



07827



NATIONS UNIES
CONSEIL ÉCONOMIQUE ET SOCIAL

Distr.
LIMITÉE
E/CN.14/SM/39
8 mai 1981

français
Original : ANGLAIS

ECONOMIQUE COMMISSION POUR L'AFRIQUE

Groupe de travail sur l'organisation
le contenu et la méthodologie des enquêtes
sur les ménages

Addis Abéba, 29 juin 3 juillet 1981

ANALYSE DES DONNEES D'UNE ENQUETE

TABLE DE MATIERES

	<u>Paragrap</u> hes	<u>Pages</u>
I. INTRODUCTION	1 - 5	1
II. BESOINS D'ANALYSE ET RESPONSABILITES	6 - 17	2 - 3
III. ACTIVITES PRECEDANT L'ANALYSE	18 - 24	4 - 5
IV. OPTIONS D'ANALYSES	25 - 33	5 - 6
V. APPROCHE PAR LES DONNEES DE BASE	34 - 41	6 - 8
VI. INFORMATIONS SUPPLEMENTAIRES SUR L'ANALYSE	42 - 49	8 - 9
VII. CONCLUSION	50 - 54	10

TRADUCTION NON-OFFICIELLE

I. INTRODUCTION

1. Un document abrégé sur l'analyse et l'évaluation des données d'une enquête (E/CN.14/SM/27) a été produit pour la deuxième réunion du groupe de travail sur l'organisation, le contenu et la méthodologie des enquêtes sur les ménages, en octobre 1979. Son but était d'attirer l'attention sur certaines conditions de base et les différentes manières de les satisfaire. La plupart des principales considérations seront reprises dans le présent document mais avec des procédés additionnels, découlant des réflexions et discussions ultérieures.
2. Toutefois, nous devons tout d'abord souligner que le présent document ne présentera pas l'ensemble des techniques utilisées dans l'analyse des données d'une enquête. Son but est simplement limité à l'identification des principales options d'analyse suivant les applications probables et, autant que possible, dans un contexte pratique de travail. Ce faisant, il sera nécessaire de tenir compte d'autres aspects relevant des opérations d'enquête et peut-être de considérer plus globalement les activités statistiques et de planification parce que l'analyse elle-même ne peut être considérée isolément.
3. Au cours des premières discussions, il était conclu que le sujet n'était pas assez connu pour produire un document satisfaisant en ce moment et qu'il vaudrait mieux s'en tenir aux étapes initiales de l'élaboration des données de base d'enquête et à la préparation des tableaux nécessaires à l'analyse.
4. Deux raisons principales sont à la base de la nature limitée des remarques dans le présent document. D'abord, la collecte des données dans les ménages et dans d'autres unités statistiques est de plus en plus considérée comme une opération intégrée et continue; l'analyse classique n'est plus pleinement satisfaisante dans ces conditions, et de nouvelles méthodes doivent être développées. Ensuite, l'analyse des données a été, en générale, entièrement négligée dans la plupart des pays en voie de développement; en conséquence, il existe très peu d'expériences pratiques sur lesquelles nous pouvons baser nos conclusions.
5. Néanmoins, d'importants progrès sont entrain d'être enregistré dans toute la région en ce qui concerne la collecte des données, et il s'avère nécessaire de développer sans tarder un schéma d'analyse adéquat de telle sorte que les nouvelles données puissent être pleinement utilisées pour les besoins de la planification et pour d'autres applications. Le présent document essaiera de dépasser un peu le cadre des données de base et de la tabulation, mais il n'est pas certain que des conclusions entièrement valables soient obtenues.

II. BESOINS D'ANALYSE ET RESPONSABILITES

6. Après avoir participé, en 1970, au recensement général de la population, bon nombre de pays africains ont porté leur attention sur les enquêtes auprès des ménages et les informations incomplètes dont dispose la CEA montre que plus de la moitié des pays de la région mène présentement d'importantes activités sur le terrain. Le Programme Africain concernant la mise en place de Dispositifs d'Enquête auprès des Ménages (PADEM) est une réponse à cet intérêt renouvelé pour les enquêtes et fournit de l'assistance aux pays individuellement dans l'établissement ou l'amélioration de dispositifs permanents d'enquête, en vue de la production continue des données économiques et sociales.

