



Logiciel EduStat

Distributions de fréquences

La procédure « Fréquences » permet d'obtenir, pour une ou plusieurs variables de la **base active**, la proportion que représente chacune des valeurs observées. Deux options de fréquences sont disponibles : le calcul de fréquences simples et la compilation de fréquences distribuées. Par ailleurs, la procédure « Tableaux de fréquences » permet de répartir la distribution des fréquences pour une ou plusieurs variables selon les valeurs d'une ou de deux autres variables.

Ce document comprend des renseignements utiles pour accomplir des compilations sur la distribution des fréquences. La base de données « **EchantillonnageFR.edu** » sera utilisée pour exécuter les exercices proposés. Cette base de données (elle est aussi utilisée dans la présentation de la sélection d'un échantillon) comprend 25 variables et 450 enregistrements. À noter qu'il serait aussi possible d'utiliser la base de données « **ElevesFR.edu** » pour produire des distributions de fréquences. On trouvera ces fichiers dans le dossier d'EduStat « Bases de données EduStat » qui ont probablement été transférés au moment de l'installation du logiciel.

Nous produirons des compilations statistiques reliées aux options suivantes du logiciel EduStat :

- La compilation de fréquences simples
- La compilation de fréquences distribuées
- La compilation de tableaux de fréquences

Plusieurs documents sont disponibles au moment de l'utilisation du logiciel EduStat. Si on souhaite avoir des renseignements complémentaires sur la démarche à retenir pour effectuer des distributions de fréquences, il est possible de consulter les documents suivants faisant partie de la documentation. Pour y accéder, choisir à partir du menu « **Aide** », l'option « **Documentation** ». Une fenêtre comprenant cinq options s'affiche alors. Il est possible d'ouvrir l'une des trois options permettant d'avoir des renseignements sur la production de distributions de fréquences en choisissant l'option de son choix : Diaporamas, Vidéos, Procéduriers et les Brochures. Il est aussi possible de consulter le « Guide d'utilisation ». Nous vous conseillons de prendre connaissance de certains de ceux-ci avant de réaliser les activités présentées dans ce document. Voici quelques références.

- Les pages 46 à 49 du Guide d'utilisation.
- Le visionnement des vidéos suivants : Fréquences simples; Fréquences distribuées; Tableaux de fréquences.
- Le visionnement des vidéos suivantes : Fréquences simples; Fréquences distribuées; Tableaux de fréquences.

- La consultation des procéduriers suivants : Fréquences simples; Fréquences distribuées; Tableaux de fréquences.
- La consultation des brochures suivantes : Fréquences simples; Fréquences distribuées; Tableaux de fréquences.

❖ Fréquences simples

À l'aide de cet exercice, il sera possible de produire des distributions de fréquences simples; dans un premier temps, nous utiliserons les variables suivantes de la base de données « EchantillonnageFR.edu » :

STATUT : Le statut de l'école, c'est-à-dire : école publique et école privée

ZONE : Le rattachement des écoles au milieu urbain et au milieu rural

Voici la démarche à retenir.

- Information fournie

Il est possible donc d'obtenir les renseignements suivants sur les fréquences produites :

Fréquences en **nombres** seulement

Fréquences en **pourcentage** seulement

Fréquences en **nombres et pourcentage**

- Traitement des valeurs manquantes

Dans un premier temps, nous retiendrons chaque variable par rapport à toutes les valeurs incluses.

- Choix des variables à traiter

Sélection	Type
STATUT	À
ZONE	À

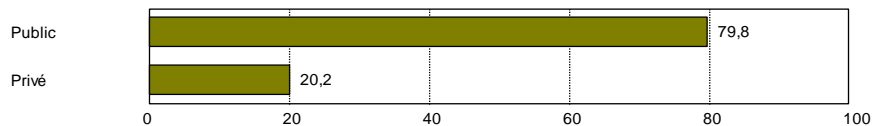
Nous souhaitons traiter les deux variables suivantes : le statut de l'école et la location par rapport à la zone.

Voici des exemples de rapport de distributions au regard du statut des écoles et les zones d'implantation; nous pourrions retenir un nombre plus grand de variables. Le logiciel a aussi produit des graphiques pour illustrer les résultats obtenus.

Statut de l'école

	N	Proportion
Public	359	79,8
Privé	91	20,2
Total	450	100,0

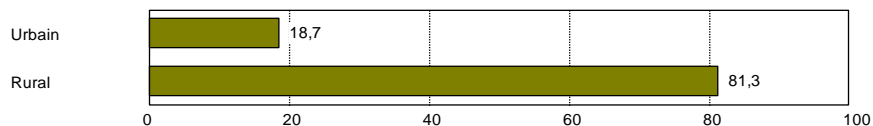
Exemples de distributions de fréquences Statut de l'école



Zone d'implantation de l'école

	N	Proportion
Urbain	84	18,7
Rural	366	81,3
Total	450	100,0

Exemples de distributions de fréquences Zone d'implantation de l'école



❖ Fréquences distribuées

Cette option de compilations statistiques permet de produire des distributions observées par rapport à une variable de format numérique et deux ou trois variables contextuelles.

