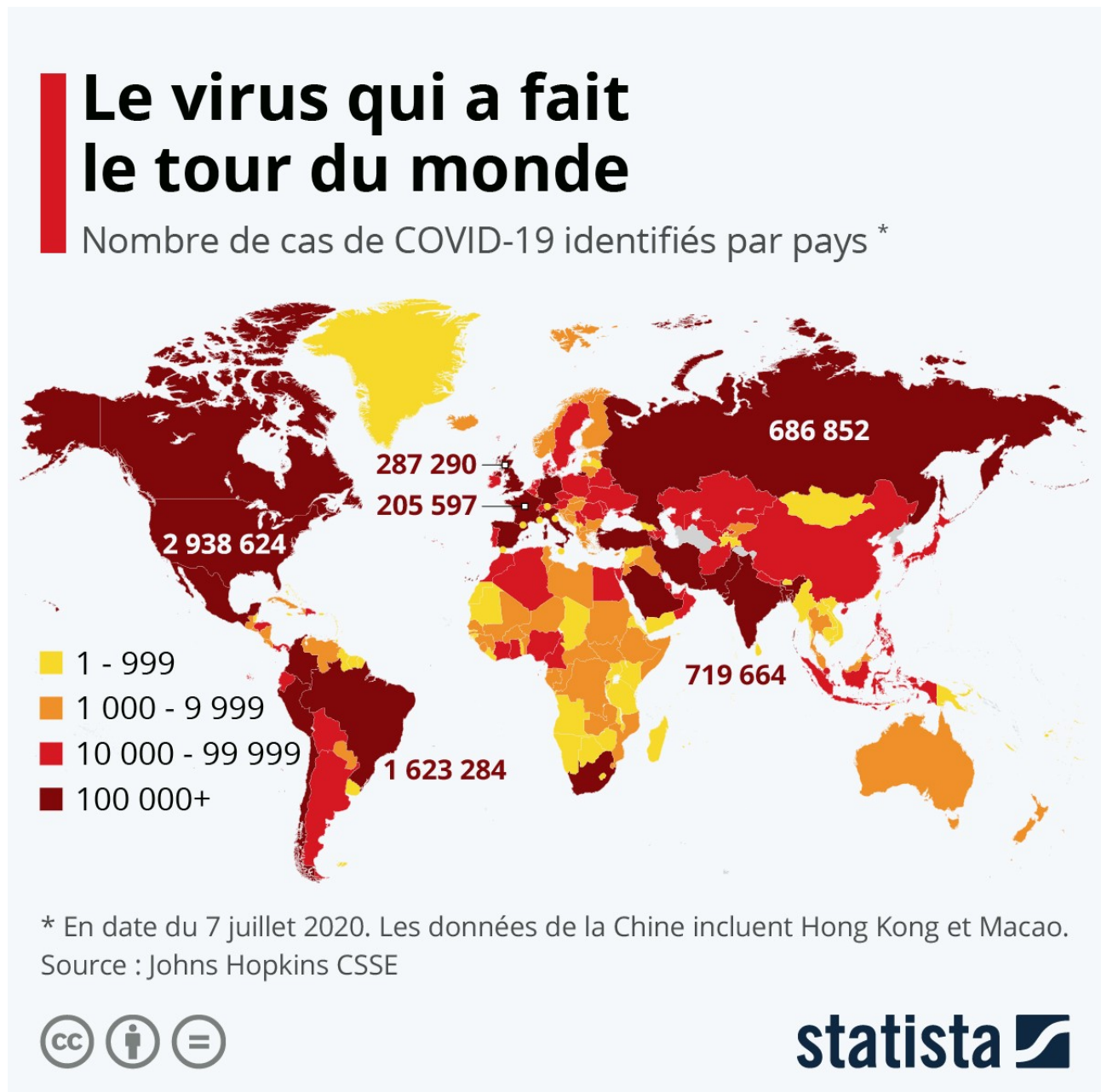


La proportionnalité pour les nuls

Révision : la règle de trois.

https://fr.wikipedia.org/wiki/R%C3%A8gle_de_trois

Travaux pratiques – 1 le Covid dans le monde



<https://fr.statista.com/infographie/20622/carte-des-cas-covid-19-confirmes-par-pays-dans-le-monde/>

Pour l'Europe, plusieurs pays sont en jaune : le Liechtenstein, Gibraltar, Andorre, Monaco, Saint-Marin, Malte, le Monténégro, le Kosovo, la Lettonie et la Géorgie.