7. Le besoin d'information sur l'emploi, toute autre activité de production, le revenu, la consommation et les dépenses, les caractéristiques démographiques et sociales des ménages, a déjà été clairement démontré. Nous avons maintenant atteint le stade où il est nécessaire de considérer plus attentivement leurs utilisations afin de déterminer de quelle façon elles peuvent être analysées.

8. La plupart des documents antérieurs sur le PADEM, et sa contrepartie au niveau national (PNDEM), ont fait remarquer que les ménages sont les fournisseurs de tout travail et qu'ils sont les bénéficiaires du produit des efforts de développement et des problèmes économiques. Il est à noter aussi qu'au cours des dernières années, il y a eu un intérêt croissant pour des questions telles que la main-d'oeuvre disponible et la pauvreté.

9. Il s'en suit que les utilisations des données sur les ménages relèvent pour une large mesure des activités économiques de la population et de leurs niveaux de vie. Une petite étude montrera que les applications de données dans ces deux domaines constituent une base essentielle pour plusieurs aspects de planification, d'administration et d'activité de l'entreprise.

10. On pourrait aussi mentionner que depuis que les programmes nationaux d'enquête fourniront beaucoup de données, démographiques, sociales et économiques intégrées, il existe un grand nombre de thèmes ou de matières qui peuvent être considérés comme sujets d'analyse. Par exemple il est possible d'obtenir, à partir des domaines d'études déjà identifiés dans certaines des propositions de projets déjà formulées, assez de données pour des sujets de recherches tels que les besoins de base, le développement rural, la pauvreté, les indicateurs économiques et socio démographiques du développement, et la réforme agraire. Le genre d'analyse entreprise à partir du programme d'enquête, sera ainsi fonction des objectifs spécifiques de la recherche.

11. Dans ces conditions, il est clair que les besoins d'analyse concernant les données sur les ménages sont très larges. Pour la convenance, nous pouvons distinguer deux types d'analyse. Primo, celles qui s'intéressent à la situation générale de la structure et aux tendances économiques et sociales, portant principalement sur les données d'ensemble. secondo, l'analyse, plus poussée de certains aspects particuliers de la situation économique et sociale, basée dans une large mesure sur les inter-relations entre les variables; Elle couvre une gamme complexes de facteurs qui déterminent la qualité de la vie et l'inter-action de ces facteurs avec l'activité économique.

12. Pour les données d'ensemble, des modèles d'analyse existent déjà. Pour ce qui a trait à l'économie, il existe le Système des Comptes Nationaux des Nations Unies, le Système de comptabilité de Produit Matériel (CPM) et beaucoup d'autres possibilités, ensuite des configurations plus détaillées ou plus spécialisées telles que la matrice de comptes sociaux et plus récemment, la matrice des comptes alimentaires. Dans le cas des données sociales il existe quelques propositions dans le cadre du Système des Statistiques Sociales et Démographiques, mais on est souvent porté vers une approche plus simple en l'absence d'une unité commune en laquelle la plupart des données pourraient être exprimées.

13. Pour des analyses approfondies basées sur l'inter-relation des données, la situation est beaucoup moins satisfaisante et des modèles sont disponibles seulement dans peu des cas. L'Enquête Mondiale de la Fécondité a produit beaucoup de bons documents de travail dans son propre domaine. Il existe aussi des techniques semblables pour l'analyse d'autres variables démographiques et sociales, et le Bureau International du Travail a produit des modèles d'analyse du revenu, de la consommation, des dépenses et des données sur la main-d'oeuvre. Toutefois, la situation générale apparaît comme ceci: lorsque des recommandations sur les données de base requises et les tabulations sont disponibles pour la plupart des domaines d'études, les manuels d'analyse existent toujours en nombre limité alors qu'ils sont très nombreux dans le domaine de la population. Le travail entrepris par l'Enquête Mondiale de la Fécondité, le BIT et d'autres agences, doit être largement étendu à d'autres domaines nouveaux.

14. Cependant la principale faille dans les manuels techniques disponibles se situe au niveau de l'inter-relation des données entre différents domaines. Ceci constitue quelque chose de nouveau dans l'effort d'élaboration de programmes d'enquête continus et intégrés, qui appelle une nouvelle méthodologie d'analyse. Les besoins, dans ce cadre, seront pris en considération dans toutes les remarques faites dans le présent document.

15. Une liste provisoire des données à collecter à des fins d'analyse, ci-dessus indiquée, a déjà été publiée dans le document sur les données requises pour l'enquête sur les ménages, E/CN.14/SM/22, qui constitue l'un des documents de référence de la présente réunion.

