant d'apprécier la qualité de la méthode utilisée. Ainsi on pourra apprécier la fiabilité des données produites du point de vue de la connaissance démographique.*

Pour obtenir des données « officielles », les étudiants pourront utiliser les publications de l'INSEE (disponibles à la bibliothèque des Deux Lions), l'Internet, les sites de l'INED et de l'INSEE… Le site de l'INSEE fournit les données du recensement pour la région Centre par exemple.

Un fichier *individus* et un fichier *enfants*

Deux fichiers ont été crées à partir des saisies réalisées par les étudiants. Le premier, le « fichier individus » est la compilation des fichiers étudiants (avec suppression des enregistrements incohérents et remplissage des cases vides). Dans ce fichier, d'autres variables ont été créées en lien avec les sujets proposés.
Le second fichier, baptisé « fichier enfants » propose un enregistrement pour chaque enfant déclaré par le ou la répondante. Il est plus adapté à certaines études portant sur les enfants.

Le fichier *individus*.

Il s'agit de l'ensemble des questionnaires saisis. Il y a en tout 3.123 individus. Le nombre maximum d'enfants enregistrés est de 12. Les variables enf1$13 à enf7$20 ont donc été supprimées du fichier.
A l'ouverture du fichier, un mot de passe est requis. Il sera communiqué en TD.
Neuf variables supplémentaires ont été créées. Elles sont notées **c01** à **c09**. Voici leur description :

Var.	Libellé	Modalités
c01	Âge par différence de millésimes à la première maternité de la mère du répondant.	= *v15-v09* 97. NC si v09>9000 ou v15>9000
c02	Âge par différence de millésimes à la première maternité ou paternité du répondant.	=*enf2$01-v01* 97. NC si enf2$01>9000
c03	Année de naissance du dernier enfant de l'enquêté.	=*enf2$i avec i=v17* 9997. NC si v17=0 ou si v17>90
c04	Intervalle entre la première naissance et la dernière naissance de l'enquêté.	=*c03-enf2$01* 97. NC si v17<2 ou v17>90 ou enf2$01>9000 ou c03>9000
c05	Intervalle intergénésique moyen sur l'ensemble des naissances de l'enquêté. (Hypothèse : les jumeaux sont négligeables)	=*c04/(v17-1)* 97. NC si c04=97 or v17>90
c06	Intervalle entre la première naissance et la dernière naissance de la mère de l'enquêté.	=*v16-v15* 97. NC si v13>90 ou v13=1 ou v16>9000 ou v15>9000
c07	Intervalle intergénésique moyen sur l'ensemble des naissances de la mère de l'enquêté. (Hypothèse : les jumeaux sont négligeables)	=*c06/(v13-1)* 97. NC si c06=97
c08	État matrimonial légal	1. Célibataire (v03=1 ou 2) 2. Marié(e) (v03=3 ou 6 ou 7 ou 8) 3. Veuf(ve) (v03=4 ou 9) 4. Divorcé(e) (v03=5 ou 10) 97. Valeur manquante (v03>90)
c09	Situation de couple	1. Vit seul (v03=1, 8, 9 ou 10) 2. Vit maritalement (v03=2, 3, 4 ou 5) 3. Marié(e) (v03=6 ou 7) 97. Valeur manquante (v03>90)

Illustration :

id1	id2	id3	ident	v01	v02	v03	v04	v05	v06	v07	v08	v09	v10	v11	v12	v13	v14	v15	v16	v17	c01	c02	c03	c04	c05	c06	c07	c08	c09
2	1	76	21 076	1 950	1	4	5	1 922	1	997	5	1 922	1	997	1	4	3	1 942	1 950	0	20	97	9 997	97	97,00	8	2,67	3	2
5	6	13	56 013	1 942	1	5	5	1 921	0	52	2	1 925	0	65	7	2	1	1 942	1 942	3	17	18	1 970	10	5,00	0	0,00	4	2
6	5	70	65 070	1 939	2	9	5	1 918	0	79	3	1 912	0	78	3	3	0	1 939	1 939	4	27	18	1 961	4	1,33	0	0,00	3	1
1	5	87	15 087	1 950	1	1	1	1 925	96	997	96	1 925	1	997	96	2	0	1 950	1 950	2	25	19	1 976	7	7,00	0	0,00	1	1
6	3	95	63 095	1 946	2	10	5	1 907	0	79	1	1 909	0	93	2	2	96	1 946	1 946	1	37	20	1 966	97	97,00	0	0,00	4	1
6	5	46	65 046	1 944	2	9	5	1 925	0	996	6	1 927	0	996	1	2	1	1 944	1 944	3	17	21	1 969	4	2,00	0	0,00	3	1
6	2	39	62 039	1 935	1	9	6	1 910	0	80	5	1 913	0	88	3	2	1	1 935	1 935	2	22	22	1 960	3	3,00	0	0,00	3	1
2	2	38	22 038	1 948	2	7	1	9 996	1	997	3	1 922	1	997	2	2	0	1 948	1 948	2	26	22	1 975	5	5,00	0	0,00	2	3
2	2	46	22 046	1 947	1	6	1	1 919	1	997	2	1 922	1	997	3	2	0	1 947	1 947	2	25	23	1 970	0	0,00	0	0,00	2	3
1	3	47	13 047	1 946	1	5	6	1 922	1	997	2	1 920	1	997	10	2	0	1 946	1 946	2	26	24	1 994	24	24,00	0	0,00	4	2
4	6	67	46 067	1 941	1	6	5	1 912	1	997	9	1 915	1	997	3	2	1	1 941	1 941	2	26	25	1 968	2	2,00	0	0,00	2	3
6	6	43	66 043	1 941	1	9	5	1 916	0	74	9	1 903	0	84	3	2	0	1 941	1 941	2	38	25	1 968	2	2,00	0	0,00	3	1
4	2	2	42 002	1 937	2	6	5	1 909	1	997	9	1 911	0	82	1	2	0	1 937	1 937	1	26	27	1 964	97	97,00	0	0,00	2	3
5	6	73	56 073	1 949	1	2	4	1 925	1	997	2	1 930	0	22	4	2	1	1 949	1 949	3	19	29	1 983	5	2,50	0	0,00	1	2

Le fichier *enfants*.

