

Communiquer au moyen des données du recensement : Visualisation des données

Sélection de sujets relatifs aux recensements internationaux¹

Publié en Juin 2017

INTRODUCTION

Cette note technique présente des pratiques efficaces de présentation et de visualisation des données de recensement et d'enquête produites par les organismes nationaux de statistique (NSO, selon l'acronyme anglais).

La visualisation des données peut aider les NSO à atteindre un public plus large, à communiquer plus efficacement des informations prioritaires et à découvrir des modèles non identifiés. Des éléments visuels peuvent être créés facilement à l'aide de logiciels de productivité communs, mais l'élaboration d'éléments visuels professionnels et bien conçus demande un effort supplémentaire. En fin de compte, les efforts supplémentaires en valent la peine puisque des éléments visuels utiles améliorent considérablement l'efficacité des produits statistiques.

Terminologie utilisée :

Élément visuel (nom) : ensemble de représentations graphiques et de texte utilisés pour transmettre des informations, comme un diagramme, une carte ou un tableau

Visualiser (verbe) : acte de créer un élément visuel

En tant qu'autorités responsables des données, les NSO doivent prioriser la création d'éléments visuels efficaces, tant dans les produits analytiques existants que dans les produits autonomes, afin de faciliter l'accès aux statistiques officielles pour les utilisateurs de données, les décideurs et le public.

Cette note technique est la première partie d'une série de deux intitulée *Communiquer au moyen des données du recensement*. Cette série montre l'importance que revêt la transmission des résultats de recensement à un large public par le biais de produits de données pertinents et attrayants.

La première partie de cette note technique aborde des concepts importants concernant la visualisation des données, tels que la valeur des éléments visuels, la manière de produire un visuel ainsi que les bonnes pratiques. La deuxième partie inclut des exemples de bonnes pratiques pour trois types de visualisation de données : tableaux, graphiques et cartes.

CRÉATION D'ÉLÉMENTS VISUELS

Les données sont présentées dans une variété de formats et de médias visuels étant tous complémentaires. Par exemple, même si les tableaux de données sont importants, ne fournir que des tableaux sans les accompagner de graphiques ou de cartes limite la valeur de la présentation des données dans son ensemble.

La majorité des éléments visuels statiques sont produits avec des logiciels de bureautique courants, alors que d'autres nécessitent des compétences et des outils avancés.

Visualisation des données

Les gens forment naturellement des **éléments visuels** dans leur esprit (Few 2012, p. 65). Que ce soit en lisant un roman, en rêvant pendant le sommeil ou en réalisant des tâches quotidiennes, le cerveau humain est capable de comprimer de grandes quantités d'informations en la représentation visuelle d'un phénomène. Un visuel est « un outil qui permet à vos yeux et à votre cerveau de percevoir ce qui dépasse leur portée naturelle » (Cairo 2013, p. 10).

Dans la **visualisation des données**, cette capacité innée est utilisée pour fournir un contexte à de grandes quantités de données à l'aide de graphiques. Par définition, « les graphiques de données présentent visuellement des quantités mesurées par l'utilisation combinée de points, de lignes, d'un système de coordonnées, de chiffres, de symboles, de mots, d'ombrages et de couleurs » (Tuft 1983, p. 9).

¹ Cette note technique fait partie de la série « Sélection de sujets relatifs aux recensements internationaux » examinant des questions d'intérêt pour la communauté statistique internationale. Le Bureau du recensement des États-Unis aide les pays à améliorer leurs systèmes statistiques nationaux en s'engageant dans le développement des capacités afin de renforcer les compétences statistiques de manière durable.

Le progrès technologique et la prolifération gratuite et ouverte des données ont amélioré notre capacité à visualiser les données (Cairo 2011, p. 14). En particulier, la visualisation des **données de recensement et d'enquête** peut ajouter de la valeur pour plusieurs raisons :

- **Montrer des modèles qui sont autrement méconnus.** Même si les données sous-jacentes sont volumineuses ou complexes, un visuel bien conçu fera ressortir les modèles les plus importants. Ces modèles sont difficiles à identifier sans l'aide d'éléments visuels.
- **Transmettre rapidement des informations.** Sans une analyse approfondie, la plupart des lecteurs ne peuvent pas interpréter utilement un tableau de données. Les éléments visuels peuvent transformer un tableau de données en un graphique de facile interprétation.
- **Fournir des preuves convaincantes.** Les visualisations de données consolident l'histoire de l'auteur en facilitant le partage des connaissances avec le public.

Les experts en la matière peuvent utiliser des éléments visuels durant la phase d'analyse exploratoire des données (Voir la section sur l'analyse exploratoire des données dans la première partie de cette série). Cependant, cette note technique portera sur l'utilisation de la visualisation des données pour présenter des résultats à un autre public.

Types et médias visuels

L'un des premières décisions à prendre dans la production d'un élément visuel est de déterminer son moyen de diffusion ou de publication. Les produits **imprimés** ont des exigences différentes de celles des produits Web. Dans le cas du papier, par exemple, il faut tenir compte du poids, de l'épaisseur, du couchage et de la luminosité du papier lui-même. De leur côté, les produits **numériques** exigent généralement une connexion Internet fiable et doivent pouvoir être visualisés sur plusieurs appareils (ordinateurs personnels, téléphones intelligents et tablettes).

Un élément visuel peut être **statique** ou **interactif**. Les visuels statiques sont les plus courants puisqu'ils sont les plus simples à créer. De nombreux éléments visuels statiques peuvent être créés à l'aide de logiciels de bureautique communs.

Les éléments visuels interactifs sont généralement associés aux produits destinés à Internet. Lorsqu'il est bien conçu, un visuel interactif peut montrer une plus grande quantité d'informations et retenir l'attention du public pendant plus longtemps. Toutefois, la conception d'éléments visuels interactifs exige une plus grande capacité technique et peut requérir une consultation externe.

En raison de sa facilité d'adoption et de son utilité pour les NSO, cette note technique se concentrera sur les éléments visuels statiques.

Éléments visuels statiques

Les produits analytiques des NSO fournissent des informations sur les modèles importants et les tendances à travers le temps et l'espace. Dans la plupart des produits analytiques des recensements et des enquêtes nationales, les principales formes d'éléments visuels statiques seront

des **tableaux**, des **diagrammes**, et des **cartes**. Ces trois formes graphiques sont populaires en raison de leur efficacité et de leur simplicité à transmettre des informations.

Chaque type d'élément visuel a ses points forts et ses points faibles. En somme, le choix de l'élément visuel est un compromis qui permet d'accepter certains points faibles inhérents au format de présentation s'ils sont compensés par les points forts. Ces considérations seront examinées plus en détail ultérieurement.