16. La deuxième question à considérer dans cette partie du document est de savoir qui sera le responsable de l'analyse des données d'enquête. A sa première session en 1980, la Conférence commune des planificateurs, statisticiens et démographes africains a examiné les moyens de renforcer les relations entre statisticiens et planificateurs, et a abouti à la conclusion que le fait que l'analyse des données soit négligée dans plusieurs pays provient, dans une certaine mesure, d'un manque d'union. La Conférence a été d'accord que l'analyse devrait engager et les producteurs, et les utilisateurs des données. Ceci aurait des avantages certains sur la qualité de l'analyse, l'efficacité de la planification et d'autres applications.

17. La façon dont statisticiens et utilisateurs des données pourront collaborer dans l'analyse des données d'une enquête devra être déterminée à la lumière de l'expérience pratique et de la compétence des personnes disponibles. Tout le travail d'exploitation des données de base et plusieurs calculs destinés à l'analyse devront, bien sûr, revenir aux statisticiens; mais, à un niveau initial minimum, les utilisateurs devront être consultés lorsqu'un programme de tabulation ou une analyse est planifié, et ces derniers devront participer à l'interprétation des résultats.

III. ACTIVITES PRECEDANT L'ANALYSE

18. Ils existe bon nombre d'activités d'enquête, antérieures à l'analyse, qu'il faudra mentionner parce qu'elles influent sur l'analyse elle-même. Elles ont trait à la qualité des données, et il est aussi utile d'avoir une idée sur le moment où commence l'analyse après ces étapes préliminaires.

19. Les procédures normales de contrôle de qualité sont, en effet, plus importantes à tous les stades d'une enquêtes mais on pourrait noter qu'il y a une certaine différence entre les enquêtes à court passage unique et celles qui sont étendues sur une longue période et qui nécessitent des visites simples ou multiples. La précision des résultats d'une enquête de courte durée dépend principalement du soin avec lequel elle a été préparée, et d'une supervision stricte au cours du travail de terrain. Durant l'enquête, le service central ne dispose plus de temps pour intervenir.

20. Dans les enquêtes qui sont étalées dans le temps, il existe des possibilités supplémentaires pour contrôler les résultats au cours du déroulement du travail sur le terrain, et rectifier des erreurs avant que celles-ci ne causent de dégât significatif. Dans ce cas, il est nécessaire de faire des résumés des données enregistrées, si possible tous les mois, qui contribue à améliorer continuellement la qualité de l'enquête et la cohérence des données. De plus il serait avantageux de pouvoir comparer ces résumés avec des données d'autres sources indépendantes telles que les données d'un recensement général, des offices de commercialisation et des exportations. Un procédé de test continu au cours d'une enquête de longue durée peut améliorer énormément la qualité des données. Il est surprenant que bien de statisticiens africains prennent le risque inutile de ne pas surveiller la progression d'une enquête d'une manière systématique.

21. Les enregistrements d'une enquête doivent faire l'objet d'un contrôle manuel et automatique comme dans le cas des recensements de la population. Toutefois la quantité de données collectées lors d'une enquête est nettement plus petite et, de ce fait, le contrôle ne devrait pas apporter des corrections importantes si la supervision des travaux sur le terrain a été bien assurée. En particulier il est souhaitable de faire le moins possible d'imputation des données qui manquent et de correction des biais, car, il n'est jamais certain que de tels ajustements amélioreront les résultats. Chaque fois qu'on apporte des modifications aux données originales, il est important que celles-ci soient notées dans le détail.

22. L'évaluation des données est certainement le plus essentiel des étapes préliminaires à l'analyse. Elle arrive surtout vers la fin du stade du traitement des données, mais pourrait être considérée comme faisant partie intégrante du travail tout au long de l'enquête. Un aspect important a déjà été mentionné à propos du contrôle de la qualité et les procédures de contrôle contribuent aussi à la précision des données. Le calcul de l'erreur de sondage est un autre pas dans cet exercice d'évaluation. D'autres procédés d'évaluation comprennent l'utilisation de sous-échantillons enchevêtrés, de réinterview et d'un sous-échantillon de ménages, mais ils doivent être intégrés dans le plan de l'enquête, en particulier lorsqu'il y a un programme continu d'enquêtes. Toute l'évaluation conduit finalement à décider de la bonne qualité des données de base en vue de leurs utilisations pour une analyse détaillée.

