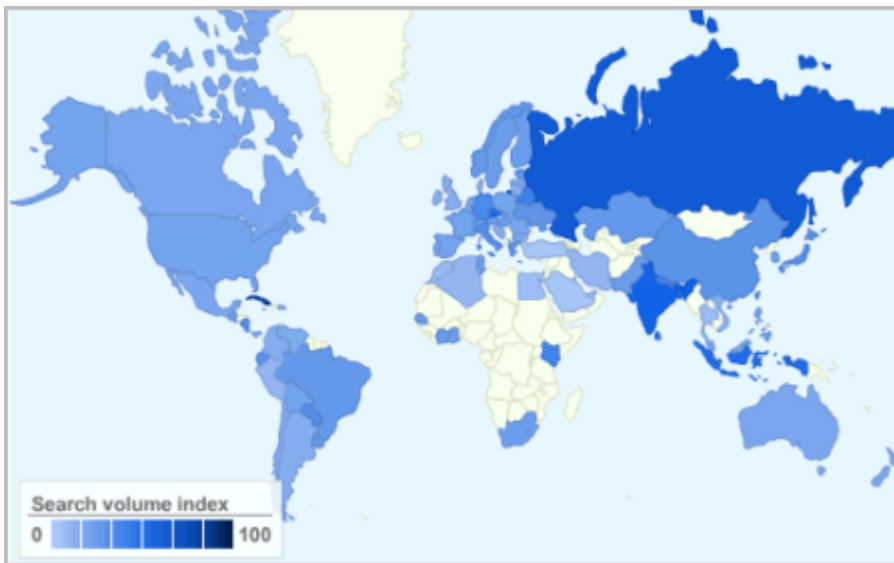


Let's play with Google Insights Search

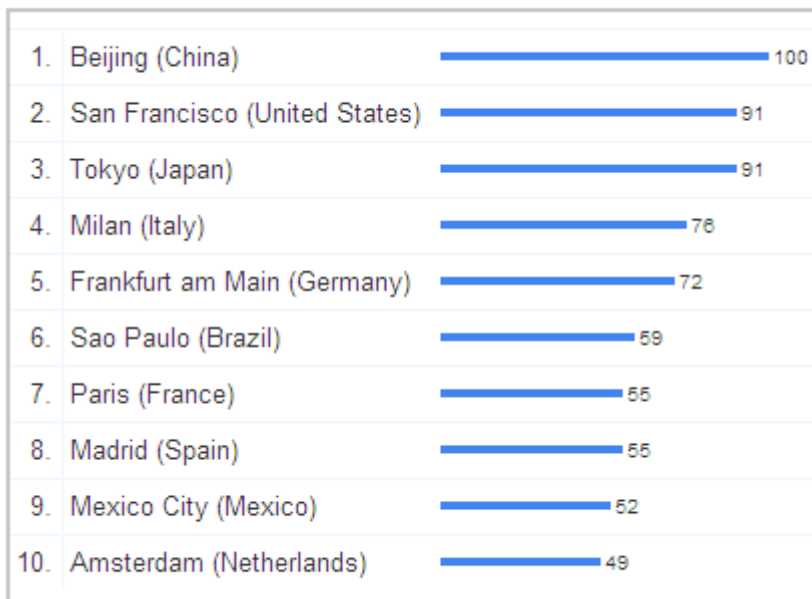


Vous voyez la carte ci-dessus ? Elle correspond à la distribution mondiale de recherches Google sur le mot-clé « linux » ces douze derniers mois.

Les dix premiers pays sont les suivants :

| | | |
|-----|------------------------------------|-----|
| 1. | Cuba | 100 |
| 2. | Russian Federation | 75 |
| 3. | India | 69 |
| 4. | Indonesia | 61 |
| 5. | Czech Republic | 61 |
| 6. | Bangladesh | 61 |
| 7. | Sri Lanka | 58 |
| 8. | Taiwan | 51 |
| 9. | Belarus | 49 |
| 10. | Kenya | 47 |

Et les dix premières villes sont les suivantes :



Ces informations proviennent de Google Insights Search, nouveau service d'analyse des requêtes des internautes mis à disposition par Google (et ne concernant donc que les requêtes effectuées via le moteur Google).

Il fait suite, en l'améliorant, au service Google Trends que j'avais utilisé dans un précédent billet pour évoquer l'évolution de certaines distributions GNU/Linux.

