

SOCIO532.2  
STATISTIQUES & INFORMATIQUE APPLIQUÉES AUX SCIENCES SOCIALES  
© El Hadj Touré, Ph.D.

LABO SPSS #5

Analyse de tableaux multivariés

Les statistiques bivariées permettent de répondre à des questions relationnelles en mettant en évidence une relation entre deux variables. Toutefois, « une association statistique ne peut être interprétée, sans de grandes précautions, comme une relation causale » (Boudon, 2001: 29). Une fois qu'une relation est établie, les chercheurs considèrent communément le rôle d'une troisième variable (Lazarsfeld, 1955). En apportant de l'information additionnelle, les statistiques multivariées aident à clarifier la nature d'une relation bivariée de façon à dégager le modèle causal dans lequel elle s'insère. Pour conclure de façon convaincante à un lien de causalité, il faut s'assurer qu'il n'est pas dû à une variable tierce et l'inscrire dans un raisonnement théorique.

Dans le but d'examiner une relation entre deux variables en incluant une variable contrôle (VC), l'analyse de tableaux multivariés (ATM) constitue une première technique adéquate à cet effet. Mécaniquement, l'**élaboration de tableau** consiste à scinder l'échantillon en deux groupes ou plus définis par la variable-contrôle, et à analyser la relation entre une variable indépendante (VI) et la variable dépendante (VD) au sein de chacun des groupes. On obtient ainsi des **tableaux partiels**. Des tests du chi-deux et des mesures d'association (dont le gamma et le gamma partiel) permettent d'approfondir l'analyse tabulaire multivariée aux fins de diagnostic causal.

La mécanique est identique pour les modèles d'élaboration de tableau que sont la spécification, la reproduction, l'explication, l'interprétation et la dissimulation. Il est facile de mettre en pratique ces modèles à l'aide de SPSS. Au préalable, il faut toujours identifier la VI, la VD et la VC. Précisément, une des variables est identifiée comme le facteur explicatif (X), l'autre comme l'effet présumé (Y) que l'on cherche à modéliser, et une troisième comme la variable que l'on garde constante (Z) ou pour laquelle on tient compte des différences. Cette identification des variables et de leur positionnement dans la chaîne causale est d'autant plus importante conceptuellement qu'elle détermine la façon dont le modèle causal est interprété en vue de tirer une conclusion appropriée. Je veux donc attirer votre attention sur l'importance des modèles conceptuels dans l'analyse statistique. En particulier, il est important de préciser la façon dont la VC est conceptualisée : antécédente confondante, intermédiaire, antécédente modératrice, dissimulatrice.

Dans ce labo, nous allons apprendre au moyen d'une démonstration et d'exercices pratiques, à produire un **tableau bivarié d'ordre zéro** et surtout des **tableaux partiels d'ordre un**, déterminer la signification statistique des relations bivariées et partielles à l'aide des tests du chi-carré, déterminer la signification réelle d'une relation multivariée

à l'aide des Gamma et V de Cramer, et schématiser un modèle initial et un modèle causal en y inscrivant les résultats et les interpréter de façon appropriée.

## 1. La spécification (modération) : modèle typique de l'élaboration de tableau

La modération (causalité interactive) est le modèle par excellence de l'analyse de tableaux multivariés. Lorsque la relation bivariée initiale se révèle seulement dans un tableau partiel (un niveau de la VC) et s'annule dans un autre tableau partiel (un autre niveau de la VC), on parle de modération ou spécification. La **spécification** est un terme technique approprié, puisque la variable tierce (VC) spécifie ainsi à quel niveau se trouve une relation de départ entre deux variables. La relation initiale devient donc une **relation modérée, spécifiée** par la VC. En revanche, lorsque la relation bivariée initiale se révèle dans tous les tableaux partiels (tous les niveaux de la VC), il y a **reproduction**, ou absence de spécification, puisque chacun des tableaux partiels reproduit la relation de départ. Dans ce cas de figure, la relation initiale devient une **relation véritable ou directe**. Reproduction et relation véritable sont des termes techniques appropriés pour décrire cette situation d'absence de spécification.

### 1.0. Étapes de l'élaboration de tableau pour un modèle de spécification

#### Étape 1 : Analyse du tableau bivarié d'ordre zéro

L'association entre la VI (X) et la VD (Y) est-elle statistiquement significative ? Quelle est la force et quelle est la direction éventuelle de cette relation initiale ou primitive (X-Y) ? Au départ, la relation initiale est significative statistiquement et réellement !

#### Étape 2 : Analyse des tableaux partiels pour apporter plus d'informations

L'association entre la VI (X) et la VD (Y) est-elle statistiquement significative au sein de chaque niveau de la variable-contrôle (Z) modératrice ? Quelle est la force et la direction éventuelle de la relation initiale contrôlée pour chaque niveau de la variable-contrôle Z (Z1 : X-Y; Z2 : X-Y)? Ou plutôt la signification statistique (chi-carré) et la signification réelle (gamma) sont-elles variables d'un tableau partiel à l'autre ?

#### Étape 3 : Schématisation des modèles initial et causal aux fins de diagnostic causal

Après avoir schématisé les résultats, dites quel diagnostic causal est révélé par l'analyse des schémas ? Spécification ou non !

Résumé de l'élaboration de tableau (Fox: 311)			
Variable-contrôle VC	Relation VI-VD primitive?	Relation VI-VD contrôlant l'effet de VC (tableaux partiels)	Conclusion à tirer
Antécédente ou plutôt modératrice	Oui	G varie d'un tableau à l'autre	La relation VI-VD existe pour une valeur de VC (spécification)

Légende : G= Gamma pour chacun des tableaux partiels

NB : Pour inscrire les résultats statistiques sur les deux schémas, une façon consiste à placer les mesures d'association (Gamma ou V de Cramer) sur les flèches, mettre des astérisques sur les Gammas ou V de Cramer si les tests du chi-carré sont significatifs (par exemple \* $p < 0,05$ ; \*\* $p < 0,01$ ; \*\*\* $p < 0,001$ ). Si le test n'est pas significatif, la flèche est barrée.

