



# Introduction à SPSS

Atelier de formation A08  
Plateforme de recherche clinique et évaluative

[rechclinique@crchudequebec.ulaval.ca](mailto:rechclinique@crchudequebec.ulaval.ca)



# Objectifs d'apprentissage

1. Connaitre le fonctionnement de SPSS
2. Manipuler des fichiers de données
3. Créer des variables
4. Effectuer des analyses statistiques de base dans SPSS



# Plan de la présentation

1. Généralités: fenêtres, importation, exportation
2. Manipulation de données
3. Analyses descriptives et graphiques
4. Tests statistiques



# Téléchargement de SPSS

## Version d'évaluation

- Essai gratuit pendant 14 jours, version 25
- <https://www.ibm.com/ca-fr/products/spss-statistics>

## License avec l'Université Laval

- <https://www.dti.ulaval.ca/offre-de-services/logiciels>
- SPSS Premium 25
- 99\$ pour une licence annuelle pour étudiants, disponible à partir de la mi-décembre jusqu'au 31 décembre de l'année suivante (ou 39\$ par session)
- 180\$, 280\$ ou 349\$ selon le nombre de modules, pour les employés, jusqu'au 31 décembre



# Exemple de données

Excel: *DonnéesSPSS.xlsx*

Contexte: Décès chez des patients atteints de cancer

Variable	Valeurs
Décès	0: Vivant, 1: Mort
Sexe	Femme, Homme
Fumeur	0: Non Fumeur, 1: Occasionnel, 2: Régulier
Diabète	Oui, Non
IMC	Nombre
Diamètre	Nombre (mm)
Suivi	Durée de suivi jusqu'au décès ou la fin de suivi (semaines)
Id	Numéro d'identification



# Section 1: Généralités



# Fenêtres

- Données
  - Vue de données
  - Vue des variables
- Syntaxe
- Sortie
- Script



# Syntaxe

- Programmation à l'aide de code SPSS pour effectuer les opérations désirées et analyses
- Pas obligatoire; les menus permettent d'effectuer la majorité des opérations et analyses
  - Les opérations dans les menus sont traduites par de la syntaxe (affichées avant chaque résultat)
  - + d'options disponibles dans la syntaxe

*\* Dans l'atelier, les menus seront utilisés \**





# Saisie de données

1. Déterminer les caractéristiques des variables dans «Vue des variables»
2. Saisir les valeurs dans « Vue de données »



# Saisie de données

## Caractéristiques des variables

**Nom:** court, sans symbole, représentatif

**Type:** Chaîne (caractères), Numérique, Date, etc.

**Largeur:** Attention aux caractères pour ne pas les tronquer

**Décimales:** Contrôler l'affichage

**Libellé:** Identifier plus clairement la variable

**Valeurs:** Pour les variables codifiées

**Manquant:** Indiquer quels codes correspondent à une valeur manquante

**Mesure:** Échelle, Ordinale, Nominale -> Influence les opérations et analyses possibles sur la variable



# Importation de données

- Ouvrir -> Données
- Fichiers de plusieurs sources: .sav, .sas7bdat, .xlsx, .txt, etc.
- Excel: Sélection de la plage de données

*\* Toujours valider les caractéristiques des variables et faire des statistiques descriptives pour s'assurer que les données ont bien été lues\**



# Exportation de données

- Enregistrer sous
- Exporter vers la base de données: ajouter les données à un fichier déjà existant (nouvelles observations, variables ou feuilles)
- Fichiers de plusieurs sources: .sav, .sas7bdat, .xlsx, .txt, etc.
- Noms de variables + libellés



# Présentation de résultats

Dans la sortie:

- Table des matières
- Couper les sections inutiles
- Ajout de titres et de texte (Onglet « Insérer »)
- Alignement (Onglet « Format »)



# Exportation de résultats

- Copier/Coller les tableaux et graphiques: image, tableau, texte
- Exporter
  - Objets à exporter: Tout visible ou Sélectionnés
  - Format Word, PDF, Excel, HTML, etc.



