

SCI6007 Méthode de recherche en sciences de l'information

Cours 9 : Statistiques descriptives (partie 1 de 2)

Exemples

Caractères quantitatifs continus ou discrets avec beaucoup de modalités [1/4]

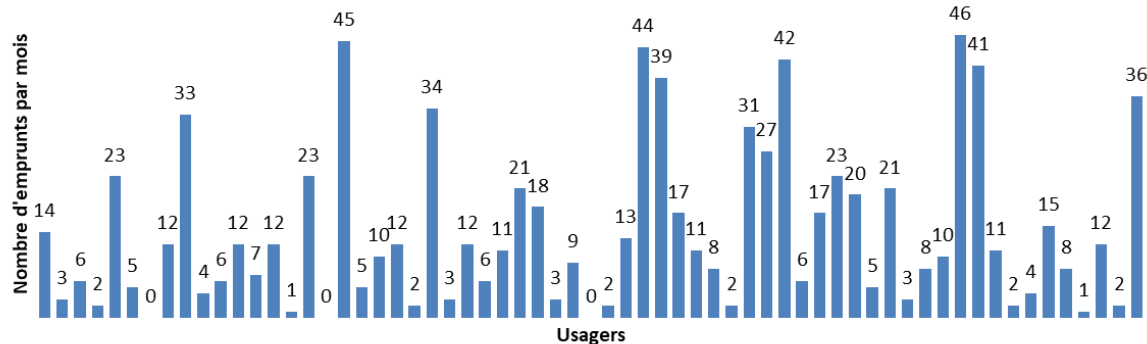
Caractères : nombre de documents empruntés sur une période de 3 mois par des usagers

Réponses obtenues (63 répondants) :

14	4	5	21	17	23	11
3	6	10	18	11	20	2
6	12	12	3	8	5	4
2	7	2	9	2	21	15
23	12	34	0	31	3	8
5	1	3	2	27	8	1
0	23	12	13	42	10	12
12	0	6	44	6	46	2
33	45	11	39	17	41	36

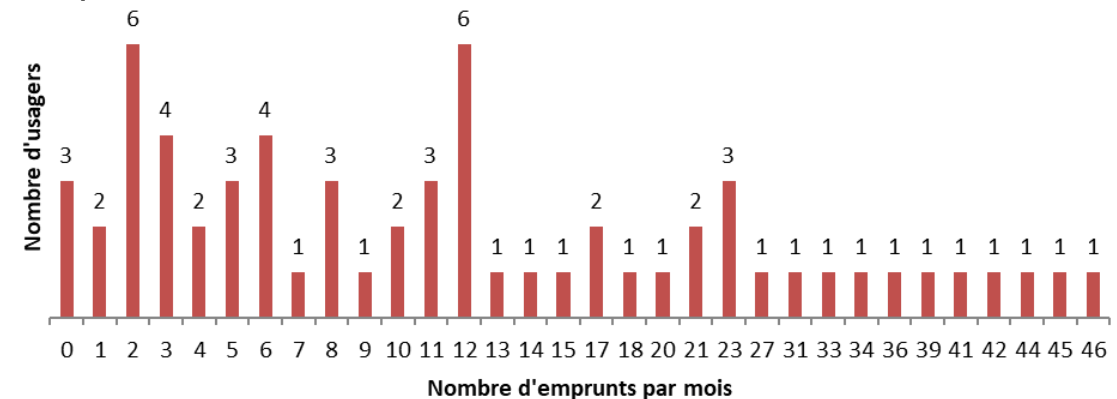
Problème 1 :

63 données ... Trop détaillé ...



Problème 2 :

63 données prenant 32 valeurs différentes ... Encore trop détaillé ...



Caractères quantitatifs continus ou discrets avec beaucoup de modalités [2/4]

Solution? Calculer non pas les fréquences (effectifs) par rapport aux modalités mais les calculer par rapport à des intervalles (classes)

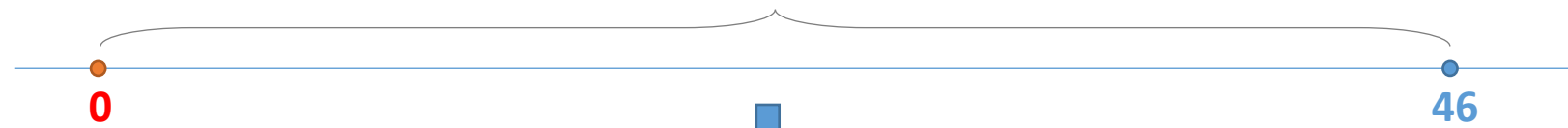
Caractères : nombre de documents empruntés sur une période de 3 mois par des usagers

14	4	5	21	17	23	11
3	6	10	18	11	20	2
6	12	12	3	8	5	4
2	7	2	9	2	21	15
23	12	34	0	31	3	8
5	1	3	2	27	8	1
0	23	12	13	42	10	12
12	0	6	44	6	46	2
33	45	11	39	17	41	36

minimum

maximum

Longueur (étendue) = 46



Étendue à découper en intervalle (classes)

Par ex., si on veut découper exactement en 7 * morceaux (=classes), chaque morceau mesurera $46/7 = 6,5714$



Longueur de la classe = 6,5714

* choix arbitraire (entre 5 et 15) ou basé sur une méthode mathématique (par exemple, la règle de Sturges : la valeur entière la plus proche de $1 + 3,3 \log_{10} n$ où n est la taille de l'échantillon)

Caractères quantitatifs continus ou discrets avec beaucoup de modalités [3/4]

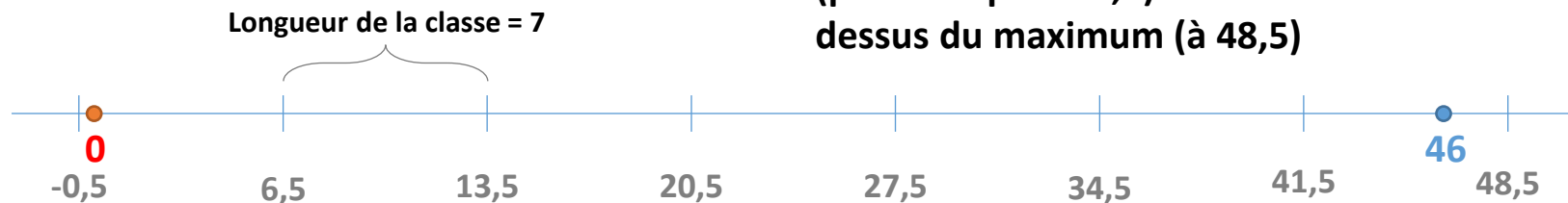
Solution? (suite)

Caractères : nombre de documents empruntés par mois par les usagers

14	4	5	21	17	23	11
3	6	10	18	11	20	2
6	12	12	3	8	5	4
2	7	2	9	2	21	15
23	12	34	0	31	3	8
5	1	3	2	27	8	1
0	23	12	13	42	10	12
12	0	6	44	6	46	2
33	45	11	39	17	41	36



Pour bien englober les extrêmes, on peut « surestimer » la longueur de classe, par exemple ici à 7 au lieu de 6,5714, ce qui permettra de commencer la première classe légèrement en dessous du minimum (par exemple à -0,5) et terminer la dernière classe légèrement au-dessus du maximum (à 48,5)



Ne reste ensuite qu'à calculer combien de modalités se retrouvent dans chacune des classes!

Caractères quantitatifs continus ou discrets avec beaucoup de modalités [4/4]*Solution? (suite et fin)***Caractères** : nombre de documents empruntés par mois par les usagers

14	4	5	21	17	23	11
3	6	10	18	11	20	2
6	12	12	3	8	5	4
2	7	2	9	2	21	15
23	12	34	0	31	3	8
5	1	3	2	27	8	1
0	23	12	13	42	10	12
12	0	6	44	6	46	2
33	45	11	39	17	41	36

Classes	Fréquences absolues	Fréquences relatives
-0,5 – 6,5	24	38,10%
6,5 – 13,5	17	26,98%
13,5 – 20,5	6	9,52%
20,5 – 27,5	6	9,52%
27,5 – 34,5	3	4,76%
34,5 – 41,5	3	4,76%
41,5 – 48,5	4	6,35%

Principes généraux pour la présentation

Principes visuels [1/2]

« Tout ce qui ne sert pas, nuit! »

Principes généraux pour la présentation

Principes visuels [2/2]

Illustration par un contre-exemple

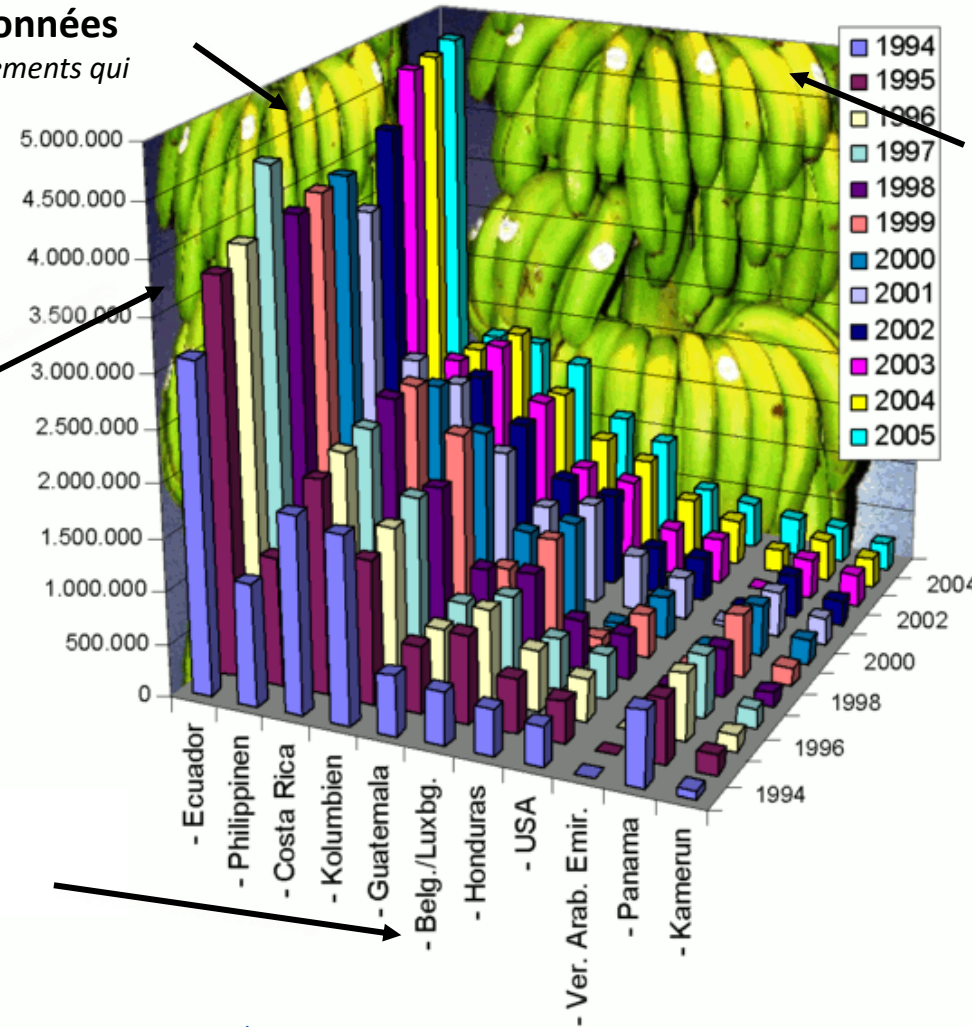
1. Montrer les données

Ce ne sont pas les ornements qui comptent...