Son mot de passe est le même que celui du ficher *individus*.
A chaque ligne, correspond un des enfants déclarés dans les variables enf1$01 à enf7$12.
Lorsque l'on s'intéresse à des données concernant les enfants des enquêtés (pour calculer l'âge moyen à la première naissance pour les enfants des enquêtés par exemple), nous avons au dénominateur un nombre d'enfants et non un nombre d'individus enquêtés. D'où la nécessité d'avoirs un recours à un fichier d'enfants.
Ce fichier comporte 6.957 enfants.

Les variables spécifiques à ce fichier sont les suivantes, *enf1* à *enf7* étant directement obtenues à partir des variables *enf1$01* à *enf7$12* ; et *cenf01* à *cenf06* ayant été calculées par nos soins.

Var.	Libellé	Modalités
rang	Rang de l'enfant	
identenf	Identifiant enfant	*=ident*100+rang*
enf1	Sexe de l'enfant i	1. Homme 2. Femme 96. NSP 98. AI 99. Refus
enf2	Année de naissance de l'enfant	*Sur quatre chiffres* 9996. NSP 9998. AI 9999. Refus
enf3	Enfant toujours en vie ?	0. Non 1. Oui 96. NSP 98. AI 99. Refus
enf4	Âge au décès	*Sur deux chiffres* 97. NC si enf3<>0 96. NSP 98. AI 99. Refus
enf5	A eu des enfants ?	0. Non 1. Oui 96. NSP 98. AI 99. Refus
enf6	Nombre d'enfants qu'il a eu	*Sur deux chiffres de 1 à 20.* 97. NC si enf5<>1 96. NSP 98. AI 99. Refus
enf7	Année de naissance de l'aîné	*Sur quatre chiffres.* 9997. NC si enf5<>1 9996. NSP 9998. AI 9999. Refus
cenf01	Année de naissance de l'enfant précédant	***Sur quatre chiffres*** 9997. NC si rang=1
cenf02	Année de naissance de l'enfant suivant	***Sur quatre chiffres*** 9997. NC si rang=v17
cenf03	Naissance multiple (supposée multiple si né la même année que l'enfant précédent ou l'enfant suivi d'où une surestimation de la gémellité)	0. Non 1. Oui (enf2=cenf01 OR enf2=cenf02) 97. NC ((cenf01>9000 AND cenf02>9000) OR enf2>9000)
cenf04	Intervalle intergénésique avec naissance précédente (NC si naissance multiple).	*=enf2-cenf01* 97. NC (enf2>9000 OR cenf01>9000 OR cenf01=enf2)
cenf05	Intervalle intergénésique avec naissance suivante (NC si naissance multiple).	*=cenf02-enf2* 97. NC (enf2>9000 OR cenf02>9000 OR cenf02=enf2)
cenf06	Âge à la première paternité ou première maternité.	*=enf7-enf2* 97. NC (enf2>9000 OR enf7>9000)

De plus, dans le fichier *enfants*, vous disposez de l'ensemble des variables concernant l'enquêté, de *id1* à *c09*.

ATTENTION : soyez très vigilant quant au fichier à utiliser selon ce que vous désirez calculer. Si vous vous intéressez à la structure par sexe et âge des enquêtés, utilisez le fichier *individus*. En effet, dans le fichier *enfants*, **les individus n'ayant pas eu d'enfants n'apparaissent pas**, mais ceux en ayant eu plusieurs apparaissent autant de fois qu'ils ont eu d'enfants. Le fichier *enfants* ne doit donc être utilisés QUE pour calculer des indicateurs portant sur les enfants.

UNIVERSITE FRANCOIS RABELAIS
DEUG 1 Sociologie – Enseignement de DEMOGRAPHIE
Laurent NOWIK / Aurélie GODARD / Joseph LARMARANGE

Fiche Technique

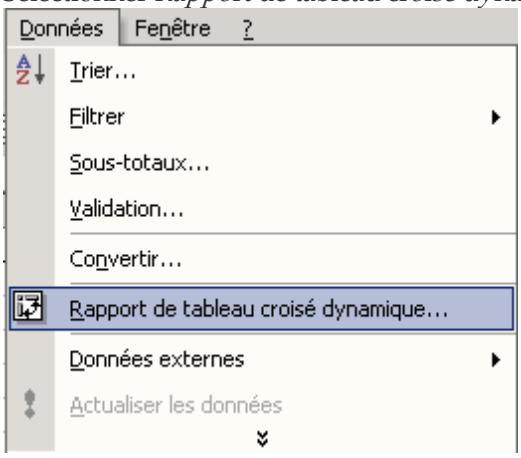
Réalisation d'un tableau croisé dynamique sous Excel

ATTENTION : Les exemple affichés ont été calculés sur une partie du fichier (1.500 questionnaires) et diffèrent donc des résultats que l'on obtiendrait avec l'ensemble de la base.