## Section 2 : Manipulation de données



## Onglet «Données»

- Fusionner des fichiers
  - Ajouter des variables: Identifier les variables-clés
  - Ajouter des observations: Spécifier les variables à conserver et indiquer la provenance des variables sous forme de variable





## Onglet «Données»

- Scinder un fichier
  - Analyser toutes les observations (défaut)
  - Comparer les groupes (K résultats par tableau, selon une variable à K catégories)
  - Organiser la sortie par groupe (K tableaux de résultats, selon une variable à K catégories)



## Onglet «Données»

- Sélectionner des observations
  - Toutes les observations (défaut)
  - Selon une condition logique: Sous-échantillon ayant une caractéristique particulière
  - Par échantillonnage aléatoire: % ou nombre d'observations sélectionnées aléatoirement
  - Filtrer les observations non sélectionnées (variable indicatrice de sélection)
  - Copier les observations dans un nouvel ensemble (sous-échantillon enregistré ailleurs)
  - Supprimer les observations non sélectionnées (définitif!) <sup>17</sup>



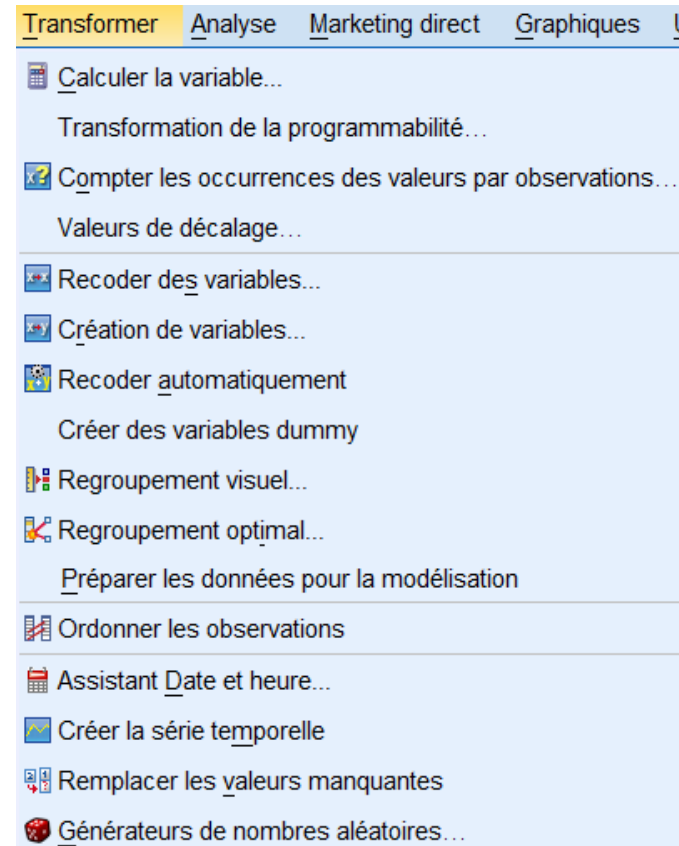
## Onglet «Données»

- Pondérer les observations
  - Ne pas pondérer les observations (défaut)
  - Pondérer les observations par:
    - Variable de pondération à sélectionner
    - Exemple: Données de sondage



# Onglet «Transformer»

- Calculer la variable
  - Opération mathématiques
  - Fonctions statistiques, dates, valeurs manquantes, arithmétique, etc.
    - Leurs arguments et définitions apparaissent lorsqu'on clique sur une fonction





## Onglet «Transformer»

- Recoder des variables
  - Regrouper des catégories
  - Créer une variable catégorique à partir d'une variable continue
  - « Anciennes et nouvelles valeurs »



## Onglet «Transformer»

- Création de variables
  - Semblable à « Recoder des variables », mais permet de ne pas écraser la variable originale