4. Éviter rayures, carreaux, etc.

et les bananes...

5. Éviter les abréviations



2. Maximiser le ratio données-encre

Minimiser l'encre pour ce qui n'est pas une donnée...

3. Éviter la redondance

Le design du graphique doit en lui-même indiquer comment lire le graphique

Principes généraux pour la présentation

Principes organisationnels [1/2]

Tableau 1. Effectifs des différents corps d'emploi rémunérés à la Grande Bibliothèque de 2019 à 2022 (nombre de personnes à temps complet) *

Corps d'emploi	2019	2020	2021	2022
Bibliothécaire	81	81	81	82
Technicien.ne en documentation	49	49	49	50
Autre personnel rémunéré	178	178	160	164
Total	308	308	290	296

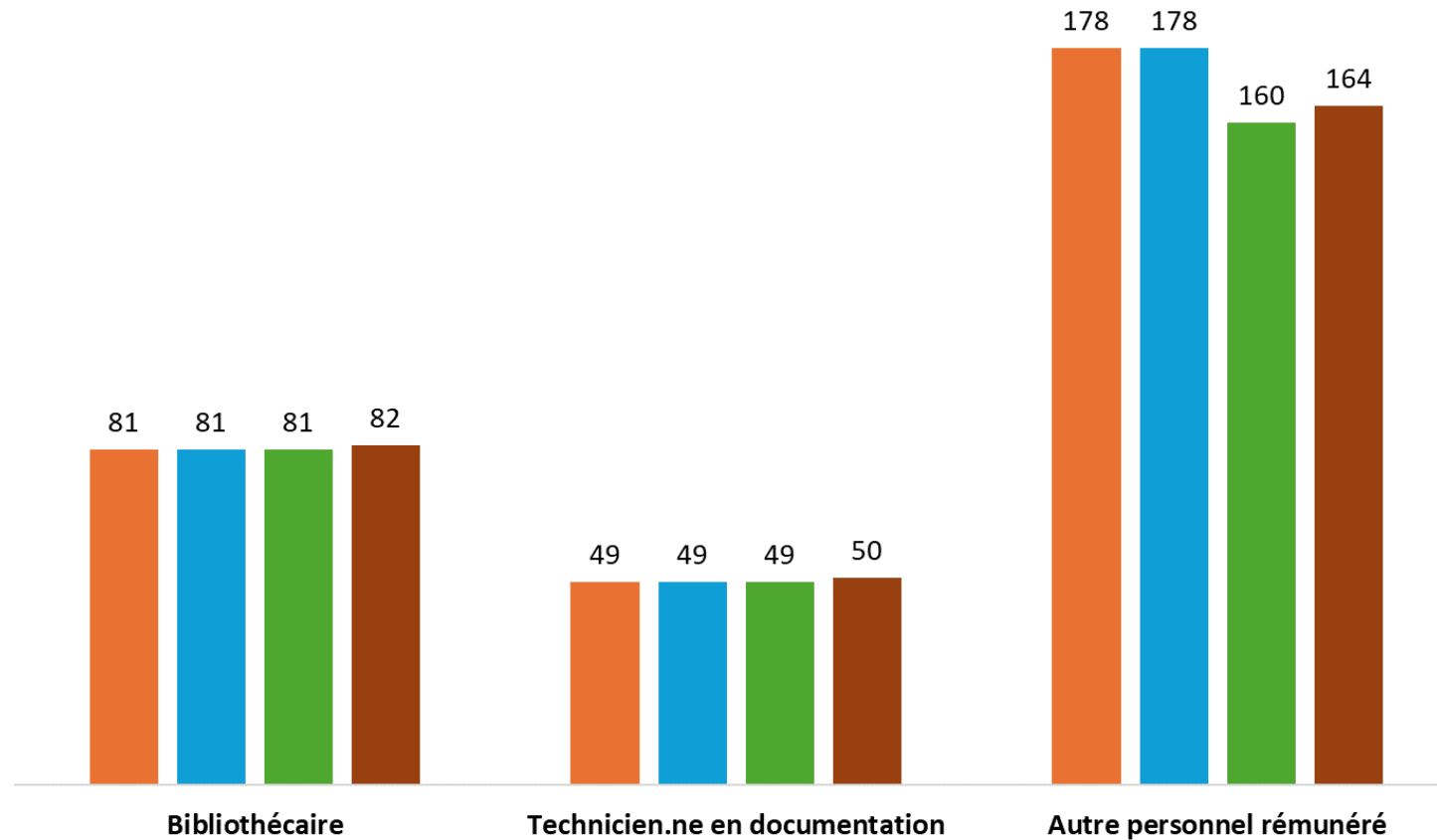
* Source des données : StatBib (<https://applications.banq.qc.ca/apex/f?p=306:1>)

Bien titrer les
tableaux

Principes généraux pour la présentation

Principes organisationnels [2/2]

Bien titrer les graphiques



Graphique 1. Effectif des différents corps d'emploi rémunérée à la Grande Bibliothèque de 2019 à 2022 (nombre de personnes à temps complet) *

Mise en forme des tableaux

Exemple * [1/2]

Tableau problématique

Prêts	Janvier	Février	Mars	Avril	Mai	Juin
Livres	3256	6523	5412	3526	4578	3598
CD	652	587	326	547	628	493
DVD	896	874	589	986	659	785
Total	4804	7984	6327	5059	5865	4876

Solution possible – Étape 1 : Éliminer le quadrillage inutile

Prêts	Janvier	Février	Mars	Avril	Mai	Juin
Livres	3256	6523	5412	3526	4578	3598
CD	652	587	326	547	628	493
DVD	896	874	589	986	659	785
Total	4804	7984	6327	5059	5865	4876

* Données fictives

Mise en forme des tableaux

Exemple * [2/2]

Tableau problématique

Prêts	Janvier	Février	Mars	Avril	Mai	Juin
Livres	3256	6523	5412	3526	4578	3598
CD	652	587	326	547	628	493
DVD	896	874	589	986	659	785
Total	4804	7984	6327	5059	5865	4876

Solution possible – Étape 3 : Aligner les chiffres à droite et utiliser le séparateur de milliers pour faciliter la lecture

Prêts	Janvier	Février	Mars	Avril	Mai	Juin
Livres	3 256	6 523	5 412	3 526	4 578	3 598
CD	652	587	326	547	628	493
DVD	896	874	589	986	659	785
Total	4 804	7 984	6 327	5 059	5 865	4 876

* Données fictives

(premier niveau : aspect visuel)

Exercice : Qu'est-ce qui ne va pas?

Tableau 2. Utilisation d'Internet, par groupe d'âge et par province (%)

Province	Groupe d'âge	2018	2020	2022
Alberta	15 à 24 ans	100	100	100
Alberta	25 à 44 ans	97,3	97,6	98,4
Alberta	45 à 64 ans	94,8	96,8	95,9
Alberta	65 ans et plus	79,5	83	84,3
Colombie-Britannique	15 à 24 ans	96,7	99,5	99
Colombie-Britannique	25 à 44 ans	99,3	98,9	99,9
Colombie-Britannique	45 à 64 ans	96,5	95,9	98,4
Colombie-Britannique	65 ans et plus	81,1	81,8	89,6
Île-du-Prince-Édouard	15 à 24 ans	100	96,3	100
Île-du-Prince-Édouard	25 à 44 ans	98,7	98,5	100
Île-du-Prince-Édouard	45 à 64 ans	88,1	94,7	94,2
Île-du-Prince-Édouard	65 ans et plus	69	74,5	77,8
Manitoba	15 à 24 ans	99,5	95,3	100
Manitoba	25 à 44 ans	97,2	97,4	99,3
Manitoba	45 à 64 ans	92,8	91,7	93
Manitoba	65 ans et plus	70,2	76,5	76,6
Nouveau-Brunswick	15 à 24 ans	100	100	100
Nouveau-Brunswick	25 à 44 ans	98,1	99	99,4
Nouveau-Brunswick	45 à 64 ans	91,7	93,1	95
Nouveau-Brunswick	65 ans et plus	69,3	73,8	73,4

Note: Le visuel n'est pas celui du site de Statistique Canada

Exercice : Exemple de solution aux problèmes visuels

Tableau 2. Utilisation d'Internet, par groupe d'âge et par province (%)

Province	Groupe d'âge	2018	2020	2022
Alberta	15 à 24 ans	100,0	100,0	100,0
Alberta	25 à 44 ans	97,3	97,6	98,4
Alberta	45 à 64 ans	94,8	96,8	95,9
Alberta	65 ans et plus	79,5	83,0	84,3
Colombie-Britannique	15 à 24 ans	96,7	99,5	99,0
Colombie-Britannique	25 à 44 ans	99,3	98,9	99,9
Colombie-Britannique	45 à 64 ans	96,5	95,9	98,4
Colombie-Britannique	65 ans et plus	81,1	81,8	89,6
Île-du-Prince-Édouard	15 à 24 ans	100,0	96,3	100,0
Île-du-Prince-Édouard	25 à 44 ans	98,7	98,5	100,0
Île-du-Prince-Édouard	45 à 64 ans	88,1	94,7	94,2
Île-du-Prince-Édouard	65 ans et plus	69,0	74,5	77,8
Manitoba	15 à 24 ans	99,5	95,3	100,0
Manitoba	25 à 44 ans	97,2	97,4	99,3
Manitoba	45 à 64 ans	92,8	91,7	93,0
Manitoba	65 ans et plus	70,2	76,5	76,6
Nouveau-Brunswick	15 à 24 ans	100,0	100,0	100,0
Nouveau-Brunswick	25 à 44 ans	98,1	99,0	99,4
Nouveau-Brunswick	45 à 64 ans	91,7	93,1	95,0
Nouveau-Brunswick	65 ans et plus	69,3	73,8	73,4

