

Chiffres, statistiques et interprétation, une formation de la Société Royale Belge de Statistique

G. Haesbroeck

DÉPARTEMENT DE MATHÉMATIQUE – UNIVERSITÉ DE LIÈGE

Congrès de la SBPMef, août 2021

Contexte

La Société Royale Belge de Statistique a été créée en 1937 et compte aujourd'hui de l'ordre de 250 membres.

Les objectifs de la Société sont de

- contribuer au progrès scientifique en incitant à la collaboration entre les statisticiens belges
- aider le public au sens large à mieux comprendre la place de la statistique dans le monde moderne.

Un projet concret a été lancé par le Conseil d'Administration en 2020: créer une formation en statistique pour les journalistes.

Comment bien rapporter et interpréter des données

FORMATION À DESTINATION DES
JOURNALISTES

ORGANISÉE PAR LA SOCIÉTÉ ROYALE BELGE
DE STATISTIQUE (SRBS)



Pourquoi les journalistes?

De nombreux articles de la presse présentent des données brutes ou des résultats d'analyses statistiques. Il n'est pas rare de constater que

- la présentation n'est pas optimale (voire est biaisée)
- l'interprétation n'est pas correcte/complète/assez prudente

De nombreux citoyens acceptent ces informations sans esprit critique.

Globalement, le “data journalism” se développe (notamment dans les pays anglosaxons), ce qui implique le besoin de compétences additionnelles.

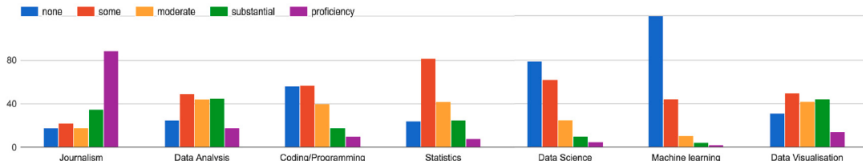


Figure 4. Level of formal training in related knowledge fields (%), N = 181.

Source: Data Journalism practices globally: skills, education, opportunities and values, B.R. Heravi et M. Lorenz,



Supports de cette formation

La première version du contenu et des supports de cette formation a été développée par la plateforme technologique de Support en Méthodologie et Calcul Statistique ([SMCS](https://uclouvain.be/smcs)) de l'Université catholique de Louvain sous la coordination de la Société Royale Belge de Statistique (SRBS).



Souhaits du Conseil d'Administration de la Société Royale Belge de Statistique



HOME PROJECT ACTIVITIES

RSSB-COURSES

Our courses on data visualization and interpretation are available as 'in-house' training. The course can be tailored to the individual needs, for example by using your own publications as a basis for practical exercises. If you are interested you can contact the president of the RSSB via email at president@rssb.be.

TEACHERS

The RSSB supports the teaching of statistics at secondary schools. We offer half-day courses (in Dutch or French) to teachers who are interested in improving their statistical literacy. It is important that our secondary school students can interpret correctly data, and that we learn them to look critically at tables and graphs. The course consists of a series of recent examples seen in the media.

JOURNALISTS



Contenu de la formation préparée par l'équipe de l'UCLouvain

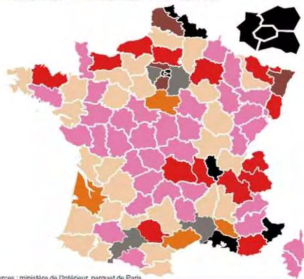
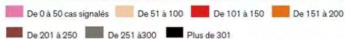
- Présentation et comparaison de chiffres
- Calcul et présentation de pourcentages
- Corrélation
- Graphiques

Aujourd'hui, quelques extraits de cette formation vont être exploités, et quelques exemples complémentaires vont être ajoutés.

Présentation et comparaison de chiffres

- Comparaison à l'aide de chiffres :
 - ▶ Radicalisation en France (Journal du dimanche du 21 juin 2017)
 - ▶ Les chiffres Covid du journal Le Soir du 11 août 2021
 - ▶ La Belgique en tête du bilan de l'AFP (RTBF, 22 mai 2020)
- “Raccourci” de présentation :
 - ▶ Perte d'argent des belges (7 sur 7, 4 juillet 2020)
 - ▶ Coronavirus en Angleterre (Futura Santé, 28 juillet 2021)

La carte de France de la radicalisation



Sources : ministère de l'intérieur, parquet de Paris

Ce graphe est-il facile à lire ?

Peut-on conclure que les régions noires
sont les plus radicalisées ?



TAUX D'INCIDENCE VS NOMBRE ABSOLU

Une échelle avec gradation des couleurs est plus facile à lire.

Le choix des couleurs n'est pas judicieux.

Il est plus facile d'interpréter un taux d'incidence (par rapport à une référence) qu'un nombre absolu.

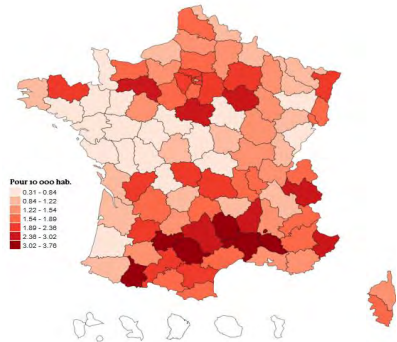
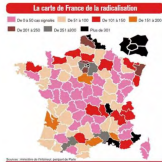
Attention, il faut que le nombre de personnes radicalisées et le nombre d'habitants par région aient été relevés à une même date.

Bien que le graphique soit **correct** (et utile pour certaines analyses),

le nombre total de personnes radicalisées n'est pas toujours un bon indicateur de l'ampleur de la radicalisation d'une région (car les régions ont des densités de population différentes) !