*\*Les trois dernières manipulations peuvent être effectuées conditionnellement*

*(Si: Condition facultative de sélection de l'observation)\**

*\*Valider les variables obtenues\**



# Section 3 :

## Analyses descriptives et graphiques



# Onglet « Analyse/ Statistiques descriptives »

- Fréquences
  - Tables de fréquences
  - Statistiques: Fractiles, Tendance centrale, Dispersion, Distribution
  - Graphiques: Barres, circulaire, histogramme

*\* Attention: Tous ces éléments peuvent être obtenus pour les variables qualitatives et quantitatives, ce qui ne fait pas toujours du sens! \**





# Onglet « Analyse/ Statistiques descriptives »

- Explorer
  - Statistiques descriptives pour les variables quantitatives
  - Sur toutes les valeurs ou selon des facteurs
  - Statistiques: Descriptives
  - Tracés: Boîtes à moustaches, Histogramme, Tracés de répartition gaussiens avec tests



# Onglet « Analyse / Statistiques descriptives »

- Tableaux croisés
  - Mettre en relation deux variables qualitatives
  - Spécifier les variables en Ligne(s) et Colonne(s)
  - Statistiques: Plusieurs tests (voir section 4)
  - Cellules: Effectifs, Pourcentage Position,
  - Graphiques à barres en cluster



## Onglet « Graphiques »

- Générateur de graphiques
  - Sélectionner le type de graphique désiré
  - Glisser les variables sur les axes du graphique
  - Modifier les propriétés du graphique
- Éditeur de graphiques
  - Double clic sur un graphique
  - Changer les couleurs, les formats, modifier la légende, ajouter du texte, etc.
  - Enregistrer le modèle de graphique






# Section 4 : Tests statistiques



## Onglet « Analyse »

- Statistiques descriptives: Tableaux de fréquences, Khi-deux, Fisher, Mc Nemar, mesures d'association et d'accord
- Comparer les moyennes: Tests T
- Modèle linéaire général: ANOVA
- Corrélation :Pearson et Spearman
- Régression: Linéaire et Logistique
- Tests non paramétriques: Wilcoxon, Rangs signés de Wilcoxon, Kruskal-Wallis
- Survie: Kaplan-Meier

Analyse	Marketing direct	Graphiques	U
Rapports			▶
Statistiques descriptives			▶
Tableaux			▶
Comparer les moyennes			▶
Modèle linéaire général			▶
Modèles linéaires généralisés			▶
Modèles Mixtes			▶
Corrélation			▶
Régression			▶
Log Linéaire			▶
Réseaux neuronaux			▶
Classification			▶
Réduction des dimensions			▶
Echelle			▶
Tests non paramétriques			▶
Prévisions			▶
Survie			▶
Réponses multiples			▶
 Analyse des valeurs manquantes			
Imputation multiple			▶
Echantillons complexes			▶
 Simulation...			
Contrôle de qualité			▶
 Courbe ROC...			



# Exemple 1

- Est-ce que la proportion de patients décédés est plus grande chez les diabétiques?
  - VD = Décès, VI = Diabète, Données non corrélées
  - Test -> Khi-deux ou Fisher
  - Rapport de risque ou de cotes pour quantifier l'association
- Est-ce que la proportion de patients diabétiques après 2 ans est la même qu'après 10 ans?
  - (Données non disponibles)
  - Deux variables de diabète par participant -> Données corrélées
  - Test -> Mc Nemar



# Onglet «Statistiques descriptives / Tableaux croisés»

- Khi-deux
  - Association entre 2 variables qualitatives
  - Affiche aussi le résultat du test de Fisher
- Mc Nemar
  - Comparaison de proportions appariées
- Risque
  - Rapport de risques, rapport de cotes
- Onglet « Exact »
  - Valeur p exacte (non asymptotique), utile lorsque faibles effectifs.
  - Très long lorsque beaucoup de catégories présentes.



