

J. M. DETHOOR

J. E. FORGE

Programmes et langages

Revue française d'automatique, informatique, recherche opérationnelle. Recherche opérationnelle, tome 6, n° V1 (1972), p. 95-104

http://www.numdam.org/item?id=RO_1972__6_1_95_0

© AFCET, 1972, tous droits réservés.

L'accès aux archives de la revue « Revue française d'automatique, informatique, recherche opérationnelle. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

PROGRAMMES ET LANGAGES

par J. M. DETHOOR ⁽¹⁾ et J. E. FORGE ⁽¹⁾

Nous ouvrons aujourd'hui notre rubrique à huit programmes de dépouillement d'enquête.

- ASKING (ISS)
- DAPHNE (SIA)
- DETECT (ISS)
- G.T.S. (CAP)
- N.L.T. (3 I)
- PERSEE (STAD)
- PIST (SERTI)
- PRALINE (IBM)

Nos précédentes rubriques ont traité de programmes d'ordonnancement et de gestion prévisionnelle.

Nos prochaines rubriques concerneront la gestion du personnel, le contrôle de données, la gestion des stocks, la gestion financière à court terme.

Comme d'habitude nous invitons les sociétés de service et les maisons de « software » à nous faire parvenir de brefs résumés de leurs notices (au maximum deux pages dactylographiées).

*
*
*

Programme : ASKING (ISS)

ASKING est un programme d'interrogation et de dépouillement de fichier dont les possibilités vont du simple tri en fonction des classes à l'étude statistique où à la recherche de corrélations.

(1) Centre de Calcul de la Chambre de Commerce et d'Industrie de Paris.

Description

De structure modulaire, ASKING est très souple et peut être utilisé aussi bien pour le dépouillement d'un bordereau d'enquête que pour l'exploitation statistique d'enregistrements complexes tels que des dossiers médicaux.

ASKING interroge le fichier en procédant de manière séquentielle. Il sélectionne, compte et calcule au moyen de cartes « question » rédigées dans un langage simple qui autorise néanmoins toutes les opérations susceptibles d'intéresser n'importe quel exploitant de fichiers sur ordinateur.

Les résultats sont édités sur imprimante en langage clair avec les titres et les libellés choisis et fournis par l'utilisateur.

Possibilités

En une seule lecture du fichier il est possible de répondre à plusieurs questions, leur nombre dépendant essentiellement de la taille de l'ordinateur utilisé.

Une question comprend une opération de sélection suivie de traitements.

ASKING comporte les fonctions suivantes :

- comptages,
- pourcentages,
- histogrammes,
- listes,
- moyennes,
- écart types,
- coefficients de corrélation.

Fichiers à interroger

Avec ASKING il est possible de travailler sur des enregistrements de longueur fixe ou variable.

ASKING peut exploiter indifféremment un fichier de caractères ou un fichier de mots.

Le fichier à interroger peut être sur cartes perforées ou sur un support magnétique quelconque.

Aspects techniques

Écrit en FORTRAN, la taille-mémoire nécessaire à ASKING est fonction de l'application traitée et dépend du nombre de questions posées.

Disponible en service-bureau ASKING fonctionne sur différents types d'ordinateurs (IBM-64 K minimum).

UTILISATION

Pour toute information écrire à ISS, 6, rue Firmin-Gillot, Paris (15^e).

*
* *

Programme : DAPHNÉ (S.I.A.)

DAPHNÉ est un langage spécialisé pour le dépouillement d'enquêtes et l'analyse de fichiers; il permet la mise en œuvre des opérations essentielles suivantes :

- prise en charge d'un ou plusieurs fichiers contenant les données ;

— transformation éventuelle de l'information contenue dans le fichier, grâce à des opérations arithmétiques et logiques, des sélections, des groupements de code, des conditions, etc...;

— stockage, en mémoire du calculateur, des tableaux obtenus par croisement d'informations, et édition agréable et variée (valeurs brutes et pourcentages) de ces tableaux.

Domaines d'applications

DAPHNÉ est utilisé dans de nombreux domaines de l'activité économique et notamment :

1. Etude de marché et d'opinion.
2. Urbanisme.
3. Etudes du trafic urbain.
4. Analyse des ventes.
5. Statistiques sur panels de consommation.
6. Statistiques médicales (par exemple analyse du coût d'une journée d'hôpital).
7. Statistiques agricoles.
8. Statistiques sur fichiers de recensement.