TAUX D'INCIDENCE VS NOMBRE ABSOLU

Avec un taux d'incidence (et une gradation des couleurs), le même graphe devient :



Utilisation de chiffres bruts

Il n'est pas rare de "montrer" des chiffres bruts dans la presse:

Les chiffres du covid en Belgique

Données au 10/08 - Evolution sur une semaine

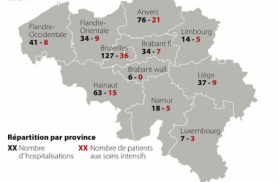


Avancement de la vaccination chez les + de 18 ans

Personnes ayant reçu une première dose de vaccin (red) Non-vaccinées (grey)



Occupation des hôpitaux



Répartition par province

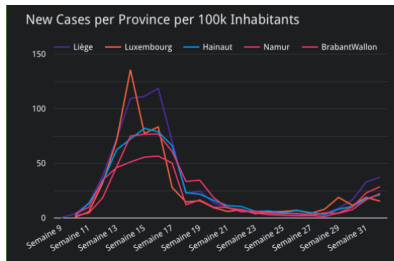
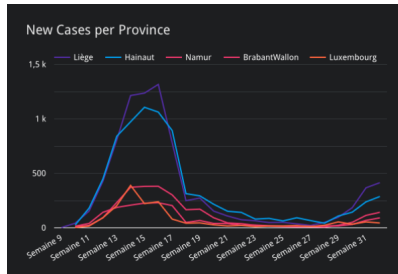
XX Nombre d'hospitalisations XX Nombre de patients aux soins intensifs

C'est l'utilisation de ceux-ci pour effectuer des comparaisons qui peut être problématique.

Comparaison à l'aide d'une base commune

Pour comparer des chiffres obtenus sur des entités de masse différente, il convient de travailler en pourcentages ou taux.

Exemple: les statistiques de Sciensano (le “ranking” des provinces varie si on parle en termes de chiffres bruts ou en termes de taux):



Cependant, la difficulté d'opérer une comparaison honnête dépasse le simple fait d'imposer une base commune.



La Belgique a le taux de mortalité lié au Covid-19 le plus élevé au monde

Selon un bilan établi par l'AFP, la Belgique est le pays au monde avec le plus grand nombre de décès par rapport à sa population, soit 84 décès pour 100.000 habitants. Le Royaume-Uni et l'Italie suivent.

Rédaction 04-07-20, 22:13 Source: Belga



247



3

Quels facteurs pourraient, selon vous, expliquer que la Belgique figure en tête de ce classement le 04/07/2020 ?

- 4 ministres de la santé et 4 ministres de la prévention de la santé ?
- Pénurie de matériel ?
- Consommation excessive de chocolat ?

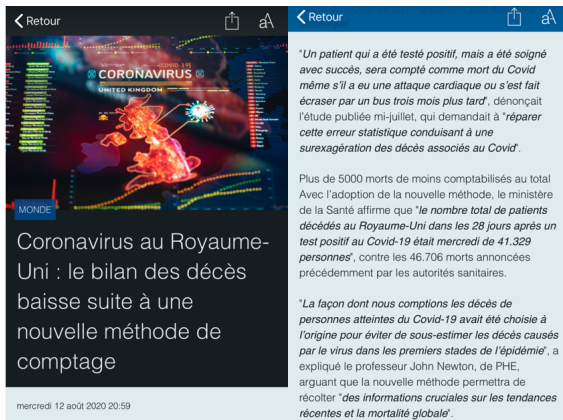


COMPARABILITÉ

Bien que des taux (plutôt que des chiffres bruts) soient utilisés, **ces chiffres ne sont quand même pas tous comparables :**

- La **densité de population** est différente d'un pays à l'autre.
- La méthode de **comptage** et la **fiabilité** des chiffres rapportés peuvent varier d'un pays à l'autre.
 - Certains pays ne comptabilisent que :
 - les décès dans les hôpitaux (et non en maisons de retraite), ou
 - les décès de personnes testées positives.
 - Le **dépistage** peut ne pas être effectué de la même manière : les pays pauvres manquent de tests, d'autres ne testent que les cas graves, etc.
- Les données ne correspondent pas au même **stade de l'épidémie**.
 - A un même instant, certains pays ont des cas sur leur territoire depuis plusieurs semaines tandis que d'autres n'ont eu leurs premiers patients que quelques jours auparavant !

Illustration concrète de la difficulté des comparaisons “spatiales”





COMPARABILITÉ

Dans une comparaison entre plusieurs groupes (comme des zones géographiques), il faut s'assurer que les chiffres rapportés soient obtenus de la même manière et représentent bien la même chose !

Il faut donc éviter autant que possible de comparer ce qui n'est pas comparable...

Il est plus correct de comparer la surmortalité que le nombre de cas liés au covid-19 pour les raisons énoncées précédemment.

Des chiffres sans contexte ne veulent rien dire !

Et/ou bien nuancer les propos !

Avertir le lecteur sur les dangers des comparaisons, ne pas hésiter à utiliser le conditionnel, expliquer pourquoi les chiffres comparés doivent être interprétés avec prudence...

Erreurs de présentation/interprétation des résultats

Des “raccourcis” (titre accrocheur, calculs trop simplistes...) peuvent amener des fausses impressions aux lecteurs.



Économie

Selon Test Achats, les ménages belges ont perdu en moyenne 1626 euros durant la crise du coronavirus

RTBF

🕒 Publié le vendredi 22 mai 2020 à 05h05



527 ➦

Plus de 60% des ménages belges déclarent **avoir perdu de l'argent**, pour un montant moyen de 1626 euros, lors de **la crise sanitaire du coronavirus**. C'est ce qui ressort d'un sondage réalisé par Test Achats les 14 au 15 mai auprès d'un échantillon de 1000 Belges, âgés de 18 à 74 ans. Au total, cette perte financière représente 4,9 milliards d'euros.