Tableau croisé Diabète ^ Deces

		Deces		Total
		0	1	
Diabete Non	Effectif	71	16	87
	% dans Deces	89,9%	76,2%	87,0%
Oui	Effectif	8	5	13
	% dans Deces	10,1%	23,8%	13,0%
Total	Effectif	79	21	100
	% dans Deces	100,0%	100,0%	100,0%

Estimation du risque

	Valeur	Intervalle de confiance à 95%	
		Plus bas	Supérieur
Rapport des cotes pour Diabete (Non / Oui)	2,773	,801	9,604
Pour cohorte Deces = 0	1,326	,853	2,062
Pour cohorte Deces = 1	,478	,211	1,083
N d'observations valides	100		

*Attention à l'ordre des modalités utilisé pour calculer les rapports*

Tests du khi-deux

	Valeur	ddl	Sig. approx. (bilatérale)	Sig. exacte (bilatérale)	Sig. exacte (unilatérale)
khi-deux de Pearson	2,746 <sup>a</sup>	1	,097	,139	,102
Correction pour continuité <sup>b</sup>	1,670	1	,196		
Rapport de vraisemblance	2,423	1	,120	,139	,102
Test exact de Fisher				,139	,102
N d'observations valides	100				

a. 1 cellules (25,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 2,73

b. Calculée uniquement pour une table 2x2



*Le test du Khi-Deux asymptotique suppose que l'effectif théorique de chaque cellule est supérieur à 5.*





## Exemple 2

- Est-ce que le diamètre des tumeurs des patients décédés diffère de celui des patients encore vivants?
  - VD = Diamètre (continue), VI = décès (dichotomique)
  - Données non corrélées
  - Test -> Test T (ou son équivalent non-paramétrique)
- Est-ce que le diamètre des tumeurs lors de leur découverte diffère de celui 6 mois plus tard?
  - (Données non disponibles)
  - Deux variables de diamètre (continues) par participant -> Données corrélées
  - Test -> Test T Paired (ou son équivalent non-paramétrique)



# Onglet «Comparer les moyennes»

- Test T pour échantillons indépendants
  - Variable(s) à tester: VD continue
  - Variable de regroupement: VI dichotomique
- Test T pour échantillons appariés
  - Chaque mesure appariée doit être dans une variable différente (une ligne par patient, pas une ligne par mesure)



### Statistiques de groupe

Décès		N	Moyenne	Ecart-type	Erreur standard moyenne
Diamètre	Vivant	79	8,9063	5,70439	,64179
	Mort	21	13,9238	7,19471	1,57002

### Test d'échantillons indépendants

		Test de Levene sur l'égalité des variances		Test-t pour égalité des moyennes				
		F	Sig.	t	ddl	Sig. (bilatérale)	Différence moyenne	Différence écart-type
Diamètre	Hypothèse de variances égales	3,698	,057	-3,384	98	,001	-5,01748	1,48253
	Hypothèse de variances inégales			-2,958	27,049	,006	-5,01748	1,69613

 *Le test T suppose aussi l'homogénéité des variances et la normalité des résidus*

Intervalle de confiance 95% de la différence	
Inférieure	Supérieure
-7,95952	-2,07544
-8,49735	-1,53761



## Exemple 3

- Est-ce que le diamètre des tumeurs diffère selon la consommation de tabac?
  - VD = Diamètre (continue)
  - VI = Fumeur (qualitative)
  - Données non corrélées
  - Test -> ANOVA (analyse de la variance, ou son équivalent non paramétrique)



# Onglet «Modèle linéaire général»

- Univarié -> Permet d'effectuer une ANOVA
  - Variable dépendante: VD continue
  - Facteurs Fixes: VI qualitative
  - Post Hoc: Comparaisons multiples
  - Enregistrer (pour validation de modèle):  
Prévisions non standardisés, Résidus de Student
  - Options: Afficher les moyennes pour la VI,  
Estimations d'effet de taille, (pour validation de  
modèle: Tests d'homogénéité, tracé résiduel), etc.