## 1.1. Analyse bivariée de la relation initiale ou primitive (X-Y)

Intéressons-nous à la relation entre l'intérêt pour la politique (interet2) et l'intention de vote à la prochaine élection présidentielle (vote) chez les étudiants.

Intérêt pour la politique (X) → Intention de vote (Y)

La question de recherche est la suivante : *Y a-t-il une relation d'association entre l'intérêt pour la politique et l'intention de vote chez les étudiants de L2 inscrits en sociologie à l'UGB ?* On peut penser que plus les gens s'intéressent à la politique, plus ils sont susceptibles de participer aux élections. Pour vérifier cette hypothèse de travail, construisons un tableau croisé bivarié, procédons au test du chi-carré et calculons une ou des mesures d'association appropriées.

Grâce au menu « Utilitaires » (variables) et surtout en sortant les fréquences et pourcentages, nous pouvons voir que les deux variables sont catégorielles et leur mise en relation se justifie en utilisant les techniques susmentionnées. Puisque tout semble correct, passons à l'analyse bivariée descriptive et inférentielle.

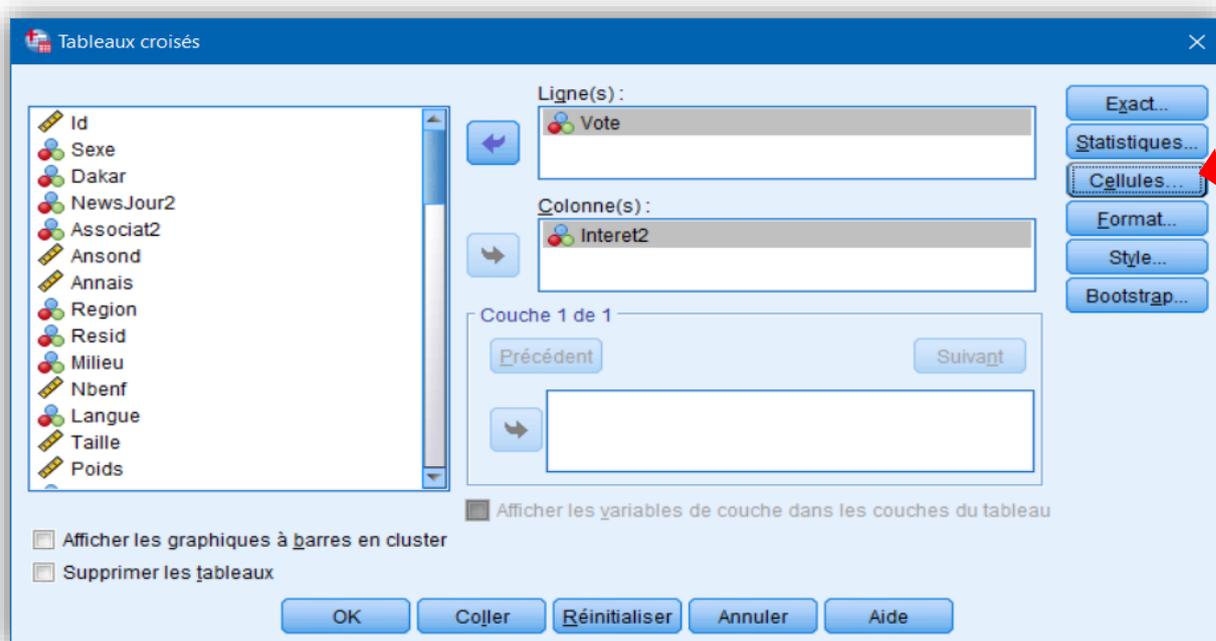
### Analyse

#### Statistiques descriptives

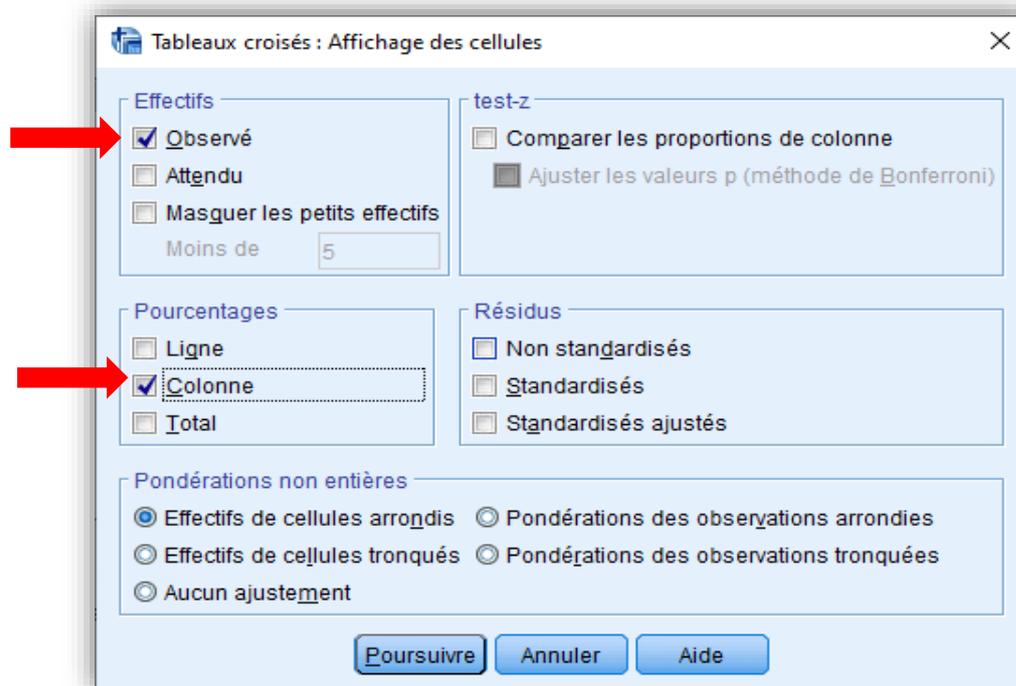