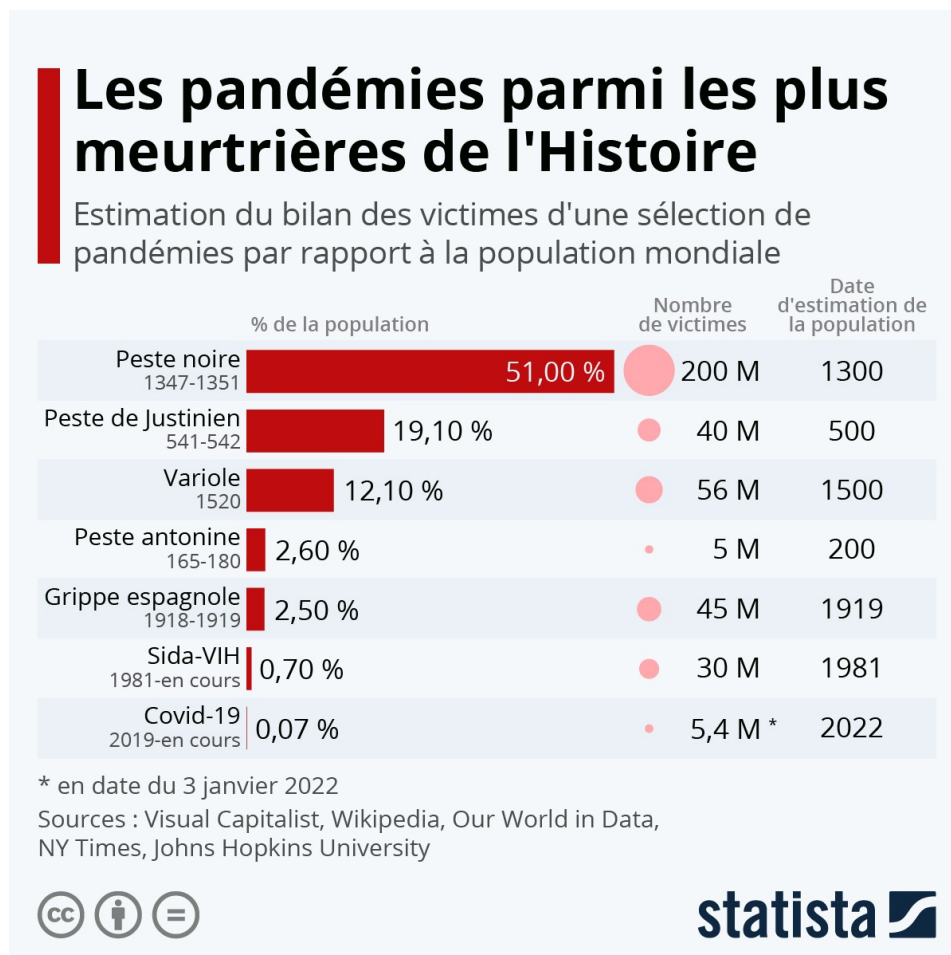
À l'inverse, plusieurs pays européens sont en rouge foncé : le Royaume-Uni, la France, l'Allemagne, l'Espagne et l'Italie.

Pouvez-vous en déduire quels pays européens sont les plus touchés par le Covid en 2020 ?

Selon ces statistiques officielles, quel est le pays le plus affecté par le Covid, entre la France (68 millions d'habitants) et l'Inde (1326 millions d'habitants) ?

Comparez la situation en France, et à Saint Marin (1100 cas, 33607 habitants).

Travaux pratiques – 2 les pandémies dans l'Histoire



<https://fr.statista.com/infographie/26184/bilan-victimes-pandemies-par-rapport-population-mondiale/>

Quelles sont les épidémies qui ont fait le plus de victimes ? (1) en **valeur absolue**, et (2) en **valeur relative**.

Travaux pratiques – 3 en linguistique

sur la position d'un point d'inflexion. Pour évaluer l'impact général des HN (Svennson (2010)) a recensé le nombre de réponses fournies par Google à la requête *Digital Humanities* et cela à deux intervalles de temps séparés de 18 mois : le 10-3-2008 et le 6-10-2009. Nous avons, de notre côté, soumis cette même requête un an plus tard : le 22-9-2010 et puis récemment, 1er février 2014. Le résultat confirme parfaitement le boom comme le montre le tableau 1 :

TAB. 1 – *Visibilité des HN sur Internet*

Date de la requête	Nombre de références données par Google
10-3-2008	98 800
6-10-2009	180 000
22-9-2010	8 930 000
1-02-2014	45 000 000

Voyez-vous le problème ?

Voyez-vous l'autre problème ?

Voyez-vous le troisième problème ?

Bilan

En science, on compare souvent des entités entre elles. Ces entités peuvent être de types très variés : par exemple des pays (TP 1), des épidémies (TP 2), ou la l'ensemble des pages Web indexées par Google (TP 3). En linguistique, ces entités sont souvent des textes, ou groupements de textes.

Généralement, si on crée ces entités, c'est pour compter des choses dedans. Par exemple, le nombre de malades dans chaque pays (TP 1), le nombre de morts pour chaque épidémie (TP2) ou le nombre de résultats retournés par Google à chaque année (TP3).