La case 1 illustre un processus de création d'un élément visuel. Même si les auteurs peuvent être tentés d'utiliser un groupe d'éléments visuels répétitifs pour l'ensemble d'un produit et pour de multiples variables, certaines variables n'ont pas besoin d'être accompagnées d'un élément visuel. Les auteurs doivent peser les avantages et les inconvénients de la visualisation des données pour déterminer si un élément visuel valorise un produit ou s'il détourne l'attention du public du message principal.

Case 1.

Processus de création d'une visualisation

- 1) **Quelles données doivent être visualisées?** Au moment de choisir un indicateur à visualiser, tenez compte de la variable principale. Celle-ci constituera généralement l'axe des x dans une représentation graphique en deux dimensions, par exemple, les années de recensement, les catégories de réponse.
- 2) **Pourquoi est-il important de visualiser ces données?** La visualisation doit avoir un objectif. Pensez aux modèles intéressants dans les données qui ne peuvent être montrés qu'avec une visualisation.
- 3) **Quel type de visualisation doit être créé?** Le choix de la visualisation dépend des données. Par exemple, si l'objectif est de montrer un schéma spatial, une carte sera le type de visualisation à privilégier. En revanche, si l'objectif est de montrer un modèle sur une période donnée, un graphique linéaire pourrait constituer la visualisation idéale. Toute visualisation implique un compromis.
- 4) **Comment cette visualisation complète-t-elle les autres visualisations figurant dans le produit?** À moins qu'une visualisation ne soit publiée comme un produit autonome, elle doit être choisie en fonction des autres visualisations, du moyen de publication, de l'espace disponible et du sujet traité.
- 5) **Comment cette visualisation sera-t-elle créée?** Un logiciel est nécessaire pour créer des visualisations. Les cartes sont habituellement générées par un logiciel de système d'information géographique (GIS, selon l'acronyme anglais), alors que les diagrammes et les tableaux peuvent être générés par un logiciel de productivité.

Source : U.S. Census Bureau.

Compétences et logiciels requis

La création d'éléments visuels professionnels exige des compétences en **conception graphique**, en **conception Web**, en **programmation**, en **cartographie** ou en **gestion des données**. Il existe d'autres compétences plus difficiles à maîtriser, comme les capacités artistiques et analytiques.

Plusieurs logiciels permettent de créer des éléments visuels professionnels et attrayants. La figure 1 présente certains outils couramment utilisés, y compris des options de logiciels payants et libres lorsque possible. Il convient de noter que certains logiciels exigent une formation avancée ou une expérience approfondie, notamment en matière de conception graphique, de développement Web, de cartographie et d'analyse de données.

Figure 1.

Logiciels pour la création d'éléments visuels

Des applications spécialisées peuvent être nécessaires pour produire certains éléments visuels.

Catégorie	Exemples	Fonctionnalité	Difficulté
Logiciels de productivité	Microsoft Excel, OpenOffice	Créer des éléments visuels simples	Novice
Outils de visualisation dédiés	Tableau, WEAVE	Créer des éléments visuels plus complexes	Intermédiaire
Analyse de données	Stata, R	Traitement et analyse de données	Avancé
Conception graphique	Adobe Illustrator, Inkscape	Création et finalisation des graphiques	Avancé
Cartographie	ArcGIS, QGIS	Création de cartes et données géospatiales	Avancé
Développement Web	D3, Leaflet	Création de visuels interactifs pour le Web	Avancé

Remarque : Plusieurs des exemples fournis dans ce tableau sont des logiciels avancés qui requièrent une formation et de l'expérience pour être utilisés de manière professionnelle.

Source : U.S. Census Bureau.

Dans la plupart des cas, les outils de visualisation des données fournis par les logiciels de productivité courants permettent de créer des éléments visuels de données, tels que des tableaux et des diagrammes, avec un minimum de formation additionnelle. Les éléments visuels créés dans ces logiciels sont hautement personnalisables et peuvent être de qualité professionnelle.

Toutefois, toutes les fonctionnalités de ces logiciels ne sont pas souhaitables. Les auteurs doivent s'assurer que leurs graphiques sont clairs et faciles à interpréter. Les bonnes pratiques ainsi que celles à éviter pour des types de visuels spécifiques sont abordées plus en détail dans les pages suivantes de cette note technique.

Chacun de ces logiciels peut être maîtrisé au moyen d'une formation. Cependant, une formation ne fournit qu'une introduction limitée du sujet. Une longue expérience - peut-être sur plusieurs années - sera nécessaire pour devenir un expert en visualisation de données. Ainsi, pour éviter de perdre le cap en matière d'expertise, certains organismes peuvent engager des graphistes professionnels ou faire appel à des fournisseurs pour les aider à produire des éléments visuels plus complexes.

BONNES PRATIQUES

Le principal objectif d'un élément visuel est d'améliorer la compréhension. Pour atteindre cet objectif, les auteurs devraient suivre les conseils suivants. Cette note technique utilise des exemples du Bureau du recensement des États-Unis pour illustrer des bonnes pratiques.

Utiliser des données structurées

Ce sont les données sous-jacentes qui constituent la base d'un élément visuel. Pour créer un élément visuel, les données doivent généralement être organisées dans un format qui peut être lu par un ordinateur, comme le format (Comma Separated Values) (CSV, selon l'acronyme anglais), le langage XML (eXtensible Markup Language) (XML, selon l'acronyme anglais), le format (JavaScript Object Notation) (JSON, selon l'acronyme anglais) ou une base de données relationnelle. Ces formats offrent des **données structurées**, ce qui signifie qu'un ordinateur peut facilement les lire pour les analyser. Ces formats diffèrent des données non structurées, comme le format (PDF, selon l'acronyme anglais) (Portable Document Format), qui n'est pas facilement lisible par les logiciels

d'analyse de données. Un exemple de la distinction entre données structurées et non structurées est présenté en figure 2.

De manière générale, les NSO devraient rendre disponibles leurs produits de données en un format structuré afin de faciliter leur utilisation à des fins de visualisation tant à l'interne qu'à l'externe. Ces formats sont adaptés aux efforts mondiaux visant à améliorer la transparence et la disponibilité des données gouvernementales ainsi qu'à accroître l'utilité des données statistiques nationales (United Nations 2013).

Les données structurées peuvent sous la forme de fichier plat ou de tabulation en croix. Les cellules individuelles d'un tableau plat représentent une seule variable, tandis que les cellules d'une tabulation en croix représentent la combinaison de plusieurs variables. Un élément visuel peut être créé en utilisant un des deux formats, ce qui dépendra de l'objectif de l'auteur.