#### Tableaux croisés



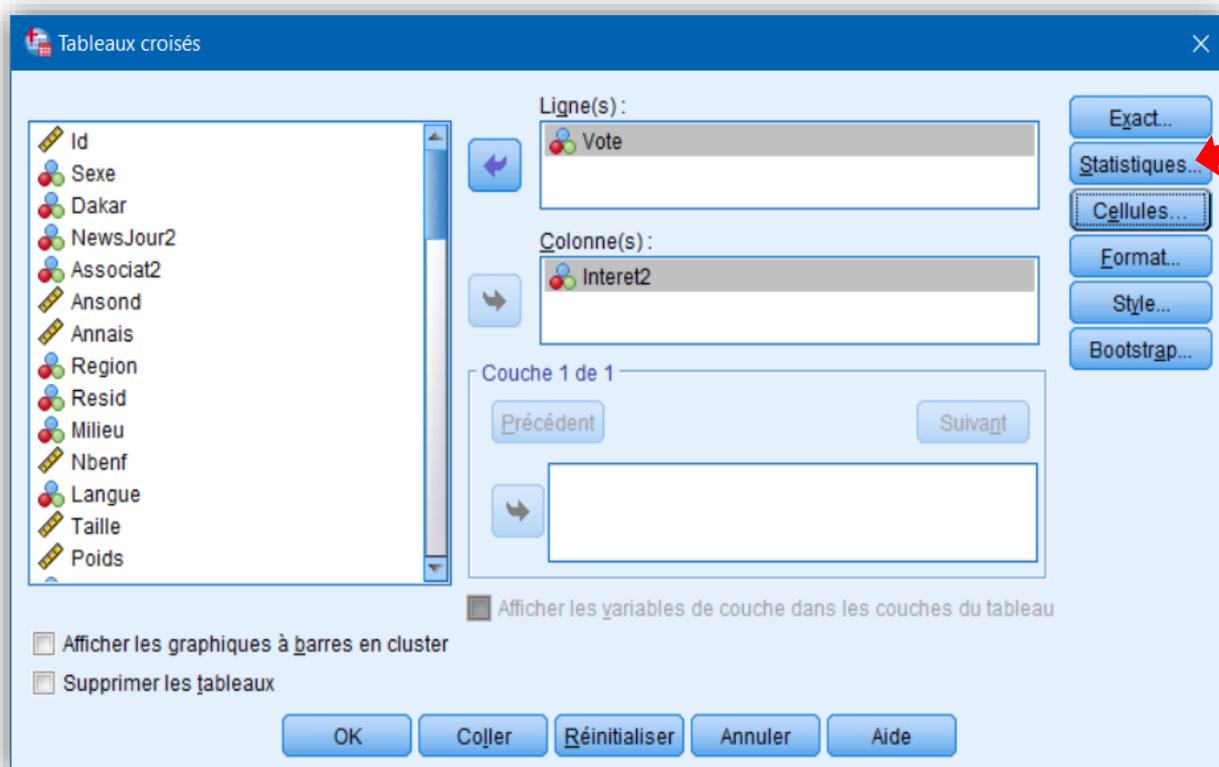
Cliquez sur la commande « **Tableaux croisés** » ! Sélectionnez la VD que vous souhaitez soumettre à l'analyse (vote) et faites-la passer dans le rectangle Lignes(s). Faites de même avec la VI (interet2) de sorte qu'elle se loge, cette fois-ci, dans le rectangle Colonne(s). La capture d'écran ci-dessous en donne l'illustration.



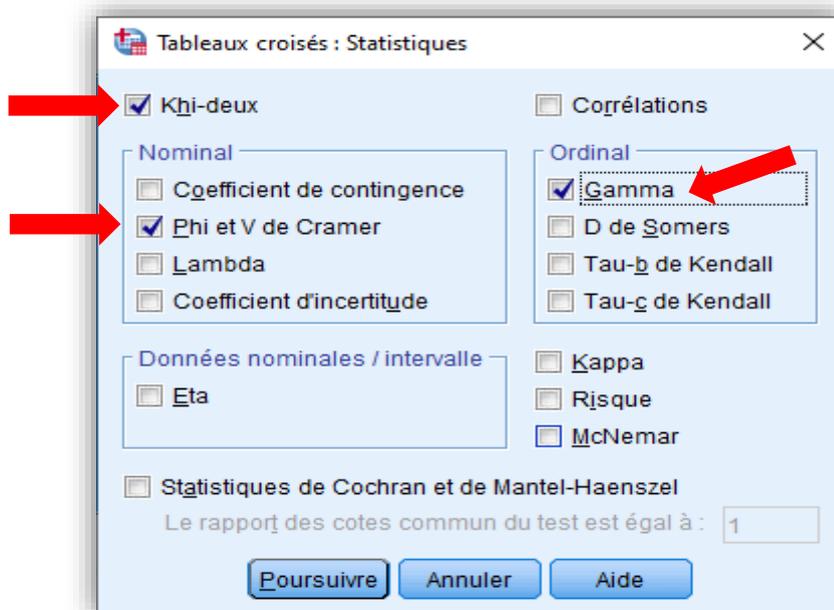
Cliquez sur l'option **Cellules** pour sélectionner les informations que vous désirez voir apparaître dans les cellules du tableau croisé. Un écran s'affiche!



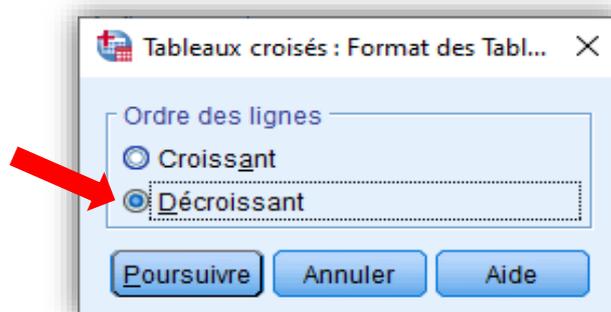
Poursuivez pour revenir à l'écran principal de l'analyse de tableaux croisés.



Cliquez sur l'icône « Statistiques » pour sélectionner le test d'hypothèse « Khi-deux », et la mesure d'association V de Cramer. Puisque les deux variables sont dichotomiques et codées 0/1, il est possible de les conceptualiser comme ordinales, de telle sorte qu'on peut aussi sortir le Gamma.



Poursuivez et cliquez sur Format pour sélectionner Décroissant.



Poursuivez et validez le tout pour obtenir les résultats.

