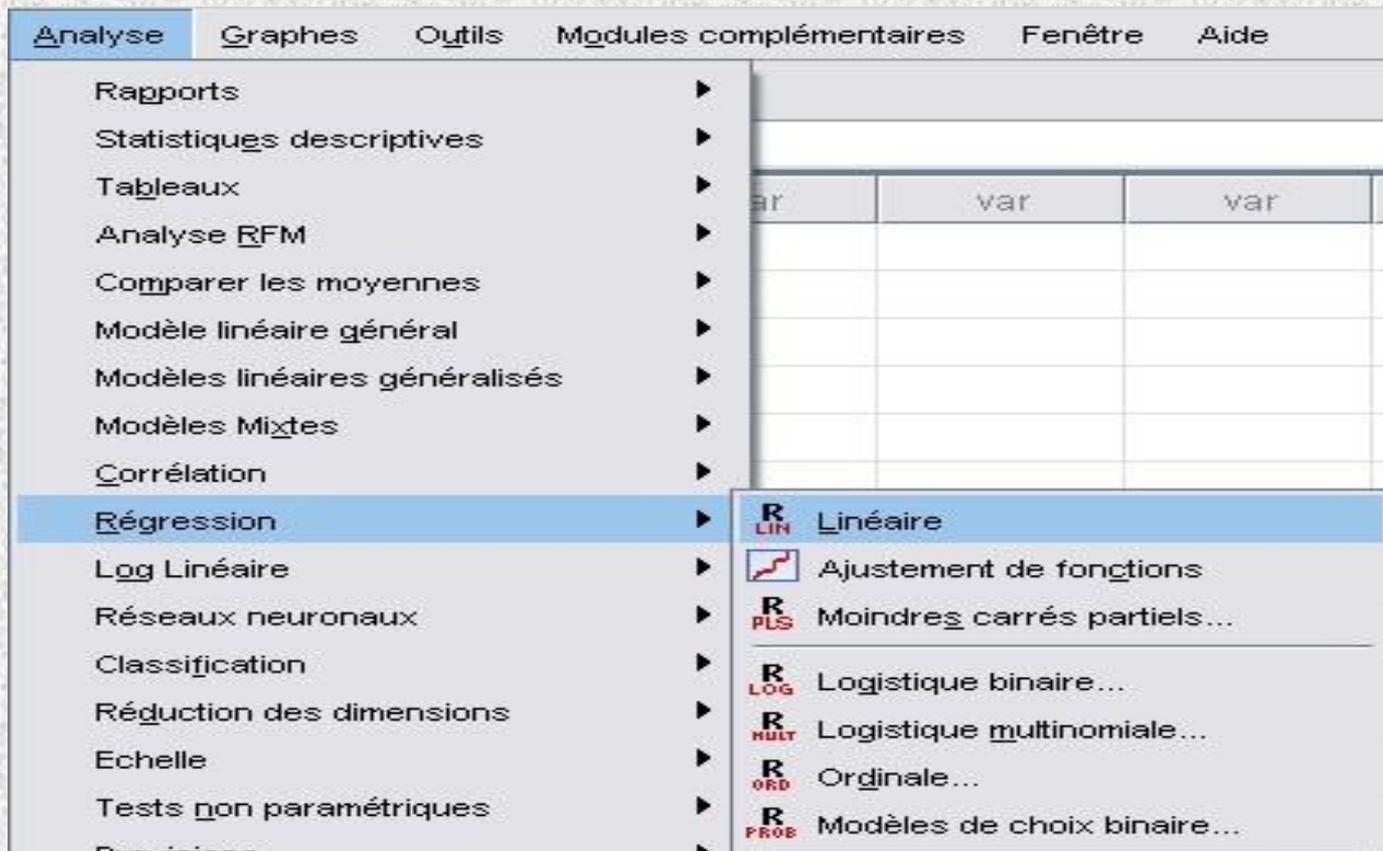


Régression linéaire multiple

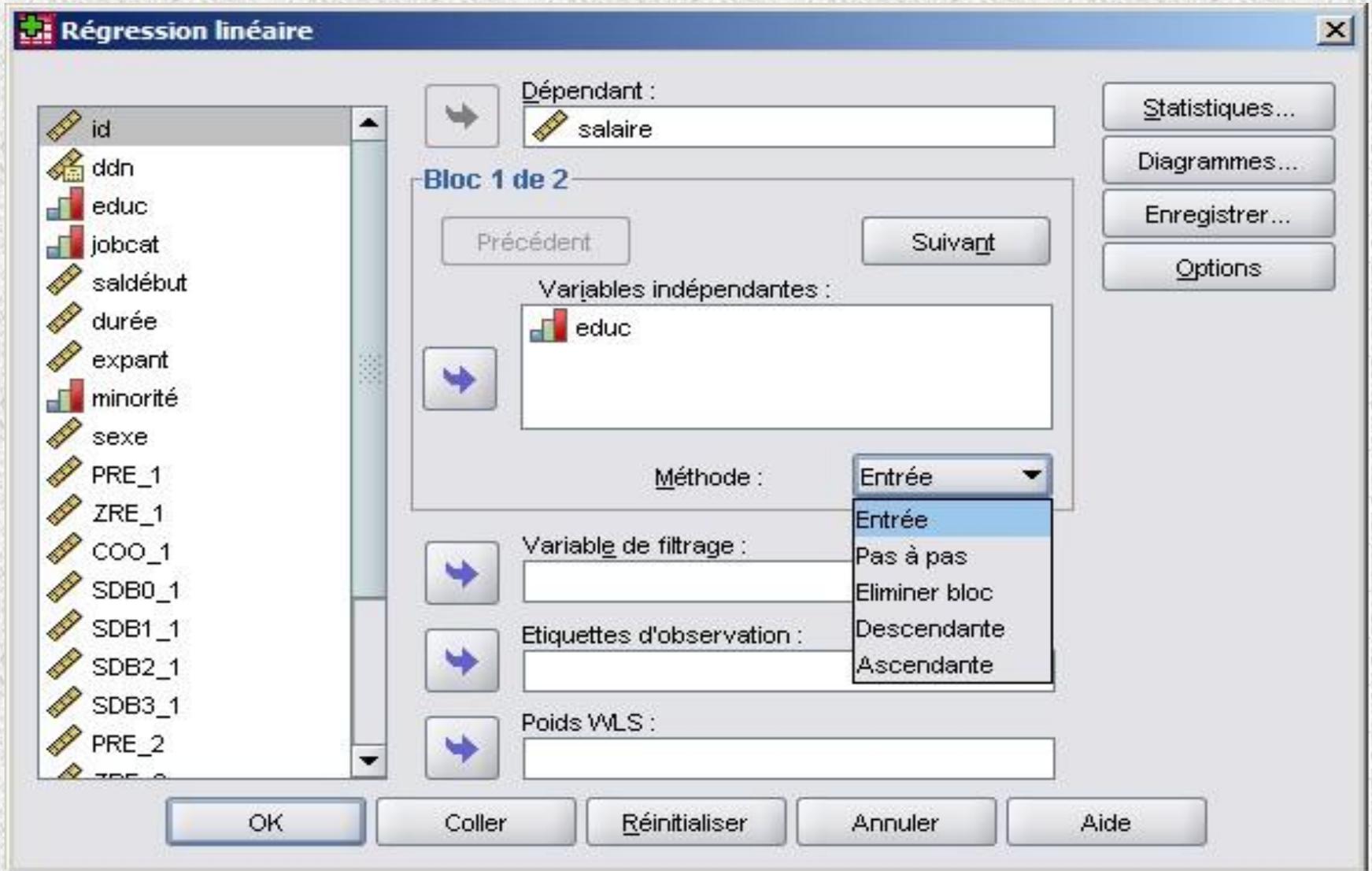
Application de la régression multiple sous SPSS

1. Pour réaliser l'analyse, cliquez sur **Analyse, Régression**, puis **Linéaire**.



2. En cliquant sur  , insérez la variable dépendante et la ou les variable(s)

indépendante(s) dans les boîtes appropriées.