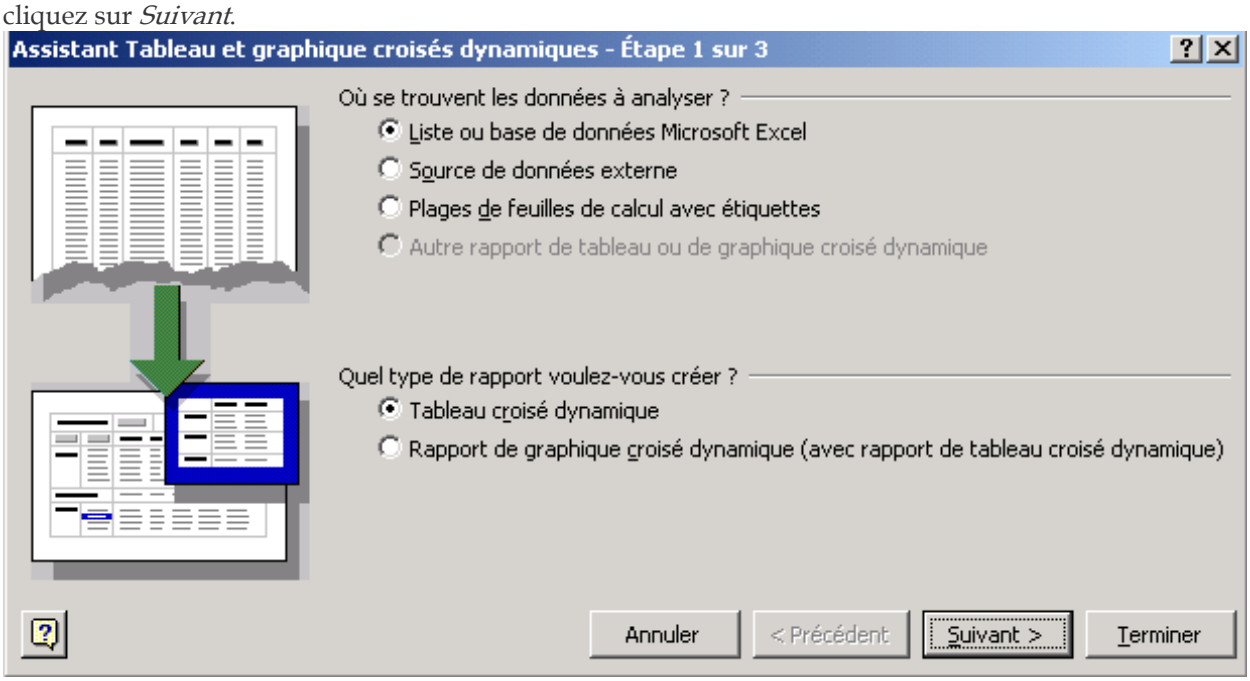
1. Création d'un tableau croisé d'effectifs à deux variables

Nous allons croiser le sexe (v02) avec l'activité actuelle (v07).

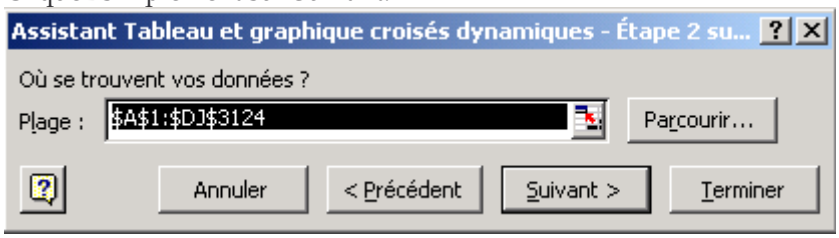
Étape 1 :
Sélectionner *Rapport de tableau croisé dynamique* dans le menu *Données*.



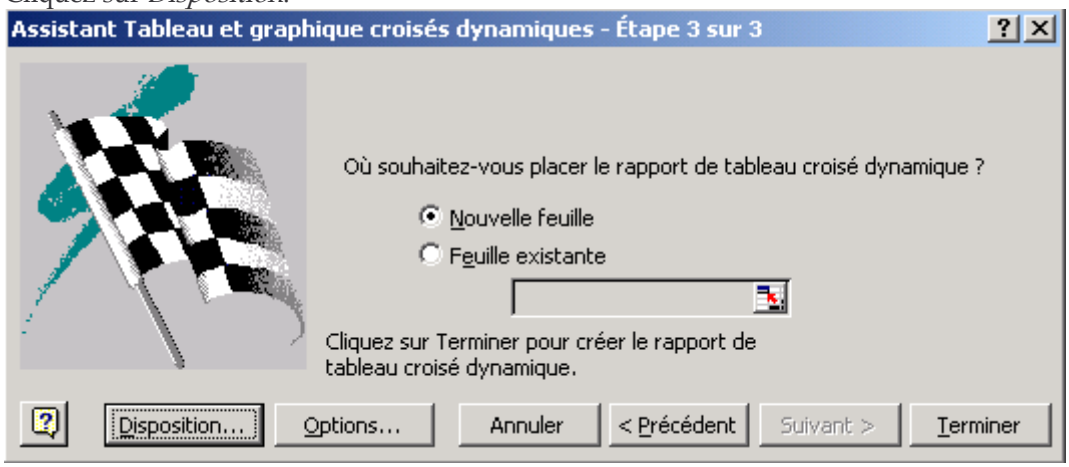
Étape 2 :
Vérifiez que *Liste ou base de données Microsoft Excel* et *Tableau croisé dynamique* sont sélectionnés et cliquez sur *Suivant*.