Note: Le visuel n'est pas celui du site de Statistique Canada

Exercice : Autre exemple de solution aux problèmes visuels

Tableau 2. Utilisation d'Internet, par groupe d'âge et par province (%)

Province	Groupe d'âge	2018	2020	2022
Alberta	15 à 24 ans	100,0	100,0	100,0
Alberta	25 à 44 ans	97,3	97,6	98,4
Alberta	45 à 64 ans	94,8	96,8	95,9
Alberta	65 ans et plus	79,5	83,0	84,3
Colombie-Britannique	15 à 24 ans	96,7	99,5	99,0
Colombie-Britannique	25 à 44 ans	99,3	98,9	99,9
Colombie-Britannique	45 à 64 ans	96,5	95,9	98,4
Colombie-Britannique	65 ans et plus	81,1	81,8	89,6
Île-du-Prince-Édouard	15 à 24 ans	100,0	96,3	100,0
Île-du-Prince-Édouard	25 à 44 ans	98,7	98,5	100,0
Île-du-Prince-Édouard	45 à 64 ans	88,1	94,7	94,2
Île-du-Prince-Édouard	65 ans et plus	69,0	74,5	77,8
Manitoba	15 à 24 ans	99,5	95,3	100
Manitoba	25 à 44 ans	97,2	97,4	99,3
Manitoba	45 à 64 ans	92,8	91,7	93,0
Manitoba	65 ans et plus	70,2	76,5	76,6
Nouveau-Brunswick	15 à 24 ans	100,0	100,0	100,0
Nouveau-Brunswick	25 à 44 ans	98,1	99,0	99,4
Nouveau-Brunswick	45 à 64 ans	91,7	93,1	95,0
Nouveau-Brunswick	65 ans et plus	69,3	73,8	73,4

Suite de l'exercice : Qu'est-ce qui ne va pas? (deuxième niveau : aspect organisationnel)

Note: Le visuel n'est pas celui du site de Statistique Canada

Source des données : Statistique Canada. [Tableau 22-10-0135-01 Utilisation d'Internet, par province et par groupe d'âge](https://doi.org/10.25318/2210013501-fra)

DOI : <https://doi.org/10.25318/2210013501-fra>

Exercice : Exemple de solution aux problèmes organisationnels

Tableau 2. Utilisation d'Internet, par groupe d'âge et par province (%)

Groupe d'âge	Province	2018	2020	2022
15 à 24 ans	Alberta	100,0	100,0	100,0
	Colombie-Britannique	96,7	99,5	99,0
	Île-du-Prince-Édouard	100,0	96,3	100,0
	Manitoba	99,5	95,3	100,0
	Nouveau-Brunswick	100,0	100,0	100,0
25 à 44 ans	Alberta	97,3	97,6	98,4
	Colombie-Britannique	99,3	98,9	99,9
	Île-du-Prince-Édouard	98,7	98,5	100,0
	Manitoba	97,2	97,4	99,3
	Nouveau-Brunswick	98,1	99,0	99,4
45 à 64 ans	Alberta	94,8	96,8	95,9
	Colombie-Britannique	96,5	95,9	98,4
	Île-du-Prince-Édouard	88,1	94,7	94,2
	Manitoba	92,8	91,7	93,0
	Nouveau-Brunswick	91,7	93,1	95,0
65 ans et plus	Alberta	79,5	83,0	84,3
	Colombie-Britannique	81,1	81,8	89,6
	Île-du-Prince-Édouard	69,0	74,5	77,8
	Manitoba	70,2	76,5	76,6
	Nouveau-Brunswick	69,3	73,8	73,4

Note: Le visuel n'est pas celui du site de Statistique Canada

Exercice : Autre solution aux problèmes organisationnels

Tableau 2. Utilisation d'Internet, par groupe d'âge et par province (%)

Visuel du site de
Statistique Canada

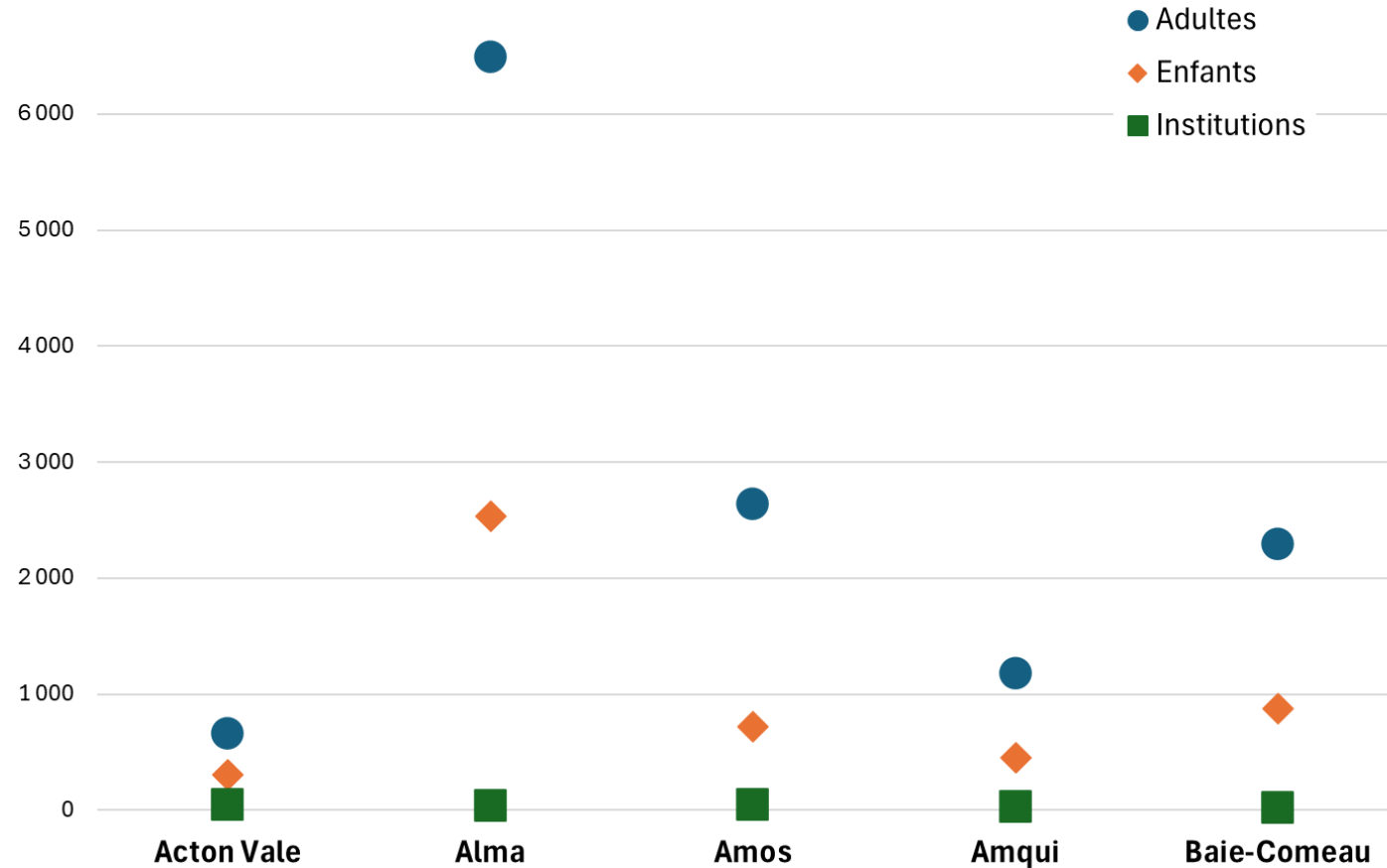
Utilisation d'Internet à partir de n'importe quel endroit ^{3, 4}		Utilisation d'Internet		
Géographie	Groupe d'âge	2018	2020	2022
		Pourcent		
Canada (carte)	Total, 15 ans et plus	91,3 ^A	92,3 ^A	94,5 ^A
	15 à 24 ans	98,6 ^A	98,0 ^A	99,2 ^A
	25 à 44 ans	98,2 ^A	98,2 ^A	99,2 ^A
	45 à 64 ans	93,9 ^A	94,2 ^A	96,4 ^A
	65 ans et plus	71,2 ^A	76,3 ^A	82,6 ^A
Terre-Neuve-et-Labrador (carte)	Total, 15 ans et plus	86,0 ^A	89,8 ^A	91,7 ^A
	15 à 24 ans	92,5 ^A	100,0 ^A	100,0 ^A
	25 à 44 ans	97,0 ^A	99,7 ^A	99,2 ^A
	45 à 64 ans	89,4 ^A	92,9 ^A	93,7 ^A
	65 ans et plus	64,7 ^A	70,9 ^A	79,2 ^A
Île-du-Prince-Édouard (carte)	Total, 15 ans et plus	88,5 ^A	91,3 ^A	92,6 ^A
	15 à 24 ans	100,0 ^A	96,3 ^A	100,0 ^A
	25 à 44 ans	98,7 ^A	98,5 ^A	100,0 ^A
	45 à 64 ans	88,1 ^A	94,7 ^A	94,2 ^A

Source des données : Statistique Canada. [Tableau 22-10-0135-01 Utilisation d'Internet, par province et par groupe d'âge](#)