**Tableau croisé Vote Seriez-vous oui ou non prêt(e) à voter lors de la prochaine élection présidentielle ? \* Interet2 Intérêt pour la politique**

			Interet2 Intérêt pour la politique		Total
			,00 Non	1,00 Oui	
Vote Seriez-vous oui ou non prêt(e) à voter lors de la prochaine élection présidentielle ?	1,00 Oui	Effectif	19	65	84
		% dans Interet2 Intérêt pour la politique	63,3%	91,5%	83,2%
	,00 Non	Effectif	11	6	17
		% dans Interet2 Intérêt pour la politique	36,7%	8,5%	16,8%
Total	Effectif		30	71	101
	% dans Interet2 Intérêt pour la politique		100,0%	100,0%	100,0%

	Valeur	ddl	Signification asymptotique (bilatérale)	Sig. exacte (bilatérale)	Sig. exacte (unilatérale)
khi-deux de Pearson	11,994 <sup>a</sup>	1	,001		
Correction pour continuité <sup>b</sup>	10,063	1	,002		
Rapport de vraisemblance	10,989	1	,001		
Test exact de Fisher				,001	,001
Association linéaire par linéaire	11,875	1	,001		
N d'observations valides	101				

a. 0 cellules (0,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 5,05.  
b. Calculée uniquement pour une table 2x2

		Valeur	Erreur asymptotique standard <sup>a</sup>	T approximatif <sup>b</sup>	Signification approximative
Nominal par Nominal	Phi	,345			,001
	V de Cramer	,345			,001
Ordinal par Ordinal	Gamma	,725	,135	2,902	,004
N d'observations valides		101			

a. L'hypothèse nulle n'étant pas considérée.  
b. Utilisation de l'erreur asymptotique standard en envisageant l'hypothèse nulle.

Récapitulons les résultats dans un tableau synthétique, qui inclut les informations clés de l'analyse tabulaire bivariable, du test du chi-carré et des mesures d'association.

**Tableau 1.1.** *Intention de vote à la prochaine élection présidentielle selon l'intérêt pour la politique chez les étudiants*

Intention de vote	Intérêt pour la politique	
	Non	Oui
Oui	63	92
Non	37	8
Cas (n)	(30)	(71)
Chi-carré	11.99***	
V de Cramer	0.35***	
Gamma	0.73**	

Notes. Les entrées correspondent à des % (n=101). \*\*\*p<0,001 ; \*\*p<0,01.  
Source. Sondage\_EtudiantsL2Socio\_2021.

**Interprétation statistique (analyse succincte) :** L'analyse tabulaire bivariable montre que 92% des étudiants intéressés par la politique ont l'intention de voter lors de la prochaine élection présidentielle, comparativement à seulement 63% des étudiants

non-intéressés par la politique. Il existe une différence de 29 points en pourcentages. La relation entre l'intérêt pour la politique et l'intention de vote est statistiquement significative dans la population étudiante étudiée ( $\chi^2=11.99$ ;  $p<0.001$ ). Elle s'avère forte (V de Cramer= 0.35 ; Gamma=0.73).

**Interprétation théorique (sociologique) :** Il ressort de l'analyse tabulaire bivarié la conclusion suivante : les gens intéressés par la politique sont plus susceptibles de participer aux élections. Cela peut s'expliquer, entre autres, par le fait que les gens intéressés par la politique sont souvent des citoyens informés mus par le souci de participer à la gestion des affaires publiques en choisissant les politiques selon leur préférence.

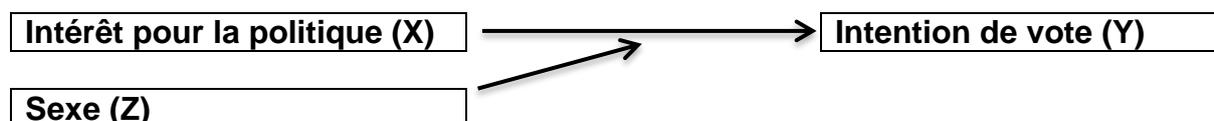
## 1.2. Analyse multivariée : la relation bivariée contrôlée (X-Y selon Z)

Certes l'intérêt pour la politique est statistiquement associé à l'intention de vote. Mais, il s'avère important de mieux comprendre les mécanismes sous-jacents à cette relation. Aussi, se pourrait-il qu'une variable additionnelle, comme le sexe, influe sur la relation d'association initiale? Autrement dit, les férus de politique ne sont-ils pas plus susceptibles de voter, comme nous venons de le voir, simplement parce qu'ils sont surtout des hommes? La question de recherche est la suivante :

***L'intérêt pour la politique influe-t-il sur l'intention de vote, en contrôlant l'effet modérateur du sexe ? Précisément, le sexe modère-t-il l'effet de l'intérêt pour la politique sur l'intention de vote chez les étudiants de L2 inscrits en sociologie ?***

En effet, on peut supposer que les personnes intéressées par la politique sont plus susceptibles de voter lorsqu'elles sont des hommes plutôt que des femmes. Cette hypothèse impliquerait que l'on considère le sexe comme une variable qui contribue à modérer, spécifier la relation entre intérêt pour la politique et intention de vote. Le modèle causal ou d'élaboration ainsi suggéré est la **modération** ou **spécification**. Pour vérifier cette hypothèse de travail, la mécanique consiste à observer la relation entre **Interet2** et **Vote** à l'intérieur des tableaux partiels créés à partir des deux catégories de la variable **Sexe**. Ci-dessous, une représentation schématique du modèle du point de vue conceptuel et statistique.