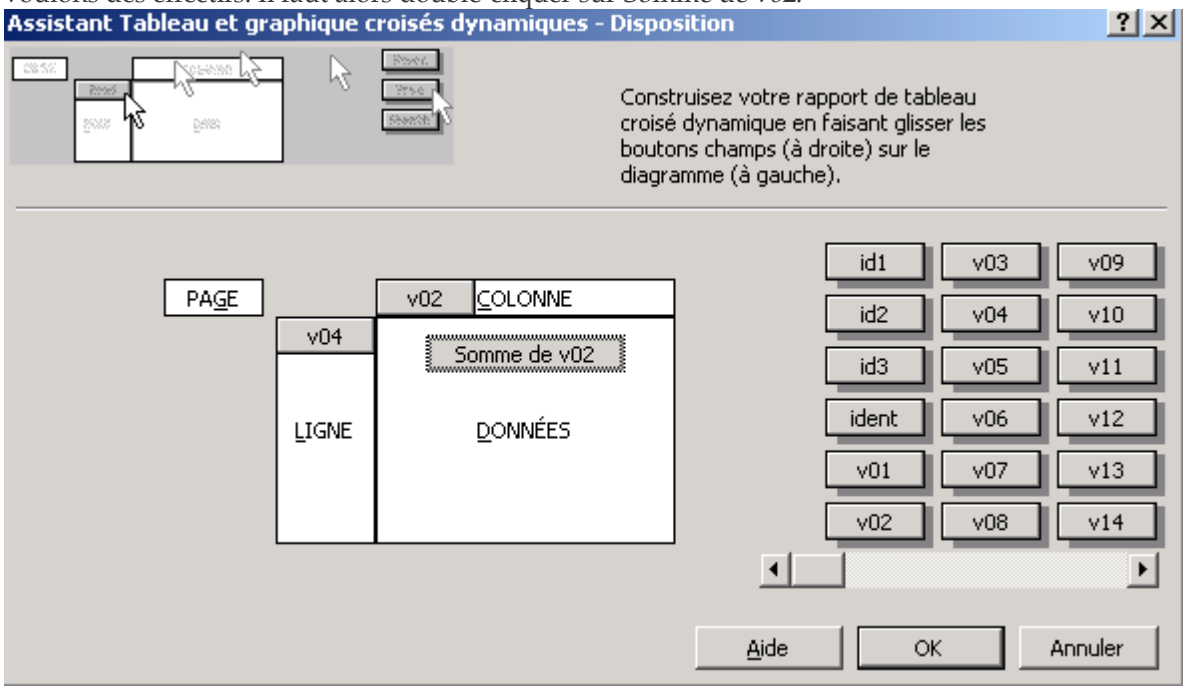
Étape 3 :
L'assistant détecte automatiquement les cellules contenant les données (à condition d'être dans la feuille contenant les données lorsque vous sélectionnez *de tableau croisé dynamique* dans le menu *Données)*. Cliquez simplement sur *Suivant*.



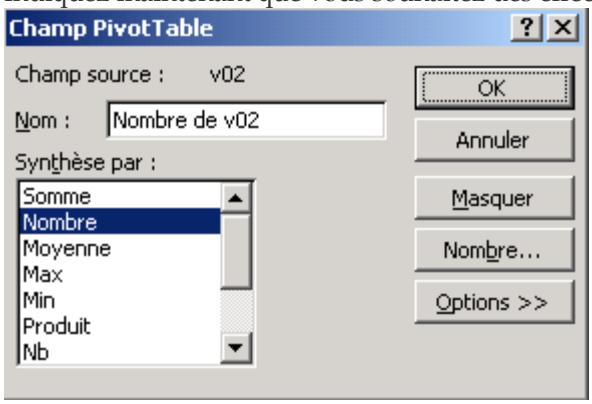
Étape4 :
Cliquez sur *Disposition*.



Étape 5 :
Par glisser-déplacer avec la souris, indiquez les variables qui seront utilisées pour la construction du tableau. Dans notre cas, nous allons mettre le sexe (v02) en colonnes et l'activité actuelle (v04) en lignes. Il faut également indiquez une variable qui va être utilisées pour le comptage des observations. Ici n'importe laquelle fera l'affaire, par exemple v02 que l'on va de nouveau glisser et déposer avec la souris mais cette fois-ci dans la case donnée. Par défaut, Excel choisi de représenter la somme des v02 en données. Or nous voulons des effectifs. Il faut alors double-cliquer sur *Somme de v02*.