23. Il existe un autre aspect qu'il vaut la peine de mentionner, qui concerne les procédures d'estimation plutôt que la qualité des données de l'enquête. Pour diverses raisons pratiques, les unités de sondage dans bon nombre d'enquêtes en Afrique n'ont pas été choisies en accord avec les objectifs de l'enquête. Cette façon de s'écarter des probabilités de sélection prévues contribue à biaiser les estimations si l'erreur n'est pas perçue, ou à introduire un travail supplémentaire de correction des pondérations dans la réévaluation des résultats de l'enquête. Une telle situation est un handicap sérieux pour l'analyse, elle est à éviter.

24. Ainsi qu'il a déjà été mentionné, le programme de tabulation pour le traitement des données doit être élaboré en collaboration avec les utilisateurs des résultats de l'enquête; autrement, ladite collaboration ne sera pas possible au niveau de l'analyse. Les tableaux devront normalement être reproduits dans le rapport de l'enquête. Le texte du rapport décrit l'enquête et ses résultats, en plus il essaie d'interpréter les résultats, ce qu'on peut déjà considérer comme une analyse. Il est par conséquent difficile de dire précisément où finit le traitement des données et où commence leur analyse. Dans le cadre du présent document, l'analyse des données est considérée comme tout ce qui se fait pour disposer des résultats de l'enquête après le stade de la collecte.

IV. OPTIONS D'ANALYSES

25. Comme indiqué dans le chapitre sur les besoins, il convient de distinguer deux sortes d'analyse : celles concernant l'ensemble des données et celles traitant les interrelations entre les variables. Le second groupe peut être subdivisé selon que les données proviennent de la même enquête ou d'enquêtes différentes.

26. Il n'est pas du tout nécessaire de parler ici de l'analyse en termes de données d'ensemble. Les résultats de l'enquête sont simplement utilisés comme source de données pour l'obtention des composantes appropriées d'un cadre général de travail. A la limite, différentes enquêtes peuvent contribuer au même cadre de travail, lequel peut lui-même servir de moyen d'intégration des données et peut aider à montrer des relations entre les thèmes.

27. L'inter-relation des variables étudiées dans la même enquête est aussi quelque chose de simple. Des tableaux élaborés au stade du traitement des données pourront suggérer de possibles relations qu'il est besoin d'étudier plus en détail, et l'analyse pourra être de nature très variée. En plus, il y a une analyse spécialisée, appliquée à plusieurs domaines d'études, qui a été déjà mentionnée.

28. Le problème qui se pose concerne l'inter-relation des variables issues de différentes enquêtes et des expériences pratiques manquent à ce sujet. Quelques méthodes qui ont été mises au point depuis longtemps sont mentionnées ci-après.

29. Il est nécessaire d'utiliser pour toutes les enquêtes, des concepts standards, des définitions et des classifications appropriées aux sujets étudiés. Ceci est une prédisposition à toute sorte de comparaison de données plutôt qu'une méthode d'analyse.

30. Une liste provisoire de questions de base a été publiée dans le document E/CN.14/SI/22. Si ces questions sont posées dans toutes les enquêtes indépendamment du sujet étudié, elles fournissent de variables communes qui permettent de classer les résultats de l'enquête de la même façon, chose qui favorise l'étude des interrelations.

31. Un cas spécial est celui des "enquêtes types" pour l'étude continue de l'agriculture et d'autres sujets importants et qui étudient des sujets divers sous formes de modules. Ces cas aussi ont besoin des concepts de base du genre de ceux indiqués plus haut, si l'échantillon au niveau du ménage doit changer assez fréquemment.

32. Dans le cas où les échantillons de différentes enquêtes ont les mêmes unités à plus d'un degré, il est possible de lier entre eux l'ensemble des résultats des enquêtes pour ces unités. Il n'est pas souhaitable que ceci soit fait au niveau du ménage excepté lorsque des procédures de remplacement partiel ont été effectuées.

33. Dans toutes ces méthodes du type de celles indiquées plus haut, il est nécessaire d'avoir à l'idée le calendrier des différentes enquêtes. Dans le cas des "enquête-types", le principal point d'intérêt serait d'expliquer les changements intervenus, dans le temps, dans les mêmes variables et cela ne pose aucun problème. Lorsque différentes variables sont étudiées par rapport à une autre, les périodes pendant lesquelles la collecte des données a eu lieu doivent être soigneusement prises en considération au moment de l'interprétation des résultats de l'analyse. Des considérations analogues s'appliquent aussi lorsque les estimations de différentes enquêtes sont mises ensemble ou lorsqu'on entreprend une analyse de séries chronologiques.