MOYENNE GLOBALE VS DANS UN GROUPE

Ne pas confondre moyenne générale et moyenne dans un groupe !

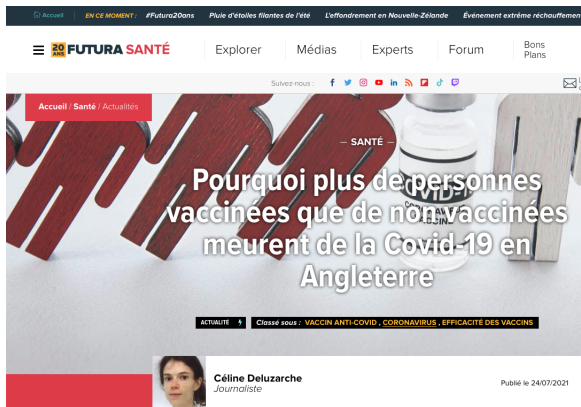
Selon Test Achats, les ménages belges ont perdu en moyenne 1626 euros durant la crise du coronavirus.
[Titre de l'article]

Cela voudrait dire que *tous les ménages belges (100 %)* ont perdu en moyenne 1626 € pendant la crise du coronavirus. Ce chiffre est énorme et incorrect !

Plus de 60 % des ménages belges déclarent avoir perdu de l'argent, pour un montant moyen de 1626 euros, lors de la crise sanitaire du coronavirus. C'est ce qui ressort d'un sondage réalisé par Test Achats les 14 au 15 mai auprès d'un échantillon de 1000 Belges, âgés de 18 à 74 ans. Au total, cette perte financière représente 4,9 milliards d'euros. [Contenu de l'article]

Cela veut dire que *parmi les ménages belges ayant perdu de l'argent (> 60 %)*, le montant moyen perdu est de 1626 €.

Article dans le Futura Santé du 27 juillet 2021



C'est une statistique assez perturbante dont se nourrissent les antivax : entre le 1^{er} février et le 21 juin 2021 en Angleterre, il y a eu davantage de morts de la Covid-19 parmi les personnes vaccinées et testées positives au variant Delta que parmi les personnes non vaccinées, atteste un rapport du gouvernement. Plus exactement, sur les 257 décès, 45 personnes avaient reçu une dose, 118 avaient reçu deux doses, et 92 n'avaient reçu aucune injection. La conclusion hâtive qu'il pourrait être tiré de ces statistiques est que le vaccin n'est pas efficace, voire dangereux.

Données du rapport du gouvernement

Table 5. Attendance to emergency care and deaths by vaccination status among Delta confirmed cases (sequencing and genotyping) including all confirmed Delta cases in England, 1 February 2021 to 21 June 2021

	Age group (years)**	Total	Cases with specimen date in past 28 days	Vaccination status unknown	<21 days post dose 1	≥21 days post dose 1	Received 2 doses	Unvaccinated
Delta cases	All cases	123,620	63,707	14,359	8,562	17,933	10,834	71,932
	<50	111,008	57,673	12,900	8,453	13,391	5,600	70,664
	≥50	12,404	5,957	1,252	109	4,542	5,234	1,267
Deaths within 28 days of positive specimen date	All cases	257	N/A	2	1	44	118	92
	<50	26	N/A	-	-	3	2	21
	≥50	231	N/A	2	1	41	116	71

Que voit-on “de prime abord” ?

Globalement, risque plus élevé parmi les vaccinés

En ne prenant que les catégories Vaccinés 2 doses et Non Vaccinés:

	Vaccinés 2 doses	Non Vaccinés
Nombre de cas	10834	71932
Nombre de décès	118	92
Risque	1%	0,1%

Mais un *facteur confondant* important n'est pas pris en compte dans ce calcul de risque: l'âge...

Les groupes des personnes vaccinées ou non vaccinées ne sont pas homogènes selon cette caractéristique.

En tenant compte de l'âge

	Vaccinés 2 doses		Non Vaccinés	
	< 50	≥ 50	< 50	≥ 50
Nombre de cas	5600	5234	70644	1267
Nombre de décès	2	116	21	71
Risque	0,03%	2,2%	0,03%	5,6%

Dans chaque “tranche d'âge”, les risques de décès sont égaux ou plus importants pour les non vaccinés par rapport aux vaccinés. Cependant, globalement, la catégorie plus âgée est “sous-représentée” dans la catégorie des non vaccinés, ce qui change l'effet (paradoxe de Simpson).

Utilisation de pourcentages (de variation)

- Choix de la base
 - ▶ Evolution des cas Covid (RTBF, 18 juillet 2020), réaction de B. Rentier (7 sur 7, 5 août 2020)
 - ▶ Féminisation des noms de rues en Flandre (Radio 1, 3 mars 2020) ou à Bruxelles (RTBF, 3 mars 2020)
- Erreur ou imprécision
 - ▶ Taux de réussite à la hausse
 - ▶ Evolution du leasing
 - ▶ Pourcentages de chômage par tranche d'âge

Rappelez-vous en juillet/août 2020



Les pourcentages d'augmentation annoncés quotidiennement dans la presse sont “vertigineux”



Les calculs sont corrects

Le pourcentage de variation caractérisant l'évolution (augmentation ou diminution) d'une variable (à valeur positive) au cours du temps est donné par

$$\% \text{ de variation} = \frac{\text{valeur finale} - \text{valeur initiale}}{\text{valeur initiale}} \cdot 100\%,$$

et les résultats annoncés dans la presse et calculés à partir des chiffres bruts étaient corrects.