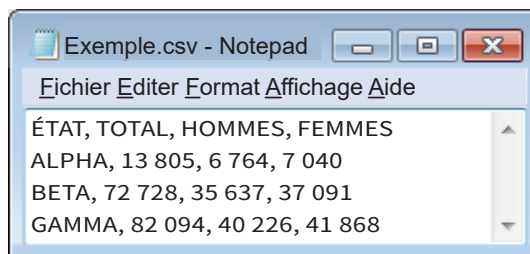
Figure 2.

Données structurées et. Non structurées

Le tableau d'information suivant peut paraître structuré. Néanmoins, lorsque présenté dans un format tel que ce document (PDF), les numéros ne sont pas facilement lisibles par ordinateur. Il est donc non structuré.

État	Total	Masculin	Féminin
Alpha	13 805	6 764	7 040
Beta	72 728	35 637	37 091
Gamma	82 094	40 226	41 868

Par contre, l'image ci-dessous présente le même tableau sauvegardé dans un fichier de valeurs séparées par des virgules (CSV) et visualisé avec le logiciel Notepad. Ce format de données est donc structuré puisque chaque colonne de numéros est clairement séparée par un délimiteur (ici, une virgule) et que chaque registre est placé sur une ligne distincte.



Source : U.S. Census Bureau.

Fournir un contexte

Le **contexte** est essentiel dans la visualisation des données. Puisque les lecteurs sont susceptibles de ne regarder que les éléments visuels et d'ignorer le texte dans un produit volumineux, chaque élément doit contenir suffisamment d'informations pour que le public puisse l'interpréter rapidement.

Le **titre** est la première chose que la majorité des lecteurs verront et il doit donc être descriptif. Il convient d'indiquer clairement dans le titre la ou les variables présentées dans l'élément visuel et la période couverte. Les titres peuvent aussi être utilisés pour faire passer le message principal du graphique statistique.

Citez les **sources** utilisées pour la création de l'élément visuel. Pour les rapports publiés, la citation doit apparaître dans l'élément visuel lui-même et être convenablement référencé au bas de la page ou aux citations complètes à la fin du document. Pour les éléments visuels autonomes qui apparaissent sur des affiches ou comme produits destinés à Internet, la citation complète doit figurer sur le visuel. Les références des citations comprennent l'auteur ou l'organisme source (par exemple, le NSO lui-même); l'opération spécifique, le rapport et l'année (par exemple, Census, Summary File 1, 2011); et un hyperlien si disponible sur Internet.

Incluez également des instructions **d'utilisation des données** dans l'élément visuel. Pour les recensements et les enquêtes nationales, ces instructions incluent habituellement des liens vers des documents didactiques fournissant des définitions de variables, des erreurs d'échantillonnage et des erreurs non dues à l'échantillonnage (comme les erreurs de couverture et les erreurs de réponse).

En ce qui concerne les tableaux de données, un hyperlien est normalement inclus pour consulter **l'ensemble de données structurées (lisibles par un ordinateur)**. Comme mentionné antérieurement, les organismes doivent s'efforcer de fournir un accès aux ensembles de données structurées afin d'étendre leur utilisation par d'autres organismes gouvernementaux, le milieu universitaire et le secteur privé.

Utilisez des éléments graphiques

Le **graphisme** est un art. Les graphistes sont souvent nécessaires pour créer des éléments visuels attrayants, professionnels et efficaces. Néanmoins, même les utilisateurs de données ayant peu d'aptitudes artistiques peuvent créer des éléments visuels de bonne qualité en utilisant efficacement les éléments graphiques présentés ci-dessous.

La prudence est de mise pour les auteurs qui utilisent ces éléments graphiques. Par exemple, comme un pourcentage de la population est incapable de distinguer certaines couleurs en raison du daltonisme (souvent le rouge et le vert), il convient d'utiliser des palettes de couleurs adaptées aux daltoniens pour la plupart des produits. Une couleur peut aussi avoir une signification culturelle positive ou négative. Ces éléments peuvent aussi être surutilisés, ce qui résulte en un visuel distrayant, et par conséquent, moins fonctionnel.

Utilisez la **couleur** pour attirer l'attention sur des éléments spécifiques du visuel ou pour nuancer d'autres éléments moins importants. Par exemple, un graphique linéaire peut comporter une série de lignes et une seule ligne de tendance. En utilisant le rouge pour la ligne de tendance

et le gris clair pour les autres lignes, l'auteur peut attirer l'attention vers l'élément le plus important du visuel.

Les **symboles**, par exemple, peuvent mieux exprimer les proportions que les diagrammes circulaires, ces derniers étant difficiles à interpréter avec précision. Les cartes de symboles gradués constituent une autre utilisation courante des symboles, qui sont de taille différente pour refléter une valeur absolue (une taille plus grande reflète une valeur plus grande, et vice versa).

Le choix de la **police d'écriture** affectera aussi l'interaction du public avec un élément visuel. Les cartes, par exemple, peuvent utiliser des polices différentes pour désigner des caractéristiques naturelles comme des masses d'eau et des caractéristiques socioculturelles comme des noms de lieux. À titre d'exemple, les cartes peuvent utiliser des polices différentes pour désigner les caractéristiques naturelles telles que les masses d'eau et les caractéristiques socioculturelles comme les noms de lieux.

D'autres exemples d'éléments graphiques à envisager dans les éléments visuels sont : l'orientation, la forme, la longueur et la largeur des lignes, les dimensions, les courbes, les marqueurs ajoutés, la délimitation, l'intensité, la disposition spatiale et le mouvement (Knaflitz 2015, p. 105).

EXEMPLES PAR TYPE

Le reste de cette note technique inclut des conseils pour la création d'éléments visuels efficaces, attrayants et significatifs pour retraiter les données de recensement ou d'enquête en trois types principaux : diagrammes, cartes et tableaux. Ces recommandations complètent les bonnes pratiques mentionnées antérieurement et qui s'appliquent à tous les éléments visuels.

DIAGRAMMES

Des types de diagrammes d'une complexité variable sont utiles pour visualiser les statistiques nationales. Par souci de simplicité, cette note technique n'abordera que quatre types de diagrammes : **les diagrammes à barres, les diagrammes linéaires, les pyramides des âges, et les pictogrammes.**

DIAGRAMMES À BARRES

Les diagrammes à barres peuvent montrer l'importance relative des variables discrètes / nominales de recensement et d'enquête pour une même période ou au fil du temps. Les diagrammes à barres peuvent être présentés sous la forme de barres simples par catégories, de barres empilées les unes au-dessus des autres, ainsi que de barres doubles qui permettent une comparaison côte à côte. Les barres peuvent aussi être présentées verticalement ou horizontalement.

Dans la figure 3, par exemple, le ratio de dépendance total est séparé en ratios de dépendance des jeunes et ceux des personnes plus âgées par l'utilisation d'un diagramme de barres empilées. Cette technique démontre que la dépendance des personnes âgées augmente au fil du temps par rapport à celle des jeunes.