Caractéristiques techniques

LE LANGAGE : DAPHNÉ est écrit en COMPASS, le langage d'assemblage de l'ordinateur CDC 6600.

LES DIMENSIONS :

- Tableaux : jusqu'à 100 000 cases peuvent être réservées pour un seul passage.
- Capacité d'un tableau : un tableau peut avoir jusqu'à 16 000 cases.
- Dimensions d'un tableau : 3 et plus éventuellement.
- Nombre des variables : pas de limites.
- Grandeur des variables : sans limites.
- Nombre de réponses multiples par variable : jusqu'à 59.
- Nombre de fichiers d'entrée : jusqu'à 9.
- Capacité des fichiers : pas de limites.
- Nombre de cartes par questionnaire : pas de limites.
- Types de cartes par questionnaire : à l'intérieur de chaque questionnaire toute combinaison de types de cartes est à mise.

Pour des informations plus détaillées, s'adresser à : S.I.A. Metra International
Direction Commerciale, 35, boulevard Brune, Paris (14^e).

*
* *
*

Programme : DETECT (ISS)

DETECT est un programme généralisé de dépouillement d'enquêtes par édition de tableaux comparatifs, mis au point et commercialisé par ISS.

CONFIGURATION REQUISE :

— CDC 6 600, 64K de mémoire centrale, 110 K ECS (pour élaboration de tableaux),

- 1 lecteur de cartes ou 1 bande et 1 imprimante,
- IBM 360-256K version O.S.

LANGAGE SOURCE : FORTRAN

Description :

- Traitement de données d'enquête et regroupement par tableaux libellés en clair, directement exploitable grâce à un macro-langage programmé pour chaque dépouillement.

- Les tableaux édités par DETECT comportent :

- des cumuls de valeur,
- des comptages,
- des moyennes,
- des pourcentages.

- Le macrolangage permet de calculer pour chaque tableau et série de données, les valeurs statistiques classiques.

- Création de l'ensemble des tableaux en un seul passage ordinateur.

- Prise en charge par le programme de cartes données, ordinaires et de cartes perforées.

- Traitement d'enregistrements de longueur variable, de la multiperforation.

Le programme comporte le programme FORTRAN proprement dit et un jeu de données qui se subdivise en :

- une suite de cartes paramètres perforées conformément au langage conventionnel spécifique du programme de dépouillement d'enquête. Elles représentent une vingtaine de macro-instructions qui permettent la création et l'édition des tableaux souhaités;

- le lot de cartes de données.

DETECT peut être soit vendu, soit exploité par ISS en service-bureau.

En cas d'achat du package, ISS assure la formation des utilisateurs, la mise en place dans l'entreprise, ainsi que la maintenance du programme.

En service-bureau, ISS se charge du dépouillement complet : création des bordereaux d'enquête, vérification des cartes, traitement des données et édition des résultats et ce, dans un délai de quelques jours.

Le coût de l'opération est fonction des données fournies et des résultats demandés, c'est-à-dire matériellement du nombre de cartes lues et de pages éditées.

Utilisation

Pour toute information écrire à ISS, 6, rue Firmin-Gillot, Paris (15^e).

*
* *

Programme : GTS (CAP FRANCE) : (Générateur de Tableaux statistiques)

G.T.S., générateur de tableaux statistiques, a été conçu et mis au point par l'Institut National de Statistiques et Sciences Economiques (INSEE) pour le dépouillement des recensements de la Population Française.

1. Les tableaux caractéristiques obtenus par G.T.S.

Ils sont décrits grâce aux paramètres suivants :

- *les croisements à effectuer :*

G.T.S. croise, dans un tableau, jusqu'à 8 critères. *Ce sont les 8 dimensions de ventilation du tableau ;*

- *les quantités à totaliser :*

G.T.S. peut, à l'intersection des 8 critères de ventilation, cumuler 14 quantités différentes, dites critères de totalisation.