### Tests des effets inter-sujets

Variable dépendante: Diamètre

Source	Somme des carrés de type III	ddl	Moyenne des carrés	D	Sig.	Eta au carré partiel
Modèle corrigé	123,629 <sup>a</sup>	2	61,815	1,550	,217	,031
Ordonnée à l'origine	5140,259	1	5140,259	128,924	,000	,571
Fumeur	123,629	2	61,815	1,550	,217	,031
Erreur	3867,431	97	39,870			
Total	13911,220	100				
Total corrigé	3991,060	99				

Estimations

Variable dépendante: Diamètre

Fumeur	Moyenne	Erreur standard	Intervalle de confiance à 95%	
			Borne inférieure	Limite supérieure
Non Fumeur	9,530	,706	8,129	10,931
Occasionnel	10,255	1,904	6,476	14,033
Régulier	13,422	2,105	9,245	17,600

 *L'ANOVA suppose aussi l'homogénéité des variances et la normalité des résidus*

### Comparaisons multiples

Variable dépendante: Diamètre

Bonferroni

(I) Fumeur	(J) Fumeur	Différence des moyennes (I-J)	Erreur standard	Sig.	Intervalle de confiance à 95%	
					Borne inférieure	Limite supérieure
Non Fumeur	Occasionnel	-,7245	2,03051	1,000	-5,6713	4,2222
	Régulier	-3,8922	2,22001	,248	-9,3006	1,5162
Occasionnel	Non Fumeur	,7245	2,03051	1,000	-4,2222	5,6713
	Régulier	-3,1677	2,83807	,801	-10,0818	3,7464
Régulier	Non Fumeur	3,8922	2,22001	,248	-1,5162	9,3006
	Occasionnel	3,1677	2,83807	,801	-3,7464	10,0818



## Exemple 4

- Est-ce qu'il y a une association entre l'IMC et le diamètre de la tumeur?
  - IMC + Diamètre = continues
  - Test -> Corrélacion
- Est-ce que le suivi est plus court pour les grosses tumeurs?
  - VD = Suivi, VI = Diamètre (continues)
  - Test -> Corrélacion (ou régression)



# Onglet «Corrélation»

- Bivariée
  - Pearson (linéaire) et Spearman (monotone, basée sur les rangs)





### Corrélations

			IMC	Diamètre	Suivi
Rho de Spearman	IMC	Coefficient de corrélation	1,000	-,116	,038
		Sig. (bilatérale)	.	,249	,706
		N	100	100	100
	Diamètre	Coefficient de corrélation	-,116	1,000	-,193
		Sig. (bilatérale)	,249	.	,055
		N	100	100	100
	Suivi	Coefficient de corrélation	,038	-,193	1,000
		Sig. (bilatérale)	,706	,055	.
		N	100	100	100



## Exemple 5

- Est-ce que l'IMC permet de prédire le diamètre de la tumeur?
  - VD = Diamètre (continue)
  - VI = IMC (continue)
  - Test -> Régression (ou corrélation)



# Onglet «Régression»

- Linéaire
  - Dépendant: VD continue
  - Variables indépendantes: VI continue(s)
  - Méthode: Toutes les variables (entrez) ou sélection (étape par étape, précédent, suivant)
  - Statistiques: Estimations des coefficients, Qualité de l'ajustement ( $R^2$ ,  $R^2_a$ ), Tests de colinéarité
  - Tracés: Validation du modèle (résidus)
  - Enregistrer: Validation du modèle (résidus et autres mesures d'influence)