Avantages : Condense de vastes ensembles de données dans un format simple; compare plusieurs catégories en même temps; est facile d'accès pour le public en raison de sa popularité.

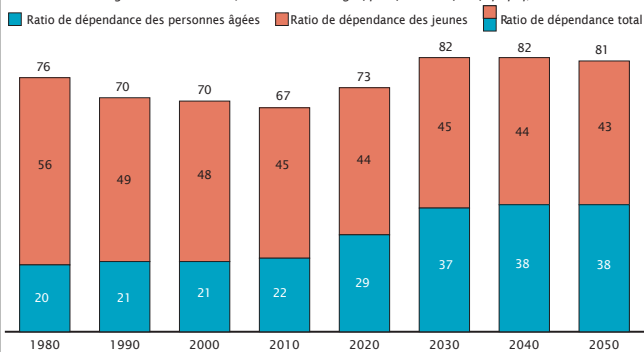
Inconvénients : Subjectif, dépendamment de l'ordre des catégories et du formatage des axes; ne peut afficher qu'un nombre limité de barres sans devenir surchargé.

Figure 3.
Exemple d'élément visuel : Diagramme à barres

Les diagrammes à barres sont utiles pour illustrer les proportions. Cet exemple est un diagramme à barres empilées avec un ensemble de barres (dépendance des jeunes) placé au-dessus d'un autre (dépendance des personnes âgées). Ce diagramme à barres illustre parfaitement le fait que la dépendance des personnes âgées a augmenté par rapport à la dépendance des jeunes, confirmant ainsi l'affirmation de l'auteur que la population américaine vieillit.

Ratios de dépendance : De 1980 à 2050

(Pour obtenir des informations sur la protection de la confidentialité, l'erreur de non-échantillonnage et des définitions, voir www.census.gov/prod/cen2010/doc/sf1.pdf)



Remarque : Le ratio de dépendance total est le nombre de personnes âgées de 0 à 19 ans et de 65 ans et plus pour 100 personnes âgées de 20 à 64 ans. Le ratio de dépendance des jeunes est le nombre de personnes âgées de 0 à 19 ans pour 100 personnes âgées de 20 à 64 ans. Le ratio de dépendance des personnes âgées est le nombre de personnes âgées de 65 ans et plus pour 100 personnes âgées de 20 à 64 ans.
Sources : 1980, U.S. Bureau of the Census, 1983; 1990, U.S. Bureau of the Census, 1992; 2000, U.S. Census Bureau, 2001; 2010, U.S. Census Bureau, 2011; 2020 to 2050, U.S. Census Bureau, 2012a; 1980 to 2010, decennial census; 2020 to 2050, 2012 National Population Projections, Middle series.

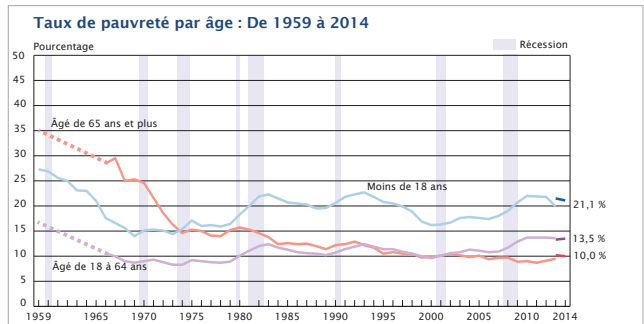
Source : U.S. Census Bureau, 2014.

DIAGRAMMES LINÉAIRES

Les diagrammes linéaires sont utiles pour illustrer les tendances des données de recensement et d'enquête au fil du temps. Par exemple, le diagramme linéaire dans la figure 4 montre le taux de pauvreté mesuré par les recensements et les enquêtes qui varie au fil du temps. L'inclinaison de la ligne présentée doit avoir une signification, que celle-ci soit utilisée pour des données temporelles ou non temporelles.

Figure 4.
Exemple d'élément visuel : Diagramme linéaire

Les diagrammes à barres sont utiles pour illustrer les tendances. Ici, les taux de pauvreté sont désagrégés par âge, montrant comment les taux pour certains groupes ont augmenté ou diminué au fil du temps. Des informations supplémentaires sont ajoutées en soulignant les années durant lesquelles l'économie nationale était en récession. Cependant, ces informations sont atténuées pour éviter de détourner l'attention des points de données primaires. Des remarques additionnelles sont également présentées pour expliquer la parution des séries de données de 2013-2014, lorsque les questions de l'enquête ont été reformulées.



Remarque : Les données de 2013 reflètent l'application des questions reformulées sur le revenu. Pour plus d'informations, voir l'annexe D. Les points de données sont placés aux valeurs médianes des années respectives. Pour la période de 1960 à 1965, les données concernant les personnes âgées de 18 à 64 ans et de 65 ans et plus ne sont pas disponibles. Pour des informations sur les récessions, voir l'annexe A. Pour des informations sur la protection de la confidentialité, les erreurs d'échantillonnage, les erreurs non imputables à l'échantillonnage et les définitions, voir <http://ftp2.census.gov/programs-surveys/cps/techdocs/cpsmar15.pdf>.
Source : Bureau du recensement des États-Unis, Current Population Survey, 1960 to 2015 Annual Social and Economic Supplements.

Source : U.S. Census Bureau, 2015.

Avantages : Permet de visualiser les tendances au fil du temps; utile pour présenter plusieurs variables d'une même catégorie; peut être combiné avec d'autres éléments visuels pour mettre en évidence les aspects importants de la tendance.

Inconvénients : Ne peut utiliser que quelques lignes avant d'être surchargé (bien qu'il puisse utiliser plus de lignes que de barres dans un diagramme à barres); exige un grand nombre de points de données pour illustrer des tendances significatives au fil du temps.

PYRAMIDES DES ÂGES

La pyramide des âges (par exemple, la figure 5) est l'un des moyens les plus efficaces de visualiser la structure de la population établie par un recensement ou une enquête. Par exemple, une pyramide ayant une base plus large que le sommet (c'est-à-dire une population plus jeune) correspondra à un pays dont les besoins sont différents de ceux qui ont une pyramide de dimensions égales. La forme d'une pyramide des âges peut renseigner rapidement le public sur la tendance démographique d'un pays dans les années à venir.

Avantages : Permet de visualiser la structure de la population; de voir les disparités par âge/sexe; d'effectuer une désagrégation par sous-groupes spécifiques comme par ethnies, langues ou géographie.

Inconvénients : Perte de granularité si de grandes tranches d'âge sont utilisées; difficulté d'intégrer des variables supplémentaires.

