



Econometric analysis of the determinants of socioeconomic development projects' data quality in the Democratic Republic of the Congo

Analyse économétrique des déterminants de la qualité des données des projets de développement socioéconomique en République Démocratique du Congo

Olivier Mumbere Muhongya, PhD

Professeur Associé à l'Institut Supérieur de la Gombe
Kinshasa – République Démocratique du Congo

Résumé: Déterminer la qualité des données est un pas important vers l'amélioration de la qualité des données au sein d'une entreprise ou d'un programme/projet. Néanmoins, identifier les facteurs qui influent sur la qualité des données est encore mieux. Cet article part des données puisées des rapports d'évaluation de la qualité des données des projets et programmes socioéconomiques mis en œuvre en République Démocratique du Congo et financés par l'Agence Américaine de Développement (USAID). Il tente d'élaborer un modèle économétrique explicatif de la qualité des données. Les résultats montrent qu'aucun des déterminants puisés dans la pratique actuelle en RDC ne s'avère pertinent. Il faut interroger, plus en amont, l'approche d'évaluation même de la qualité des données. Celle-ci intègre l'approche qualitative qui est susceptible de rendre le jugement sur la qualité des données fort subjectif. Or ceci risque de rendre inopérant tout effort de modélisation des déterminants de la qualité des données.

Mots-clés : qualité des données ; modélisation ; critères de qualité des données.

Abstract: Determining data quality is an important step toward improving data quality within an enterprise or a program/project. However, identifying factors that influence data quality is deemed to be better. This article uses data drawn from data quality assessment reports for socio-economic projects and programs implemented in the Democratic Republic of Congo and funded by the United States Development Agency (USAID). It builds up an econometric model attempting to explain data quality. The results show that none of the determinants drawn from practice prove to be relevant. One should, therefore, question, further upstream the very approach to evaluating data quality. This integrates the qualitative approach which is likely to make the judgment on the quality of the data very subjective. This risks rendering any efforts to model the determinants of data quality ineffective.

Keywords: data quality; modelization; data quality criteria set.

Digital Object Identifier (DOI): <https://doi.org/10.5281/zenodo.11518971>

1. Introduction

Les chiffres demeurent la référence la plus importante lorsqu'il s'agit d'évoquer la performance d'une entreprise (chiffre d'affaires) ou celle d'un projet de développement économique dans divers domaines comme la santé, l'éducation, la sécurité alimentaire et même la gouvernance.

Cependant, autant tout le monde fait confiance aux chiffres autant ceux-ci nous parviennent avec une qualité souvent douteuse. Il est dès lors légitime de se poser la question de savoir si l'on n'a plus confiance aux chiffres dans quelle autre variable pourrions-nous rassembler notre confiance ? Le chiffre est incontournable aussi bien pour la communication des résultats ou performance que pour la planification des activités futures d'une entreprise ou d'un projet de développement socioéconomique. En outre, ce qui est vrai d'un projet de développement l'est encore plus d'un pays ou sa macroéconomie. Pourtant, les différents chercheurs et praticiens de la gestion ont du mal à s'accorder sur le concept qualité. Dans l'entre-temps, il faut reconnaître la difficulté de s'accorder sur le concept qualité des données. Ce concept est intrinsèquement difficile à appréhender. Ainsi, pour Julien TIÈCHE la notion de qualité est aussi fortement liée à son contexte et à la façon dont elle est appréhendée. La qualité associe ainsi toujours le sujet (le client ou l'utilisateur) et l'objet (le produit ou le service). Elle est donc constamment relative à un contexte, une attente (Chaumette 2006, p. 41). Ce concept est perçu désormais tantôt comme objet et tantôt comme objectif d'étude ou de pratique de gestion. D'où la nécessité de s'accorder sur les critères ou dimensions de la qualité des données. Evaluer la qualité des données revient de rapprocher les critères de qualité des données de la réalité des données en face.

Voilà pourquoi lorsque Laure – Bertille parle de la modélisation de la qualité des données relevant d'un système, elle revient sur les différentes dimensions de la qualité des données telles que perçues par différents auteurs comme Brodie, Delen et Rijsenbri ; Wang ; Redman ; Clabreto, Pinon, Pouillet et Richez. Elle ne présente pas les déterminants de la qualité des données.