DOI : <https://doi.org/10.25318/2210013501-fra>

Graphiques

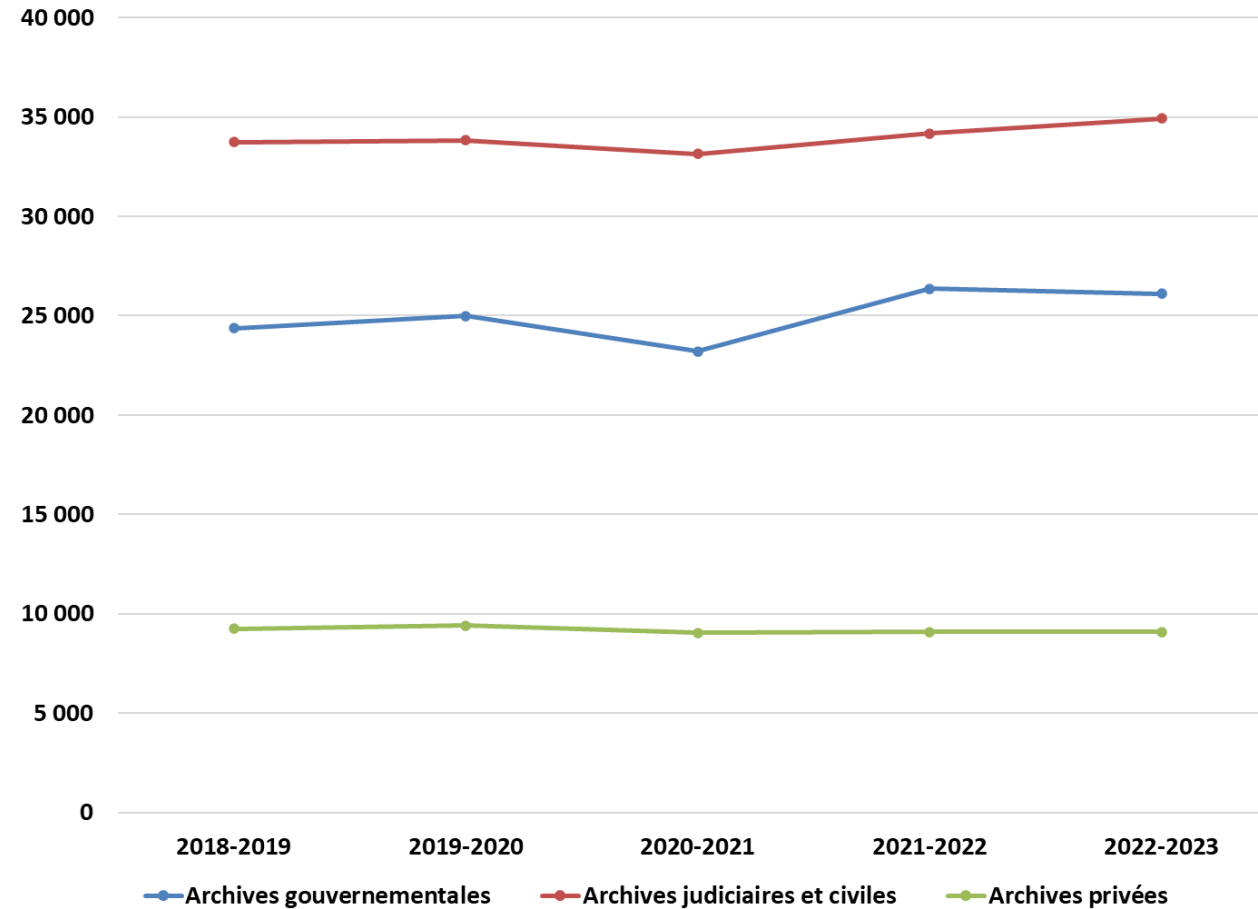
Types : Graphique en points



Graphique 2. Nombre d'inscription dans les bibliothèques publiques autonomes de 5 villes du Québec (2022)

Graphiques

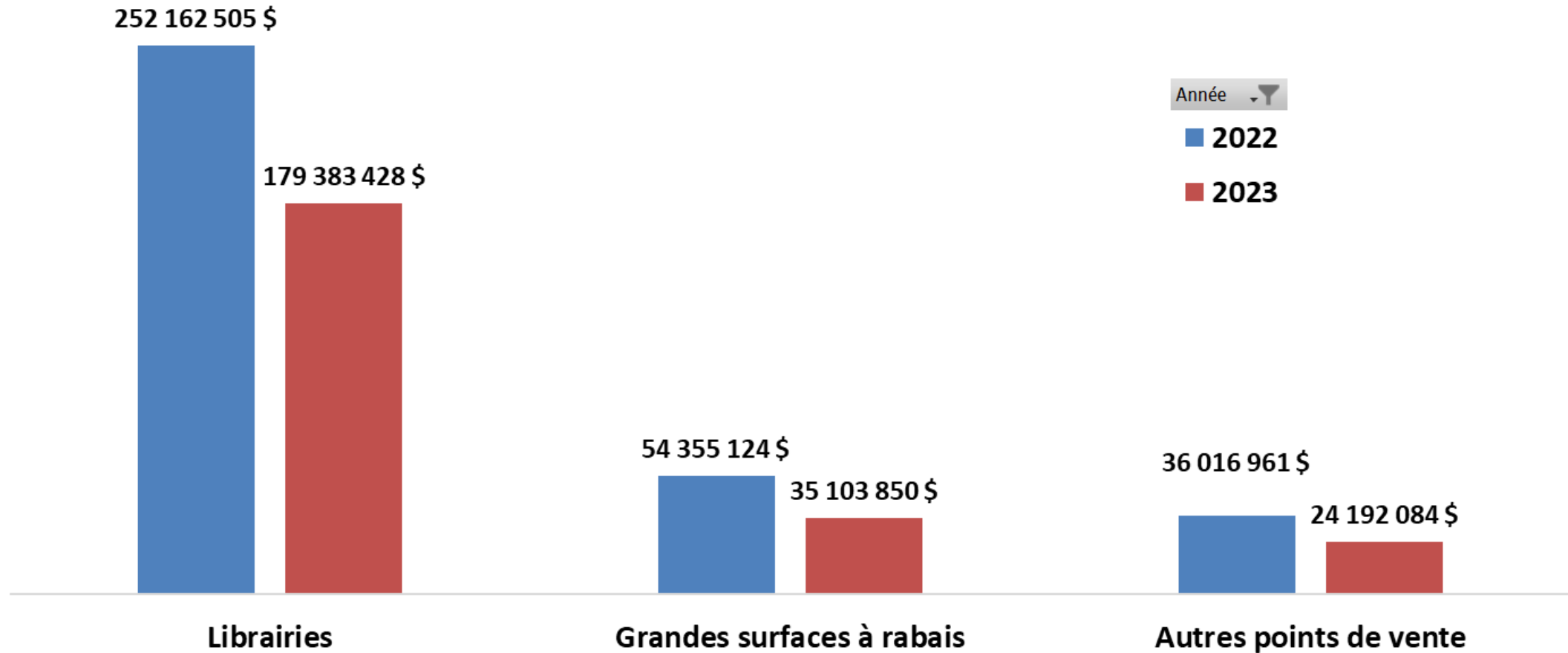
Types : Graphique en lignes (courbes)



Graphique 3. États des collections des centres et services d'archives au Québec pour certains types de collections (mètres linéaires)

Graphiques

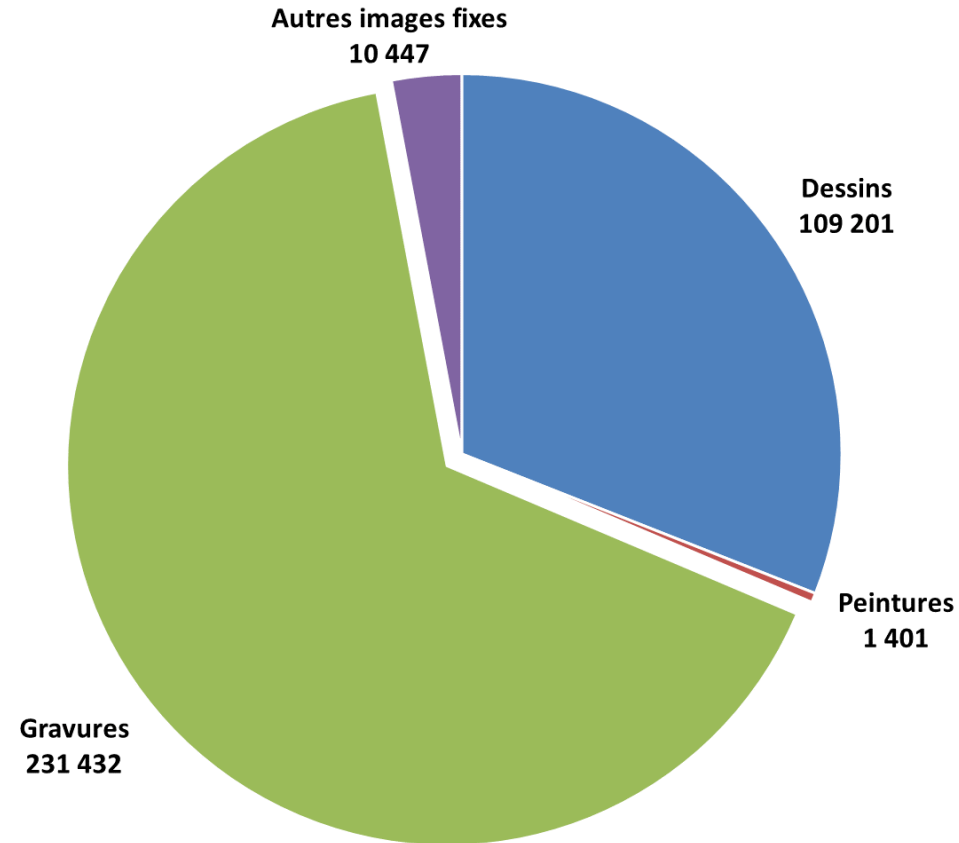
Types : Graphique en bâtons



Graphique 4. Ventes de livres neufs au Québec selon différents types de points de vente, 2022-2023 (\$)

