
Traitement des données

Le traitement des données

Avant de pouvoir saisir l'information recueillie, il est nécessaire d'avoir préparé la description de la base de données qui contiendra les renseignements recueillis. Il peut donc exister une base de données « vide » préalablement à l'entrée des données. Cette base servira ordinairement à la définition de la fenêtre de saisie de l'information ou du lien entre le mécanisme d'entrée des données et les fichiers à constituer. À la suite des opérations de saisie des données, plusieurs activités prennent ordinairement place dans le but de préparer les bases de données qui seront utilisées pour effectuer les compilations requises. Le logiciel EduStat peut être utilisé pour exécuter le traitement des données par rapport aux étapes habituellement retenues; il s'agit de celles-ci :

On peut identifier les étapes suivantes :

- l'entrée des données par lot;
- l'importation des lots dans la base de données;
- le contrôle de qualité;
- la première phase de compilation;
- le jumelage d'information provenant de différentes sources;
- la création de bases de données de différents formats.

A. Entrée des données par lot

Les données sont ordinairement entrées sur support informatique par « lot », c'est-à-dire un ensemble de documents répartis en plusieurs regroupements de taille modeste à chacun desquels correspondra un fichier de saisie. Il sera plus facile d'exercer le suivi de chacun de ces ensembles de formulaires tout au long des opérations qui conduiront à la constitution des bases de données nécessaires à l'exécution des compilations requises. Dès qu'un lot a été codifié et que la notation a été réalisée, il est possible d'en confier la saisie aux opérateurs concernés.

B. Importation des lots dans la base de données

Au fur et à mesure de la disponibilité des fichiers lots, il sera possible et souhaitable d'intégrer ces données dans la base de données en préparation. La constitution graduelle de la base de données permettra d'exercer un contrôle de qualité sur une base quotidienne. Si des problèmes sont détectés (par exemple, présence de valeurs non prévues, format de valeur non conforme aux attentes), il sera plus facile de résoudre les problèmes dès leur identification. De premières compilations (par exemple, des distributions de fréquences) pourront être exécutées; ces opérations pourront fournir des renseignements préliminaires sur les données recueillies et ajuster les étiquettes de valeurs mieux adaptées à la réalité.

C. Contrôle de qualité

Certaines activités peuvent être entreprises pour contrôler la qualité des données saisies et intégrées à la base de données en préparation. En plus de produire périodiquement des distributions de fréquences pour détecter de possibles anomalies, il peut être utile de réaliser des compilations qui rendent compte de la rigueur des activités de codification et de correction. Nous pouvons citer ici le mécanisme de comparaison des valeurs de la base de données en préparation avec une base formée d'un échantillon d'enregistrements saisis par les soins d'une autre équipe d'entrée de données.

D. Première phase de compilations

Après avoir intégré l'ensemble des fichiers lots prévus, il est possible de réaliser certaines compilations préliminaires touchant les variables qui serviront à regrouper les enregistrements à des fins d'analyse. Par exemple, des renseignements contextuels sur les élèves, les enseignants ou autres répondants doivent être disponibles pour exécuter certains traitements. S'il y a absence d'information ou information non prévue pour les variables dites « sensibles », il sera approprié de consulter les épreuves ou questionnaires recueillis et apporter les corrections si cela s'avère être possible. Il est parfois possible de faire des liens entre certains éléments d'information pour vérifier la cohérence des renseignements fournis.

L'exécution de différentes procédures de traitement de l'information permettra de finaliser la base de données. Il peut être nécessaire de codifier certaines variables pour rendre l'information plus cohérente et plus facile à traiter lorsque l'on établira des comparaisons avec d'autres séries de données.

E. Jumelage d'information provenant de différentes sources

Les renseignements recueillis dans le cadre d'une activité d'évaluation proviennent ordinairement de plusieurs sources. On peut, par exemple, administrer des épreuves à des élèves, leur demander de répondre à un questionnaire complémentaires, sonder leurs enseignants ou autre personnel. On peut aussi recueillir des renseignements contextuels en provenance de différentes sources. Au moment de l'analyse des données, il sera souhaitable de mettre en relations les renseignements provenant de ces différentes sources. Il devra donc y avoir « jumelage » d'information. Pour cela, il faudra avoir prévu des numéros d'identification permettant d'établir les liens nécessaires.

F. Création de bases de données de formats variés

Enfin, les données recueillies sont ordinairement traitées à l'aide d'outils informatiques divers. Il est rare qu'un seul logiciel soit en mesure d'exécuter l'ensemble des compilations souhaitées. Aussi faut-il être en mesure de produire des formats de fichiers compatibles aux logiciels qui serviront à traiter toutes les compilations prévues.

Adresse courriel : info@mcpythagore.com