En République Démocratique du Congo, chaque Agence de Développement a son approche sur la qualité des données. Celle qui semble la mieux formalisée sur le terrain est celle de l'Agence Américaine de Développement (USAID). Fondée sur la pratique de l'évaluation de la qualité des données auprès de plusieurs Agences de Développement international financées par l'Agence de Développement Américaine (USAID) entre 2017 et 2020, cette étude cherche à comprendre les déterminants de la qualité

des données. Quels sont les facteurs déterminants de la qualité des données en constitue la question centrale.

Il s'agit d'une approche essentiellement qualitative sous la forme d'une étude de cas. Celle-ci a fait recours à un modèle économétrique.

1.1. Revue de la littérature

La littérature abonde d'articles et publications autour de la qualité des données. Les choses deviennent plus intéressantes lorsqu'on réalise la pertinence des publications sur la corrélation entre la qualité des données et les coûts de l'entreprise. Markus Eppler et al font la revue et la catégorisation de tous les coûts potentiels qui sont entraînés par la faiblesse de la qualité des données et en examine leur évolution. Et David Loshin d'enfoncer le clou dans son article lorsqu'il affirme qu'il existe une preuve irréfutable sur les durs impacts (sur l'activité) directement liés à la faiblesse de la qualité des données. En procédant par une étude de cas Loshin souligne quatre catégories d'impacts à savoir les impacts d'ordre financier ; de confiance et de satisfaction ; de productivité et de risque et compliance.

En allant plus loin, Even et Adir ont inauguré des outils de modélisation de l'inégalité et son mesurage dans de très grandes bases des données. Nous montrons, poursuivent-ils, que les évaluations quantitatives de l'inégalité, ensemble avec les évaluations qui reflètent la présence et l'impact des défauts peuvent informer sur l'état actuel de la qualité ; orienter les politiques de suivi de la qualité des données et aider à prioriser les efforts d'amélioration.

Les deux auteurs finissent par conclure que le lien entre les défauts de qualité, les efforts d'amélioration et les gains économiques ont jusqu'à présent été insuffisamment examiné.

Surfant sur le même raisonnement, mais de manière plus pragmatique, cette étude se fonde sur un nombre important d'évaluation

Plus pragmatique, et se fondant sur un nombre important des rapports d'évaluation de la qualité des données réalisées entre 2017 et 2020, et article part d'une analyse du modèle de régression multiple où la qualité des données est expliquée par la qualité des indicateurs ; la formation des agents de collecte des données et le nombre d'évaluation de la qualité des données réalisées sur un indicateur.

1.2. Méthodologie

Cette étude est qualitative. C'est une étude de cas qui s'appuie, outre sur la revue de la littérature existante, sur les données collectées via des rapports d'évaluation de la qualité des données réalisées par l'auteur de l'article sur la période allant entre 2017 et 2020. L'analyse des données a été réalisée par le logiciel SPSS.

1.3. Résultats

Il s'agissait d'établir les déterminants de la qualité des données (Variable qualitative ordinale) en rapport avec le nombre de fois que les agents de collecte des données ont été formés (variable quantitative); la qualité des indicateurs (variable qualitative); la qualité des indicateurs (variable qualitative)

Après une analyse descriptive à travers les tableaux d'effectifs, nous allons procéder à une analyse de la corrélation à travers le test du Khi-carré et le calcul du coefficient de contingence. En cas de dépendance significative entre la qualité des données et l'un ou l'autre déterminant retenu a priori, il

sera procédé à l'estimation d'un modèle de choix multiples ou de choix binaire après avoir ramené la qualité des données à une variable binaire.

Il l'issue des calculs statistiques décrits dans l'annex 2, il s'avère qu'il n'existe aucune corrélation entre la variable qualité des données d'une part et les variables qualité de l'indicateur ; formation des agents de collecte des données et nombre d'évaluation de qualité des données.