V. APPROCHE PAR LES DONNEES DE BASE

34. Comme le traitement des données de bon nombre d'enquêtes est maintenant informatisé, la plupart des pays voudront étudier la mise en place et l'utilisation des bases de données permanentes d'enquête. Elle comprendra probablement un fichier de données élémentaires pour chacune des enquêtes entreprises.

35. A la lumière des remarques faites dans paragraphes précédents, les principales fonctions de chaque fichier de données seront les suivantes :

- 1) Préparation des tableaux de l'enquête au stade du traitement des données.
- 2) Calcul des estimations, des erreurs d'échantillonnage etc... selon la méthode prévue lors de l'échantillonnage.
- 3) Analyse spécialisée supplémentaire recommandée pour le sujet étudié.
- 4) Analyse à plusieurs variates des variables étudiées dans l'enquête.
- 5) Classification des variables de l'enquête suivant les principales variables de base choisies.
- 6) Agrégation des résultats d'enquête à tous les niveaux pour chaque unité-échantillon.
- 7) Fourniture des données pour tous travaux et recherches ad hoc.

36. Le point intéressant des fichiers à sept fonctions est que toutes ont trait aux opérations effectuées sur des fichiers individuels. La seule situation dans laquelle deux ou plusieurs fichiers doivent être utilisés simultanément ou fusionnés est celle dans laquelle différentes enquêtes utilisent le même échantillon de ménages, ce qui, comme indiqué plus haut, n'est possible que si l'on effectue un remplacement partiel de l'échantillon. Même pour les fonctions 5) et 6), qui ont directement trait à l'inter-relation des variables étudiées dans différentes enquêtes, les données nécessaires peuvent être obtenues des fichiers individuels avant que l'analyse proprement dite ne soit faite.

37. Ceci implique qu'il n'y a aucune contrainte rigide qui imposerait une structure entièrement standardisée du fichier pour toutes les enquêtes. Cependant, pour permettre une manipulation plus aisée des données et pour simplifier le problème de la préparation de logiciel il est certainement souhaitable que tous les fichiers soient aussi semblables que possible. Le point important qui se dégage ici, est que la possibilité d'une certaine flexibilité dans la construction de fichiers individuels d'enquête est une chose importante si l'on se met à l'idée que différents plans de sondage, différents espaces et quantités des données doivent être mis ensemble et que les idées sur la méthode à utiliser pour réaliser un tel travail évolue nécessairement dans le temps.

38. Il est envisagé que les fichiers individuels d'une enquête donnée soient séquentiels, chaque enregistrement renfermant toutes les données jusqu'à la dernière unité de sondage et que le moyen de base de stockage de données soit la bande magnétique. Mais ceci ne s'applique pas aux données provenant du recensement préliminaire des unités de l'avant-dernier degré de sondage qui nécessitent évidemment des fichiers distincts; ils sont importants mais ne peuvent pas être considérés dans le présent document.

39. Le fait que la base de données elle-même soit une source permanente d'information, appelle aussi quelque commentaire. Actuellement la plupart des résultats d'enquête sont publiés sous forme d'un rapport formel d'enquête qui présente un intérêt limité pour des analyses ultérieures. Il est à souhaiter que l'informatisation des bases de données évoluera progressivement et accroîtra les fonctions des rapports. Les méthodes d'élaboration de telles bases de données seront discutées ultérieurement.

40. Les fichiers de base d'une enquête comportent une très grande quantité de données élémentaires dont on a besoin pour les fonctions 1) à 6) indiquées plus haut, et peut-être pour d'autres fonctions de même nature. Ces données ne peuvent pas sortir du service statistique à cause de la clause du secret statistique contenue dans la législation régissant la statistique dans la plupart des pays. Ainsi, pour toute recherche détaillée, entreprise par des agences extérieures telle qu'envisagée dans la fonction 7), il est nécessaire de préparer des fichiers nettoyés de données élémentaires qui respectent leur caractère confidentiel.