Mais

- Il faut “relativiser” ou interpréter avec prudence les pourcentages (de variation) lorsque ceux-ci sont calculés sur des chiffres très réduits
- Il faut s'assurer du caractère “invariable” des conditions de calcul.

Réaction de B. Rentier (ex-Recteur de l'ULiège)

77 ACTU SHOW SPORT LIFESTYLE VIDÉO Q

Belgique Monde Faits divers Insolite Ecologie Tech Santé Economie Sciences

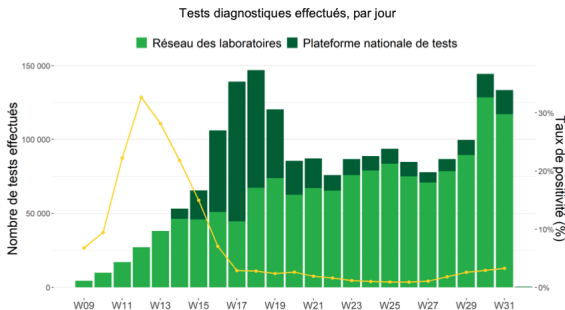


▲ Bernard Rentier © belga

L'ex-recteur de l'ULiège et virologue Bernard Rentier veut présenter les chiffres du Covid-19 autrement: "Ce n'est pas +60%, mais +1% de cas confirmés"

Définition du taux de positivité

Il faut relativiser les nombres de cas détectés en fonction du nombre de tests effectués → le “taux de positivité” est maintenant ajouté par défaut dans les rapports (de la RTBF notamment).



Proximus 09:42 99 %

Cas détectés¹ : entre le 01/08 et le 07/08, 1755 nouvelles infections au coronavirus ont été détectées en moyenne chaque jour. C'est une hausse de 12% par rapport à la semaine précédente.

Tests : entre le 01/08 et le 07/08, une moyenne de 50.973 tests ont été effectués quotidiennement, un total en baisse de 6% par rapport à la semaine précédente.

Admissions : elles s'élèvent en moyenne à 47,4 entre le 04/08 et le 10/08. C'est une hausse de 22% par rapport à la semaine précédente.

Personnes hospitalisées² : 467 patients sont actuellement hospitalisés en lien avec le Covid-19, dont 121 patients traités en soins intensifs.

Taux de positivité³ : sur la base des résultats des tests obtenus entre le 01/08 et le 07/08, il est de 3,8%, en hausse de 0,6% par rapport à la semaine dernière.



Actualité > Belgique



Réagir

Publié le lundi 6 Juillet 2020
à 22h45
Par Sudinfo avec Belga

J'aime 9

Partager

Coronavirus: malgré les craintes, le taux de réussite à la hausse cette année dans l'enseignement supérieur

Ainsi, dans l'ensemble des six universités francophones, le taux de réussite lors de cette session de juin progresse de pas moins de 9% en moyenne.

Selon des chiffres livrés parallèlement par le Conseil des recteurs francophones (Cref), le taux d'examens réussis les universités en juin est de 77%, alors qu'il tournait autour de 68% les trois années précédentes.

Quelque chose vous surprend-il dans ce texte ?

<https://www.sudinfo.be/id217851/article/2020-07-06/coronavirus-malgre-les-craintes-le-taux-de-reussite-la-hausse-cette-annee-dans>

Calculs “incorrects”

Afin de relativiser les chiffres bruts ou calculer des pourcentages, encore faut-il

- choisir la bonne base;
- effectuer le calcul “dans le bon sens”



MEER VROUW OP STRAAT

"Van de 3500 straten in Antwerpen, zijn er amper 70 naar vrouwen vernoemd"

3 maart 2020



Cela représente 2 % des rues seulement.

On sous-entend :

- que **les femmes sont largement sous-représentées** dans les noms des rues.
- qu'il y aurait une **discrimination basée sur le genre** dans les noms des rues.

Que pensez-vous de ces deux suppositions ?

Même type d'articles à la RTBF

INFO Rechercher sur l

À la une Fil Info **Coronavirus** Belgique Régions Monde Économie Société Média

Régions Bruxelles Brabant Wallon Hainaut Liège Namur Luxembourg Flandre

A peine 6 % des rues bruxelloises portent des noms de femmes

A map of Brussels, Belgium, showing a dense network of streets. Most streets are highlighted in yellow, while a small number of streets are highlighted in purple, representing the 6% of streets named after women. The map includes labels for various districts and landmarks.

Seules 6 % des rues bruxelloises portent des noms de femmes © Tous droits réservés



POSTENPAKKER MEER VROUW OP STRAAT

Dat krijgt nog een straatje

In Antwerpen zijn het er maar 70 van de 3.469, zei Sofie Lemaire aan het begin van *Meer vrouw op straat*. Dat zijn 70 straten die naar een vrouw zijn genoemd. Omdat er bij de 3.469 ook wel Kerkstraten en Grote Banen zullen zijn, loopt het weetje een beetje mank: we weten niet hoeveel Antwerpse straten naar mannen zijn genoemd. Meer dan 70, nemen we aan.

... il y a également des noms comme "Kerkstraat" et "Grote Baan" parmi les 3 469 rues... nous ne savons pas combien de rues d'Anvers portent un nom d'hommes. Plus de 70, sans doute...

Si on tient compte du fait que certaines rues n'ont un nom ni d'homme, ni de femme... **Ça change le sens de l'information !**



■ CALCULER UN POURCENTAGE

Quand on mentionne une proportion, bien s'assurer que le total considéré est pertinent pour l'analyse que l'on en fait !