2. Discussion

Les gestionnaires des projets et programmes considèrent actuellement qu'en veillant au choix de l'indicateur, il y a de fortes chances d'améliorer la qualité des données. D'où non seulement les bailleurs des fonds insistent sur le choix le plus méticuleux de l'indicateur (suivant les critères ad hoc). Une fois l'indicateur bien choisi, l'Agence de mise en œuvre du projet ou programme forment les agents de collecte des données en veillant particulièrement à faire comprendre la définition de l'indicateur ainsi que les limitations éventuelles dans la collecte des données. La formation des agents de collecte (et souvent de tous les agents du programme) est une recommandation récurrente particulièrement à l'issue des évaluations de la qualité des données soit pour prévenir soit pour maintenir les standards.

Pendant que ces ingrédients d'amélioration de la qualité des données sont considérées comme pratiquement incontournables, il s'avère que la qualité des données comme l'exactitude dépend de plusieurs autres facteurs difficiles à cerner notamment la bonne volonté du gestionnaire car très souvent le responsable de suivi et évaluation est un membre du personnel du projet ou programme. Ceci fait que ce membre du personnel, qui est supervisé par le gestionnaire principal, demeure à la merci de ce dernier. Ceci laisse la qualité des données dans un espace peu protégé car les modifications demeurent toujours possibles à tout moment.

Il conviendra donc d'explorer d'autres facteurs susceptibles d'influer sur la qualité des données et les intégrer dans un modèle d'analyse multi variée. C'est ici où nous tombons d'accord avec Even et Adir lorsqu'ils concluent que *des politiques d'amélioration de la qualité et des décisions de priorisation de la qualité doivent également tenir compte d'autres contraintes comme les engagements d'activités ainsi que la disponibilité des ressources financières et humaines dans les efforts de correction*¹.

Par ailleurs, il convient de souligner aussi le fait que le recours à l'approche qualitative peut inviter au processus une certaine subjectivité sur le jugement de la qualité des données. Johannan Levesque dans son mémoire conseille en amont de l'évaluation de la qualité des données *d'ajouter une mesure concernant l'appréciation personnelle de l'expert qui utilise le système MUM. Cette mesure sera souvent la seule qui pourra être insérée de manière explicite. Elle aura comme effet d'influer sur la rigueur de l'analyse en y ajoutant une certaine subjectivité*². Dans un tel contexte, la subjectivité peut être déterminante dans la détermination de la qualité des données. Il serait donc judicieux de combiner à la fois la méthode qualitative et celle quantitative pour garantir une décision optimale quant à la qualité des données.

¹ Adir Even et al. : Comparative Analysis of Data Quality and Utility Inequality Assessments, t: <http://aisel.aisnet.org/ecis2008> , p. 13.

² Johann Levesque : *l'évaluation de la qualité des données spatiales (approche top-down et gestion de l'évaluation de la qualité des données géo spatiales ; approche top-down et gestion de la méta qualité*, département des sciences géomatiques ; faculté de foresterie et de géomatique ; Université de Laval ; Québec ; 2007, p. 123

3. Conclusion

Aucune des variables explicatives retenues (qualité de l'indicateur ; formation des agents de collecte des données et fréquence de conduite d'évaluations de la qualité des données) ne permet pas d'expliquer la qualité des données. C'est dire en d'autres termes que les variables choix de l'indicateur ; formation des agents de collecte des données et même la fréquence de conduite d'évaluations de la qualité des données ne sont pas suffisantes pour expliquer la qualité des données. Il faut explorer d'autres variables. La meilleure piste à explorer est celle qui expliquerait la véritable chaîne des décisions sur le reportage ou la diffusion des données au sein des systèmes de gestion des données.

Enfin, il sera toujours recommandable que les évaluations de qualité des données utilisent à la fois l'approche qualitative et quantitative pour garantir une décision sur la qualité qui soit optimale.