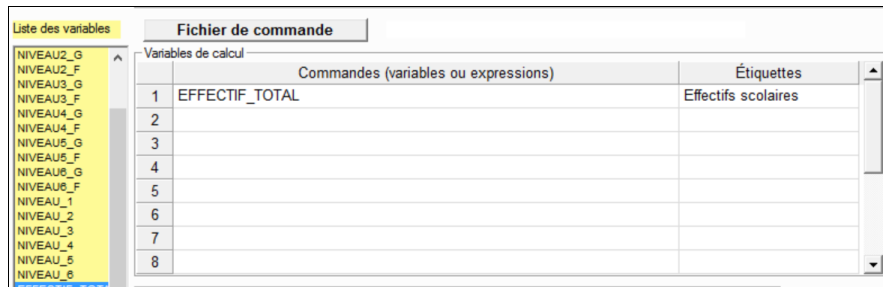
Voici la démarche à retenir.

- Sélection des variables contextuelles (maximum : 3)

Sélection	Étiquettes
REGION	Les régions
STATUT	Le statut des écoles

Nous avons sélectionné deux variables contextuelles : la région et le statut.

- Identification de la variable de calcul



La variable de calcul concerne l'effectif des écoles faisant partie de la base de données sélectionnée. Voici le rapport produit :

Les régions	Le statut des écoles	Effectifs scolaires
Région numéro 10	Public	5435
	Privé	1016
	Total	6451
Région numéro 11	Public	4403
	Privé	199
	Total	4602
Région numéro 20	Public	3982
	Privé	169
	Total	4151
Région numéro 16	Public	5267
	Privé	883
	Total	6150
Région numéro 17	Public	10518
	Privé	273
	Total	10791
Région numéro 14	Public	3354
	Privé	110
	Total	3464
Région numéro 21	Public	3367
	Privé	759
	Total	4126
Région numéro 18	Public	6245
	Total	6245
Région numéro 12	Public	8070
	Privé	3467
	Total	11537
Région numéro 22	Public	5008
	Privé	4068
	Total	9076
Région numéro 15	Public	6765
	Privé	969
	Total	7734
Région numéro 19	Public	8917
	Privé	1139
	Total	10056
Région numéro 13	Public	3296
	Privé	340
	Total	3636
Total général		88019

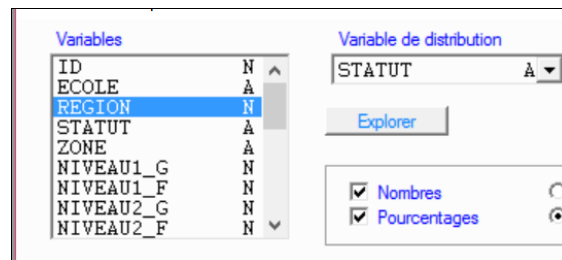
Ce rapport répartit les 88 019 élèves faisant partie de la base de données par rapport à chaque région et en fonction du statut des écoles. Par exemple, on constate qu'il y a 3 636 élèves dans les écoles de la « Région numéro 13 »; il y a 3 296 élèves dans les écoles publiques et 340 dans les écoles privées.

❖ Tableaux de fréquences

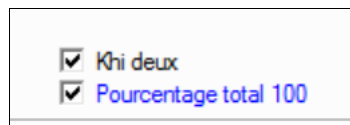
Cette option permet d'établir les liens qui s'établissent entre deux ou trois variables. Un indice statistique – le khi-deux – permet de connaître l'importance statistique de ces liens.

Nous compilerons un tableau en utilisant les deux variables suivantes : les régions administratives et le statut des écoles. Voici la démarche à suivre :

- Identification de la première variable et de la deuxième variable



- Il faut aussi indiquer que l'on souhaite le calcul du khi-deux.



Voici le rapport produit

Région administrative	Public		Privé		Total	
	Nombre	%	Nombre	%	Nombre	%
Région numéro 10	26	7,2	7	7,7	33	7,3
Région numéro 11	37	10,3	4	4,4	41	9,1
Région numéro 12	18	5,0	3	3,3	21	4,7
Région numéro 13	43	12,0	0	0,0	43	9,6
Région numéro 14	32	8,9	18	19,8	50	11,1
Région numéro 15	41	11,4	12	13,2	53	11,8
Région numéro 16	27	7,5	3	3,3	30	6,7
Région numéro 17	24	6,7	1	1,1	25	5,6
Région numéro 18	18	5,0	1	1,1	19	4,2
Région numéro 19	18	5,0	7	7,7	25	5,6
Région numéro 20	33	9,2	9	9,9	42	9,3
Région numéro 21	16	4,5	21	23,1	37	8,2
Région numéro 22	26	7,2	5	5,5	31	6,9
Total	359	100,0	91	100,0	450	100,0

Khi deux: 62,636; Degrés de liberté:12; Probabilité 0,000

La valeur obtenue au khi-deux (62,636) indique qu'il y a un lien statistiquement significatif entre les deux variables (indice de probabilité inférieur à 0,05). La proportion d'écoles publiques et privées varie donc en fonction des régions administratives.