Calcul effectué

% de nom de femme dans les rues =
Nombre de rues ayant un nom de femme / Nombre total de rues

Interprétation : **Un pourcentage par rapport à l'ensemble des rues.**

Calcul à effectuer

% de nom de femmes dans les rues ayant un nom de personne =
Nombre de rues ayant un nom de femme / Nombre de rues ayant un nom de personne

Interprétation: **Un pourcentage en rapport aux rues ayant un nom de personne.**

Erreur de “sens”

JT de la RTBF (12 janvier 2018)

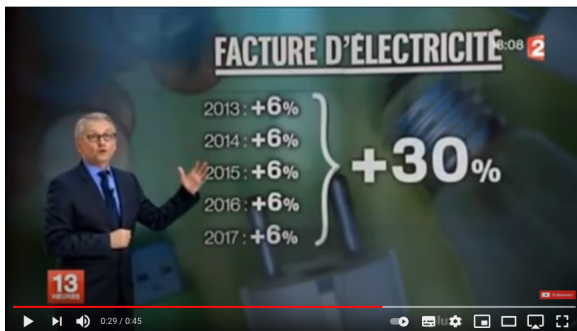


Annonce d'une augmentation de 30% d'utilisateurs du système de leasing en un an sur base des éléments suivants:

- Janvier 2016: 7000 utilisateurs
- Janvier 2017: 10000 utilisateurs

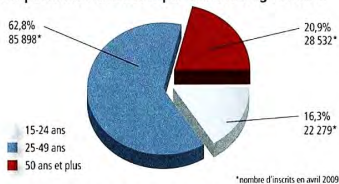
D'autres erreurs/raccourcis sont possibles

JT de FR2 (janvier 2013,
<https://www.youtube.com/watch?v=z6OyOpDqhwk>)





Proportion de chômeurs par tranche d'âge en avril



Les chiffres les plus récents démontrent et confirment que la classe d'âge la plus touchée reste celle des 25-49 ans.

Sachant que, dans ce diagramme, on se réfère à la population des chômeurs, **auriez-vous tiré la même conclusion?**

Parmi tous les chômeurs, 16,3 % ont entre 15 et 24 ans.

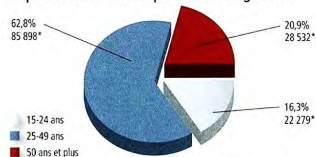
Parmi tous les chômeurs, 62,8 % ont entre 25 et 49 ans.

Parmi tous les chômeurs, 20,9 % ont 50 ans ou plus.



■ POURCENTAGE GLOBAL ET AU SEIN DE CATÉGORIES

Proportion de chômeurs par tranche d'âge en avril



* nombre d'inscrits en avril 2009

Les chiffres les plus récents démontrent et confirment que la classe d'âge la plus touchée reste celle des 25-49 ans.

Sachant que dans ce diagramme on se réfère à la population des chômeurs, auriez-vous tiré la même conclusion?

Parmi tous les chômeurs, 16,3 % ont entre 15 et 24 ans.
Parmi tous les chômeurs, 62,8 % ont entre 25 et 49 ans.
Parmi tous les chômeurs, 20,9 % ont 50 ans ou plus.

Ce qui nous intéressait vraiment !

Parmi les 15-24 ans, quel est le pourcentage de chômeurs ?

Parmi les 25-49 ans, quel est le pourcentage de chômeurs ?

Parmi les ≥ 50 ans, quel est le pourcentage de chômeurs ?

Graphiques

Des graphiques de “plus en plus sophistiqués” apparaissent dans les médias, avec, dans certains cas des notions sous-jacentes non triviales.

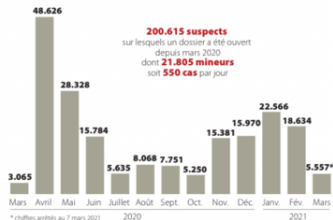
Ce sont les diagrammes en barres et en secteurs qui sont les plus utilisés mais des cartogrammes, des courbes temporelles,... se retrouvent fréquemment dans les journaux.

Quelques exemples

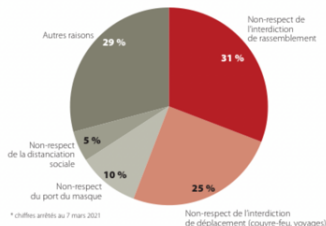
Le journal Le Soir des 13 et 14 mars 2021

Nombre de suspects impliqués dans des affaires ouvertes pour « infraction corona »

Au niveau des parquets correctionnels, des parquets de la jeunesse et des audiorats du travail



La motivation du P.-V.

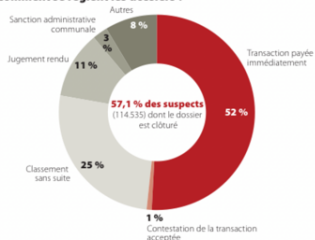


une page sur deux

* chiffres arrêtés au 7 mars 2021

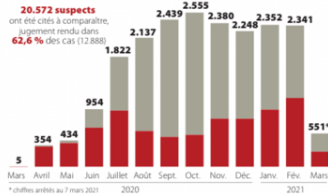
Non-respect de l'interdiction de déplacement (couvre-feu, voyages)

Comment se règlent les dossiers ?