REFERENCES

1. Adir Even et al. : Comparative Analysis of Data Quality and Utility Inequality Assessments, t: <http://aisel.aisnet.org/ecis2008>
2. Johann Levesque : *l'évaluation de la qualité des données spatiales (approche top-down et gestion de l'évaluation de la qualité des données géo spatiales ; approche top-down et gestion de la méta qualité*, département des sciences géomatiques ; faculté de foresterie et de géomatique ; Université de Laval ; Québec ; 2007

Annex 1. : Données tirées des rapports d'évaluation de la qualité des données

N°	Indicateurs	Année	Qualité des données	Data quality issues	Formation des agents de collecte	Nombre d'EQD réalisées sur l'indicateur	Qualité d'EQD	Qualité de l'Indicateur
1	Nombre de ménages bénéficiaires	2017	1	Validity, reliability, and integrity are still issues	0	0	0	1
2	Nombre d'individus bénéficiaires de l'aide visant à améliorer la productivité	2017	1	Validity, reliability, and integrity are still issues	0	0	0	1
3	Nombre d'agriculteurs qui ont appliqué de nouvelles technologies ou des pratiques de gestion	2017	1	Validity, reliability, and integrity are still issues	0	0	0	0
4	Nombre d'hectares de terres arables sur lesquelles sont appliquées de nouvelles technologies	2017	1	Validity, reliability, and integrity are still issues	0	0	0	1
5	Nombre d'agents de santé formé dans la prise en charge de la malaria avec la combinaison de la thérapie fondée sur l'artémisinisme	2017	1	severe data quality limitations relating to validity and reliability.	1	2	0	1
6	Nombre de personnes formés dans n'importe quelle composante de la stratégie de l'OMS "Mettre fin à la tuberculose"	2017	1	severe data quality limitations relating to validity and reliability.	1	2	0	1
7	Nombre d'agents de santé formés au diagnostic de malaria par le laboratoire	2017	1	severe data quality limitations relating to validity and reliability.	1	2	0	0
8	Nombre d'agents de santé formés sur le traitement préventif en période de grossesse	2017	1	severe data quality limitations relating to validity and reliability.	1	2	0	1

9	Nombre de nouveaux nés qui reçoivent un traitement antibiotique pour soigner une infection	2017	3	moderate data quality limitations relating to timeliness and data integrity.	1	2	0	1
10	s Nombre d'enfants de moins de 5 ans suspectés de souffrir de pneumonie qui reçoivent des traitements antibiotiques d'agents de santé formés	2017	1	severe data quality limitations relating to validity, timeliness and data integrity.	1	2	0	0
11	Nombre de litres d'eau désinfectés aux produits disponibles à des endroits d'utilisation	2017	4	low data quality limitations relating to data validity.	1	2	0	1
12	s Nombre de ceux qui acceptent les nouvelles méthodes contraceptives	2017	1	severe data quality limitations relating to validity and reliability.	1	2	0	1
13	Nombre de cas de tuberculose multi-résistante	2017	5	no data quality limitations relating to validity, reliability, timeliness or integrity.	1	2	0	1
14	Nombre de cas de tuberculose multi-résistante qui ont commencé leur traitement de seconde ligne.	2017	5	no data quality limitations relating to validity, reliability, timeliness or integrity.	1	2	0	1
15	Nombre de femmes en couche ayant reçu les utérotoniques durant le troisième stade de travail	2017	3	moderate data quality limitations relating to timeliness and data integrity.	1	2	0	1

16	Number of newborns not breathing at birth who were resuscitated in USG-supported programs Nombre de nouveaux-nés ayant perdu connaissance qui reprennent vie	2017	3	moderate data quality limitations relating to timeliness and data integrity.	1	2	0	1
17	Nombre de cas de diarrhée d'enfant traités	2017	1	severe data quality limitations relating to validity, timeliness and data integrity.	1	2	0	1
18	Nombre d'années de protection de couples	2017	4	low data quality limitations relating to reliability	1	2	0	1
19	Nombre d'agents de santé qui partagent l'information sur la planification familiale et/ou des services au cours de l'année	2017	4	low data quality limitations relating to reliability	1	2	0	0
20	Nombre de personnes qui ont accès au service sanitaire élémentaire	2017	1	severe data quality limitations relating to data reliability, timeliness and precision.	1	2	0	1
21	Nombre de lieux d'aisance offerts dans des bâtiments institutionnels.	2017	4	moderate data quality limitations relating to timeliness and data integrity.	1	2	0	1