### Récapitulatif des modèles

Modèle	R	R-deux	R-deux ajusté	Erreur standard de l'estimation
1	,166 <sup>a</sup>	,027	,017	6,29358

a. Valeurs prédites : (constantes), IMC

### ANOVA<sup>a</sup>

Modèle		Somme des carrés	ddl	Moyenne des carrés	D	Sig.
1	Régression	109,363	1	109,363	2,761	,100 <sup>b</sup>
	Résidu	3881,697	98	39,609		
	Total	3991,060	99			

a. Variable dépendante : Diamètre

b. Valeurs prédites : (constantes), IMC

### Coefficients<sup>a</sup>

Modèle		Coefficients non standardisés		Coefficients standardisés	t	Sig.
		A	Erreur standard	Bêta		
1	(Constante)	14,762	2,958		4,991	,000
	IMC	-,177	,107	-,166	-1,662	,100

a. Variable dépendante : Diamètre



*La régression suppose l'indépendance des observations, la linéarité entre VD et VI, ainsi que l'homogénéité des variances et la normalité des résidus*



## Exemple 6

- Est-ce que le diamètre, le diabète et la consommation de tabac permettent de prédire le décès des patients?
  - VD = Décès (dichotomique)
  - VI = Diamètre, Diabète, Tabac (Continue + Nominale)
  - Analyse multivariée
  - Test -> Régression logistique



# Onglet «Régression»

- Logistique binaire
  - Dépendant: VD dichotomique
  - Covariables: VI qualitative(s) ou quantitative(s)
  - Méthode: Toutes les variables (entrez) ou sélection (↓,↑)
  - Nominale: Définir la modalité de référence pour les VI nominales
  - Enregistrer: Validation du modèle
  - Options: Hosmer-Lemeshow, CI pour  $\exp(B)$ , tracés de classement



### Codage de variables dépendantes

Valeur d'origine	Valeur interne
Vivant	0
Mort	1

### Tests de spécification du modèle

		Khi-Chi-deux	ddl	Sig.
Etape 1	Etape	15,417	4	,004
	Bloc	15,417	4	,004
	Modèle	15,417	4	,004

### Codages des variables nominales

		Fréquence	Codage des paramètres	
			(1)	(2)
Fumeur	0	80	,000	,000
	1	11	1,000	,000
	2	9	,000	1,000
Diabete	Non	87	,000	
	Oui	13	1,000	

### Variables dans l'équation

		A	E.S.	Wald	ddl	Sig.	Exp(B)	IC pour Exp(B) 95%	
								Inférieur	Supérieur
Etape 1 <sup>a</sup>	Diabete(1)	1,469	,705	4,341	1	,037	4,344	1,091	17,294
	Diametre	,127	,042	9,104	1	,003	1,136	1,046	1,234
	Fumeur			1,872	2	,392			
	Fumeur(1)	-,505	,959	,278	1	,598	,603	,092	3,953
	Fumeur(2)	,946	,794	1,418	1	,234	2,576	,543	12,221
	Constante	-3,070	,633	23,520	1	,000	,046		



*La régression logistique suppose aussi l'indépendance des observations et la linéarité entre  $\text{logit}(p)$  et  $V_i$*



## Exemple 7

- Est-ce que le diamètre des tumeurs des patients décédés diffère de celui des patients encore vivants?
- Est-ce que le diamètre des tumeurs diffère selon la consommation de tabac?
  - Comme les exemples 2 et 3
  - Alternatives non paramétriques:
    - Mann-Whitney (Wilcoxon) -> Test T
    - Wilcoxon en séries appariées -> Test T Pairé
    - Kruskal-Wallis -> ANOVA





# Onglet «Tests non paramétriques»

- Échantillons indépendants
  - Champs
    - Champs de test: VD continue
    - Groupes: VI dichotomique ou nominale
  - Paramètres
    - Personnaliser les tests
    - U de Mann-Whitney (2 échantillons) si VI dichotomique (Équivalent du test de Wilcoxon)
    - Kruskal-Wallis ANOVA à un facteur (k échantillons) si VI nominale, Comparaisons multiples “Tous par paire”



**Récapitulatif du test d'hypothèse**

	Hypothèse nulle	Test	Sig.	Décision
1	La distribution de Diamètre est identique sur les catégories de Décès.	Test U de Mann-Whitney à échantillons associés	,003	Rejeter l'hypothèse nulle.