Un critère de totalisation est soit :

- une constante,
- une variable du fichier,
- une expression calculée à partir des variables du fichier,

- *les filtres :*

Avant de prendre un individu en compte dans un calcul, G.T.S. vérifie qu'il satisfait :

- 1° les conditions d'appartenance au tableau,
- 2° les conditions d'appartenance au critère de totalisation.

Il y a donc deux types de test avant tout calcul.

• *la partition :* (au sens mathématique du terme) de la population sur laquelle doit s'effectuer la ventilation.

Les statistiques fournies par G.T.S. portent sur plusieurs sous-ensembles de la population.

Une 9^e classification est possible suivant la ligne générale d'intérêt du statisticien. Elle est dite « géographique » (car le plus souvent c'est un découpage géographique qui s'impose) ; elle est par essence différente des autres dimensions car elle est liée à un tri préalable du fichier.

2. Les principes du système

• A partir d'un squelette du programme, G.T.S. sélectionne les seules instructions nécessaires à l'élaboration des tableaux demandés.

- La place occupée en mémoire par un tableau dépend :

- du nombre de cases de ce tableau ; ce nombre est proportionnel au nombre de postes de chacun des critères de ventilation et au nombre de critères de totalisation,
- du nombre d'instructions nécessaires à l'élaboration du tableau.

Cette place ne dépend absolument pas de la partition géographique choisie.

La taille maximum d'un tableau, avec une partition mémoire de 200 K est de 30 000 cases environ,

- Différents tableaux sont élaborés en même temps.

G.T.S. attribue automatiquement un ordre de passage à chacun des tableaux demandés de façon à assurer une occupation maximale de la mémoire centrale.

3. La structure du fichier d'entrée

Deux organisations du fichier sont superposées :

- une hiérarchie en niveaux d'information (4 au maximum);
- un tri du fichier, condition essentielle à l'obtention de la « dimension géographique ».

4. Les traitements ultérieurs. Edition

G.T.S. offre les diverses options de traitement :

- *Expression de résultats* sous forme de pourcentages.
- *Calcul de totaux partiels.*
- *Suppression de un ou plusieurs critères de ventilation.*
- *Récapitulation* d'un tableau.

Un tableau est formé de sous-tableaux portant chacun sur un sous-ensemble de la population.

Une « récapitulation » permet le cumul de plusieurs sous-tableaux.

EXEMPLE : un tableau porte sur les 95 départements français. Il peut être « récapitulé » sur les 21 régions de programme, chaque région étant composée de plusieurs départements.

- *Edition*

Elle peut se faire sous la forme standard ou dans un cadre préimprimé.

Il est possible de ne pas imprimer un tableau qui sert uniquement d'intermédiaire dans un traitement.

- *Répertoire des tableaux*

On peut demander un répertoire des tableaux. Il permet un archivage rapide des résultats.

5) Caractéristiques informatiques

G.T.S. écrit en PL 1 et Assembleur 360 est exploité sur IBM 360 fonctionnant sous O.S., release 17 et suivantes en MFT, MVT ou PCP. 200 K minimum.

UTILISATION

G.T.S. est diffusé par CAP France.

Le prix de vente comprend :

- l'installation du produit-programme,
- la formation des utilisateurs,
- la documentation,
- la maintenance gracieuse pendant 6 mois.

Le client a la possibilité de souscrire un contrat de maintenance moyennant une redevance annuelle égale à 8 % du prix de vente du produit. Par ce contrat, le client bénéficie de la mise à jour du système lors du changement de l'OS ainsi que des améliorations qui peuvent être apportées au produit lui-même.

*
* *

Programme : NLT (3 I) : Nouveau Langage de Tabulation**A) Types de travaux effectués**

1) Prise en compte d'informations relatives à une ou plusieurs unités statistiques hiérarchisées entre elles (9 niveaux maximum) à partir d'un ou plusieurs fichiers cartes, bandes ou disques traités séquentiellement (9 fichiers maximum).

2) Organisation et mise en forme de ces informations, construction de nouvelles informations par calculs arithmétiques et logiques et consultation de tables. Les informations peuvent être unitaires, vectorielles ou matricielles. Archivage automatique de ces informations et des libellés associés.

3) Utilisation des informations mises en forme comme paramètres de dimension, de filtrage et de pondération de tableaux de contingence de 1 à 4 dimensions à n'importe quel niveau de comptage et de rupture.