*Modèle conceptuel*



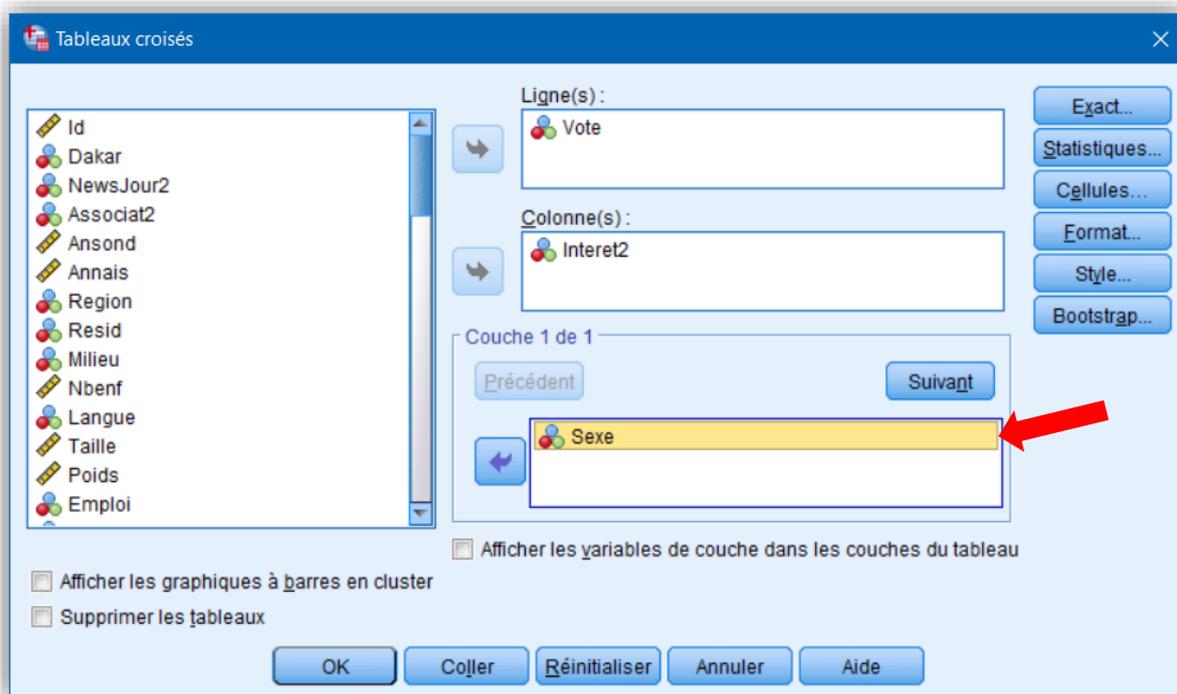
*Modèle statistique*

Sexe	Homme $Z_1$	<b>Intérêt pour la politique (X)</b>	?	<b>Intention de vote (Y)</b>
	Femme $Z_2$	<b>Intérêt pour la politique (X)</b>	?	<b>Intention de vote (Y)</b>

La nouvelle variable SEXE étant qualitative et correctement recodée, sortons les **tableaux partiels d'ordre un** selon la procédure ci-dessous :

**Analyse**

**Statistiques descriptives**  
**Tableaux croisés**



Tel qu'illustré ci-dessous, la variable-contrôle modératrice « Sexe » est placée dans la boîte Couche ou strate (1 de 1). Validez !

**Tableau croisé Vote Seriez-vous oui ou non prêt(e) à voter lors de la prochaine élection présidentielle ? \* Interet2 Intérêt pour la politique \* Sexe Indiquez votre sexe.**

Sexe Indiquez votre sexe.			Interet2 Intérêt pour la politique		Total	
			,00 Non	1,00 Oui		
2 Masculin	Vote Seriez-vous oui ou non prêt(e) à voter lors de la prochaine élection présidentielle ?	1,00 Oui	Effectif	9	32	41
			% dans Interet2 Intérêt pour la politique	75,0%	94,1%	89,1%
		,00 Non	Effectif	3	2	5
			% dans Interet2 Intérêt pour la politique	25,0%	5,9%	10,9%
		Total	Effectif	12	34	46
	% dans Interet2 Intérêt pour la politique	100,0%	100,0%	100,0%		
1 Féminin	Vote Seriez-vous oui ou non prêt(e) à voter lors de la prochaine élection présidentielle ?	1,00 Oui	Effectif	10	33	43
			% dans Interet2 Intérêt pour la politique	55,6%	89,2%	78,2%
		,00 Non	Effectif	8	4	12
			% dans Interet2 Intérêt pour la politique	44,4%	10,8%	21,8%
		Total	Effectif	18	37	55
	% dans Interet2 Intérêt pour la politique	100,0%	100,0%	100,0%		
Total	Vote Seriez-vous oui ou non prêt(e) à voter lors de la prochaine élection présidentielle ?	1,00 Oui	Effectif	19	65	84
			% dans Interet2 Intérêt pour la politique	63,3%	91,5%	83,2%
		,00 Non	Effectif	11	6	17
			% dans Interet2 Intérêt pour la politique	36,7%	8,5%	16,8%
		Total	Effectif	30	71	101
	% dans Interet2 Intérêt pour la politique	100,0%	100,0%	100,0%		

Tableau partiel "Homme"

Tableau partiel "Femme"

Tableau d'ordre 0

## Tests du khi-deux

Sexe	Indiquez votre sexe.	Valeur	ddl	Signification asymptotique (bilatérale)	Sig. exacte (bilatérale)	Sig. exacte (unilatérale)
2 Masculin	khi-deux de Pearson	3,346 <sup>d</sup>	1	,067		
	Correction pour continuité <sup>b</sup>	1,664	1	,197		
	Rapport de vraisemblance	2,919	1	,088		
	Test exact de Fisher				,103	,103
	Association linéaire par linéaire	3,273	1	,070		
	N d'observations valides	46				
1 Féminin	khi-deux de Pearson	8,030 <sup>c</sup>	1	,005		
	Correction pour continuité <sup>b</sup>	6,180	1	,013		
	Rapport de vraisemblance	7,627	1	,006		
	Test exact de Fisher				,012	,008
	Association linéaire par linéaire	7,884	1	,005		
	N d'observations valides	55				
Total	khi-deux de Pearson	11,994 <sup>a</sup>	1	,001		
	Correction pour continuité <sup>b</sup>	10,063	1	,002		
	Rapport de vraisemblance	10,989	1	,001		
	Test exact de Fisher				,001	,001
	Association linéaire par linéaire	11,875	1	,001		
	N d'observations valides	101				

Chi-carré du tableau partiel "Homme"

Chi-carré du Tableau partiel "Femme"

Chi-carré du tableau d'ordre 0

- a. 0 cellules (0,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 5,05.  
 b. Calculée uniquement pour une table 2x2  
 c. 1 cellules (25,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 3,93.  
 d. 2 cellules (50,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 1,30.