Les significations asymptotiques sont affichées. Le niveau de signification est ,05.

**Récapitulatif du test d'hypothèse**

	Hypothèse nulle	Test	Sig.	Décision
1	La distribution de Diamètre est identique sur les catégories de Fumeur.	Test de Kruskal-Wallis à échantillons indépendants	,237	Retenir l'hypothèse nulle.

Les significations asymptotiques sont affichées. Le niveau de signification est ,05.



*Ces tests non-paramétriques supposent l'homogénéité des variances, même si une distribution normale n'est pas requise*



# Onglet «Tests non paramétriques»

- Échantillons liés
  - Champs
    - Champs de test: Chaque mesure appariée (continue) doit être dans une variable différente (une ligne par patient, pas une ligne par mesure)
  - Paramètres
    - Personnaliser les tests
    - Wilcoxon en séries appariées (2 échantillons) -> Rangs signés de Wilcoxon



## Exemple 8

- À quelle vitesse les décès surviennent-ils?
- Est-ce que la survie est affectée par la consommation de tabac?
  - VD = Suivi (continue) + Décès (dichotomique)
  - VI = Tabac (nominale)
  - Test -> Courbe de Kaplan-Meier + test du Log Rank



## Onglet «Survie»

- Kaplan-Meier
  - Heure: VD continue = Temps de suivi
  - Statut: VD dichotomique = Évènement / censure
  - Critère: VI nominale (facultative)
    - Comparer les facteurs: Log Rank
  - Options:
    - Tracés: Survie