3. Si vous désirez absolument que la première variable indépendante soit incluse, privilégiez la méthode **Entrée**.

4. Pour créer des blocs (groupes) de variable(s) indépendante(s) dans le cadre d'une régression hiérarchique, cliquez sur

A rectangular button with a light gray background and a thin border, containing the word "Suivant" in a dark gray font.

lorsque le premier bloc est construit, puis insérez les variables indépendantes des autres blocs en répétant cette procédure. La méthode de régression (Entrée, Pas à pas, etc.) peut être déterminée pour chaque bloc. Habituellement, la méthode **Entrée** est utilisée à moins d'a priori théoriques particuliers.

- id
- ddn
- educ
- jobcat
- saldébut
- durée
- expant
- minorité
- sexe
- PRE_1
- ZRE_1
- COO_1
- SDB0_1
- SDB1_1
- SDB2_1
- SDB3_1
- PRE_2
- ZRE_2

Dépendant :
salaire

Bloc 2 de 2

Précédent Suivant

Variables indépendantes :
sexe
durée

Méthode : Ascendante

Variable de filtrage :
Règle...

Etiquettes d'observation :

Poids WLS :

- Statistiques...
- Diagrammes...
- Enregistrer...
- Options

OK Coller Réinitialiser Annuler Aide

5. Vous pouvez choisir une variable de filtrage pour limiter l'analyse à un sous-échantillon formé par les participants ayant obtenu une ou des valeur(s) particulière(s) à cette même variable.
6. Vous pouvez aussi spécifier une variable qui permettra d'identifier les coordonnées sur le graphique (**Étiquettes d'observation**).
7. Enfin, vous pouvez choisir une variable numérique pondérée (**Poids WLS**) pour effectuer l'analyse des moindres carrés. Par cette analyse, les valeurs sont pondérées en fonction de leurs variances réciproques, ce qui implique que les observations avec de larges variances ont un impact moins important sur l'analyse que les observations associées à de petites variances.
8. Assurez-vous d'avoir sélectionné les options nécessaires (par exemple, sous le bouton Statistiques).
9. Pour procéder à l'analyse, cliquez sur 

Une présentation détaillée de toutes les options est disponible dans le procédurier de la régression simple.

Le bouton

Statistiques...

Pour la régression multiple, nous suggérons de cocher les options suivantes :

Régression linéaire : Statistiques

Coefficients de régression

- Estimations
- Intervalles de confiance
Niveau (%) :
- Matrice de covariance
- Qualité de l'ajustement
- Variation de R-deux
- Caractéristiques
- Mesure et corrélations partielles
- Tests de colinéarité

Résidus

- Durbin-Watson
- Diagnostic des observations
 - Points atypiques : écarts-types
 - Toutes les observations

Poursuivre Annuler Aide

L'encadré Coefficients

Estimations : valeurs b pour chaque VI et son test de signification

Intervalles de confiance : intervalle pour chaque coefficient dans la population

L'encadré Résidus

Durbin-Watson : évaluation de l'indépendance des erreurs

Diagnostic des observations : valeur de la VD observée, prédite, du résiduel et du résiduel standardisé pour chaque observation. Indique quelles observations ont un résiduel standardisé de plus de 2 ou 3 é.-t. (au choix de l'utilisateur)

Les autres statistiques

Qualité de l'ajustement : fournit le test pour évaluer l'ensemble du modèle (F), le R multiple, le R^2 correspondant et le R^2 ajusté

Variation de R-deux : changement du R^2 après l'ajout d'un nouveau bloc de VI

Caractéristiques : moyenne, é.-t. et N pour toutes les variables du modèle

Mesure et corrélations partielles :

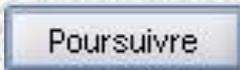
Corrélation entre chaque VI et la VD

Corrélation partielle entre chaque VI et VD en contrôlant pour les autres VI

Corrélation « partie » ou semi-partielle entre chaque VI et la variance non expliquée de la VD par les autres VI

Test de colinéarité : évaluation de la multicolinéarité dans le modèle (VIF).

Cliquez sur



pour revenir à la boîte de dialogue principale.

Le bouton

A rectangular button with a light gray background and a thin border. The text 'Diagrammes...' is centered on the button in a dark gray font.

Les graphiques offerts permettent de vérifier par un examen visuel les prémisses de la régression linéaire multiple. Celui croisant les valeurs prédites (*ZPRED) et résiduelles (*ZRESID) standardisées illustre le respect (ou le non respect) de la prémisses d'homogénéité (répartition aléatoire des points autour de 0) et de linéarité (tendance des points à se concentrer autour d'une ligne).