4) Réalisation de calculs par toutes opérations arithmétiques entre les tableaux obtenus (% , moyennes, variances, etc...) et édition libellée des résultats avec mise en page automatique très souple.

5) Sortie des informations désirées pour les articles désirés sur le support désiré (carte, bande, disque, imprimante) avec maîtrise du dessin d'enregistrement.

B) Caractéristiques

1) S'applique aux enquêtes par sondage et aux fichiers des entreprises et des administrations : fichier commerciaux, recensement, économiques ou sociologiques, etc...

2) Particulièrement intéressant pour les travaux non répétitifs pour lesquels une programmation spécifique serait trop onéreuse ou trop longue.

3) Langage utilisant un seul type de bordereau. Destiné aux statisticiens, chargés d'études, sociologues, gestionnaires, etc... non informaticiens. Apprentissage en quelques jours.

4) Compilateur écrit partie en assembleur IBM 360, partie en Fortran IV. Fonctionne sous OS.

La compilation exige 160 K. Le programme compilé peut passer sur une plus petite partition.

N.L.T. est opérationnel depuis 1966, depuis 1971 dans sa dernière version.

Il est un véritable langage sur lequel sont prévus les branchements de modèles statistiques spécifiques (typologie, redressement d'enquêtes, ...).

Il peut être exploité sur nos ordinateurs, écrit par nos services ou vendu et installé chez nos Clients.

Pour tous renseignements, il convient de s'adresser à « 3 I », Institut International d'Informatique, 69, rue Legendre, Paris (17^e).

* * *

Programme : PERSEE (STAD) : Programme d'Elaboration et de Résultats Statistiques Extraits d'Enquêtes**1. Généralités**

PERSEE est un langage général de dépouillement d'enquêtes ou de fichiers, développé par Régie-Informatique, et mis en œuvre par STAD.

n° V-1, 1972.

PERSEE est un outil puissant, simple, sûr et économique.

- Puissant, car il permet de faire des traitements extrêmement variés et complexes.
- Simple à mettre en œuvre et facile à comprendre (son emploi ne nécessite aucune connaissance des ordinateurs, ni de la programmation).
- Sûr, car il protège au maximum l'utilisateur des erreurs pouvant exister dans son fichier ou son programme.
- Économique, du fait d'une part de ses performances, d'autre part, parce que l'utilisateur n'effectue qu'un minimum d'écriture pour décrire son traitement.

2. Description

PERSEE est composé de 3 programmes :

CONTROL, PREVAR et TABUL.

2.1. CONTROL est un programme permettant de contrôler la structure du fichier d'entrée, et de le mettre sous une forme admise par PREVAR.

Le fichier d'entrée peut être :

- 1 fichier-cartes (ou bande en format S.D.F.). Un questionnaire est représenté par une ou plusieurs cartes;
- 1 fichier quelconque, dans ce cas l'utilisateur doit fournir un sous-programme de lecture transmettant un questionnaire ou article à chaque appel.

2.2. PREVAR est un programme permettant :

- de décrire les variables intervenant dans le traitement (nom, dimension, type, titre, libellés, etc...).
- de donner, pour chaque questionnaire, des valeurs à ces variables en fonction des informations contenues dans le fichier d'entrée et/ou de toutes les opérations arithmétiques et logiques nécessaires.

2.3. TABUL est un programme permettant d'obtenir des tableaux de ventilation à partir de ces variables. Ces tableaux peuvent être à 1, 2 ou 3 dimensions, et peuvent inclure :

- un filtre,
- une pondération,
- une variable de rupture pour contrôler l'édition,
- des calculs de moyenne et d'écart-type,
- différents types de totaux et sous-totaux,
- des calculs de pourcentages sur des bases variées,
- le calcul du $KH_i - 2$,
- etc...

L'édition des tableaux est automatique et se fait avec une grande variété de titres et de libellés.

3. Documentation

La documentation PERSEE comprend :

- Une brochure de description du langage. Elle suffit à l'utilisateur pour rédiger son programme.

— Un manuel technique, donnant des informations complémentaires concernant en particulier :

- Les cartes-contrôle,
- Les limites du programme qu'il n'est pas nécessaire de connaître dans le cadre d'une utilisation normale de PERSEE. (Les messages d'erreur nécessaires étant imprimés.)
- Les traitements spéciaux (par exemple la forme des fichiers utilisés).