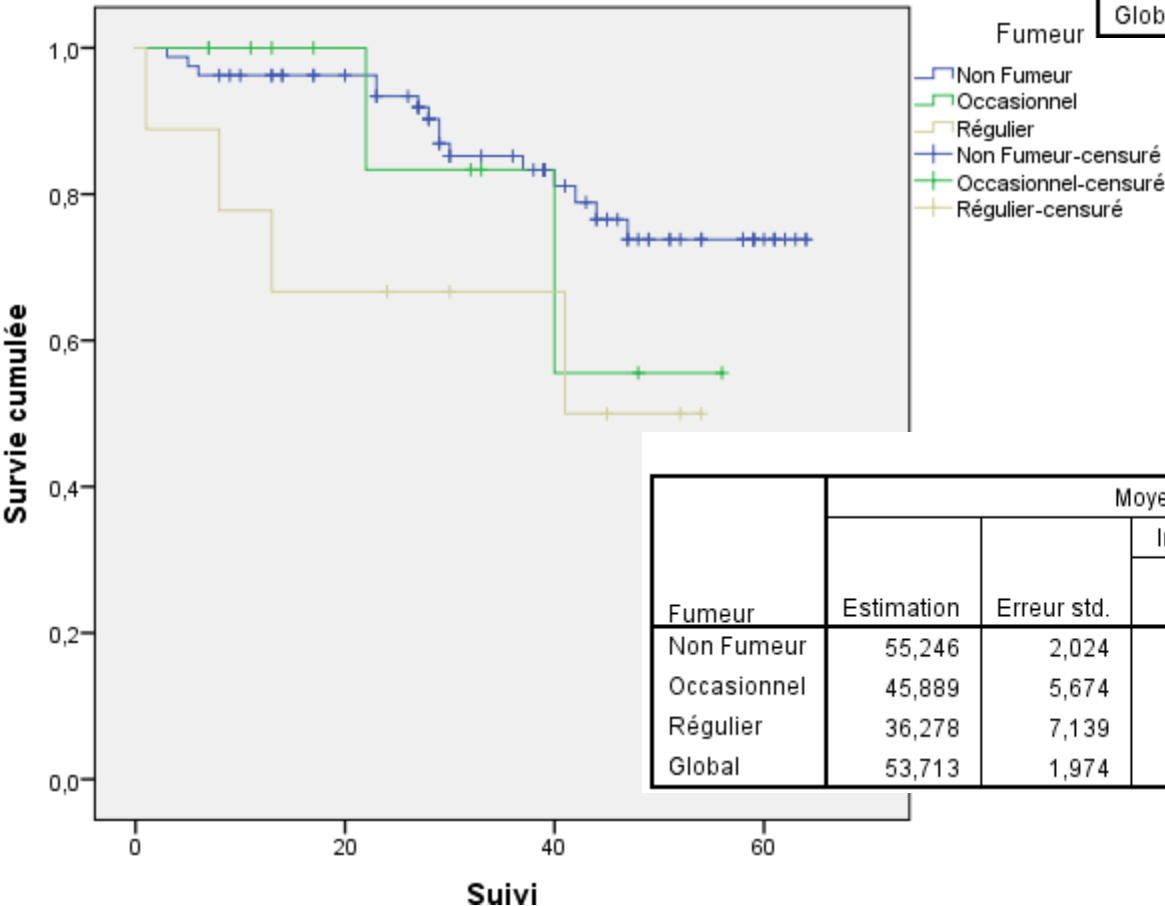


## Recherche clinique et évaluative

### Récapitulatif de traitement des observations

Fumeur	N total	Nombre d'événements	Censuré	
			N :	Pourcentage :
Non Fumeur	80	15	65	81,3%
Occasionnel	11	2	9	81,8%
Régulier	9	4	5	55,6%
Global	100	21	79	79,0%

### Fonctions de survie



### Comparaisons globales

	Khi-deux	ddl	Sig.
Log Rank (Mantel-Cox)	3,772	2	,152

Test d'égalité des distributions de survie pour les différents niveaux de Fumeur.

### Moyennes et médianes du délai de survie

Fumeur	Moyenne <sup>a</sup>				Médiane		
	Estimation	Erreur std.	Intervalle de confiance à 95 %		Estimation	Erreur std.	Interv
			Limite inférieure	Limite supérieure			Lir infé
Non Fumeur	55,246	2,024	51,279	59,213	.	.	.
Occasionnel	45,889	5,674	34,768	57,009	.	.	.
Régulier	36,278	7,139	22,286	50,270	41,000	.	.
Global	53,713	1,974	49,845	57,582	.	.	.



# Annexe: Validation des modèles

- Plus de détails sur la validation de modèles dans l'atelier 9
- Éléments pouvant invalider un modèle
  - Données influentes
  - Données manquantes
  - Multicollinéarité (analyses multivariées)
- Postulats sur lesquels des modèles sont basés
  - Homogénéité des variances des résidus
  - Normalité des résidus
  - Linéarité entre VD et VI
  - Indépendance des observations
- Pouvoir prédictif du modèle



# Annexe: Validation des modèles

- Données influentes
  - Examiner la distribution des résidus studentisés et autres mesures d'influence: distance de Cook, DfFits, DfBetas, CovRatio, etc.
- Données manquantes
  - Mécanismes MCAR, MAR, MNAR
- Multicolinéarité
  - Présence si  $VIF > 10$ , Index de condition  $> 30$  et Proportions de variance  $> 60\%$





# Annexe: Validation des modèles

- Homogénéité des variances des résidus
  - Test de Levene ou Bartlett
  - Graphique résidus studentisés \* valeurs prédites
- Normalité des résidus
  - Coefficients d'asymétrie + aplatissement
  - Histogramme, QQ Plot, Diagramme à moustaches
  - Test de Shapiro-Wilk, Kolmogorov-Smirnov



# Annexe: Validation des modèles

- Linéarité
  - Graphique résidus \* valeurs prédites
- Indépendance
  - Graphique résidus \* numéro d'observation
  - Durbin-Watson  $\approx 2$
- Pouvoir prédictif
  - $R^2$ ,  $R^2$  ajusté
  - Courbe ROC, Tableau de classement



*Merci pour votre attention!*

*Des questions?*