- DEPENDNT
- *ZPRED
- *ZRESID
- *DRESID
- *ADJPRED
- *SRESID
- *SDRESID

Dispersion 1 de 1

Précédent Suivant

Y : *ZRESID

X : *ZPRED

Diagrammes des résidus normalisé:

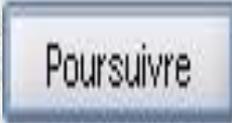
- Histogramme
- Diagramme de répartition gaussien

Générer tous les graphiques partiels

Poursuivre Annuler Aide

Pour faire plus d'un graphique, utilisez le bouton  .

L'encadré des diagrammes des résidus normalisés permet d'illustrer la distribution des résiduels (histogramme et diagrammes de répartition gaussiens), ce qui vous permet de faire un examen visuel du respect de la prémisse de normalité de la distribution des erreurs.

Cliquez sur  pour revenir à la boîte de dialogue principale.

Le bouton



Enregistrer...

Toutes les options disponibles dans ce menu permettent de créer des nouvelles variables ayant les valeurs calculées par le modèle. Il s'agit donc de choisir les variables diagnostiques permettant d'évaluer la qualité du modèle et celles qui permettent de détecter les variables ayant une importante influence sur le modèle. On choisira donc minimalement les résidus standardisés, mais on peut également ajouter les valeurs prédites non standardisées et standardisées (valeur de la VD calculée pour chaque observation) ainsi que la distance de Cook et les DfBêta(s) standardisés. Notez qu'en cochant des options dans la boîte de dialogue **Enregistrer**, vous allez obtenir un tableau de résultats de plus portant sur les statistiques des résidus et comprenant minimalement la moyenne, l'écart-type, les valeurs minimales et maximales ainsi que le N.

Prévisions

- Non standardisés
 Standardisés
 Ajustées
 Erreur standard prévision moyenne

Résidus

- Non standardisés
 Standardisés
 Studentisés
 Supprimées
 Supprimés studentisés

Distances

- Mahalanobis
 Cook
 Valeurs influentes

Intervalles de la prévision

- Moyenne Individuelle

Intervalle de confiance : %

Statistiques d'influence

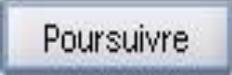
- DfBêta(s)
 DfBêta(s) standardisés
 Différence de prévision
 Dfprévision standardisée
 Rapport de covariance

Statistiques à coefficients

- Créer des statistiques à coefficients
 Créer un ensemble de données
Nom de l'ensemble de données :
 Ecriture d'un nouveau fichier de données

Exporter les informations du modèle dans un fichier XML

Inclure la matrice de covariance

Cliquez sur  pour revenir à la boîte de dialogue principale.

Le bouton 

La dernière fenêtre vous permet de déterminer les paramètres de sélection des méthodes d'entrée progressives (Ascendante ou descendante - *stepwise*). Vous pouvez utiliser la valeur de la probabilité associée à la valeur F (soit la valeur de p) ou encore la valeur de la statistique F pour introduire ou retirer des variables. Idéalement, vous conservez les valeurs par défaut à moins que vous ne vouliez que les critères d'entrée ou de retrait des variables de votre modèle soient plus sévères ou plus inclusifs.

 **Régression linéaire : Options** X

Paramètres des méthodes progressives

Choisir la probabilité de F
Entrée : Suppression :

Choisir la valeur de F
Entrée : Suppression :

Inclure terme constant dans l'équation

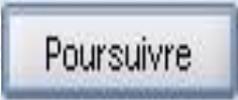
Valeurs manquantes

Exclure toute observation incomplète

Exclure seulement les composantes non valides

Remplacer par la moyenne

Évidemment, vous laissez aussi la constante dans l'équation. Vous pouvez finalement spécifier ce que vous désirez faire avec les valeurs manquantes. Encore une fois, l'option par défaut est à privilégier puisque le retrait de toute observation incomplète permet de conserver toujours le même nombre d'observations, ce qui favorise la cohérence du modèle.

Cliquez sur  pour revenir à la boîte de dialogue principale.