

INSTITUT DU SAHEL

EP. 1 530

B A M A K O

EXPLOITATION INFORMATIQUE DES DONNEES DEMOGRAPHIQUES ;  
UTILISATION DU PERSONNAL COMPUTER.-

Baba TRAORE -  
Conseiller en Collecte et  
Exploitation des Données Démographiques -  
INSAH/BAMAKO.-

EXPLOITATION INFORMATIQUE DES DONNEES DEMOGRAPHIQUES

- 1/ Aspect informatique de certains projets du PRDS/INSAH
- 2/ Exploitation de l'enquête pilote migration/BKO/ZEG-OUA
- 3/ Analyse des données de Mortalité - Recensement du Mali 1976.

## 1/ ASPECT INFORMATIQUE DE CERTAINS PROJETS DU PRDS

Au sein de l'Unité Socio-Economique et de Démographie (USED) de l'INSAH, le programme de Recherches Démographiques pour le SAHEL (PRDS) comporte les projets-suivants parmi d'autres qui sont en cours d'exécution ou presque terminés dans les différents pays sahéliens. Ce sont :

- l'exploitation et l'analyse des données démographiques (EXADD)
  - . au Sénégal
  - . au Mali
  - . au Niger
  - . en Haute-Volta
  - . en Mauritanie
  - . en Gambie
  - . au Cap-Vert
  
- l'enquête sur la Mortalité Infantile dans le SAHEL (EMIS)
  - . dans 4 villes secondaires de la Haute-Volta
  - . en milieu rural au Sénégal
  - . dans la ville de Bamako
  
- un projet de recherche méthodologie sur la collecte des données en matière de migration
  - . enquête pilote BKO
  - . enquête pilote Zégoua.

L'USED étant dotée de micro-ordinateurs "IBM-personnal computer" intervient directement dans l'exploitation informatique de certaines opérations précédemment citées, en l'occurrence EXADD/MALI (Mortalité) ; EMIS/BOBODIOLASSO/BAMAKO et les enquêtes migration de Bamako et Zégoua.

A la création du petit centre informatique de l'Unité, des démographes ont reçu une brève formation d'1 mois à la programmation en basic.

Avec cette formation de base et à l'aide de documents disponibles sur le langage basic nous programmons relativement bien maintenant nos travaux de saisie, de tabulation et d'analyse.

Concernant l'EMIS/BOBOÏDI OULASSO un important effort a été accompli à l'USED pour la saisie, la vérification, l'apurement de fichiers. Les notes techniques écrites par Etienne Van de Walle sont très instructives à ce sujet.

Nous comprenons que pour une bonne observation suivie, les micro-ordinateurs constituent des outils très efficaces. Aussi, savoir programmer pour un démographe est très utile à plus d'un titre car ça raccourcit énormément le délai de l'exploitation et de plus on sait mieux faire ce que l'on veut obtenir soi-même, qu'en l'exprimant à quelqu'un d'autre pour exécution (dialogue entre démographe et informaticien par exemple).

Parmi les travaux informatiques que nous effectuons actuellement à l'Institut, nous parlerons uniquement de l'exploitation des enquêtes migration Bamako-Zégoûa et de l'analyse des données de Mortalité du Mali 1976.

## 2/ ENQUÊTES PILOTES MIGRATION - BAMAKO - ZÉGOÛA

L'exploitation de ces enquêtes a été exposée à la 5e réunion des démographes sahéliens à Niamey en Juin 1982.

C'est dans le cadre de la Recherche d'une méthodologie pour la collecte des données de migration dans les pays du SAHEL que l'Institut du SAHEL a été invité par les différents états de mener ces enquêtes.

Bien que la méthodologie adoptée pour le test <sup>ait</sup> / été longuement critiquée à Niamey aussi bien que les résultats auxquels elle a abouti, l'aspect informatique était jugé satisfaisant.

Deux programmes ont servi à exploiter les questionnaires ménages des enquêtes : le programme de saisie et de correction des données (nigr) ; le programme de lecture et de tabulation (lire).

2-1) Le programme "nigr" avait deux buts : saisir les informations collectées sur disquettes ; corriger les erreurs dont on se rend compte après la saisie. Toutes les erreurs ne pouvant être décelées par simple affichage des données saisies à l'aide de "nigr", certaines étaient détectées par "EDLIN", ou par "system" qui sont des langages

machine du DOS (Disk operating system).

Le programme "migr" ainsi élaboré n'a pas nécessité l'établissement au préalable d'une grille de chiffrement. Les réponses aux questions ont été pour la plupart directement saisies comme l'opérateur lisait sur le questionnaire. Il est écrit de telle sorte que les informations saisies sont automatiquement codées. Nous avons aussi énormément réduit les possibilités d'erreurs par le refus par la machine de toute réponse qui n'est pas une modalité de la question à saisir ou s'il existe une incompatibilité avec la réponse d'une autre question.

Les questions comme le type d'enseignement et la branche d'activité ont été codées au préalable. Pour ces questions c'est donc le code qui a été saisi. Pour l'âge et les durées, l'agent de saisie met le chiffre "9" devant toute durée en années et le chiffre "1" devant toute durée exprimée en mois.

Après le code "9" ou "1" un nombre à un chiffre est toujours précédé d'un "0".

Exemple :

1 an est codé	901
1 mois est codé	101
10 ans sont codés	910
10 mois " "	110

Pour les lieux on précède toujours le nom de la localité par le numéro d'ordre de la région administrative ou le code affecté à un pays étranger.