Pictogrammes

Les pictogrammes sont des représentations visuelles de données utilisant de petites icônes que le public connaît bien. Chaque icône est dimensionnée de façon proportionnelle (par exemple, une icône représente 10 000 personnes) et plusieurs icônes sont utilisées pour illustrer les tendances en général. La figure 6 donne un exemple de la proportion d'hommes et de femmes parmi les vétérans de l'armée en utilisant des icônes pour montrer la disparité du sex-ratio.

Avantages : Facilement accessible à un large public; représentation rapide de modèles ou de tendances générales; idéal pour les produits de médias de masse.

Inconvénients : Représentation plus abstraite; ne convenant généralement pas aux utilisateurs de données avancés ou pour les rapports analytiques.

BONNES PRATIQUES

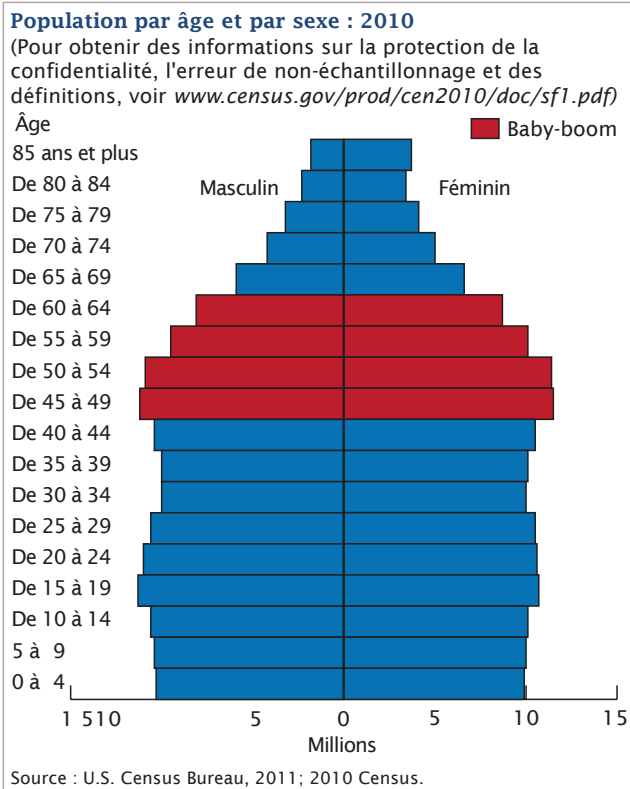
Remarque : Ces bonnes pratiques sont présentées aux figures 3 à 6.

Mettre en lumière une tendance ou un modèle : Chacun de ces tableaux souligne un élément important de l'histoire de l'auteur. Si un tableau porte à confusion ou n'est pas particulièrement utile, il est préférable de ne pas l'inclure dans le produit.

Données univoques : Chaque élément de données se distingue des autres grâce à des couleurs et des bordures de lignes différentes. À titre d'exemple, les lignes de tendance de la figure 4 sont de trois couleurs distinctes, et les barres empilées de la figure 3 sont elles aussi de couleurs distinctes et présentent des bordures autour de chaque barre. Toutes ces figures sont

Figure 5.
Exemple d'élément visuel : Pyramides des âges

Les pyramides des âges illustrent efficacement la répartition de la population par âge et par sexe. Cet élément visuel fait également ressortir un groupe spécifique, les « baby-boomers », nés entre 1946 et 1964.



Source : U.S. Census Bureau, 2015.

accompagnées de légendes pour indiquer ce que chaque couleur représente. Les figures 3 et 4 utilisent aussi des étiquettes.

Des axes précis : Les valeurs au long des axes x (horizontal) et y (vertical) doivent refléter avec précision les éléments du graphique. Pour les séries chronologiques, positionnez les points de données relatifs à la période de temps entre chaque point de données (p. ex. Figure 4). De plus, les axes quantitatifs doivent normalement commencer à zéro.

Souligner les points clés : Utilisez des éléments graphiques pour faire ressortir les points importants pour l'auteur. Par exemple, le rouge dans la figure 5 met en évidence un groupe d'âge spécifique, tandis que le gris dans la figure 4 indique les années de récession.

Connaissez votre public : Choisissez un type de graphique adapté à votre public et à votre message. La figure 6 est compréhensible par la majorité des publics. En revanche, la figure 5 peut être plus adaptée aux experts en la matière plutôt qu'au public. La figure 3 et la figure 4 sont des types de graphiques courants, mais qui peuvent quand-même être difficiles à interpréter pour un public qui n'est pas familiarisé avec les données sous-jacentes.

Pratiques à éviter

Graphiques en 3D : En général, les graphiques tridimensionnels (3D, selon l'acronyme anglais) ne sont

pas efficaces, ils distraient le lecteur et peuvent lui faire perdre de vue les données. À ne pas utiliser à moins qu'il n'existe une troisième dimension aux données.

Diagrammes à barres : L'esprit humain a du mal à interpréter les surfaces circulaires et les angles. Les diagrammes circulaires peuvent donc être trompeurs, surtout s'ils sont composés de plus de deux ou trois parts.

Complexité excessive : Évitez de communiquer trop d'idées en même temps. L'élément graphique doit être axé sur un message essentiel et être facile à comprendre pour le public.

Échelles trompeuses : Les valeurs des échelles doivent être choisies soigneusement et dépendent du type de données. Le fait de tronquer les échelles (par exemple, pour les données ayant une gamme restreinte) peut conduire à une mauvaise interprétation ou à une survalorisation de la variabilité.

Figure 6.
Exemple d'élément visuel : Pictogramme

Les pictogrammes sont généralement utilisés pour le grand public. Des icônes simples, comme celles présentées ici, sont facilement comprises par les lecteurs de tous niveaux d'éducation, y compris les écoliers. Dans cet exemple, une icône représente un million de vétérans.

Les États-Unis comptent 21,8 millions de vétérans.



Source : U.S. Census Bureau, 2012.

Cartes

Les cartes illustrent la répartition spatiale des données. Avec les cartes, il est possible de distinguer des modèles qui resteraient autrement cachés dans les tableaux de données ou les diagrammes. Lorsqu'elle est bien conçue, une carte peut appuyer l'histoire et l'argumentation de l'auteur. Inversement, une carte mal conçue peut confondre le public ou lui faire perdre tout intérêt envers le produit.

De façon générale, les cartes peuvent être divisées en deux catégories : thématiques et de référence. Les cartes thématiques, comme présentées dans la figure 7, sont utilisées pour illustrer un thème particulier, y compris la géographie sociale et physique. Les cartes de référence indiquent surtout les points de repère, les caractéristiques physiques, les lieux et d'autres informations destinés à la navigation ou au contexte. Cette note technique porte sur les cartes thématiques, puisqu'elles sont le plus souvent utilisées pour présenter des données de recensement et d'enquête.