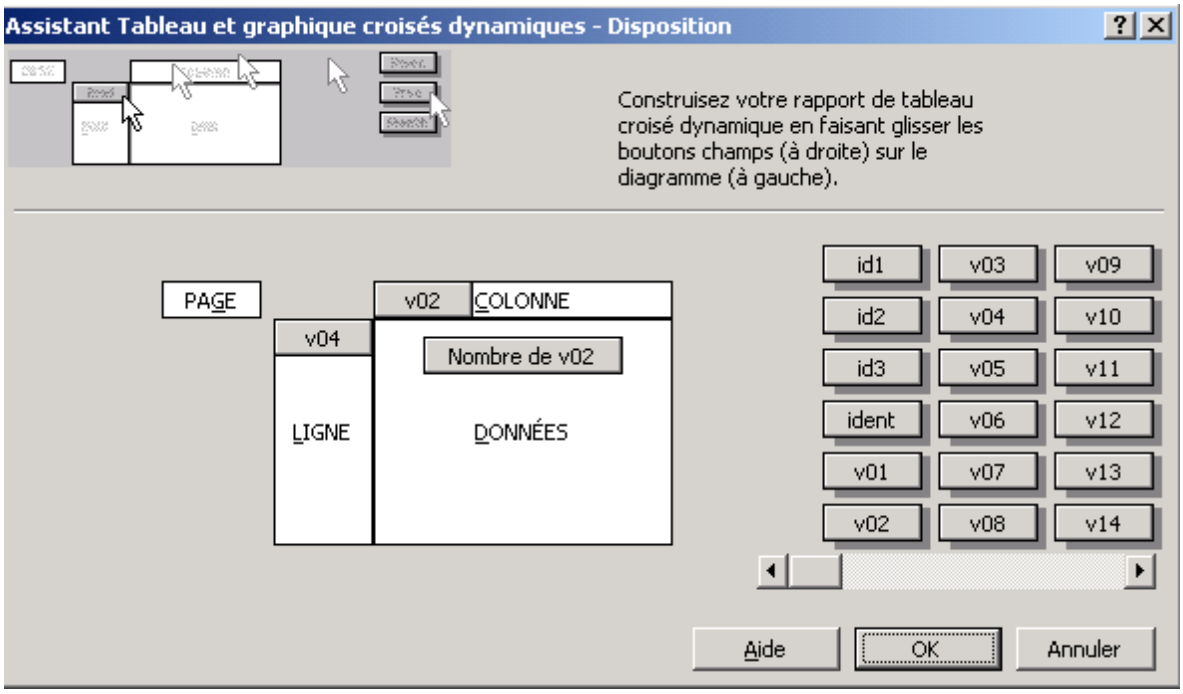


Étape 6 :
Indiquez maintenant que vous souhaitez des effectifs en choisissant *Nombre*. Puis cliquez sur *OK*

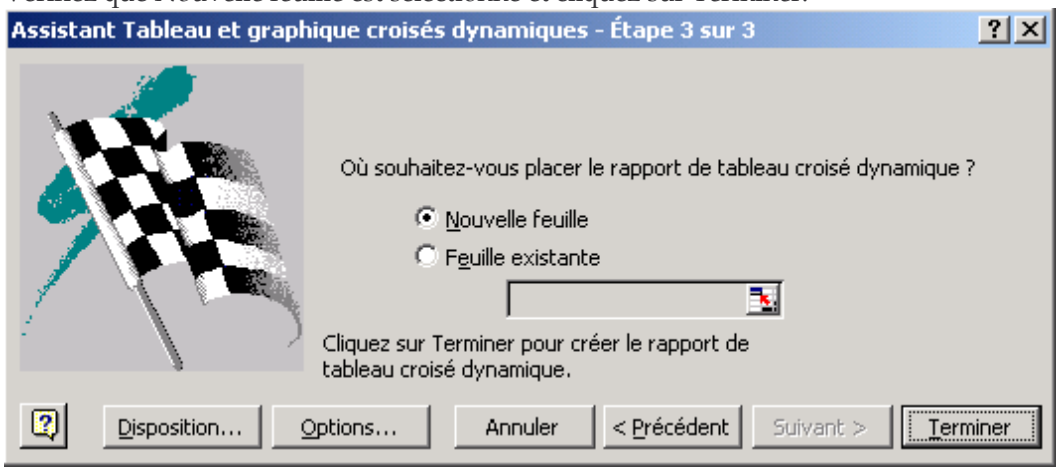


Étape 7 :
Votre tableau est maintenant correctement défini. Cliquez sur *OK*.

NB : Vous voyez que d'autres indicateurs sont aussi calculables. Des moyennes par exemple. Si vous demandez des moyennes, il vous calculera des moyennes pour chaque cellule du tableau. Cependant, cette fonction est à éviter. En effet, s'il y des ne sait pas, refus, des non concernés ou des absences d'information, ils apparaissent dans le fichier en 96, 97, 98 et 99 et sont intégrés dans le calcul de la moyenne. Il est donc préférable de sortir un tableau détaillé d'effectifs et de calculés ensuite manuellement des moyennes.
D'autre part, Excel ne vous calculera pas des pourcentages totaux, en ligne ou en colonne. Encore une fois, il est préférable de les calculer manuellement, notamment si vous décidez de ne pas tenir compte des Refus NSP et AI (on rapporte alors les effectifs observés à l'effectif valide, c'est-à-dire le nombre total d'observations moins le nombre de 96, 98 ou 99). Dans le cas de questions qui ne concernent pas tout le monde (âge au décès par exemple), il faut retirer les NC ou 97 de vos calculs. Par exemple, on ne calcule un âge moyen au décès que pour les personnes pour lesquelles on dispose d'un âge au décès (personnes ayant décédé dont on connaît l'âge ; On doit donc « retirer » pour le calcul les 996, 997, 998 et 999.
Il est possible de croiser également trois variables entre elles en rajoutant la troisième variable dans la case *page* à l'étape 5.
Le tableau que vous obtiendrez ici n'est pas forcément le tableau que vous rendrez dans votre travail final. Pour des questions de présentation, vous pouvez être amenés à procéder à des regroupements (par exemple dans une étude sur des générations regroupez les données pour des groupes de 5 années).



Étape 8 :
Vérifiez que *Nouvelle feuille* est sélectionné et cliquez sur *Terminer*.