4. Conditions d'accès et d'utilisation

1. Utilisation à la STAD

— *En libre-service*

Le client compose son programme et appelle PERSEE comme tout autre processeur existant.

— *Avec l'aide de l'Assistance technique*

STAD se charge de réaliser le programme nécessaire et d'effectuer le traitement du fichier.

2. Achat du programme

Le programme peut être cédé par Régie-Informatique. Dans le prix de cession sont inclus un cours de formation et la maintenance du programme pendant deux ans.

5. Le programme est écrit en langage assembleur pour UNIVAC (série 1100); il exige une capacité mémoire de 40 K mots. Il doit être utilisé sous EXEC 8.

6. Tout renseignement complémentaire doit être demandé à STAD, 11, rue de la Vistule, Paris (14^e). Service Promotion des Produits : ou Service Commercial).

STAD dispose également de nombreux produits complétant PERSEE, par exemple: segmentation, typologie, analyse factorielle, etc...

La plupart des sociétés spécialisées dans le dépouillement d'enquêtes utilisent PERSEE. Références sur demande.

*
* *

PIST (SERTI) : Programme général d'édition de tableaux statistiques

1. Spécifications techniques

- Configuration requise : ordinateur de 64 K octets de mémoire centrale,
- Peut traiter tout fichier consultable séquentiellement (cartes, bandes magnétiques, disques magnétiques...).

2. Principes de base et possibilités

2.1 Principe

A partir de cartes paramètres extrêmement simples, PIST constitue les tableaux en mémoire centrale (son temps de traitement est donc optimisé).

2.2 Possibilités

- Fichiers contenant des enregistrements de différents types.
- Tris à plat et tris croisés avec trois types de pourcentages.
- Tableaux de cumuls et de moyennes.
- Recodification.

- Filtrés simples et multiples.
- Pondération.
- Calcul du KHI-2.

3. Modalités d'utilisation

PIST est vendu; le prix de vente comprend la mise en place sur l'ordinateur utilisé par l'acheteur et la formation des utilisateurs (1/2 journée de formation).

La maintenance est assurée pendant deux ans.

UTILISATION

4. Pour tout renseignement complémentaire, contacter SERTI, 17, rue Monsigny, Paris (2^e).

* * *

Programme : PRALINE (IBM) : Programme d'Analyse logique et d'Interprétation numérique des enquêtes.

PRALINE, est un programme conçu et écrit par le Service Bureau de la Compagnie IBM France.

PRALINE permet d'effectuer sur tous les types d'enquêtes (études de marché, études sociologiques, démographiques, sondages d'opinion, problèmes de circulation, études médicales, agronomiques, fichiers statistiques, etc...) un dépouillement aisé à mettre en œuvre, économique, rapide et sûr.

Grâce à un langage adapté, simple à apprendre, il est possible d'épurer et de recodifier les questionnaires, d'effectuer des calculs statistiques, de ventiler les résultats sous forme de tableaux de contingence avec titres et libellés.

L'entrée des données peut se faire sur cartes perforées, bandes magnétiques, ou feuilles de lecteur optique.

POSSIBILITES

Les nombres de questionnaires et de tableaux édités sont illimités. La taille des plus grands tableaux admis est de 65 000 cases. Les tableaux peuvent avoir jusqu'à 4 dimensions. Il est possible d'introduire des pondérations afin de redresser un échantillon. En outre, des moyennes, écarts-types, pourcentages, totaux partiels sont fournis automatiquement, sur option, ainsi que les X^2 et les coefficients de corrélation. PRALINE admet jusqu'à 2 000 variables pouvant intervenir dans la ventilation, et 8 000 libellés différents.

Il fonctionne sur l'ordinateur IBM 360/75 sous le système d'exploitation Operating System/IBM du Service Bureau. Il occupe 200 à 600 K octets de mémoire, suivant les problèmes et la taille des tableaux demandés.

UTILISATION

Pour l'utiliser vous devez vous adresser :

- soit à votre Ingénieur Commercial IBM;
- soit directement au Calcul Scientifique, Service Bureau IBM, 116, avenue de Neuilly, 92-Neuilly.