AVANTAGES

Schémas spatiaux : Mettre en lumière les modèles géographiques et des disparités régionales qui ne sont pas visibles dans les graphiques ou les tableaux.

Améliorer la narration : Les modèles figurant sur les cartes peuvent améliorer l'exposé de l'auteur en fournissant des secteurs d'analyses additionnels.

Transmettre rapidement des informations : Lorsqu'elle est bien conçue, une carte peut condenser une grande quantité de données dans un graphique significatif et de facile interprétation.

Engage le public : En raison de la grande quantité d'information transmise, les cartes peuvent attirer l'attention des lecteurs et engendrer de longues discussions sur la signification des modèles présentés.

INCONVÉNIENTS

Techniquement exigeant : La création de cartes exige souvent des systèmes d'information géographique (GIS, selon l'acronyme anglais) et des logiciels de conception graphique pour obtenir un résultat professionnel.

Courbe d'apprentissage plus abrupte : Maîtriser les outils nécessaires pour cartographier efficacement demande beaucoup d'expérience. Créer des cartes Web interactives requiert aussi de l'expérience en conception Web.

Nombre limité d'indicateurs : En règle générale, une carte thématique devrait illustrer un seul indicateur. Le recours à de multiples indicateurs augmente la complexité et réduit l'efficacité de la carte elle-même, à moins qu'elle ne soit conçue par un artiste chevronné.

BONNES PRATIQUES

Remarque : Ces bonnes pratiques sont présentées dans la figure 8.

La **catégorisation** figurant dans la légende soutient la narration de l'auteur. Des points de référence spécifiques sont identifiés et reflètent les thèmes les plus importants de la carte.

La **palette de couleurs** est appropriée. Les nombres supérieurs à zéro sont catégorisés au moyen d'une gradation logique des couleurs, alors que les nombres inférieurs à zéro sont catégorisés à l'aide d'une couleur distincte.

Les **étiquettes** sont discrètes mais explicatives (abréviations postales par état). Les étiquettes sont également de couleur noire ou blanche, en fonction de la luminosité de l'arrière-plan. Les étiquettes sont aussi placées de façon à se situer approximativement au centre du polygone mais sans toucher ses extrémités. Lorsque ceci est inévitable, les étiquettes sont déplacés hors de l'état et des traits pointent vers leur emplacement exact.

Des **cartons intérieurs** sont fournis pour certains secteurs. Ces cartons fournissent des détails supplémentaires qui seraient autrement perdus.

La **moyenne nationale** est fournie à titre de référence pour le public.

Les données géospatiales (lignes et polygones) qui ont servi à produire cette carte sont **simplifiées** pour en réduire la complexité géométrique.

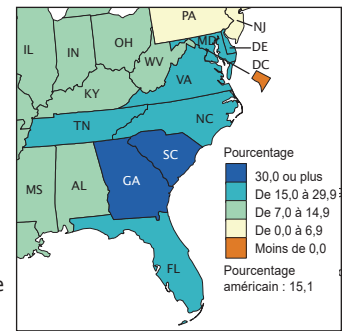
Figure 7.
Types de cartes thématiques souvent utilisées dans les statistiques nationales

Carte choroplèthe

Ces cartes contiennent des espaces grisés ou à motifs proportionnels à la variable statistique étant montrée sur la carte. Les données sont agrégés sur des unités spatiale prédéterminées (administratives ou statistiques, par exemple, la géographie du recensement).

Utilisations optimales : Préférable lorsque les données sont structurées (p. ex. taux), discrètes, et distribuées également à l'intérieur d'unités spatiales bien définies.

Considérations concernant la conception : Le nombre de catégories doit être limité (trois à sept).

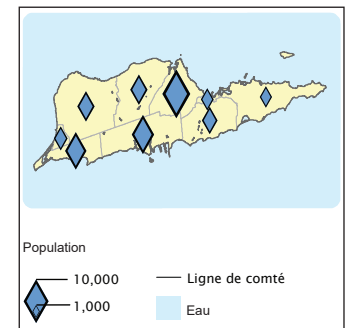


Carte de symboles gradués

Ces cartes contiennent des symboles qui varient en taille pour illustrer leur valeurs quantitatives; utilisées avec des point de données ou des données de localisation.

Utilisations optimales : Usage optimal lorsque les données comportent de grandes gammes et variations. L'objectif est d'illustrer l'amplitude relative des phénomènes dans des lieux spécifiques. Cette carte est aussi un bon choix pour compter les données.

Considérations concernant la conception : Ne doit pas être utilisé pour des données standardisées comme les taux ou pourcentages.

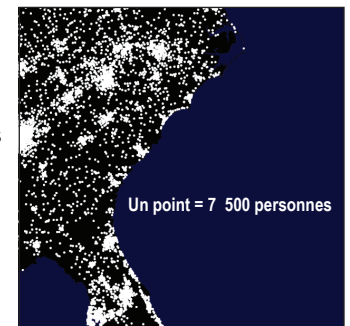


Carte de densité de point

Elle utilise des points pour illustrer la présence d'une caractéristique ou occurrence; montre un schéma spatial et la densité relative. Les points individuels peuvent représenter une ou plusieurs occurrences.

Utilisations optimales : Usage optimal avec les données de comptage; peut aussi montrer de multiples ensembles de données (en utilisant des symboles ou des couleurs différentes).

Considérations concernant la conception : Exige des outils additionnels (p. ex. géocodage) pour localiser des points sur la carte; les difficultés de perception et les techniques de conception (p. ex. taille, valeur et arrangement des points) doivent être pris en compte.

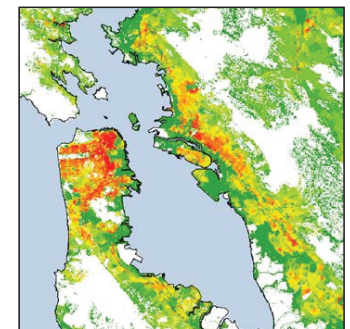


Carte dasymétrique

Une carte thématique qui utilise des symboles pour classer spatialement des données volumétriques; utilisée comme une alternative aux cartes choroplèthes.

Utilisations optimales : La carte dasymétrique est appropriée lorsque le scénario d'une distribution uniforme des phénomènes n'est pas vérifié (p. ex. répartition de population). L'interprétation peut être facilitée quand la technique de cartographie est fournie aux utilisateurs. Comme dans le cas des cartes choroplèthes, cette technique est utilisée pour illustrer des données standardisées (p. ex. taux).

Considérations concernant la conception : Cette méthode est plutôt complexe et de longue durée; l'étendue des secteurs cartographiés rarement correspond aux limites des unités de dénombrement.