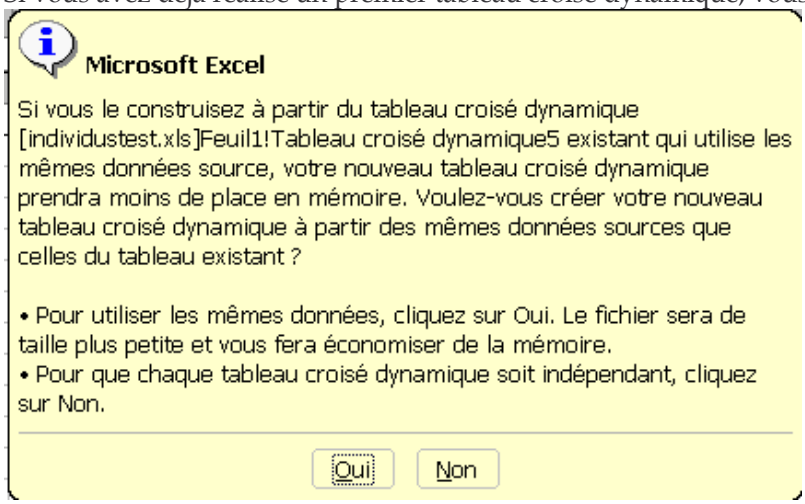
Étape 9 :
Votre tableau croisé apparaît dans une nouvelle feuille Excel.

	A	B	C	D
1				
2				
3	Nombre de v02	v02		
4	v04	1	2	Total
5	1	257	255	512
6	2	32	116	148
7	3	29	42	71
8	4	27	37	64
9	5	239	335	574
10	6	17	21	38
11	7	11	71	82
12	98	2	1	3
13	99	2	6	8
14	Total	616	884	1500

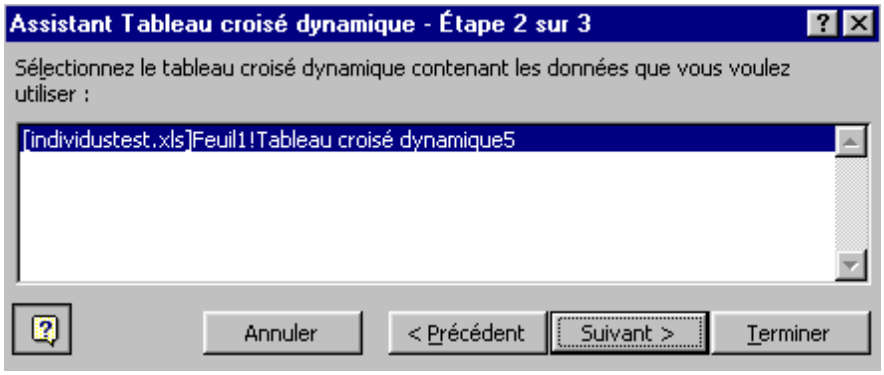
2. Age au décès du père selon le sexe.

Nous allons calculer les âges au décès du père du répondant (v07) et l'âge moyen au décès, selon le sexe (v02) des répondants. Dans le cas présent, cela ne concerne que les personnes dont le père est décédé (v06=0).

Répétez les **étapes 1 à 3**. Soyez bien vigilant à l'étape 1 à vous situez sur la feuille qui contient vos données pour que Excel détecte automatiquement la bonne plage de cellules.
Si vous avez déjà réalisé un premier tableau croisé dynamique, vous verrez apparaître le message suivant :

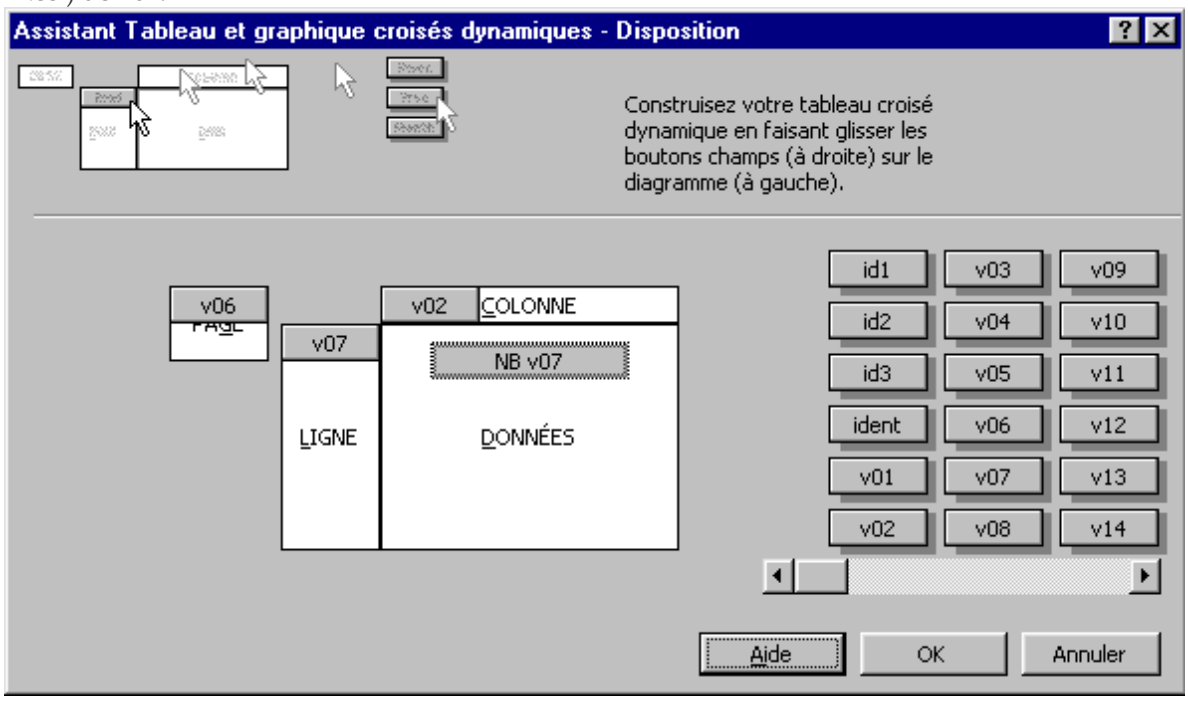


Cliquez sur *Oui* puis sur *Suivant* dans la boite de dialogue ci-dessous.