## Mesures symétriques

Sexe	Indiquez votre sexe.	Valeur	Erreur asymptotique standard <sup>a</sup>	T approximatif <sup>b</sup>	Signification approximative		
2 Masculin	Nominal par Nominal	Phi	,270		,067		
		V de Cramer	,270		,067		
	Ordinal par Ordinal	Gamma	Ordre zéro	,684	,263	1,417	,156
	N d'observations valides	46					
1 Féminin	Nominal par Nominal	Phi	,382		,005		
		V de Cramer	,382		,005		
	Ordinal par Ordinal	Gamma	Ordre zéro	,737	,162	2,547	,011
	N d'observations valides	55					
Total	Nominal par Nominal	Phi	,345		,001		
		V de Cramer	,345		,001		
	Ordinal par Ordinal	Gamma	Ordre zéro	,725	,135	2,902	,004
		Partielle d'ordre un	,722				
N d'observations valides	101						

V, Gamma du tableau partiel "Homme"

V, Gamma du tableau partiel "Femme"

V, Gamma du tableau d'ordre 0

- a. L'hypothèse nulle n'étant pas considérée.  
 b. Utilisation de l'erreur asymptotique standard

Gamma partiel que nous verrons plus tard. Pas pertinent pour le modèle de spécification!

### 1.3. Présentation et interprétation finales des résultats

Construisons un tableau synthétique unique afin d'intégrer tous les résultats de l'analyse multivariée.

**Tableau 1.2.** Intention de vote à la prochaine élection présidentielle selon l'intérêt pour la politique, contrôlant l'effet modérateur du sexe chez les étudiants

Intention de vote	Sexe			
	Homme		Femme	
	Intérêt pour la politique		Intérêt pour la politique	
	Non	Oui	Non	Oui
Oui	75	94	56	89
Non	25	6	44	11
Cas (n)	(12)	(34)		
Chi-carré	3,35		8,03**	
V de Cramer	0.27		0.38**	
Gamma	0.68		0.74*	

Notes. Les entrées correspondent à des % (n=101). \*\*p<0,01 ; \*p<0,05.

Source. Sondage\_EtudiantsL2Socio\_2021.

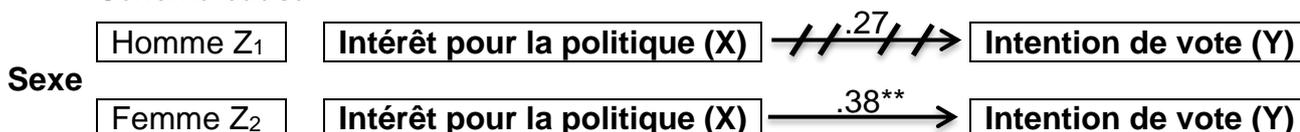
Parallèlement, afin d'interpréter de façon appropriée les résultats, schématisons-les en cohérence avec le modèle causal proposé au départ.

**Figure 1.1.** Schématisation de l'effet modérateur du sexe dans la relation entre l'intérêt pour la politique et l'intention de vote chez les étudiants

Schéma initial



Schéma causal



Notes. Les entrées correspondent à des V de Cramer. Les astérisques renvoient aux p-values des tests du chi-carré. \*\*\*p<0,001 ; \*\*p<0,01.

**Interprétation statistique finale (analyse succincte):** L'analyse tabulaire bivariée montre que 92% des étudiants intéressés par la politique ont l'intention de voter lors de la prochaine élection présidentielle, comparativement à seulement 63% des étudiants non-intéressés par la politique. Il existe une différence de 31 points en pourcentages. La relation initiale entre l'intérêt pour la politique et l'intention de vote est statistiquement significative dans la population étudiante étudiée ( $\chi^2=11.99$ ;  $p<0.001$ ). Elle s'avère forte (V de Cramer= 0.35 ; Gamma=0.73).

Lorsqu'on introduit le sexe dans la relation à des fins de contrôle, on obtient des tableaux partiels dont l'analyse montre que la relation entre l'intérêt pour la politique et l'intention de vote s'annule chez les hommes. Cependant, la relation persiste de façon statistiquement significative chez les femmes ( $\chi^2=8,03$ ;  $p<0,01$ ) et elle s'amplifie avec une intensité forte (V=0,38). Par conséquent, les résultats soutiennent l'hypothèse à l'effet que le sexe modère, spécifie la relation entre l'intérêt pour la politique et l'intention de vote. Ces deux variables sont reliées à la condition d'être femme.

**Interprétation sociologique (analyse succincte):** Le sexe permet de mieux comprendre la nature de la relation entre l'intérêt pour la politique et l'intention de vote. Toutefois, la direction de la modération est surprenante. Nous postulons l'existence d'une relation surtout chez les hommes, qui parce qu'ils sont davantage intéressés par la politique seraient plus susceptibles de voter comparativement aux femmes. Les résultats démontrent plutôt le contraire. Comment l'expliquer? En fait, même si les femmes sont probablement moins intéressées par la politique que les hommes, lorsqu'elles s'y intéressent elles pourraient être plus susceptibles de participer aux élections, car étant plus motivées. En revanche, chez les hommes, leur intérêt pour la politique ne se traduit pas forcément par un devoir de vote systématique, les hommes étant plus susceptibles de sanctionner les politiques par l'abstention.

## 2. La dissimulation

Les processus d'élaboration habituellement considérées postulent qu'il y a une relation d'ordre zéro entre la VI et la VD. Que se passe-t-il maintenant lorsque les deux variables ne sont pas reliées statistiquement? La tentation est grande d'abandonner l'analyse d'une telle relation pour explorer d'autres relations. Pourtant, dans le cas d'une relation nulle entre une VI et une VD, l'introduction d'une 3e variable peut révéler la relation. Dans ce cas, la VC dissimule la relation initiale, qu'elle révèle par la suite une fois incluse dans la relation étudiée. On parle alors de **dissimulation**, un modèle de causalité atypique qui peut être mis en pratique à l'aide de l'élaboration de tableau.