Sources : U.S. Census Bureau, CDC (selon l'acronyme anglais) and USGS (selon l'acronyme anglais).

PRATIQUES À ÉVITER

Trop de cartes : Incluez uniquement les cartes qui renforcent la narration. Les variables ne doivent pas toutes être cartographiées, en particulier les indicateurs qui suivent un modèle connu existant (par exemple, la répartition de la population). Les cartes doivent agrémenter la narration et mettre en évidence des modèles qui seraient autrement négligés.

Couleur : En général, un schéma de couleurs à plusieurs tonalités est destiné aux données catégoriques (nominales). Utilisez un schéma à tonalité unique comme le bleu pour les données d'intervalle et de ratio. Dans l'exemple, un schéma de couleur à tonalités multiples est utilisé puisque il y a un point zéro naturel (supérieur ou inférieur à 0).

Éléments de la carte : La flèche d'orientation dirigée vers le Nord et la barre d'échelle ne sont généralement pas nécessaires. Dans un tel contexte, le public comprendra sans doute que la carte est orientée de manière à ce que le nord soit en haut de la carte, et l'échelle n'est donc pas nécessaire pour interpréter les modèles. Une flèche d'orientation dirigée vers le nord et une barre d'échelle doivent toutefois être incluses dans une carte de référence.

Encombrement : Évitez d'encombrer la carte avec des informations inutiles. Les cartes thématiques doivent généralement montrer une seule variable. Il est recommandé de garder les étiquettes simples et la discrétion est de mise au moment d'étiqueter des cartes comportant de nombreuses caractéristiques.

Utilisation de diagrammes : En général, les diagrammes (par exemple, à secteurs ou à barres) ne devraient pas être placés par-dessus les données cartographiques. Les diagrammes réduisent les avantages de l'utilisation d'une carte, car les modèles deviennent difficiles à interpréter. Au lieu d'utiliser des diagrammes superposés à des données cartographiques, envisagez plutôt de diviser la carte en une série, où chaque carte représente une seule variable.

Catégorisation inadéquate : Les points de repère choisis doivent mettre en évidence les modèles importants pour l'auteur. Par contre, les points de repère ne doivent pas être choisis pour dissimuler des informations indésirables ou pour suggérer à tort une variabilité géographique inexistante.

TABLEAUX

Le point central de la présentation des données est souvent la visualisation (par exemple, les diagrammes et les cartes). Toutefois, comme le tableau de données (p. ex. figure 9) est sans doute la forme la plus courante de présentation des données de recensement et d'enquête, une attention particulière doit être accordée à sa conception.

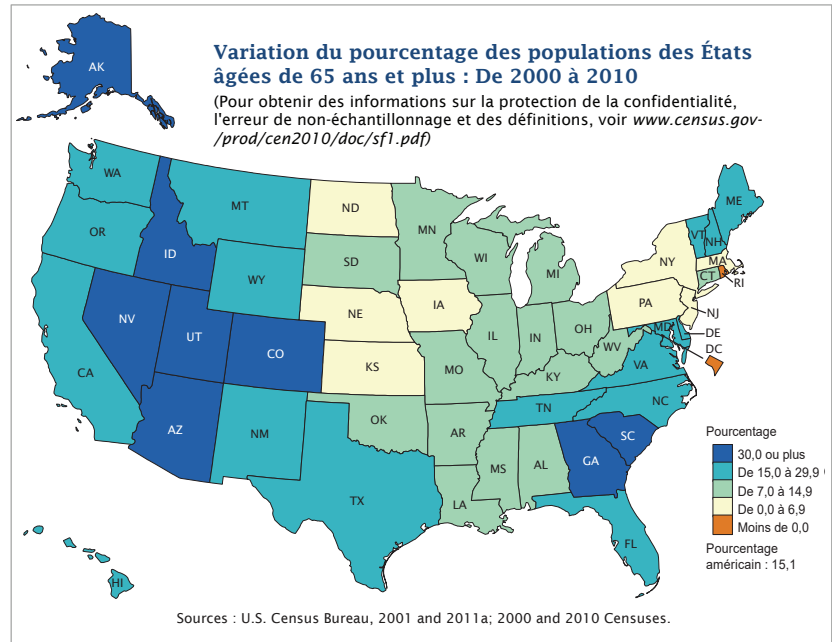
AVANTAGES

Une plus grande accessibilité : Les cartes et les diagrammes ne peuvent présenter qu'un nombre limité de variables à la fois sans accabler le lecteur. En revanche, les tableaux peuvent présenter de nombreuses variables,

Figure 8.

Exemple d'élément visuel : Cartes

Ce format de carte est utilisé par le Bureau de recensement des États-Unis dans tous ses produits, reflétant ainsi les normes d'image de marque de l'agence.



Source : U.S. Census Bureau, 2014.

des niveaux de géographie et des moments précis dans le temps. Les tableaux fournissent également des informations contextuelles plus efficacement, comme les marges d'erreur.

Plus de flexibilité pour les utilisateurs : Les utilisateurs peuvent réaliser leur propre analyse sur un tableau de données, et aucun point de données ne se voit accorder une priorité supérieure à un autre. En outre, les utilisateurs de données s'attendent à avoir accès

aux tableaux de données, et pas uniquement aux éléments visuels ou aux statistiques sommaires qui en résultent.

Plus efficace que le texte : Les tableaux peuvent afficher des données dans un espace plus réduit que si elles étaient décrites en texte, et permettent aux utilisateurs avancés de trouver facilement les valeurs les plus pertinentes à leurs besoins.

Popularité : Les tableaux sont probablement le moyen le plus commun de présentation de données pour les statistiques nationales. La plupart des publics peuvent interpréter les informations présentées dans un tableau.

INCONVÉNIENTS

Analyse approfondie : Interpréter un tableau de données peut exiger du public un temps considérable. Même si un formatage stratégique peut aider à faire ressortir les points clés (comme mettre le texte en caractères gras ou utiliser une autre couleur), les informations du tableau doivent être accompagnées de statistiques récapitulatives, de cartes et de graphiques pour souligner les découvertes importantes.

Un faible défenseur de l'histoire de l'auteur : Les tableaux ne suscitent pas de réactions immédiates auprès d'un public. Les cartes et les diagrammes sont beaucoup plus efficaces pour transmettre rapidement le message de l'auteur.

Figure 9.

Exemple d'élément visuel : Tableaux

Les tableaux peuvent présenter de grandes quantités de données, et doivent donc être conçus avec soin. Ces exemples montrent comment les éléments graphiques sont utilisés pour attirer l'attention du public sur les données statistiques. Les graphiques sont réduits au minimum : par exemple, il n'y a pas de lignes séparant les rangés, et les lignes séparant les colonnes et les en-têtes sont fines.