Graphiques

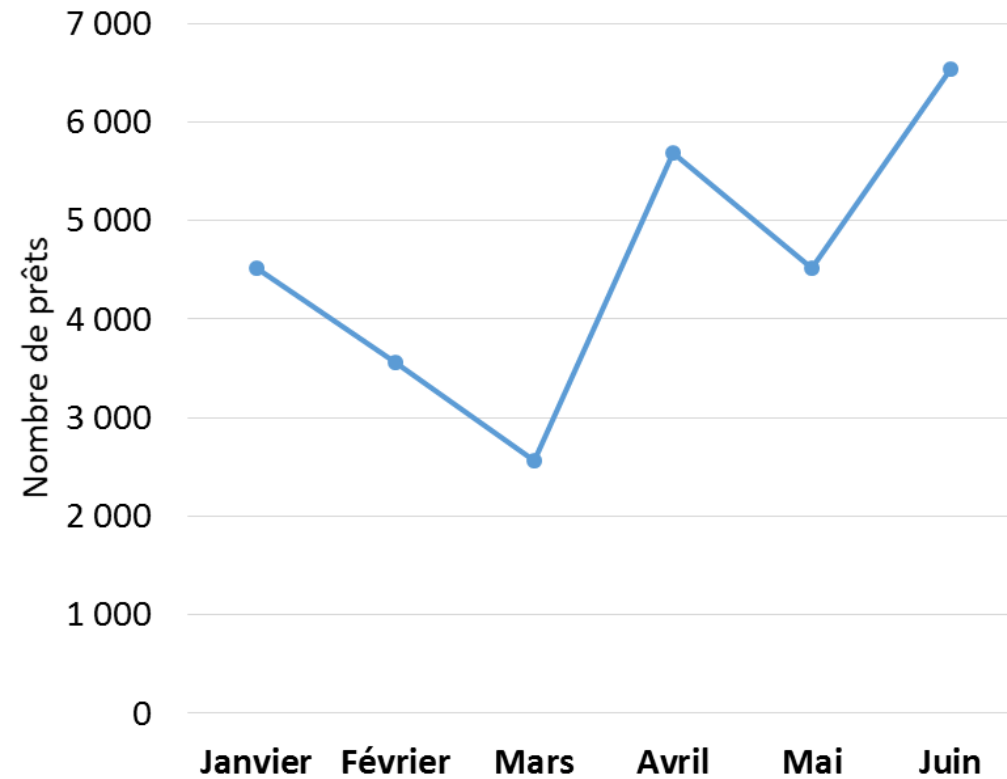
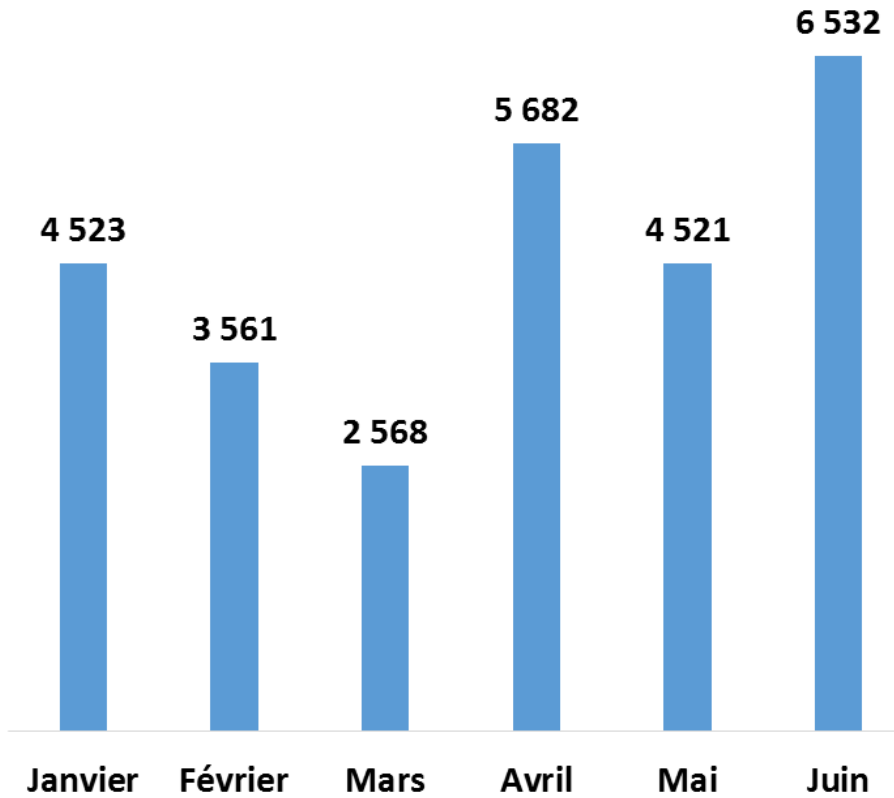
Types : Graphique en secteurs



Graphique 5. États des collections de documents iconographiques des centres et services d'archives au Québec, 2022-2023 (nombre d'items)

Graphiques

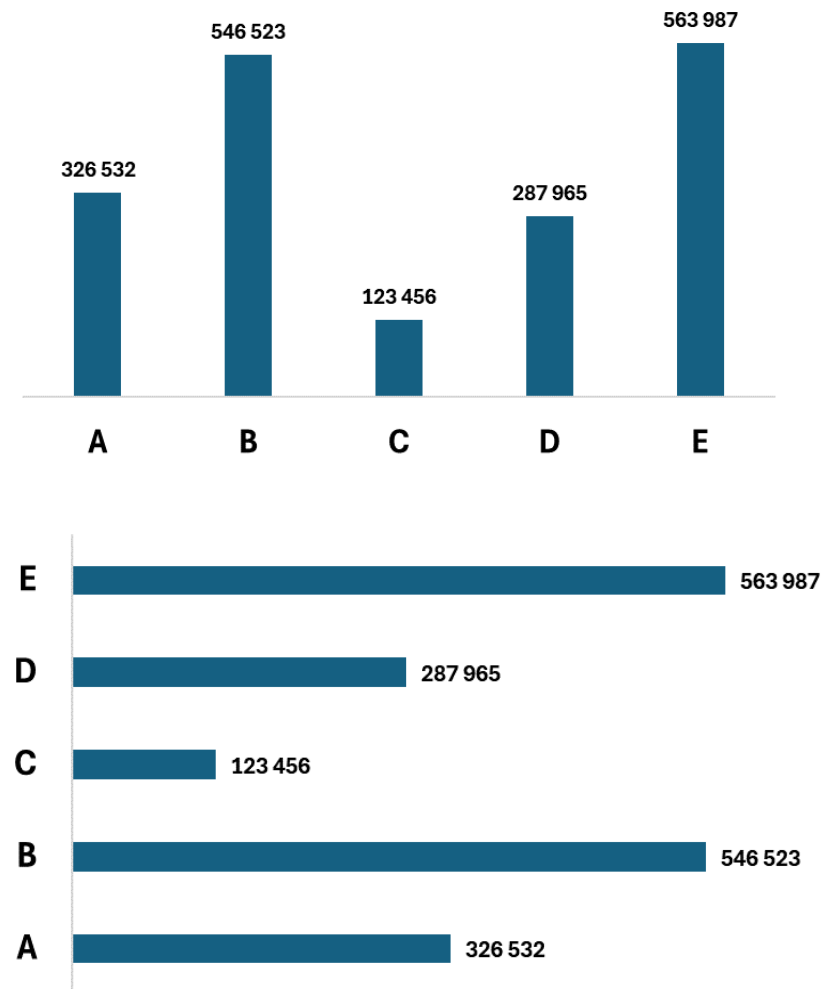
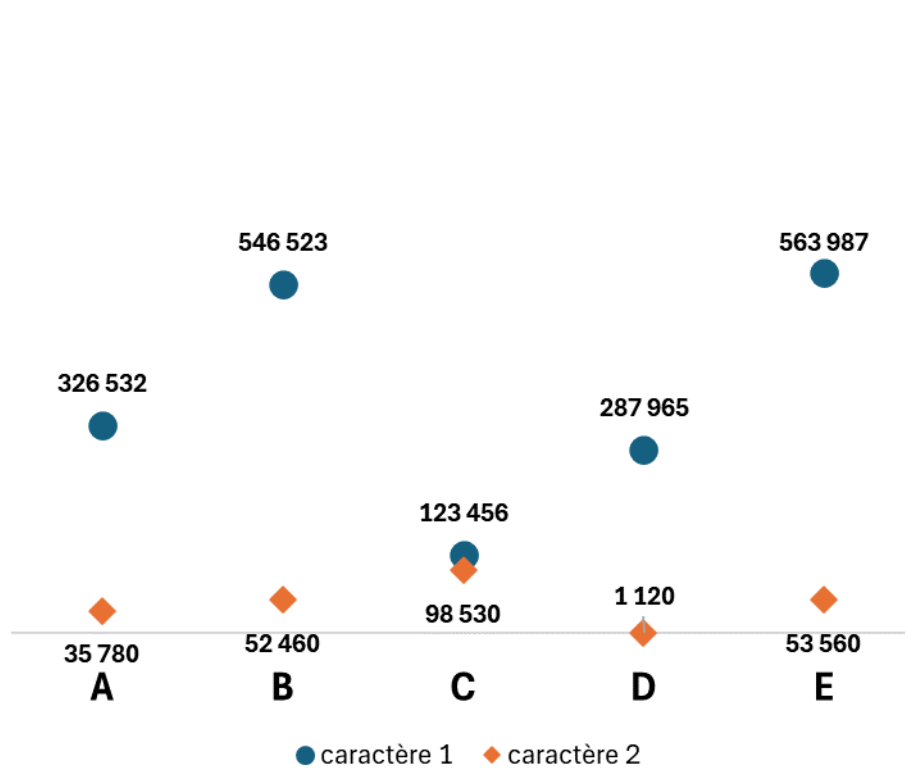
Montrer les relations : Série temporelle *



* Données fictives

Graphiques

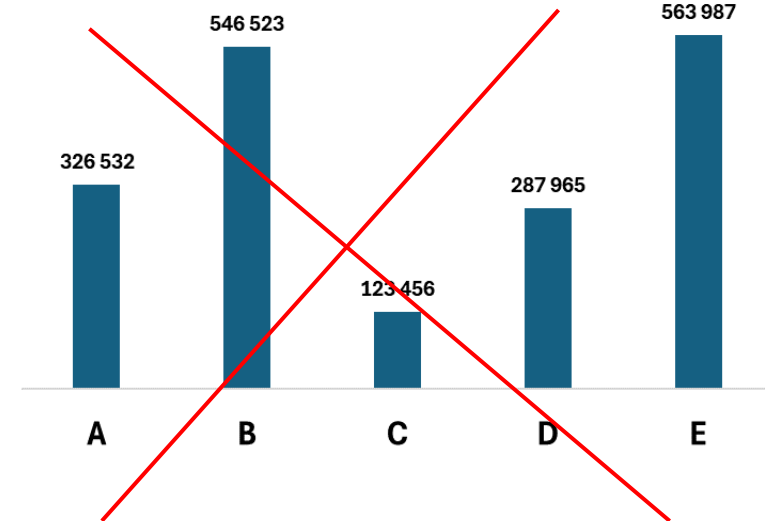
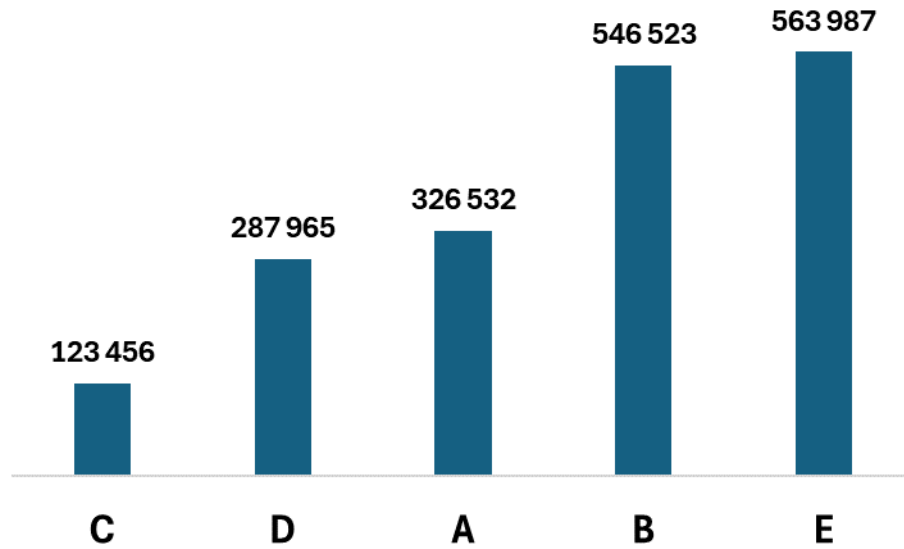
Montrer les relations : Comparaison nominale *