22	Nombre d'enfants de moins de deux ans qui sont desservis par les interventions communautaires relatives à la nutrition.	2017	1	severe data quality limitations relating to validity and reliability	1	0	0	1
23	Nombre de groupes ou d'initiatives créés dédiés à la résolution des conflits ou des causes des conflits	2014	3	Only integrity is still an issue	0	0	0	0
24	Nombre de litres d'eau désinfectés avec des produits locaux de traitement d'eaux	2014	3	Data reliability is still an issue	0	0	0	1
25	Nombre des femmes qui s'impliquent à un haut niveau à la réconciliation communautaire	2014	2	Data precision and integrity are still issues	0	0	0	1
26	Nombre d'enfants tirés des groupes armés.	2012	5	no data quality limitations relating to validity, reliability, timeliness or integrity.	0	0	0	1
27	Nombre des femmes qui participent de manière substantielle ou à une position élevée dans le processus de pacification	2012	5	no data quality limitations relating to validity, reliability, timeliness or integrity.	0	0	0	0
28	Nombre de conflits fonciers et ceux liés aux ressources naturelles résolus en faveur des populations les plus vulnérables et les communautés locales.	2012	5	no data quality limitations relating to validity, reliability, timeliness or integrity.	0	0	0	0

29	Nombre d'initiatives créées ou renforcées à travers des fonds dédiés à la résolution des conflits ou des causes de conflit.	2012	5	no data quality limitations relating to validity, reliability, timeliness or integrity.	0	0	0	0
30	Number of health workers trained in case management with artemisinin-based combination therapy with USG funds Nombre d'agents de santé formés à la prise en charge basée sur la thérapie basée sur la combinaison de l'artémisia	2012	5	no data quality limitations relating to validity, reliability, timeliness or integrity.	0	1	0	1
31	Nombre de personnes ayant bénéficié de services sociaux	2012	2	Validity and reliability are still issues	0	0	0	1
32	Nombre d'agents sociaux formés qui servent les populations vulnérables	2012	2	Validity and reliability are still issues	0	0	0	1

Annex 2. : Statistiques descriptives des données

Tableau croisé QUALITE DES DONNEES EN BINAIRE * FORMATION DES AGENTS DE COLLECTE							
			FORMATION DES AGENTS DE COLLECTE		Total		
			NON	OUI			
QUALITE DONNEES BINAIRE	DES EN	NON	Effectif	7a	9a	16	
			% compris dans QUALITE DES DONNEES EN BINAIRE	43,8%	56,2%	100,0%	
			% du total	21,9%	28,1%	50,0%	
		OUI		Effectif	7a	9a	16
				% compris dans QUALITE DES DONNEES EN BINAIRE	43,8%	56,2%	100,0%
				% du total	21,9%	28,1%	50,0%
	Total			Effectif	14	18	32
				% compris dans QUALITE DES DONNEES EN BINAIRE	43,8%	56,2%	100,0%
				% du total	43,8%	56,2%	100,0%

Chaque lettre en indice indique un sous-ensemble de FORMATION DES AGENTS DE COLLECTE catégories dont les proportions de colonne ne diffèrent pas de manière significative les unes des autres au niveau ,05.

Tableau 1 : Tableau croisé QUALITE DES DONNEES EN BINAIRE * FORMATION DES AGENTS DE COLLECTE.

Tests du Khi-deux					
	Valeur	ddl	Signification asymptotique (bilatérale)	Signification exacte (bilatérale)	Signification exacte (unilatérale)
Khi-deux de Pearson	,000a	1	1,000		
Correction pour la continuité ^b	,000	1	1,000		
Rapport de vraisemblance	,000	1	1,000		
Test exact de Fisher				1,000	,639
Nombre d'observations valides	32				
a. 0 cellules (,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 7,00.					
b. Calculé uniquement pour un tableau 2x2					

Tableau 2 : Test Khi-deux QUALITE DES DONNEES EN BINAIRE versus FORMATION DES AGENTS DE COLLECTE