41. En effet, la diffusion d'informations plus générales sur les fichiers pour information générale au niveau des données élémentaires serait très encombrante. La solution la plus simple consiste en la constitution de fichiers condensés basés sur quelques données déjà discutées et pratiques pour transmission par des terminaux. On a besoin de quelque réflexion et expérience dans la détermination du contenu du fichier, mais les centres d'intérêt semblent être les résultats globaux des enquêtes individuelles, les conclusions relatives à l'inter-relation des variables économiques, sociales et démographiques, provenant de l'analyse de toutes les enquêtes, et les données sur les plus petites unités géographiques pour lesquelles les enquêtes peuvent fournir des estimations convenables. Les détails du point 6) ne sont pas d'un grand intérêt parce qu'ils concernent

seulement un échantillon. On suppose qu'il y aurait des fichiers généraux séparés pour la comptabilité nationale et pour d'autres données économiques et sociales, mais ils ne sont pas abordés ici.

VI. INFORMATIONS SUPPLEMENTAIRES SUR L'ANALYSE

42. Selon la définition donnée au début de ce document, relative à l'analyse des données, cette section pourrait se rapporter aux fonctions 3) et 6) du fichier des données. On n'envisage pas d'aborder le point 3) car les recommandations sur des analyses spécialisées figurent dans d'autres documents et les propositions révisées concernant les données sur la main-d'oeuvre sont présentées dans un document séparé de la présente réunion.

43. Autre point à considérer : la fonction 4) relative à l'analyse à plusieurs variables des données d'enquête. L'idéal serait de commencer le travail par le calcul des corrélations entre les variables prises deux à deux, y compris les questions essentielles, retenues dans l'enquête, et ensuite de procéder, par un choix des relations les plus significatives, à l'élaboration des équations de régression de ces variables. Tous ce calcul peut heureusement être réduit car, les premiers tableaux donneront certaines indications sur les relations entre les données et quelques groupes de variables, par exemple : les relations entre le revenu et les dépenses peuvent être fournies par postes. On se demande s'il est nécessaire de pousser tout au moins au début l'analyse jusqu'à la construction de modèles, parce que le principal objectif de la recherche des interrelations est de pouvoir expliquer les conditions économiques et sociales spécifiques à des fins de politique et de planification économiques.

44. Un exemple de ce genre de travail a été effectué par l'Enquête Mondiale de la Fécondité (EMF), reporté dans son document "stratégies pour l'analyse des données de l'EMF" (Document Technique de l'EMF n°9, janvier 1977). L'EMF a classé toutes les variables contenues dans ses enquêtes comme des variables dépendantes intermédiaires et explicatives en les étudiant par rapport à la fécondité. A partir de cette base, il est possible d'élaborer un cadre pour l'étude de la fécondité. Compte tenu du fait que le domaine de l'étude est restreint, seules quelques variables explicatives peuvent être utiles pour établir des relations avec d'autres données, mais on reconnaît maintenant qu'il s'est posé des problèmes dans l'identification et l'utilisation de ces variables. Cependant le point le plus important est de savoir de quelle manière les variables peuvent être classées comme dépendantes ou explicatives dans les opérations à objectifs multiples comme celles envisagées par le PADEM. Pour certains domaines d'étude donnés, la situation est très probablement la même que celle rencontrée par l'Enquête Mondiale de la Fécondité, et une distinction satisfaisante pourrait être possible, mais ce qui serait peut-être le plus important de rechercher lorsqu'il s'agit de la question plus large de relations entre différents sujets, est de définir entre elles des relations statistiques solides. Il faut mentionner aussi que des études semblables sur les interrelations démographiques, sociales et économiques existent dans le travail de recherche entrepris par le Programme Mondial de l'Emploi du BMT, avec l'application du modèle de Bachue.

45. Le point 5) concerne la classification des variables d'une enquête suivant les variables de base retenues dans des enquêtes semblables, c'est la première étape de l'étude des inter-relations des données de différentes enquêtes. Les questions fondamentales sont les suivantes : quelles variables de base utiliser ? Leur liste est-elle satisfaisante ?

46. La liste des variables de base contenues dans le document E/CN.14/SM/22 est une liste provisoire qui renferme deux sortes de variables. Il y en a que l'on peut tenter d'utiliser comme moyens d'établir des relations entre les données de différentes enquêtes; et d'autres qui y sont insérées parce qu'elles peuvent être utiles, mais dans tous les cas fournissent des informations qui sont assez valables pour être chaque fois collectées. C'est dire qu'il reste toujours un important travail de recherche à faire, pour identifier les vraies variables de base qui peuvent effectivement être utilisées à des fins d'analyse.