Lors de ces comptages, il est très important de distinguer :

- La **fréquence absolue** : c'est le résultat brut du comptage.
- La **fréquence relative** : c'est la fréquence absolue divisée par la taille de l'entité. C'est toujours un nombre inférieur à 1 (rappel : $1\% = 0,01$; $1\text{‰} = 0,001$; $1\text{‰‰} = 0,0001$; $1\text{ppm} = 0,000001$; etc.).

Quand les entités sont de taille différente, il ne faut **jamais** comparer les fréquences absolues entre elles, car cela n'a aucun sens ; on peut uniquement comparer les fréquences relatives.

Internet pour les SdL : le *Web* comme corpus

A. Savoir lire une URL

L'adresse d'une page Web (URL) se compose de 4 parties : protocole://domaine/chemin?paramètres

1. le protocole (pour le Web : http ou https),
2. le **nom de domaine** (qui se lit de droite à gauche, et « commence » donc par un code tel quel .fr, .ca, .com, .org, etc.),
3. un chemin (facultatif),
4. divers paramètres (facultatif).

Décomposez ces URLs en entourant les 3 ou 4 parties dont elles sont constituées :

http://www.lemonde.fr/voyage/video/2010/10/21/plongee-dans-le-metro-moscovite_1424766_3546.html

http://www.tlfq.ulaval.ca/AXL/francophonie/HIST_FR_s1_Expansion-romaine.htm

<https://www.google.com/search?hl=fr&q=corpus>

B. Archiver une page Web et remonter le temps :

<https://web.archive.org>

Cherchez la version la plus ancienne du site www.univ-paris-diderot.fr

Cherchez la page qui décrit le projet de transfert de l'Université Paris Diderot sur le site Paris Rive Gauche.

C. Faire de la linguistique avec Google Search

Avant tout, le Web est-il à proprement parler un *corpus* ?

En français, dit-on « hôtel à Paris », ou bien « hôtel en Paris » ? Recherchez sur *Google Search*. Quelle conclusion en tirez-vous ?

Dans la plupart des moteurs de recherche, on peut rechercher une expression exacte en l'encadrant par des guillemets. Le caractère * peut remplacer n'importe quel mot. Le mode avancé permet de restreindre à une langue ou bien à un site Web précis.

Indice, testez ces pages : <http://www.plot-generator.org.uk/story/>
<http://www.megabambou.com/pmg/index.html>

Réessayez en vous limitant aux sites lemonde.fr, ou tripadvisor.fr, pour les expressions « à/en États-Unis », puis « à/en Isère ».

Concrètement, quelle utilisation voyez-vous pour Google Search ? Rechercher des exemples sur le Web ? Sur un site précis ? Récupérer un nombre de résultats ? Déterminer si une expression est attestée ?

Lexicographie avec Google Trends

Google Trends est un outil permettant de visualiser les requêtes effectuées par les utilisateurs de Google Search. Il est intéressant en lexicographie, mais attention, ce n'est pas un outil fait pour cela ! Il ne faut pas perdre de vue qu'il se base sur les *recherches* des utilisateurs, pas sur le contenu des sites Web.

Allez sur le site de Google Trends: <http://www.google.fr/trends/explore>

- Dans quelles régions françaises utilise-t-on le terme *chocolatine* ? Comparez avec *pain au chocolat*.
- Pour *pain au chocolat*, que se passe-t-il en octobre 2012 et 2016 ? Que pouvez-vous en déduire ?

Concrètement, Google Trends est-il utilisable pour rechercher des régionalismes ? Une évolution linguistique ?

Linguistique diachronique avec Google Ngram Viewer

Google Ngram Viewer est un outil permettant d'effectuer des recherches diachroniques dans Google Books.

Recherchez « dans » en français entre 1600 et 2000. Quelle échelle est-elle utilisée en abscisse ? En ordonnée ? Comparez avec « en ».

Comparez l'évolution de l'emploi des prépositions « à » et « en » avec Martinique, Guadeloupe et Haïti.

Corpus parallèles avec Linguee

Linguee est un moteur de recherche pour des textes multilingues, notamment des sites et des documents d'organisations internationales. Il permet de rechercher une expression (ce qu'un dictionnaire ne peut pas faire) et d'afficher une série d'exemples de traduction.

Comment traduiriez-vous « conseil de classe » en anglais ? Quelle est le degré de qualité des traductions sur lesquelles Linguee se base ? (inutile de connaître l'anglais, il y a d'autres indices dans la page...)

Essayez avec « par conséquent », « en avoir le cœur net ». Quelle traduction retenir ? Comparez « sans doute » et « sans aucun doute ». Comment traduiriez-vous « *raining cats and dogs* » (facile) et « *off the charts* » (plus difficile) en français ?

Les outils à retenir : Google Search (mode avancé), Google Scholar, Web Archive, Google Trends, Google Ngram Viewer, Linguee.