Tableau croisé QUALITE DES DONNEES EN BINAIRE * QUALITE DES INDICATEURS					
			QUALITE DES INDICATEURS		Total
			NON	OUI	
QUALITE DES DONNEES BINAIRE	NON	Effectif	3a	13a	16
		% compris dans QUALITE DES DONNEES EN BINAIRE	18,8%	81,2%	100,0%
		% du total	9,4%	40,6%	50,0%
	OUI	Effectif	5a	11a	16
		% compris dans QUALITE DES DONNEES EN BINAIRE	31,2%	68,8%	100,0%
		% du total	15,6%	34,4%	50,0%
	Total	Effectif	8	24	32
		% compris dans QUALITE DES DONNEES EN BINAIRE	25,0%	75,0%	100,0%
		% du total	25,0%	75,0%	100,0%

Chaque lettre en indice indique un sous-ensemble de QUALITE DES INDICATEURS catégories dont les proportions de colonne ne diffèrent pas de manière significative les unes des autres au niveau ,05.

Tableau n° 3 : Tableau croisé QUALITE DES DONNEES EN BINAIRE * QUALITE DES INDICATEURS.

Tests du Khi-deux					
	Valeur	ddl	Signification asymptotique (bilatérale)	Signification exacte (bilatérale)	Signification exacte (unilatérale)
Khi-deux de Pearson	,667a	1	,414		
Correction pour la continuité	,167	1	,683		
Rapport de vraisemblance	,672	1	,412		
Test exact de Fisher				,685	,343
Nombre d'observations valides	32				
a. 2 cellules (50,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 4,00.					
b. Calculé uniquement pour un tableau 2x2					

Tableau n° 4 : Tests du Khi-deux QUALITE DES DONNEES EN BINAIRE et QUALITE DES INDICATEURS

Mesures symétriques			
		Valeur	Signification approximée
Nominal par Nominal	Coefficient de contingence	,143	,414
Nombre d'observations valides		32	
a. L'hypothèse nulle n'est pas considérée.			
b. Utilisation de l'erreur standard asymptotique dans l'hypothèse nulle.			

Tableau n° 5 : mesures symétriques des tests du Khi-deux QUALITE DES DONNEES EN BINAIRE et QUALITE DES INDICATEURS

Tableau croisé QUALITE DES DONNEES EN BINAIRE * NOMBRE D'EQD REALISES SUR L'INDICATEUR						
			NOMBRE D'EQD REALISES SUR L'INDICATEUR			Total
			0	1	2	
QUALITE DONNEES BINAIRE	NON	Effectif	8a	0a	8a	16
		% compris dans QUALITE DES DONNEES EN BINAIRE	50,0%	0,0%	50,0%	100,0%
		% du total	25,0%	0,0%	25,0%	50,0%
	OUI	Effectif	6a	1a	9a	16
		% compris dans QUALITE DES DONNEES EN BINAIRE	37,5%	6,2%	56,2%	100,0%
		% du total	18,8%	3,1%	28,1%	50,0%
	Total	Effectif	14	1	17	32
		% compris dans QUALITE DES DONNEES EN BINAIRE	43,8%	3,1%	53,1%	100,0%
		% du total	43,8%	3,1%	53,1%	100,0%

Chaque lettre en indice indique un sous-ensemble de NOMBRE D'EQD REALISES SUR L'INDICATEUR catégories dont les proportions de colonne ne diffèrent pas de manière significative les unes des autres au niveau ,05.

Tableau n° 6 : Tableau croisé QUALITE DES DONNEES EN BINAIRE * NOMBRE D'EQD REALISES SUR L'INDICATEUR

Tests du Khi-deux			
	Valeur	ddl	Signification asymptotique (bilatérale)
Khi-deux de Pearson	1,345a	2	,511
Rapport de vraisemblance	1,732	2	,421
Nombre d'observations valides	32		
a. 2 cellules (33,3%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de ,50.			

Tableau n° 7 : Tests du Khi-deux QUALITE DES DONNEES EN BINAIRE versus NOMBRE D'EQD REALISES SUR L'INDICATEUR

Mesures symétriques			
		Valeur	Signification approximée
Nominal par Nominal	Coefficient de contingence	,201	,511
Nombre d'observations valides		32	
a. L'hypothèse nulle n'est pas considérée.			
b. Utilisation de l'erreur standard asymptotique dans l'hypothèse nulle.			

Tableau n° 8 : Mesures symétriques des tests du Khi-deux QUALITE DES DONNEES EN BINAIRE versus NOMBRE D'EQD REALISES SUR L'INDICATEUR.