47. L'analyse interne des enquêtes individuelles sera l'une des principales méthodes de sélection de ces variables. Lorsqu'une variable est considérée comme explicative dans plusieurs domaines d'étude, il est certain que les analyses faites montrent que sa corrélation avec leurs variables est assez forte. On peut envisager que le résultat final sera une série de listes de variables de base relatives à des groupes de sujets qui se recoupent et que ces listes seront plutôt courtes. Les questions elles-mêmes porteront probablement sur les principaux indicateurs des conditions socio-économiques et des activités des ménages, et des facteurs de base, tels que la taille et la composition des ménages, nécessaires pour la standardisation la plus poussée des données enregistrées. D'autres variables de base peuvent apparaître au cours de l'analyse indiquée ci-dessus, et il se pose aussi la question générale de savoir jusqu'à quelle limite les indicateurs socio-économiques seront utiles. Rien de plus ne peut être dit, jusqu'à ce que les pays aient expérimenté cela avec les résultats de leurs enquêtes.

48. La méthode de relier entre eux les résultats de différentes enquêtes par l'utilisation des variables de base est, en effet, une méthode indirecte, car la relation entre les variables doit être étudiée à travers leurs relations individuelles avec l'une ou plusieurs des autres variables, c'est-à-dire les questions de base. C'est la raison pour laquelle l'analyse doit se concentrer sur les relations statistiques significatives les plus fortes qui peuvent se dégager.

49. Il y a juste un commentaire à faire sur la fonction 6) qui concerne l'agrégation des résultats d'enquête par unités-échantillon individuelles. Les aggregations seront probablement plus utiles à l'un ou deux stades utilisés dans la plupart des enquêtes en Afrique et elles doivent être présentées sous forme d'estimations des valeurs de la population pour chaque unité. Ainsi que nous l'avons déjà souligné, cette méthode de relier entre eux des résultats de différentes enquêtes est applicable seulement dans le cas des enquêtes ayant les mêmes unités-échantillon aux stades concernés. Cependant les estimations par unité peuvent aussi être utilisées comme une méthode rapide pour le calcul des erreurs d'échantillonnage.

VII. CONCLUSION

50. Le présent document constitue juste une légère amélioration du document (E/CN.14/SM/27), présenté à la réunion de 1979 du Groupe de Travail africain sur les enquêtes auprès des ménages; il n'essaye pas de présenter l'analyse des données d'enquête dans le sens des différentes étapes pratiques à franchir, mais a mis plus d'accent sur l'importance d'une collaboration continue avec les utilisateurs des données.

51. Un seul nouveau concept a été introduit, il porte sur l'élaboration des bases de données nationales d'enquête. Les deux conclusions spécialement importantes à cet égard sont : 1) il n'est pas trop difficile d'élaborer des bases de données car la plupart de leurs utilisations peuvent être limitées dans un premier temps dans les fichiers d'enquêtes particulières et 2) la discipline imposée par une base de donnée et l'identification connexe des fonctions du fichier les rend très faciles pour l'organisation systématique d'un travail d'analyse.

52. Il est clair que le présent document est une large revue de la situation actuelle et certaines omissions délibérées ont été faites dans le texte. Cependant le document est probablement l'un des premiers à tenter de présenter une vue générale sur l'analyse des données de programmes continus d'enquêtes à objectifs multiples.

53. Des commentaires sur les façons de relier entre elles des variables d'enquêtes ont été probablement peu conservateurs. Il existe des relations évidentes entre des situations économiques, sociales et démographiques, qui affectent les activités productives du ménage, et les niveaux de vie; les types d'analyse proposés aideront à les identifier et les étudier. La principale lacune réside dans le fait que certains autres types d'analyse n'ont pas été abordés, tels que l'Etude sur le calcul des niveaux de vie de la Banque Mondiale et la Matrice des comptes alimentaires du Groupe d'Etude des disponibilités Alimentaires de l'Université d'Oxford. Il est nécessaire que les organismes intéressés travaillant dans ce domaine coordonnent leurs efforts de réflexion; les résultats pourraient aboutir sur des conseils sur la construction et le fonctionnement des bases de données nationales.

54. Dans la situation actuelle, le principale besoin pratique est que tous les pays africains, tentent individuellement d'analyser leurs propres données. Jusqu'à ce que quelques résultats numériques ne soient disponibles, les agences internationales, susceptibles de fournir des conseils techniques nageront dans le vide.