* Données fictives

Graphiques

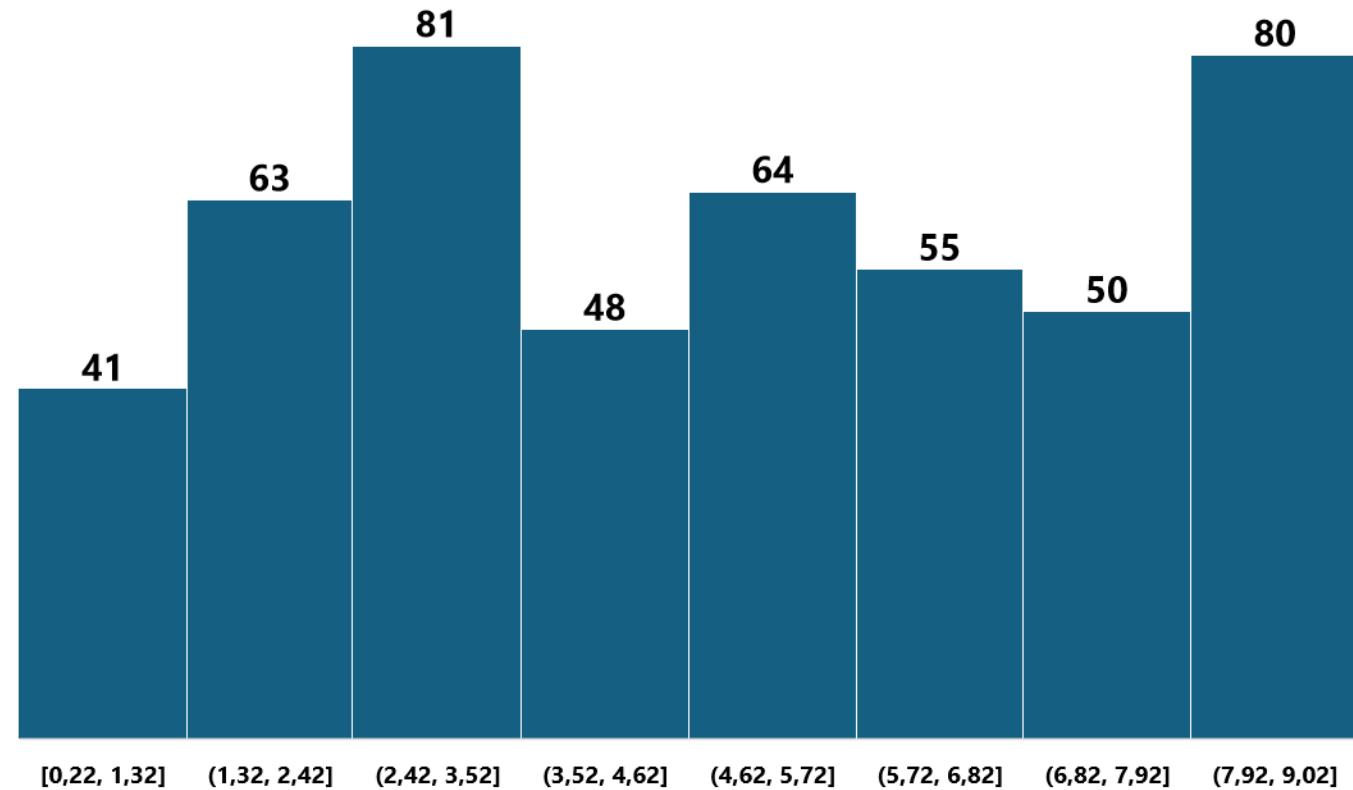
Montrer les relations : Classement *



* Données fictives

Graphiques

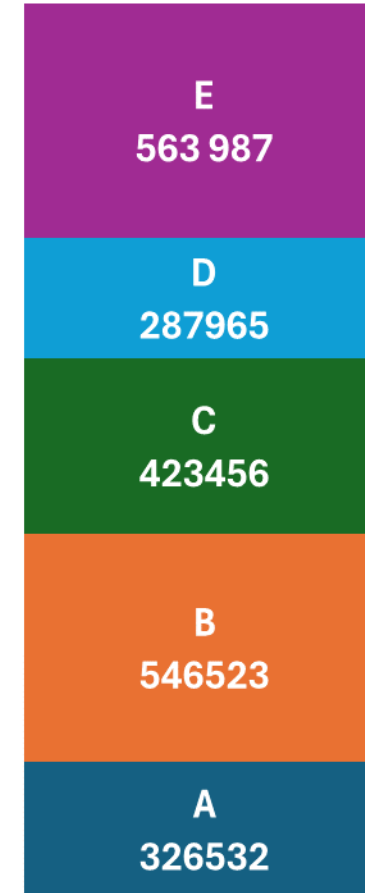
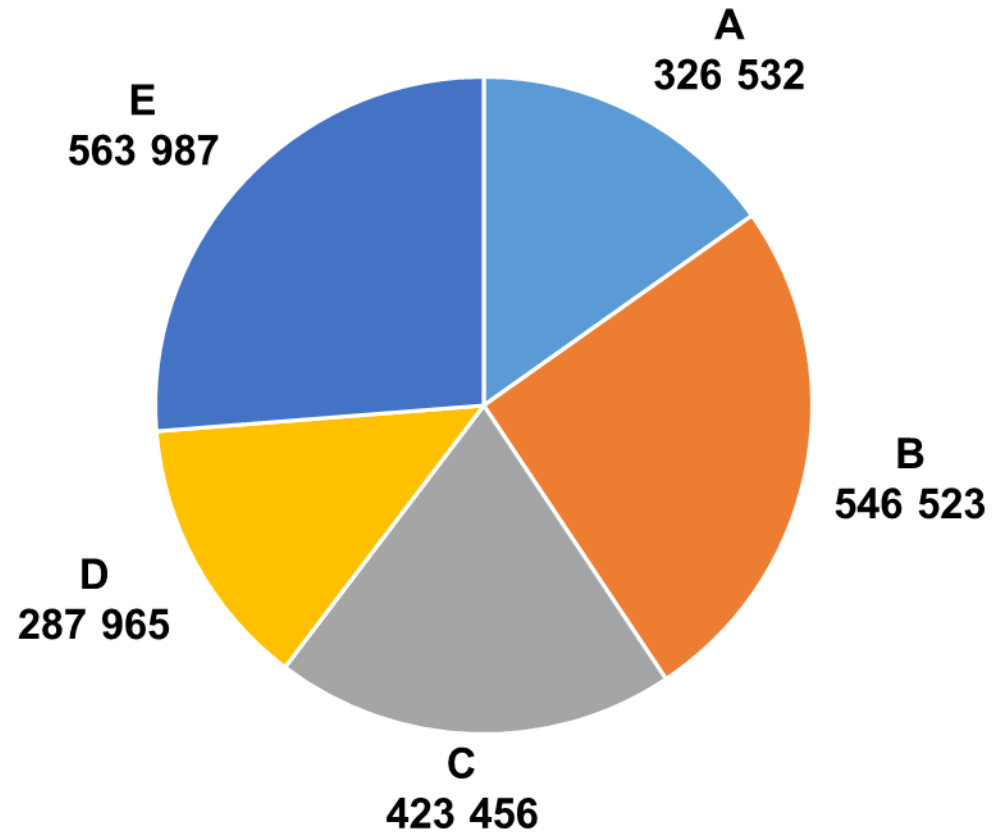
Montrer les relations : Distribution *