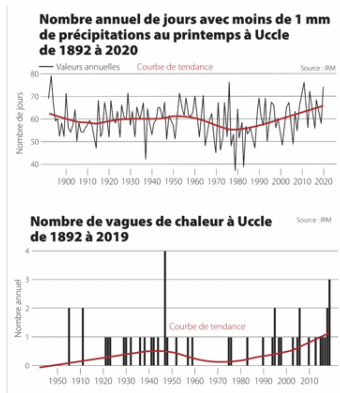
Nombre de suspects cités à comparaître devant un tribunal pour infraction aux « mesures covid »

■ Citation suite au non-paiement d'une transaction immédiate
■ Citation sans décision préalable (majoritairement pour les récidivistes)



Quelques exemples

Le journal Le Soir du 9 août 2021



Quelques exemples

Le journal Le Soir du 25 décembre 2017

Les mots les plus utilisés par Theo Francken dans ses tweets :

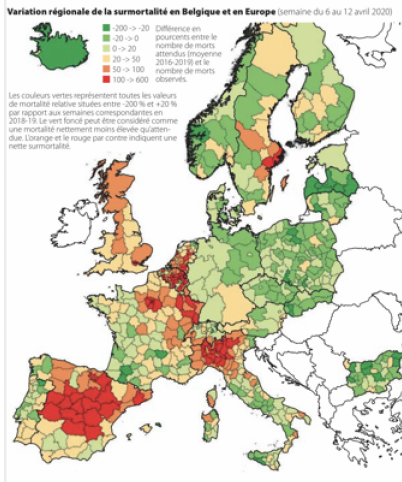


Les mots les plus utilisés par Charles Michel dans ses tweets :



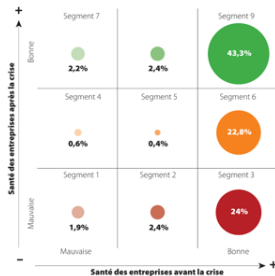
Quelques exemples

Le journal Le Soir du 28 août 2020



Quelques exemples

Le journal Le Soir du 21 janvier 2021



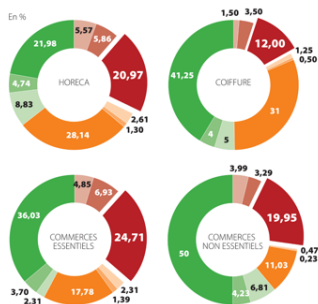
La santé des entreprises en un tableau

La comparaison des deux modèles Graydon s'exprime dans un tableau divisé en neuf segments. Comment le lire ? Sur l'axe horizontal, la santé des entreprises avant la crise. Cela se lit donc par colonne : à gauche (segments 1, 4 et 7) celles qui, en mars 2020, se trouvaient en mauvaise ou très mauvaise santé, c'est-à-dire à haut risque de faillite dans l'année. À droite (segments 3, 6 et 9) celles qui se trouvaient en bonne santé. Au milieu : les entreprises qui connaissaient des difficultés mais sans risque de faillite immédiate. Sur l'axe vertical : quel est l'impact de la crise. La ligne du bas (segments 1, 2 et 3) montre les entreprises largement impactées alors que la ligne du haut concerne les entreprises qui ont encore largement des réserves pour faire face.

Concrètement, les secteurs qui retiennent l'analyse sont d'abord les segments 3 et 9. Dans le 3, les entreprises saines avant la crise, qui ont besoin d'une injection de cash immédiate pour pouvoir survivre. Dans le 9, les entreprises qui se trouvaient en très bonne santé, avec de larges réserves avant la crise, et qui puisent dans ces réserves pour résister, sans les épuiser.

Les plus mal en point, le segment 1 : en difficulté majeure avant la crise, en situation encore plus menacée aujourd'hui. Quant au segment 7, il regroupe les (rares) entreprises en très mauvaise posture avant le covid qui, aujourd'hui, se trouvent dans une situation bien meilleure. B.JY ET VLA.

Graphiques “trompeurs”



Vingt types de commerces dans six villes

Le Soir a demandé à Graydon d'appliquer son modèle à une vingtaine de commerces (restaurants, bars, snacks, épiceries, poissonneries, boucheries, boulangeries, fromageries, cavistes, magasins de bricolage, jardinerie, librairies, marchands de journaux, magasins de vêtements, de chaussures, d'électro-ménager, de télécoms, parfumeries et salons de coiffure) en Région bruxelloise et dans cinq villes wallonnes (Charleroi, Liège, Louvain-la-Neuve, Mons et Namur). Nous avons ensuite regroupé ces commerces en quatre types (horeca, commerces essentiels et non essentiels, coiffure) avant de regrouper, en Wallonie, les cinq villes en une seule, comparable à Bruxelles. B.Y.J ET V.L.A.

Graphiques “trompeurs”

Le Soir Lundi 26 juillet 2021

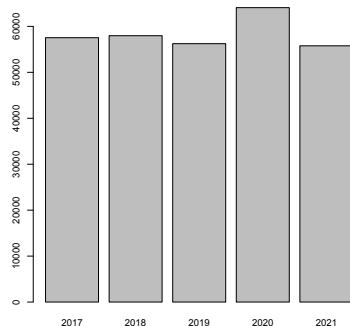
à la une 11

DÉCÈS

La mortalité à son niveau le plus bas

Les statistiques pour le premier semestre 2021 sont sorties. Elles montrent un recul des décès après le pic de 2020 provoqué par le covid. Leur nombre est le plus bas enregistré au cours des cinq dernières années.

L'évolution des décès sur les six premiers mois de l'année



Le Soir Mercredi 8 février 2017

4 LABELGIQUE

Le niveau de vie baisse pour les plus défavorisés

BIEN-ÊTRE Niveau de vie, travail, santé : les plus pauvres sont toujours plus mal lotis

► Le rapport du Bureau fédéral du Plan a été présenté mardi.

► Il n'est pas dénué de bonnes nouvelles, sauf pour les plus pauvres.

Depuis 2014, le Bureau fédéral du Plan est chargé par le gouvernement de mesurer la qualité de la vie en Belgique. Pour ce faire, il lui a été demandé d'établir une série d'indicateurs complémentaires au PIB qui, s'il s'intéresse au niveau de production d'un pays, reste assez inutile quand il s'agit d'évaluer le bien-être de sa population.

