

Base de données pour l'UAM

Format

Si vous avez un autre format que ceux listés ci-dessous, veuillez en discuter avec un membre de l'UAM.

- Excel .xls ; .xlsx
- *comma separated* .csv
- SPLUS ou R .Rdata

- STATA .dta
- SPSS .sav

Codebook

	A	B	C	D
1	ID	identifiant des patients		
2	age	âge des patient-es (en années)		
3	gender	genre: 0=homme; 1=femme		
4	hospitalisation	temps d'hospitalisation (en jours)		
5	ISS	scoer à l'ISS		
6	date_sortie	date de sortie de l'hôpital		

Inclure un codebook avec une explication des variables et de leurs modalités.

Inclure les labels des variables et des valeurs directement dans le fichier.

	Name	Type	Width	Dec...	Label	Values	Missing Co...	Measure
1	ID	String	5	0	participant's ID	None	None	5 Nominal
2	age	Numeric	18	15	age (year)	None	None	12 Scale
3	gender	Numeric	1	0	gender (0, homme)...	None	None	12 Nominal
4	hospitalisation	Numeric	1	0	days of hospital...	None	None	12 Scale
5	ISS	Numeric	1	0	ISS score	None	None	12 Scale
6	date_sortie	Date	10	0	date de sortie d	None	None	11 Scale

Données



Codées. Enlever les informations permettant l'identification des patient-es (ex. noms, préférer l'âge à la date de naissance)



Pas de caractères spéciaux



Pas de formule. Sur Excel, si les valeurs d'une variable sont le résultat d'une formule, convertir ces valeurs en numérique (1. Copier ; 2. Collage spécial ; 3. Valeur)



Format homogène. Le format de toutes les cellules d'une variable donnée doit être le même (ex. éviter de mettre '<0.1' (caractère) parmi des valeurs numériques)



Une seule information par cellule. (ex. éviter de mettre la cause du décès dans la même cellule que la date de décès)



Pas de statistiques descriptives. Ne pas inclure de statistiques descriptives (ex. moyennes, médianes, fréquences)



Regroupées. Éviter de séparer les données sur plusieurs bases de données ou onglets. Regrouper toutes les données sur un seul tableau (ex. cas et témoins dans un seul tableau, dont le statut est identifié par une variable)



Valides et finales. Vérifier la validité des données, toute anomalie doit avoir été corrigée par vos soins. Les données doivent être définitives, sans ajout ultérieur

Le nom des variables

1. noms courts
2. pas d'espace ni de caractères spéciaux (ex. 'é', '&')
3. ne comporte que des lettres, des chiffres et/ou des underscores '_'.
ne commence pas par un chiffre
4. doit commencer par une lettre (et non un chiffre)
5. un nom par variable, une variable par nom. Ne pas utiliser de couleurs pour distinguer des conditions. Chaque colonne correspond à une variable aux modalités mutuellement exclusives (ex. une variable 'sexe' (M/F), et non pas une variable 'homme' (oui/non) et une 'femme' (oui/non))
6. sur une seule ligne (la première ligne de la base de données)

Participants' ID	sexe		physio			mental	
	homme	Femme	Saturation oxygène (%)	Retrait (0=non, 1=oui)	Concentration uranium (mcg/L)	Score de dépression HADS	Score de bien-être
1	0	1	56	0	30	6	12
2	1	0	34	0	32	4	15



ID	sexe	Sat_oxygene	retrait	uranium	HADS	bien_etre
1	F	56	0	30	6	12
2	M	34	0	32	4	15

Si une base de données ne respecte pas les principes présentés dans le présent document, l'UAM se réserve le droit de refuser d'effectuer les analyses.

S'il vous est difficile de respecter l'un de ces principes ou si vous avez des questions, veuillez prendre contact avec l'UAM afin d'en discuter.

Mesures répétées

- Mesures répétées : une ligne par occurrence de mesure (plusieurs lignes par patient-e) avec au moins deux colonnes: ID et numéro de la mesure (ex. T1, T2, T3). Si vous pensez qu'un format large (une ligne par patient-e) est préférable, veuillez contacter l'UAM.

- Les données mesurées qu'une seule fois (ex. genre, âge) doivent être répétées sur chaque ligne.

ID	sexe	age	pre_operatoire			post-operatoire		
			oxygene	Retrait	Concentration	oxygene	Retrait	Concentration
1	0	65	56	0	30	51	0	28
2	1	72	34	0	32	28	1	25



ID	Time	sexe	age	Sat_oxygene	retrait	uranium
1	Pre-op	F	65	56	0	30
1	Post-op	F	65	51	0	28
2	Pre-op	M	72	34	0	32
2	Post-op	M	72	28	1	25