* Données fictives

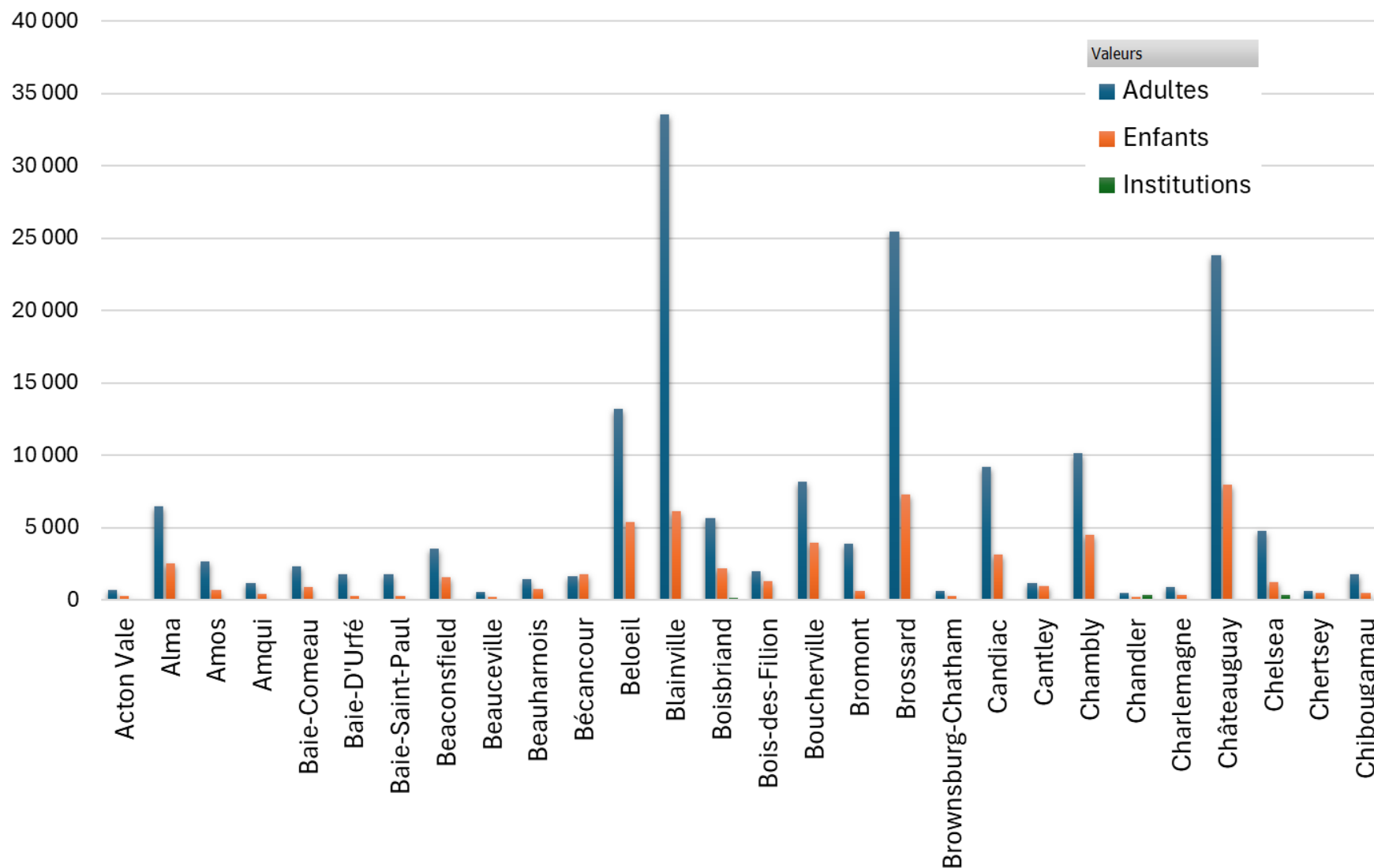
Graphiques

Montrer les relations : Proportions *



* Données fictives

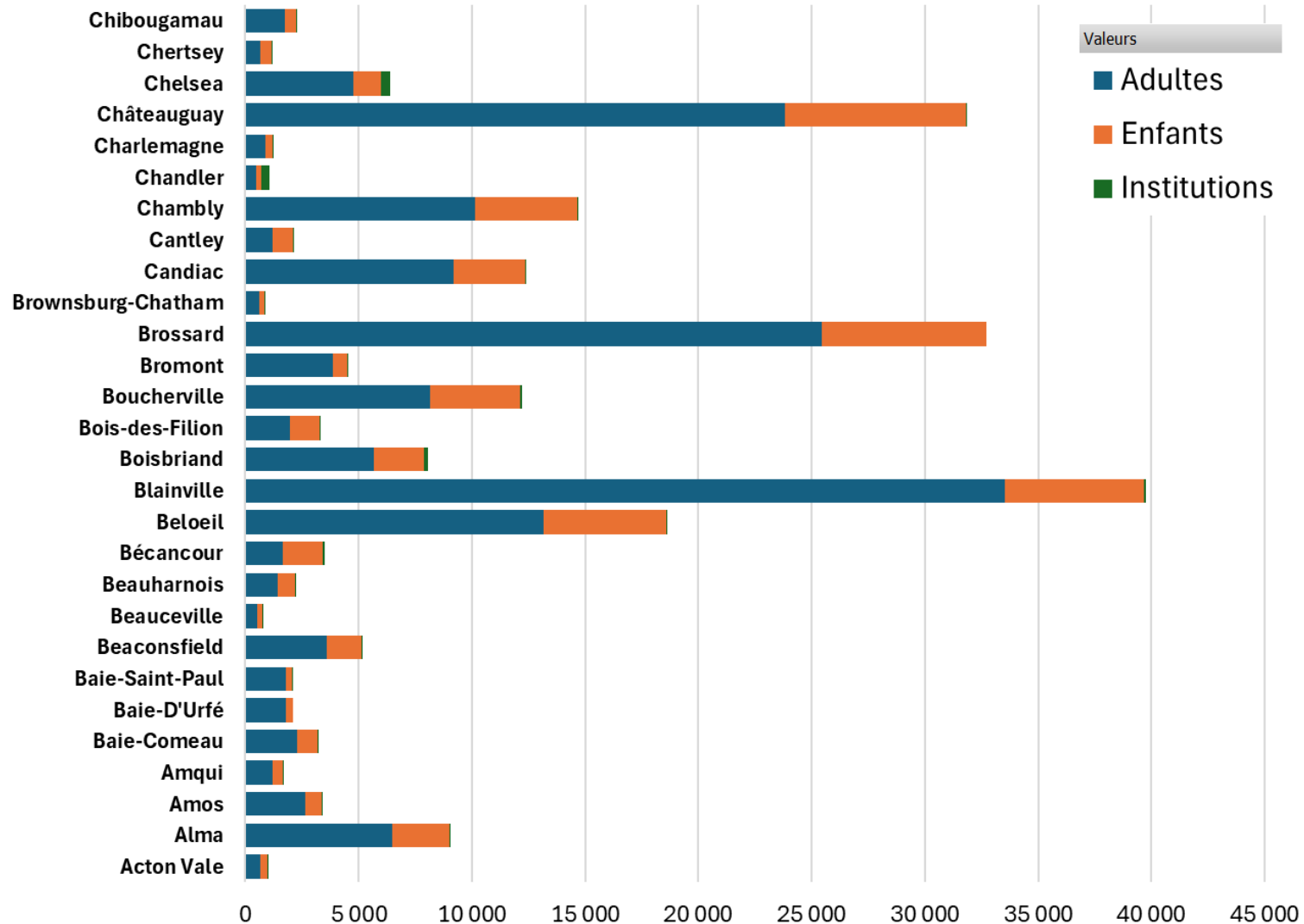
Habillage : Contre-exemple 1 [1/2]



Graphique 6. Nombre d'inscription dans les bibliothèques publiques autonomes de 28 villes du Québec (2022) *

* Source des données : StatBib (<https://applications.banq.qc.ca/apex/f?p=306:1>)

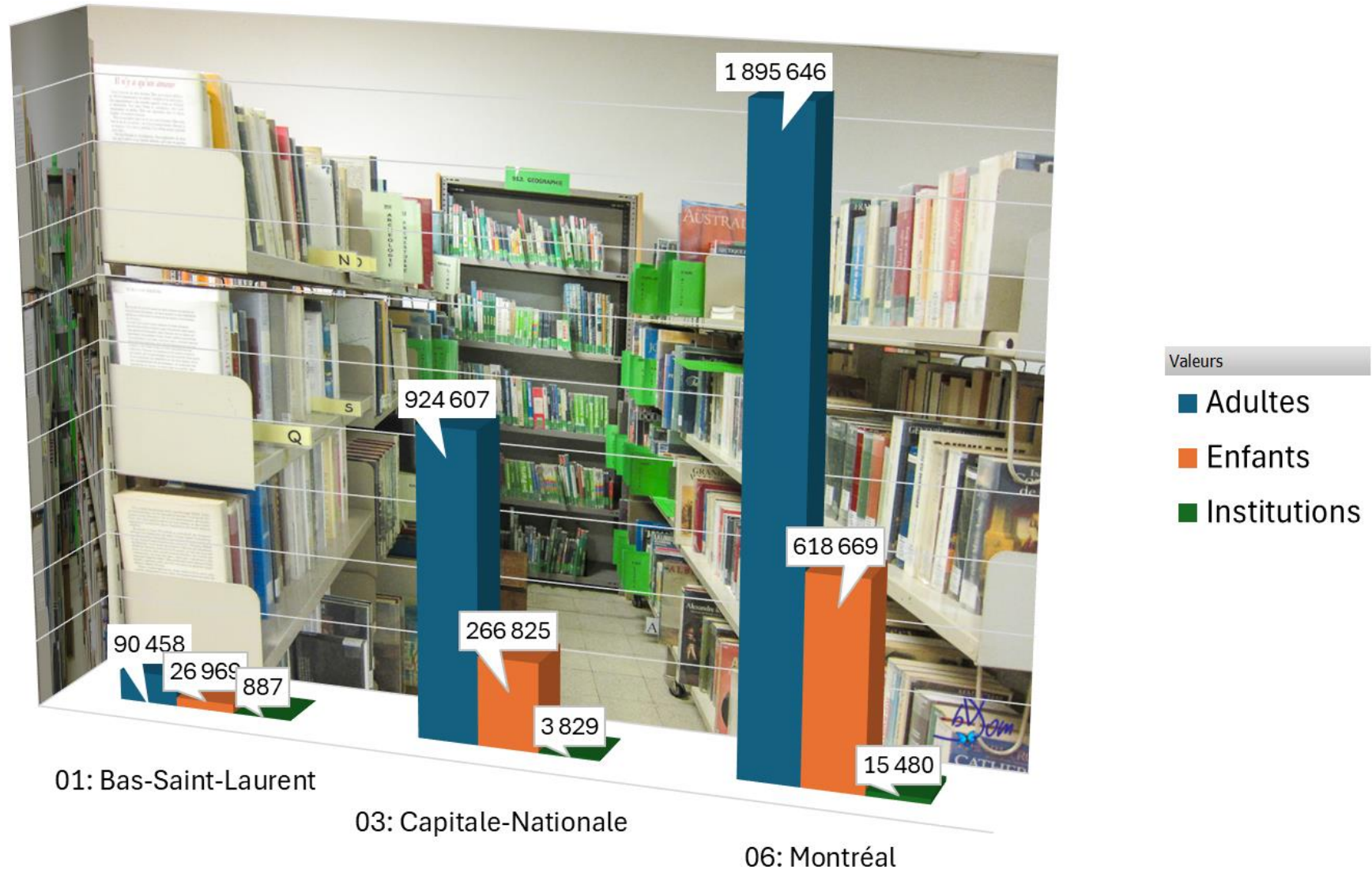
Habillage : Contre-exemple 1 / Solution



Graphique 6. Nombre d'inscription dans les bibliothèques publiques autonomes de 28 villes du Québec (2022) *

* Source des données : StatBib (<https://applications.banq.qc.ca/apex/f?p=306:1>)