Tableau croisé FORMATION DES AGENTS DE COLLECTE * QUALITE DES DONNEES					
			FORMATION DES AGENTS DE COLLECTE		Total
			NON	OUI	
QUALITES DES DONNEES	1	Effectif	4a	9a	13
		% compris dans QUALITES DES DONNEES	30,8%	69,2%	100,0%
		% du total	12,5%	28,1%	40,6%
	2	Effectif	3a	0b	3
		% compris dans QUALITES DES DONNEES	100,0%	0,0%	100,0%
		% du total	9,4%	0,0%	9,4%

	3	Effectif	2a	3a	5
		% compris dans QUALITES DES DONNEES	40,0%	60,0%	100,0%
		% du total	6,2%	9,4%	15,6%
	4	Effectif	0a	4a	4
		% compris dans QUALITES DES DONNEES	0,0%	100,0%	100,0%
		% du total	0,0%	12,5%	12,5%
	5	Effectif	5a	2a	7
		% compris dans QUALITES DES DONNEES	71,4%	28,6%	100,0%
		% du total	15,6%	6,2%	21,9%
Total	Effectif	14	18	32	
	% compris dans QUALITES DES DONNEES	43,8%	56,2%	100,0%	
	% du total	43,8%	56,2%	100,0%	

Chaque lettre en indice indique un sous-ensemble de FORMATION DES AGENTS DE COLLECTE catégories dont les proportions de colonne ne diffèrent pas de manière significative les unes des autres au niveau ,05.

Tableau n°9 : Tableau croisé FORMATION DES AGENTS DE COLLECTE * QUALITE DES DONNEES

Tests du Khi-deux			
	Valeur	ddl	Signification asymptotique (bilatérale)
Khi-deux de Pearson	10,066a	4	,039
Rapport de vraisemblance	12,706	4	,013
Nombre d'observations valides	32		
a. 8 cellules (80,0%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de 1,31.			

Tableau n° 10 : Tests de Khi-deux FORMATION DES AGENTS DE COLLECTE * QUALITE DES DONNEES

Mesures symétriques			
		Valeur	Signification approximée
Nominal par Nominal	Coefficient de contingence	,489	,039
Nombre d'observations valides		32	
a. L'hypothèse nulle n'est pas considérée.			
b. Utilisation de l'erreur standard asymptotique dans l'hypothèse nulle.			
c. Les statistiques de corrélation ne sont proposées que pour les données numériques.			

Tableau n°11 : Tests du Khi-deux FORMATION DES AGENTS DE COLLECTE versus QUALITE DES DONNEES

Tableau croisé NOMBRE D'EQD REALISES SUR L'INDICATEUR * QUALITE DES DONNEES						
			NOMBRE D'EQD REALISES SUR L'INDICATEUR			Total
			0	1	2	
QUALITES DES DONNEES	1	Effectif	5a	0a	8a	13
		% compris dans QUALITES DES DONNEES	38,5%	0,0%	61,5%	100,0%
		% du total	15,6%	0,0%	25,0%	40,6%
	2	Effectif	3a	0a	0a	3
		% compris dans QUALITES DES DONNEES	100,0%	0,0%	0,0%	100,0%
		% du total	9,4%	0,0%	0,0%	9,4%
	3	Effectif	2a	0a	3a	5
		% compris dans QUALITES DES DONNEES	40,0%	0,0%	60,0%	100,0%
		% du total	6,2%	0,0%	9,4%	15,6%
	4	Effectif	0a	0a	4a	4
		% compris dans QUALITES DES DONNEES	0,0%	0,0%	100,0%	100,0%
		% du total	0,0%	0,0%	12,5%	12,5%
	5	Effectif	4a	1a	2a	7

		% compris dans QUALITES DES DONNEES	57,1%	14,3%	28,6%	100,0%
		% du total	12,5%	3,1%	6,2%	21,9%
Total		Effectif	14	1	17	32
		% compris dans QUALITES DES DONNEES	43,8%	3,1%	53,1%	100,0%
		% du total	43,8%	3,1%	53,1%	100,0%

Chaque lettre en indice indique un sous-ensemble de NOMBRE D'EQD REALISES SUR L'INDICATEUR catégories dont les proportions de colonne ne diffèrent pas de manière significative les unes des autres au niveau ,05.