Étape 5 :

Nous allons mettre le sexe (v02) en colonnes, l'âge au décès (v07) en ligne. Il nous faut aussi tenir que nous ne voulons sélectionner que les personnes dont le père est décédé (v06=0). La variable v06 est donc mise dans la section *Page*. Pour avoir les effectifs, nous avons sélectionner *Nombre* (ou *NB* selon les versions de Excel) de v02.



On obtient alors le tableau suivant :

Microsoft Excel - individustest.xls

Fichier Edition Affichage Insertion Form

A3 =

	A	B	C	D
1	v06	(Tous)		
2				
3	NB v07	v02		
4	v07	1	2	Total
5	20		1	1
6	22	1		1
7	25		2	2
8	26		2	2
9	27		1	1
10	28	1	2	3
11	29	1	3	4
12	30	2	4	6
13	31		2	2
14	32	3	2	5
15	33	4	7	11
16	34	3	2	5
17	35	5	1	6
18	36	2	2	4
19	37	2	3	5
20	38	2	3	5
21	39	1	4	5
22	40	1	1	2
23	41	3		3
24	42	1	1	2
25	43		4	4
26	44	2	2	4
27	45	2	4	6
28	46	2		2
29	47	3	2	5
30	48	2	2	4
31	49	5	3	8
32	50	6	5	11
33	51	1	5	6
34	52	6	6	12
35	53	5	3	8

Feuil1 Feuil2 individus

Prêt

Démarrer

Microsoft Excel - individustest.xls

Fichier Edition Affichage Insertion Form

A3 =

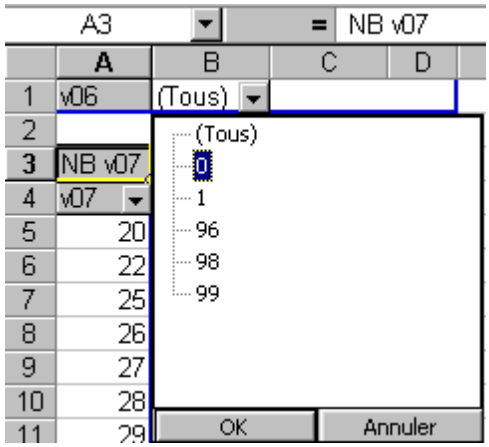
	A	B	C	D
52	70	15	29	44
53	71	13	18	31
54	72	16	24	40
55	73	14	13	27
56	74	12	15	27
57	75	16	20	36
58	76	8	15	23
59	77	8	12	20
60	78	14	15	29
61	79	8	21	29
62	80	14	26	40
63	81	8	16	24
64	82	7	18	25
65	83	10	15	25
66	84	6	15	21
67	85	5	9	14
68	86	7	11	18
69	87	7	9	16
70	88	4	4	8
71	89	3	11	14
72	90	4	12	16
73	91		3	3
74	92	5	5	10
75	93	3	2	5
76	94	1	3	4
77	95		1	1
78	96	1	1	2
79	97	1		1
80	98		1	1
81	996	6	30	36
82	997	221	289	510
83	998	3	2	5
84	999	3	4	7
85	Total	616	884	1500
86				

Feuil1 Feuil2 individus

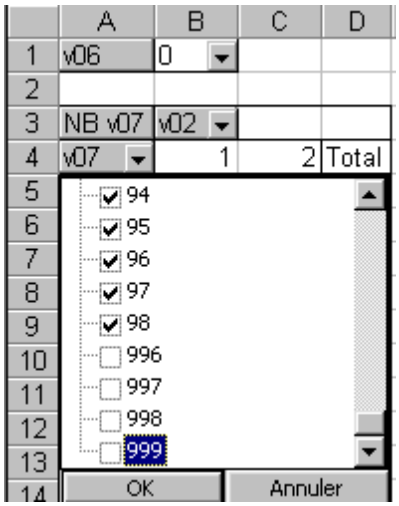
Prêt

Démarrer

Pour restreindre le tableau aux cas où v06=0, en haut à gauche, cliquez sur la petite flèche vers le bas à côté de v06 et sélectionnez *0*.

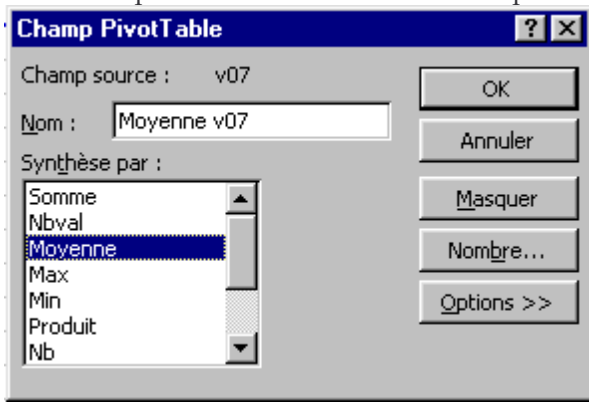


Pour ensuite retirer de l'analyse les valeurs 996, 997, 998 et 999 de la variable v07, cliquez sur la petite flèche dirigée vers le bas à coté de la cellule *v07* en haut à gauche du tableau. Désélectionnez les valeurs *996, 997, 998* et *999* puis cliquez sur *Ok*.