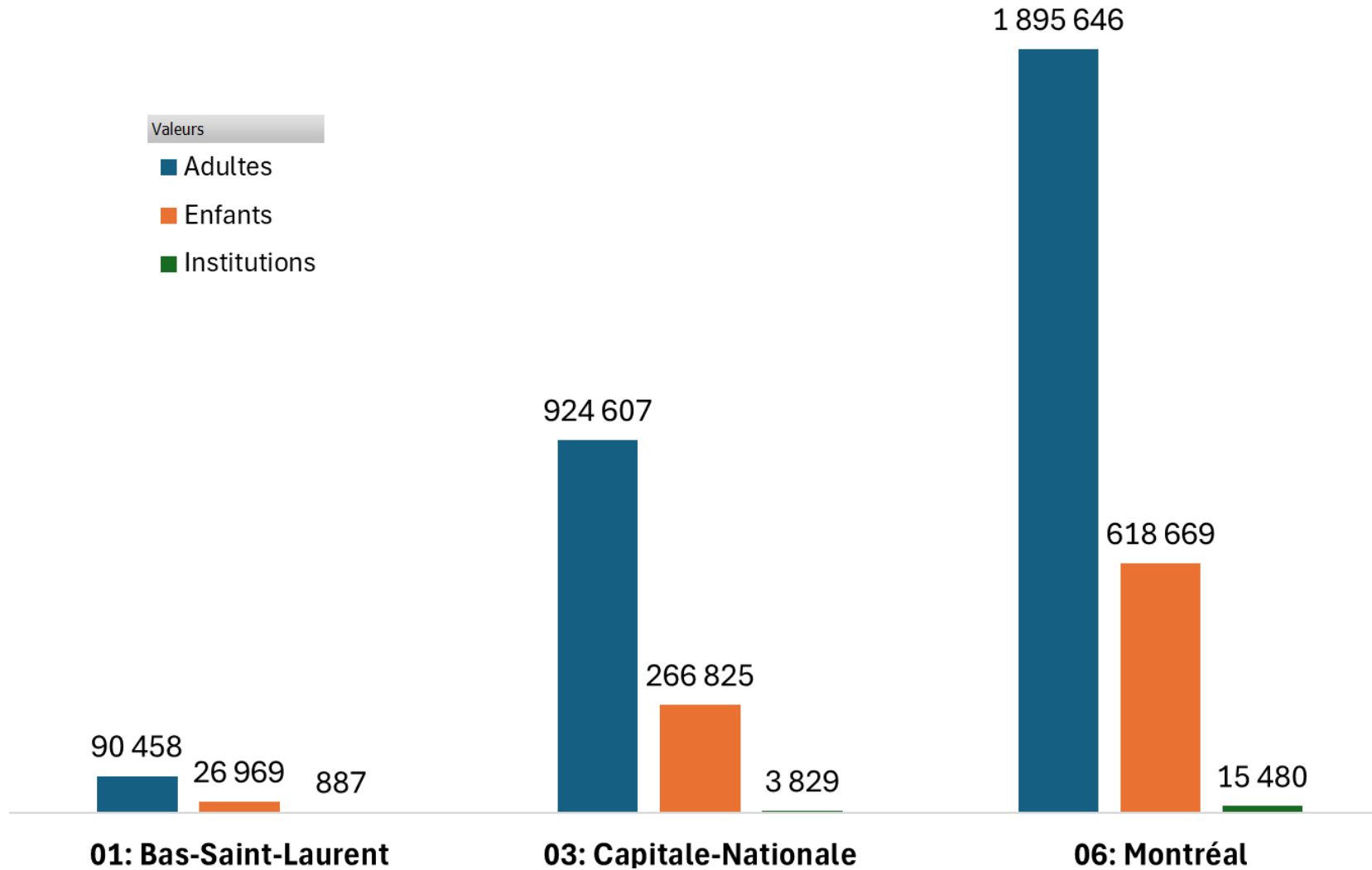
Habillage : Contre-exemple 2



Graphique 7. Nombre d'inscription dans les bibliothèques publiques autonomes de 3 régions administratives du Québec (2022) *

* Source des données : StatBib (<https://applications.banq.qc.ca/apex/f?p=306:1>)

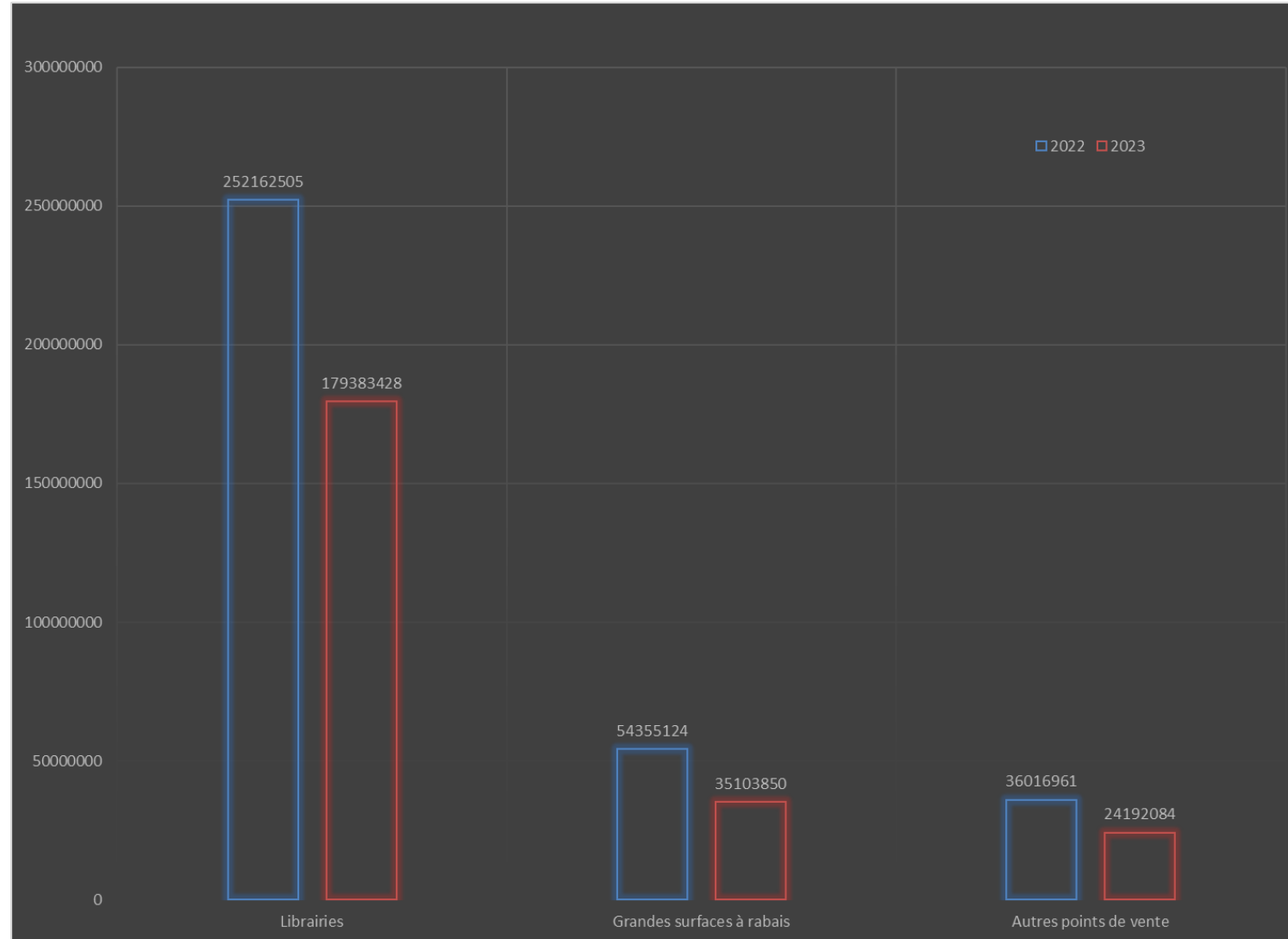
Habillage : Contre-exemple 2 / Solution



Graphique 7. Nombre d'inscription dans les bibliothèques publiques autonomes de 3 régions administratives du Québec (2022) *

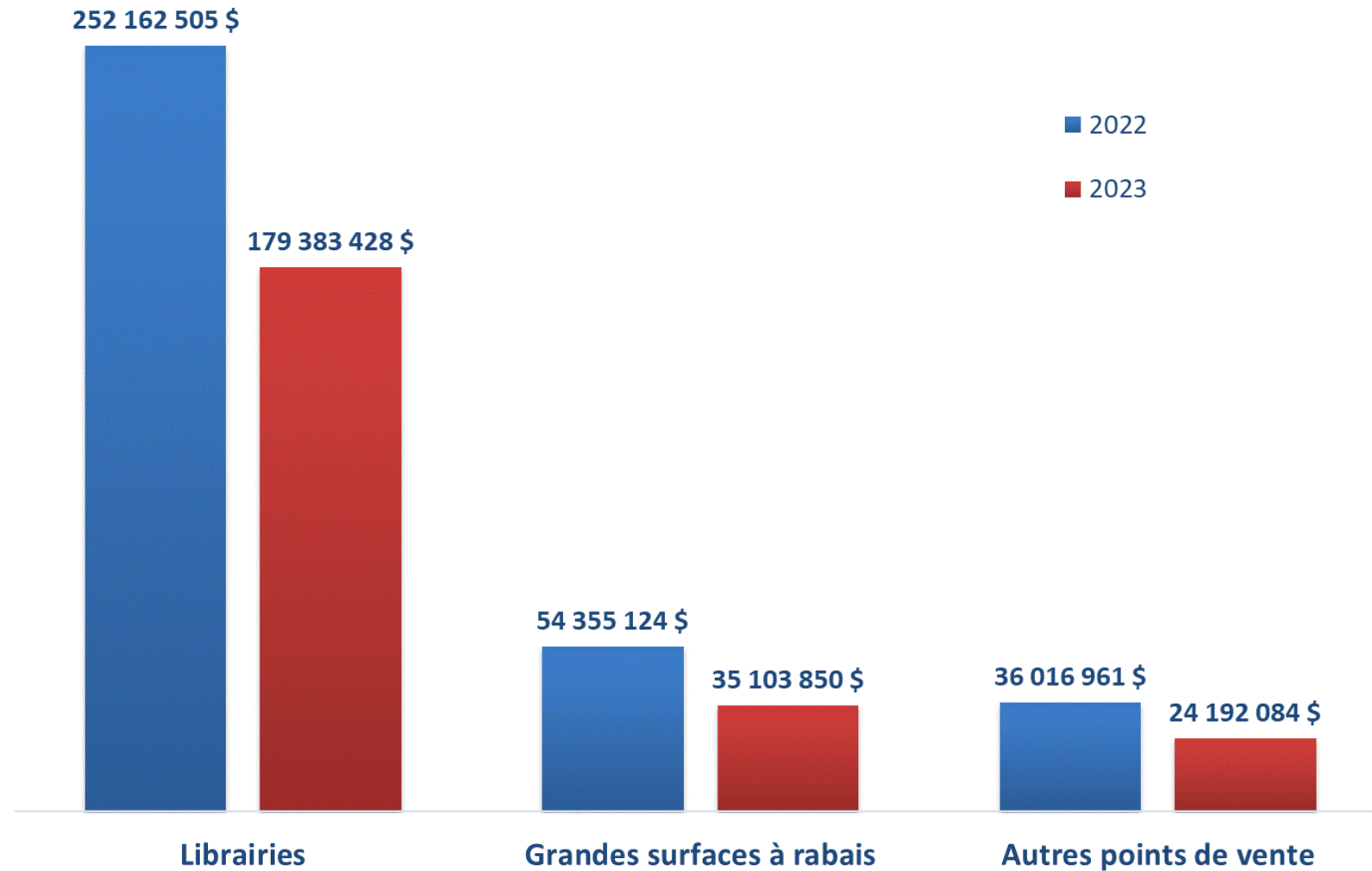
* Source des données : StatBib (<https://applications.banq.qc.ca/apex/f?p=306:1>)

Habillage : Exercice



Graphique 8. Ventes de livres neufs au Québec selon différents types de points de vente, 2022-2023 (\$)

Habillage : Exercice / Solution



Graphique 8. Ventes de livres neufs au Québec selon différents types de points de vente, 2022-2023 (\$)