Tableau n° 12 : NOMBRE D'EQD REALISES SUR L'INDICATEUR * QUALITE DES DONNEES

Tests du Khi-deux			
	Valeur	ddl	Signification asymptotique (bilatérale)
Khi-deux de Pearson	12,137a	8	,145
Rapport de vraisemblance	14,151	8	,078
Association linéaire par linéaire	,047	1	,828
Nombre d'observations valides	32		

a. 13 cellules (86,7%) ont un effectif théorique inférieur à 5. L'effectif théorique minimum est de ,09.

Tableau n° 13 : Tests du Khi-deux NOMBRE D'EQD REALISES SUR L'INDICATEUR * QUALITE DES DONNEES

Mesures symétriques					
		Valeur	Erreur standard asymptotique (a)	T approximé (b)	Signification approximée (c)
Nominal par Nominal	Coefficient de contingence	,524			,145
Intervalle par Intervalle	R de Pearson	-,039	,174	-,213	,832c

Ordinal par Ordinal	Corrélation de Spearman	-,071	,178	-,389	,700c
Nombre d'observations valides		32			
a. L'hypothèse nulle n'est pas considérée.					
b. Utilisation de l'erreur standard asymptotique dans l'hypothèse nulle.					
c. Basé sur une approximation normale.					

Tableau n°14 : Mesures symétriques Tests du Khi-deux NOMBRE D'EQD REALISES SUR L'INDICATEUR * QUALITE DES DONNEES

Tableau croisé QUALITE DES INDICATEURS * QUALITE DES DONNEES					
			QUALITE DES INDICATEURS		Total
			NON	OUI	
QUALITES DES DONNEES	1	Effectif	3a	10a	13
		% compris dans QUALITES DES DONNEES	23,1%	76,9%	100,0%
		% du total	9,4%	31,2%	40,6%
	2	Effectif	0a	3a	3
		% compris dans QUALITES DES DONNEES	0,0%	100,0%	100,0%
		% du total	0,0%	9,4%	9,4%
	3	Effectif	1a	4a	5
		% compris dans QUALITES DES DONNEES	20,0%	80,0%	100,0%
		% du total	3,1%	12,5%	15,6%
	4	Effectif	1a	3a	4
		% compris dans QUALITES DES DONNEES	25,0%	75,0%	100,0%
		% du total	3,1%	9,4%	12,5%
	5	Effectif	3a	4a	7
		% compris dans QUALITES DES DONNEES	42,9%	57,1%	100,0%
		% du total	9,4%	12,5%	21,9%

Total	Effectif	8	24	32
	% compris dans QUALITES DES DONNEES	25,0%	75,0%	100,0%
	% du total	25,0%	75,0%	100,0%

Chaque lettre en indice indique un sous-ensemble de QUALITE DES INDICATEURS catégories dont les proportions de colonne ne diffèrent pas de manière significative les unes des autres au niveau ,05.

Tableau n°15 : Tableau croisé QUALITE DES INDICATEURS * QUALITE DES DONNEES

Tests du Khi-deux			
	Valeur	ddl	Signification asymptotique (bilatérale)
Khi-deux de Pearson	2,283a	4	,684
Rapport de vraisemblance	2,881	4	,578
Nombre d'observations valides	32		

a. 8 cellules (80,0%) ont un effectif théorique inférieur à 5.
L'effectif théorique minimum est de ,75.

Tableau n°16 : Test du Khi-deux QUALITE DES INDICATEURS * QUALITE DES DONNEES

Mesures symétriques			
		Valeur	Signification approximée
Nominal par Nominal	Coefficient de contingence	,258	,684
Nombre d'observations valides		32	

a. L'hypothèse nulle n'est pas considérée.
b. Utilisation de l'erreur standard asymptotique dans l'hypothèse nulle.
c. Les statistiques de corrélation ne sont proposées que pour les données numériques.

Tableau n°17 : Mesures symétriques Test du Khi-deux QUALITE DES INDICATEURS * QUALITE DES DONNEES