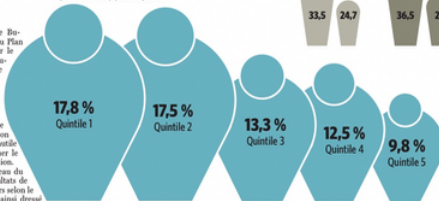
Cette année, le Bureau du Plan a ventilé les résultats de treize de ces indicateurs selon le revenu. Et le tableau ainsi dressé est peu réjouissant.

Si l'écart de bonheur entre les 20 % de Belges les plus pauvres et les 20 % les plus riches (notamment dit entre les quintiles 1 et 5, ainsi qu'on le verra dans l'infographie ci-dessous), ne pas avoir de télévision (alors qu'elle le souhaite), ne pas avoir de machine à laver (alors qu'elle le souhaite), ne

Population obèse selon la catégorie de revenu

Pourcentage de la population de 18 ans et plus

Chaque quintile représente 15 % de la population globale, du 1^{er} au 5^e, sur base de leurs revenus. Ainsi, le premier quintile représente les 20 % de Belges au revenu le plus faible. Le second quintile représente la moyenne basse. Le troisième groupe dispose d'un revenu moyen. Le quatrième quintile représente les 20 % de la population disposant d'un revenu plus élevé. Le cinquième quintile représente les 20 % de la population les plus riches.



disposer d'un poste de télévision (alors qu'elle le souhaite), ne pas avoir de téléphone (alors qu'elle le souhaite), ne pas avoir de machine à laver (alors qu'elle le souhaite), ne

plus roses. L'obésité, si elle concerne 13,7 % de la population, touche environ 17 % du premier quintile et un peu moins de 10 % des 20 % de Belges les plus riches.

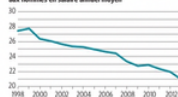
Espérance de vie à 50 ans selon le sexe et le niveau d'éducation

En années de vie après 50 ans



Ecart salarial entre les hommes et les femmes

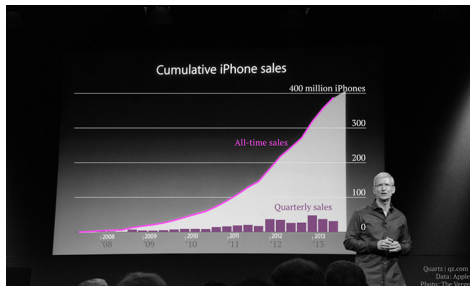
Pourcentage de différence des femmes par rapport aux hommes en salaire annuel moyen



Les inégalités hommes-femmes se réduisent

Utilisation de courbes cumulatives

En septembre 2013, Tim Cook a présenté l'évolution des ventes d'iphones à l'aide d'un graphique cumulé:

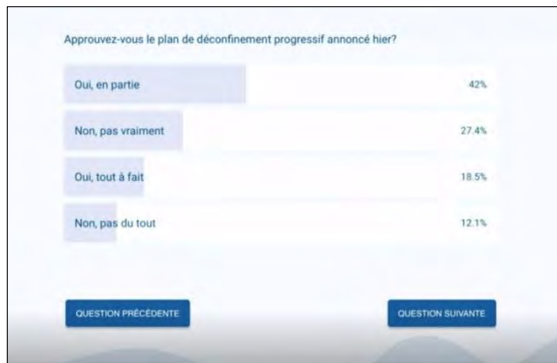


... pour ne pas (selon ses détracteurs) montrer un graphique trimestriel illustrant une baisse des ventes.

Des erreurs “flagrantes” sont aussi possibles

Reportage sur BFMTV (25 avril 2018)





5, 4, 3, 2, 1... Terminé !

Autres demandes des journalistes

A la fin de la formation, les journalistes étaient invités à suggérer des thèmes additionnels en vue de futures formations.

- Explications additionnelles sur la différence entre moyenne et médiane

Le Soir (8 et 9/05/21)

ENQUÊTE

Le salaire médian à 3.350 euros brut par mois

Le salaire médian en Belgique s'élève à 3.350 euros brut par mois, ce qui signifie qu'une moitié des Belges gagne davantage et l'autre moitié moins, selon l'enquête annuelle de Jobat. Quelque 83.000 Belges des quatre coins du pays ont participé à l'enquête du site internet d'offres d'emploi. Le salaire moyen, lui, s'élève à 3.783 euros par mois. Les secteurs de la chimie, de la pharmacie, de l'énergie, des banques et assurances sont ceux qui rémunèrent le mieux. A l'inverse, les salaires sont les plus bas dans l'horeca, le tourisme et les loisirs. Pour la première fois en quatre ans, le nombre de jours de vacances a augmenté pour les employés, de 27 à 28 en moyenne, alors qu'il est resté stable, à 24, pour les ouvriers. Enfin, le nombre d'avantages extralégaux continue à augmenter, à un peu plus de six en moyenne. Vu la généralisation du télétravail, la hausse concerne principalement les ordinateurs portables et les

Le Soir (5 et 6/12/20)

«Sauf que ce n'est pas nécessairement la faute des labos», réplique Alain Derom, président de la commission de biologie clinique. «Les chiffres varient vraiment d'un jour à l'autre. Ils tombent à 9 % parfois mais nous sommes actuellement dans une période calme, nous faisons les tests dans les temps.» Selon celui qui est aussi à la tête d'un des plus gros labos du pays, le problème vient plutôt de la logistique. «On pourrait encore renforcer la récolte des échantillons. Mais de toute façon, quand un laboratoire reçoit un lot à 17 h 30 et ferme à 18 h, il n'y aura pas d'analyse jusqu'au lendemain matin.» Alain Derom glisse aussi que, selon les données officielles, les «mauvais élèves» qui tirent la moyenne vers le bas ces derniers jours sont plutôt les laboratoires d'analyse liés aux universités, ceux de la plateforme fédérale «bis» financée par l'Etat fédéral.