Population âgée de 65 ans et plus, par âge : De 1900 à 2050

(Nombre en milliers. Pour obtenir des informations sur la protection de la confidentialité, l'erreur non imputable à l'échantillonnage et des définitions, voir www.census.gov/prod/cen2010/doc/sf1.pdf)

Source, année et date de référence	Population totale	65 ans et plus		85 ans et plus	
		Numéro	Pourcentage	Numéro	Pourcentage
Recensement					
1900 (1er juin).....	75 995	3 080	4,1	122	0,2
1910 (15 avril).....	91 972	3 950	4,3	167	0,2
1920 (1er janvier)....	105 711	4 933	4,7	210	0,2
1930 (1er avril).....	122 775	6 634	5,4	272	0,2
<Infographie tronquée à titre d'exemple>					
2000 (1er avril).....	281 422	34 992	12,4	4 240	1,5
2010 (1er avril).....	308 746	40 268	13,0	5 493	1,8
Projection					
2020 (1er juillet)....	333 896	55 969	16,8	6 693	2,0

Remarque : Les données de 1900 à 1950 excluent l'Alaska et l'Hawaï.
Sources : 1900 à 1940, and 1960 to 1980, U.S. Bureau of the Census, 1983; 1950, U.S. Bureau of the Census, 1953; 1990, U.S. Bureau of the Census, 1992; 2000, U.S. Census Bureau, 2001; 2010, U.S. Census Bureau, 2011; 2020 to 2050, U.S. Census Bureau, 2012b; 1900 to 2010, decennial census; 2020 to 2050, 2012 National Population Projections, Middle series.

Population âgée de 65 ans et plus et variation de pourcentage par Région et État : 2000 et 2010

Pour obtenir des informations sur la protection de la confidentialité, l'erreur non imputable à l'échantillonnage et des définitions, voir www.census.gov/prod/cen2010/doc/sf1.pdf)

Région et État	65 ans et plus		Variation, 2000 à 2010	
	2000	2010	Numéro	Pourcentage
États-Unis	34 991 753	40 267 984	5 276 231	15,1
Nord-est	7 372 282	7 804 833	432 551	5,9
Connecticut	470 183	506 559	36 376	7,7
Maine	183 402	211 080	27 678	15,1
Massachusetts	860 162	902 724	42 562	4,9
New Hampshire	147 970	178 268	30 298	20,5
New Jersey	1 113 136	1 185 993	72 857	6,5
New York	2 448 352	2 617 943	169 591	6,9
Pennsylvanie	1 919 165	1 959 307	40 142	2,1
Rhode Island	152 402	151 881	-521	-0,3
Vermont	77 510	91 078	13 568	17,5

Sources : U.S. Census Bureau, 2001, Census 2000 Summary File 1, Table P12, Washington, DC, available at <http://factfinder2.census.gov/>, accessed on February 20, 2012; U.S. Census Bureau, 2011, 2010 Census Summary File 1, Table PCT12, Washington, DC, available at <http://factfinder2.census.gov/>, accessed on February 20, 2012.

Source : U.S. Census Bureau, 2014.

BONNES PRATIQUES

Remarque : Ces bonnes pratiques sont présentées dans la figure 9.

Le tableau est **formaté** de manière cohérente. Par exemple, tous les nombres sont espacés de manière égale, les virgules se trouvent exactement au même endroit dans la colonne, les lignes ont la même épaisseur et tous les éléments sont bien alignés.

Tous les nombres se **cumulent** au total, comme il se doit, tant verticalement qu'horizontalement.

Les **tendances** au fil du temps sont clairement visibles, grâce à l'utilisation de pourcentages dans les deux tableaux afin de normaliser les variations dans la population totale.

Les données choisies pour ces tableaux **soutiennent l'exposé** de l'auteur indiquant que la population américaine vieillit. Seulement quelques variables sont incluses dans l'exemple de tableau pour améliorer la lisibilité.

La **série chronologique** est clairement notée. Même si le titre du tableau supérieur indique « 1900

à 2050, » les dates de recensement et de projection varient selon la décennie. Ces écarts sont indiqués à côté de chaque année dans la première colonne.

PRATIQUES À ÉVITER

Formatage distrayant : Limitez l'utilisation des lignes dans le tableau, et maintenez toutes les lignes nécessaires fines. Séparez également les en-têtes de colonne des données et ne mettez l'accent que sur quelques éléments.

Information excessive : Choisissez avec soin les données à inclure dans le tableau. Comme tout élément visuel, le tableau est censé appuyer la narration de l'auteur. Inutile d'inclure toutes les variables, puisque les tableaux complets peuvent être téléchargés à partir du site Web du NSO ou encore annexés au document principal.

Présentation incohérente : Chaque tableau devrait présenter un ensemble cohérent d'informations, par année, par géographie ou par thème. Si la présentation des données se fait par secteur géographique, les unités égales (par exemple, les provinces ou les districts) ne doivent pas être exclues sans explication.

RÉFÉRENCES

Cairo, A., *The Functional Art: an Introduction to Information Graphics and Visualization*, New Riders, 2013.

Centers for Disease Control and Prevention (CDC), *Types of Thematic Maps*, available at www.cdc.gov/dhdsp/maps/gisx/resources/maps4.html, accessed on February 22, 2017.

Few, S., *Show Me the Numbers: Designing Tables and Graphs to Enlighten*, 2nd ed., Analytics Press, 2012.

Knaflig, C.N., *Storytelling with Data: a Data Visualization Guide for Business Professionals*, Wiley, 2015.

Tufte, E.R., *The Visual Display of Quantitative Information*, 1st ed., Graphics Press, 1983.

United Nations, *Fundamental Principles of Official Statistics*, October 28, 2013, available at <https://unstats.un.org/unsd/dnss/gp/fundprinciples.aspx>, accessed on February 22, 2017.

U.S. Census Bureau, "How Do We Know? A Snapshot of Our Nation's Veterans," 2012, available at www.census.gov/library/visualizations/2012/comm/veterans.html, accessed on February 22, 2017.

_____, P23-212, *65+ in the United States: 2010*, U.S. Government Printing Office, Washington, DC, 2014, available at www.census.gov/library/publications/2014/demo/p23-212.html, accessed on February 22, 2017.

_____, Current Population Reports, P60-252, *Income and Poverty in the United States: 2014*, U.S. Government Printing Office, Washington, DC, 2015, available at www.census.gov/library/publications/2015/demo/p60-252.html, accessed on February 22, 2017.

U.S. Geological Survey (USGS), "Dasymetric Mapping: An Alternative Approach to Visually and Statistically Enhancing Population Density", 2016, available at <http://geography.wr.usgs.gov/science/dasymetric/index.htm>, accessed on February 22, 2017.