### 2.0. Étapes de l'élaboration de tableau pour un modèle de dissimulation

#### Étape 1 : Analyse du tableau bivarié d'ordre zéro

L'association entre la VI (X) et la VD (Y) est-elle statistiquement significative? Elle ne doit pas l'être au départ.

#### Étape 2 : Analyse des tableaux partiels pour apporter plus d'informations

L'association entre la VI (X) et la VD (Y) est-elle statistiquement significative au sein de chaque niveau de la variable-contrôle antécédente (Z) ou au sein d'un niveau ou de certains niveaux? Quelle est la force et la direction éventuelle de la relation initiale contrôlée pour chaque niveau de la variable-contrôle Z (Z1 : X-Y; Z2 : X-Y)?

#### Étape 3 : Schématisation des modèles initial et causal aux fins de diagnostic causal

Après avoir schématisé les résultats, dites quel diagnostic causal est révélé par l'analyse des schémas? Dissimulation ou non!

Résumé de l'élaboration de tableau (Fox: 311)			
Variable-contrôle VC	Relation VI-VD primitive?	Relation VI-VD contrôlant l'effet de VC (tableaux partiels)	Conclusion à tirer
Antécédente [Dissimulatrice]	Non	Il y a relation dans les tableaux partiels $G_{\text{tableaux partiels}} > G_{\text{d'ordre 0}}$	VC révèle la relation VI-VD masquée auparavant (dissimulation)

## 2.1. Analyse bivariée de la relation initiale (X-Y)

Intéressons-nous à la relation entre la **lecture de journaux** (newjour2) et l'**intention de vote** à la prochaine élection présidentielle (vote) chez les étudiants.



La question de recherche est la suivante : ***Y a-t-il une relation d'association entre la lecture de journaux et l'intention de vote chez les étudiants de L2 inscrits en sociologie à l'UGB ?*** On peut penser que plus les gens vont aux nouvelles en lisant les journaux, plus ils sont susceptibles de participer aux élections. Pour vérifier cette hypothèse de travail, construisons un tableau croisé bivarié, procédons au test du chi-carré et calculons une ou des mesures d'association appropriées.

Les deux variables étant qualitatives et correctement codées, sortons le **tableau bivarié d'ordre zéro**, ainsi que le test du chi-carré et des mesures d'association appropriées. Vous connaissez la procédure, laquelle est illustrée ci-dessous :

### Analyse

#### Statistiques descriptives

#### Tableaux croisés

#### Ligne

✓ Vote

#### Colonne

✓ NewJour2

#### Cellules

✓ Observé

✓ Position

#### Statistiques

✓ Khi-deux

✓ Phi et V de Cramer

✓ Gamma

#### Format

✓ Descending

Récapitulons les résultats dans un tableau synthétique, qui inclut les informations clés de l'analyse tabulaire bivariée, du test du chi-carré et des mesures d'association.

**Tableau 2.1.** *Intention de vote à la prochaine élection présidentielle selon la lecture de journaux chez les étudiants*

Intention de vote	Lecture de journaux	
	Non	Oui
Oui	79	88
Non	21	12
Cas (n)	(53)	(50)
Chi-carré	1,43	
V de Cramer	.12	
Gamma	.31	

Notes. Les entrées correspondent à des % (n=103).

Source. Sondage\_EtudiantsL2Socio\_2021.

**Interprétation statistique (analyse succincte) :** L'analyse tabulaire bivariée montre

que 88% des étudiants lecteurs de journaux ont l'intention de voter lors de la prochaine élection présidentielle, comparativement à 79% des étudiants non-lecteurs. La différence semble négligeable (9 points en pourcentage seulement). D'ailleurs, la relation entre la lecture des journaux et l'intention de vote n'est pas statistiquement significative dans la population étudiante étudiée ( $\chi^2=1,43$ ;  $p=0.232$ ).

**Interprétation théorique (sociologique) :** Il ressort de l'analyse tabulaire bivariée la conclusion suivante : les lecteurs de journaux ne sont pas plus susceptibles de participer aux élections que les non-lecteurs. Une explication plausible est que les non-lecteurs peuvent s'informer à travers d'autres canaux et être mobilisés par les politiques, de telle sorte que cela réduit l'écart avec les lecteurs de journaux qui sont naturellement plus enclins à voter du fait de leur haut niveau de conscience politique.

## 2.2. Analyse multivariée : la relation bivariée contrôlée

Certes la lecture de journaux n'est pas statistiquement associée à l'intention de vote. Mais, il est important de mieux comprendre les mécanismes sous-jacents à cette absence de relation. Aussi, se pourrait-il qu'une variable additionnelle, comme le milieu de résidence, révèle la relation initiale? La question de recherche est la suivante :

***Y a-t-il une relation entre la lecture de journaux et l'intention de vote, en tenant compte de l'effet du milieu de résidence? Précisément, le milieu de résidence révèle-t-il l'effet de la lecture de journaux sur l'intention de vote chez les étudiants de L2 inscrits en sociologie ?***

En effet, il est plausible que la relation se révèle particulièrement chez les résidents de la région urbaine de Dakar, lesquels pourraient être plus susceptibles de voter lorsqu'ils sont bien informés à travers la lecture de journaux. Cette hypothèse impliquerait que l'on considère le milieu de résidence comme une variable qui contribue à révéler la relation entre lecture de journaux et intention de vote. Le modèle causal ou d'élaboration ainsi suggéré est la **dissimulation**. Pour vérifier cette hypothèse de travail, la mécanique consiste à observer la relation entre **Lecture de journaux** et **Vote** à l'intérieur des tableaux partiels créés à partir des deux catégories de la variable **Dakar**. Ci-dessous, une représentation schématique du modèle.

Milieu	Autre Z <sub>1</sub>	Lecture de journaux (X)	?	Intention de vote (Y)
	Dakar Z <sub>2</sub>	Lecture de journaux (X)	?	Intention de vote (Y)

Sortons les **tableaux partiels d'ordre un** selon la procédure ci-dessous :

### Analyse

#### Statistiques descriptives

#### Tableaux croisés

#### Ligne

✓ Vote

#### Colonne

✓ NewsJour2

#### Couche

✓ Dakar

#### Cellules

✓ Observé

- ✓ Position
- Statistiques**
  - ✓ Khi-deux
  - ✓ Gamma
  - ✓ Phi & V de Cramer
- Format**
  - ✓ Descending

Récapitulons les résultats de l'analyse multivariée ci-dessous.