Après avoir saisi entièrement les caractéristiques d'une personne du ménage, "migr" fait afficher les données codées. A cette phase l'opérateur vérifie l'exactitude de chaque code et corrige les informations erronées. Il faut signaler cependant que par inattention, des erreurs ont échappé à la vision de l'agent de saisie. Ces erreurs si elles n'ont pas été détectées par "EDLIN" ou "SYSTEM" figureront dans le fichier définitif. S'il n'y avait pas une contrainte de temps à l'époque pour la réunion de Niamey, il serait peut être mieux de procéder à une double saisie pour confronter les deux fichiers comme on le fait actuellement pour l'enquête de Bobo-Dioulasso et de Bamako (EHS).

2.2.) Le programme "Lire" effectue les opérations suivantes :

- lecture des données saisies sur disquettes
- recodification de certaines caractéristiques (regroupement en classe d'âge d'amplitudes quinquenales par exemple)
- cumul des variables pour la tabulation
- sortie des résultats sous forme de tableaux.

Signalons qu'au niveau de la lecture des informations saisies, certaines erreurs encoere étaient détectées comme par exemple des caractères que la machine n'arrivait pas à interpréter ou des informations accolées au moment de la saisie.

Ces dernières erreurs étaient d'abord corrigées avant de passer à la recodification dont l'ampleur dépendait des caractéristiques que l'on voulait croiser.

Après la lecture du fichier et les dernières corrections tout autre blocage était essentiellement une erreur de programmation. Contrairement au programme "migr" qui a été largement inspiré du programme de saisie de l'enquête mortalité infantile de Bobo-Dioulasso, le programme "Lire" a beaucoup sollicité notre propre inspiration et un effort plus soutenu de compréhension des manuels "Basic" et "Dos" qui sont édités en anglais.

Après un travail de deux mois et demi (fin février à mai), les résultats de l'enquête test ont paru et un extrait fait objet de la présentation en annexe du présent document.

Au vu des résultats de ce travail nous avons compris beaucoup d'astuces dans la programmation et nous déduisons que la meilleure façon de savoir programmer est de traiter un cas concret comme celui-là. Signalons toutefois que c'est seulement un aspect de l'informatique que nous venons de traiter à savoir la saisie des données d'observation et la tabulation des résultats. Un aspect non moins important qu'est l'analyse est aussi très intéressant à traiter.

### 3/ ANALYSE DES DONNEES DE MORTALITE/RECENSEMENT DU MALI 1976 (EXADD/MALI)

Le traitement de ce cas sur micro-ordinateur a été un apport technique de l'Institut du SAHEL pour l'EXADD/MALI. Quatre programmes ont été écrits parmi lesquels :

- le programme dénommé "Analyse", saisit la structure par âge et par sexe des effectifs recensés et des décès observés au cours des douze derniers mois qui ont précédé le recensement aussi bien que les taux d'accroissement évalués au préalable
- le programme "NAIS" saisit les naissances survenues au cours des douze derniers mois
- le programme "Calcul" lit les différents fichiers de données, répartit les non-déclarés en âge des effectifs recensés et des décès de façon proportionnelle ; calcule les proportions en % des effectifs survivants et des décès par âge ; calcule l'effectif total par sexe de la population moyenne de l'année ; répartit cet effectif par groupes d'âge suivant la répartition par âge des effectifs recensés ; calcule les taux de mortalité par âge et les taux de mortalité infantile par sexe et par région ; <sup>crée</sup> le fichier des taux de mortalité sur disquette ; imprime enfin les résultats des calculs sur papier.
- le programme "table" lit le fichier des taux de mortalité par âge, calcule avec la formule de passage les quotients de mortalité et évalue ensuite les autres fonctions de la table. Tous ces calculs se font par sexe par région administrative du Mali et selon le type d'habitat au niveau national. (8 régions administratives, ensemble du pays, urbain (niveau national) ; rural (niveau national)).

Comme le programme "calcul", le programme "table" aussi imprime à la fin des calculs les résultats sur papier.

Pour ce qui est des deux programmes de saisie, tout comme le programme "nigr", les données saisies sont d'abord affichées et corrigées avant d'être imprimées sur disquettes. Le fichier des <sup>données</sup> ~~statistiques~~ tures par âge et sexe des effectifs recensés et des décès s'appelle "Données", celui des naissances est nommé "NAS" et celui des taux par âge est "taux".

Signalons que pour ce cas précis de l'analyse, la saisie et l'épure des fichiers n'étaient pas aussi compliquées que pour une enquête. Nous avons nous-mêmes à deux saisi ces données à moins de 3 jours ; c'est l'élaboration des programmes qui a mis du temps.

Nous donnons en annexe un extrait des résultats auxquels nous avons abouti, la liste de certains de nos programmes et les références de nos machines.



ANNEXE

CARACTERISTIQUES DES MICRO-ORDINATEURS IBM 51-50 PERSONNAL COMPUTER

Primary system : IBM 51-50 processor with 64 k bytes of storage  
251/4 "floppy disk drives

keyboard

Monochrome display (monitor)

Disk adapter card

Monitor/printer adapter card

R 5232 communications card

Epson M x 80 F/T printer

Secondary system IBM 5150 processor with 64 k bytes of storage 151/4  
"floppy disk drive

keyboard

Monochrome display (monitor)

Disk adapter card

Monitor/printer adapter card

Software basic computer language

DOS and disk basic

advanced diagnostics

Supplies and Misc.; cables, extension cords

Ground rod, strap, wire

paper, printer ribbons

spare memory chips, fuses

Tools.