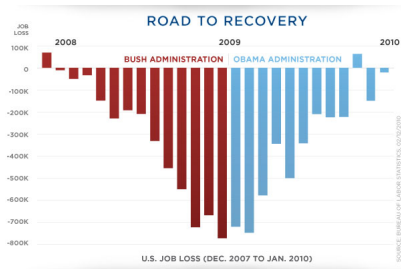
Deuxième bémol : la confiance générale dans la machine. Un labo sur cinq ne va pas «assez vite» mais une personne testée positive sur cinq ne communique toujours aucun contact à l'agent. A la fin de l'été, ce pourcentage

Autres demandes des journalistes

- Choix d'une forme de graphique appropriée

Utiliser un graphique peut être un moyen très efficace de passer un message.

Ex: courbe de Plouffe (conseiller de Obama)



En ordonnée: variation du volume d'emploi par rapport à une référence de niveau 0.

⇒ déclin économique sous la présidence de Bush; redressement économique sous Obama.

Autres demandes des journalistes

- Calcul de la marge d'erreur dans un sondage



Les pourcentages sont calculés sur base des répondants qui ont exprimé une intention de vote pour un parti, soit 839 personnes sur les 958 personnes interrogées en Wallonie (88%). Les réponses "ne votera pas" et "ne veut vraiment pas répondre" représentent 12% des personnes interrogées.

Deux marges d'erreur?

Une marge d'erreur de 3,1% et une autre de 3,4% sont citées dans l'extrait.

- Ainsi qu'indiqué explicitement dans la description de la méthodologie suivie, les pourcentages sont calculés sur les 839 personnes ayant choisi un des partis, mais le sondage proposait trois autres items, et, en comptant les réponses données à ces items, le nombre de répondants augmente (958), ce qui diminue la marge d'erreur *globale* (pas celle attachée aux pourcentages des partis).
- Il est bien mentionné dans l'article que c'est *la marge d'erreur théorique maximale* qui est donnée.

En réalité, pour pouvoir préciser une seule marge d'erreur valide pour l'ensemble des résultats (mesure théorique: valide pour toute proportion observée), c'est la variance *maximale* qui est exploitée:

Marge d'erreur maximale pour les % des

$$\text{partis} = 1,96 \cdot \frac{\frac{1}{2}(1-\frac{1}{2})}{\sqrt{839}} = 0,034$$

$$\text{Marge d'erreur maximale du sondage:} = 1,96 \cdot \frac{\frac{1}{2}(1-\frac{1}{2})}{\sqrt{958}} = 0,0316$$

Autre moyen d'attirer les élèves vers la statistique: les faire participer aux Olympiades de statistique



NL **FR** DE EN

Autres

STATBEL
La Belgique en chiffres

ACCUEIL THÈMES NOUVELLES PARTICIPANTS /

Accueil

Olympiades de Statistique 2021-2022

f in t e b

Les Olympiades de Statistique reviennent en Belgique en 2021-2022. Inscrivez-vous à notre [newsletter](#) et découvrez tout [ici](#).

- [Que sont les Olympiades de Statistique ?](#)
- [À quoi s'attendre ?](#)
- [Comment m'inscrire ?](#)
- [Newsletter](#)
- [Contact](#)
- [Éditions précédentes](#)
- [Règlements](#)



Olympiades de statistique

Que sont les Olympiades européennes de Statistique?

Les [Olympiades européennes de Statistique](#) sont une initiative d'Eurostat, l'office européen de statistique. Elles seront organisées au cours de l'année scolaire 2021-2022 dans plus de 16 pays européens. La Belgique y participera à nouveau. Statistiek Vlaanderen, l'Iweps (Wallonie) et Statbel (office belge de statistique) unissent leurs forces et organisent ensemble les Olympiades en Belgique.

L'ensemble de la compétition se déroule en ligne et comprend deux étapes : une étape nationale et une étape européenne. L'étape belge se fait en deux épreuves. Nous proposons d'abord à vos élèves un test de connaissances. Ils disposent de plusieurs semaines pour le réaliser. Les participants qui réussissent pourront alors consolider leurs compétences créatives et communicationnelles lors de la réalisation d'une vidéo. Les équipes gagnantes seront sélectionnées pour l'épreuve européenne. Plus de détails. Un calendrier et les détails sont décrits dans la rubrique : [À quoi s'attendre ?](#)

Pourquoi des Olympiades de Statistique?

L'amélioration des connaissances statistiques des jeunes et des adultes est un point d'attention important des offices de statistique à travers le monde. Eurostat, en tant qu'office européen de statistique, veut également y contribuer. C'est pourquoi a été lancée en 2017-2018 la Compétition européenne de Statistique ou Olympiades européennes de Statistique.

Lors de l'année scolaire 2020-2021, plus de 11.500 étudiants de 16 pays européens y ont participé.

Les Olympiades européennes de Statistique en Belgique sont une coopération entre Statbel, Statistiek Vlaanderen, IWEPS, INE Espagne et Eurostat soutenu par le département flamand pour l'économie, la science et l'innovation, le département de l'enseignement en Communauté flamande, le département de pédagogie du ministère de la Communauté germanophone et la Société Royale Belge de Statistique.

Calendrier

- Octobre - novembre: inscriptions (en ligne sur le site de Statbel)
- Janvier: test en ligne (y compris analyse d'un fichier de données)
- Février - mars: production d'une vidéo sur un thème déterminé
- Avril: sélection de 5 vidéos par degré par un jury régional
- Fin avril: sélection d'un gagnant par degré (participation au tour suivant: Olympiades européennes)