Le tableau ne porte plus alors que sur les personnes dont le père est décédé (v06=0) et dont on connaît l'âge au décès (v07 différent de 996, 997, 998 et 999). Vous avez alors les effectifs désirés.

Il est possible de demander à Excel de calculer l'âge moyen, maintenant que le tableau ne comporte plus que les personnes pour lequel ce calcul est justifié.
Double-cliquez sur *NB v07*. Dans le menu qui s'ouvre, choisissez *Moyenne v07*.



Vous avez alors l'âge moyen dans la ligne *Total* tout en bas du tableau. L'âge moyen au décès du père du répondant, dans notre exemple, est de 68,9 ans.

74	92	92	92	92
75	93	93	93	93
76	94	94	94	94
77	95		95	95
78	96	96	96	96
79	97	97		97
80	98		98	98
81	Total	67,76501305	69,6744186	68,89808917

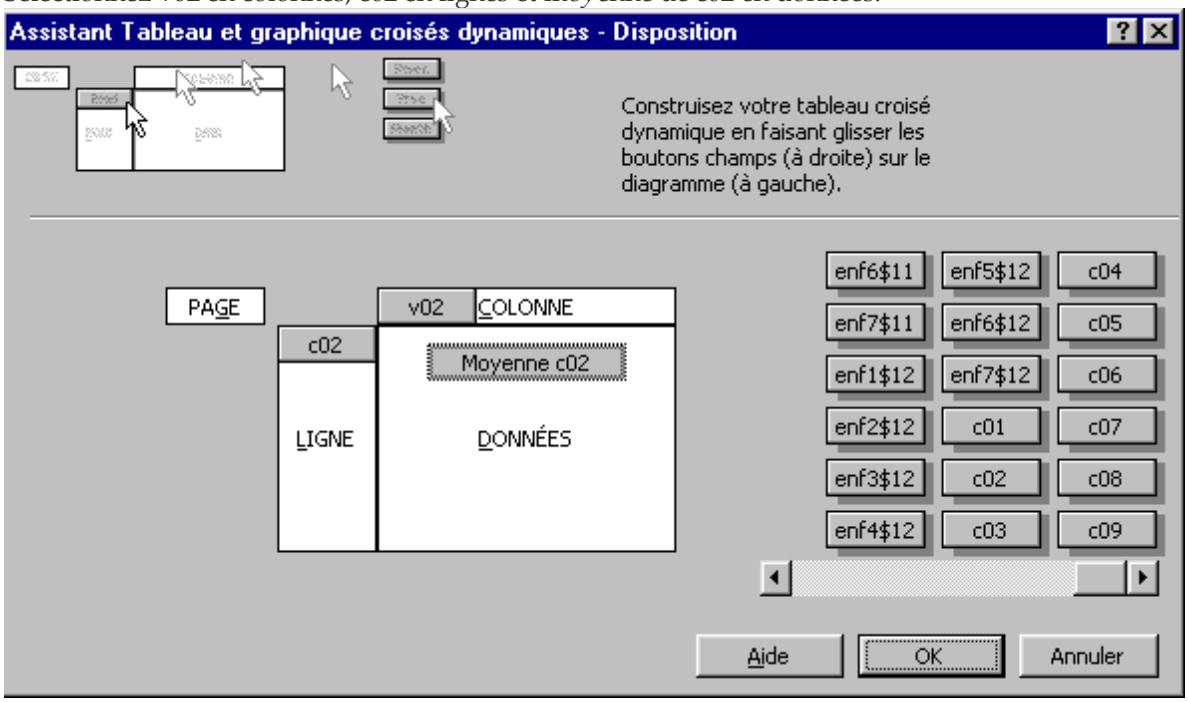
NB : La distinction selon le sexe du répondant (v02) n'est pas pertinente ici (puisqu'on étudie la mortalité du père du répondant). Nous l'avons fait afin de donner un exemple pour sa réalisation sous Excel.

3. Âge moyen à la première paternité ou maternité du répondant

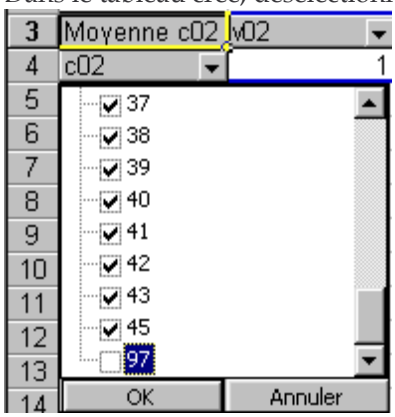
Il s'agit d'étudier l'âge à la 1ère paternité ou 1ère maternité (c02) du répondant selon son sexe (v02).

Étape 5 :

Sélectionnez v02 en colonnes, c02 en lignes et *moyenne de c02* en données.



Dans le tableau créé, désélectionnez la valeur 97 pour c02.



Dans la ligne *Total*, vous obtenez ainsi l'âge moyen à la première paternité (25,9 ans), à la première maternité (24,5 ans) et à la première naissance (25,1 ans – hommes et femmes ensembles)

31	37	37	37	37
32	38	38	38	38
33	39	39		39
34	40		40	40
35	41	41	41	41
36	42	42		42
37	43	43		43
38	45	45		45
39	Total	25,88191882	24,54891995	25,09255079