### 2.3. Présentation et interprétation finales des résultats

Construisons un tableau synthétique unique afin d'intégrer tous les résultats de l'analyse multivariée.

**Tableau 1.2.** Intention de vote à la prochaine élection présidentielle selon l'intérêt pour la politique, contrôlant l'effet modérateur du sexe chez les étudiants

Intention de vote	Milieu de résidence			
	Autre		Dakar	
	Lecture de journaux		Lecture de journaux	
	Non	Oui	Non	Oui
Oui	87	88	29	86
Non	13	12	71	14
Cas (n)	(45)	(42)	(7)	(7)
Chi-carré	0,04		4,67*	
V de Cramer	0.02		0.58*	
Gamma	0.07		0.88**	

Notes. Les entrées correspondent à des % (n=101). \*\*p<0,01 ; \*p<0,05.

Source. Sondage\_EtudiantsL2Socio\_2021.

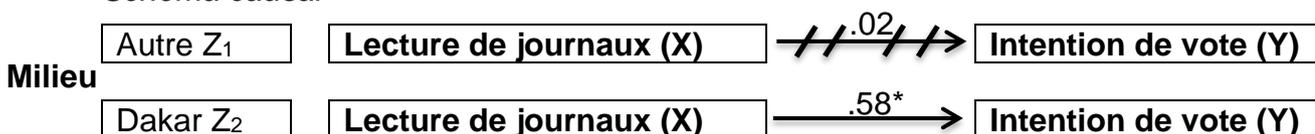
Parallèlement, afin d'interpréter de façon appropriée les résultats, schématisons-les en cohérence avec le modèle causal proposé au départ.

**Figure 1.1.** Schématisation de l'effet dissimulateur du milieu de résidence sur la relation entre la lecture de journaux et l'intention de vote chez les étudiants

Schéma initial



Schéma causal



Notes. Les entrées correspondent à des V de Cramer. Les astérisques renvoient aux p-values des tests chi-carré. \*\*p<0,01.

**Interprétation statistique (analyse succincte):** L'analyse tabulaire bivariée montre que 88% des étudiants lecteurs de journaux ont l'intention de voter lors de la prochaine élection présidentielle, comparativement à 79% des étudiants non-lecteurs. La différence semble négligeable (9 points en pourcentage seulement). D'ailleurs, la

relation entre la lecture des journaux et l'intention de vote n'est pas statistiquement significative dans la population étudiante étudiée ( $\chi^2=1,43$ ;  $p=0.23$  ;  $V=0.11$ ).

Lorsqu'on introduit le milieu de résidence dans la relation à des fins de contrôle, on obtient des tableaux partiels dont l'analyse montre que la relation entre la lecture de journaux et l'intention de vote s'annule toujours chez les résidents des autres milieux, mais se révèle de façon significative chez les résidents dakarois ( $\chi^2=4,67$ ;  $p<0,01$ ) avec une intensité forte ( $V=0,58$ ). Par conséquent, les résultats soutiennent l'hypothèse à l'effet que le milieu de résidence dissimule la relation entre la lecture de journaux et l'intention de vote. Il y a donc dissimulation. Cependant, l'effet dissimulateur est dû seulement aux Dakarois et non aux autres.

**Interprétation sociologique (analyse succincte):** Le milieu de résidence permet de mieux comprendre l'absence de relation entre la lecture de journaux et l'intention de vote. Aussi surprenant que cela puisse être, la relation est révélée chez les citoyens dakarois. Comment l'expliquer? Les Dakarois sont connus pour leur haut niveau de conscience politique. De surcroît, lorsqu'ils vont aux nouvelles à travers la lecture de journaux, leur conscience politique s'accroît, augmentant ainsi les chances de voter.

### 3. Exercice d'application :

---

Une sociologue étudie les facteurs associés au temps passé sur Internet chez les étudiants. La question de recherche est la suivante : **Le sexe (Sexe) influe-t-il sur le temps passé sur Internet (Internet3), en contrôlant l'effet du milieu de résidence (Dakar)? Précisément, l'effet du sexe sur le temps sur Internet diffère-t-il selon le milieu de résidence chez les étudiants de L2 inscrits en sociologie à l'UGB ?**

Pour élucider cette question de recherche, répondez aux interrogations ci-dessous :

#### Étape 1 : Cadre opératoire et théorique

- Quelles sont la variable dépendante, la variable indépendante, et la variable-contrôle dans ce contexte?
- Quel modèle causal est-il suggéré par la question de recherche (explication ou spécification) ?
- Proposez une explication théorique de l'élaboration.

#### Étape 2 : Description des variables

- Faites sortir les fréquences et % pour chacune des trois variables impliquées dans l'étude : **Sexe, Internet3, Dakar**.
- Selon la façon dont chacune des trois variables est distribuée, peut-on élucider leur relation à l'aide de l'analyse tabulaire multivariée? Justifiez.

#### Étape 3 : Analyse bivariée de la relation initiale ou primitive (X-Y)

- Sortez le tableau bivarié en % de la relation entre le sexe (sexe) et le temps sur internet (Internet3), le chi-carré, le V de Cramer et le Gamma;
- Interprétez statistiquement les résultats en comparant les pourcentages, et décrivant la signification statistique et réelle de la relation étudiée;

#### Étape 4 : Analyse multivariée de la relation bivariée contrôlée (X-Y selon Z)

- Sortez les tableaux partiels de la relation entre le sexe (sexe) et le temps sur

internet (Internet3), en contrôlant l'effet du milieu de résidence (Dakar), les chi-carrés, les V de Cramer et Gammas;

- b) Interprétez statistiquement les résultats en comparant les pourcentages, et en décrivant la signification statistique et réelle de la relation étudiée pour chacun des niveaux de la variable-contrôle;

*Étape 5 : Schématisation des modèles initial et causal aux fins de diagnostic causal*

- a) Schématisez les résultats du modèle initial et du modèle causal complet en y inscrivant les V de Cramer (ou gammas) et la signification statistique;
- b) Quel est le diagnostic causal de l'élaboration? Justifiez.