L'idée c'est de comparer et contextualiser certaines requêtes dans le temps et dans l'espace mais aussi par rapport au nombre total des toutes les requêtes effectuées.

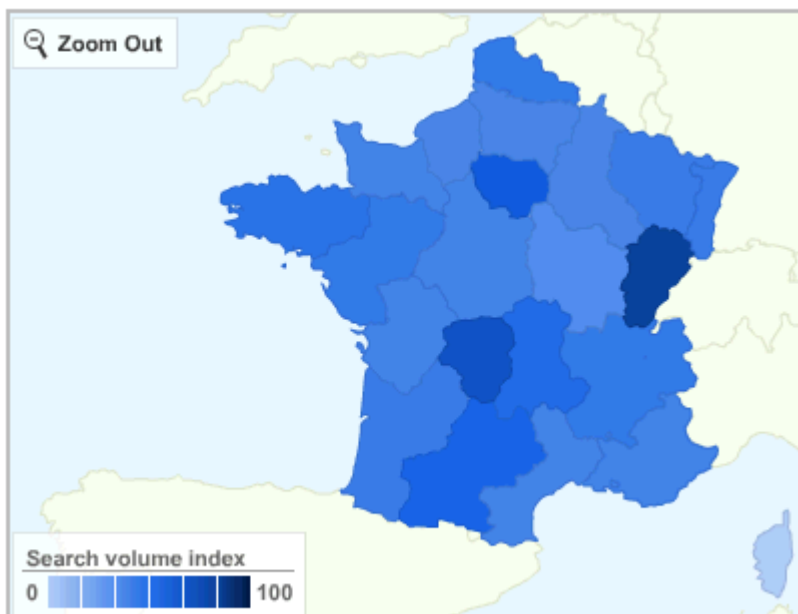
Linux en France

Pour illustrer mon propos conservons l'exemple « linux » mais cette fois-ci restreint au territoire français (toujours pour les douze derniers mois, ce qui est bien entendu paramétrable).

La première chose qui est proposée c'est une courbe d'évolution dans le temps considéré (ici donc pour « linux » en France c'est à peu près stable).



Puis la repartition région par région :



Avec le détail des dix premières régions :

| | | |
|-----|------------------------------------|-----|
| 1. | Franche-Comte | 100 |
| 2. | Limousin | 82 |
| 3. | Ile-de-France | 73 |
| 4. | Midi-Pyrenees | 65 |
| 5. | Auvergne | 61 |
| 6. | Brittany | 58 |
| 7. | Nord-Pas-de-Calais | 53 |
| 8. | Pays de la Loire | 53 |
| 9. | Rhône-Alpes | 53 |
| 10. | Alsace | 52 |

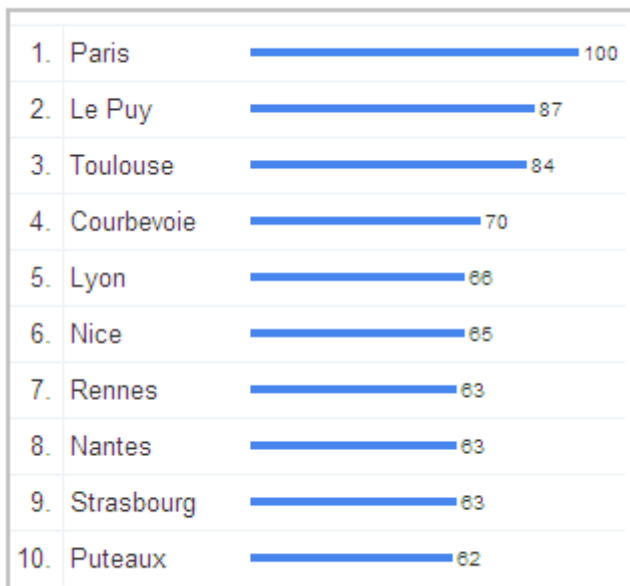
Que signifient ces chiffres à droite (qui n'apparaissent que si vous êtes connecté à un compte Google, eh oui, on a rien sans rien) ?

La première région étant la Franche-Comté (bravo pour elle !) elle est automatiquement affublée de la centaine. Cette première place veut dire, si j'ai bien compris, que parmi toutes les requêtes Google effectuées en France, c'est dans cette région que la proportion de recherches « linux » est la plus grande. Notez que cela ne nous indique pas la proportion (ou pourcentages) de requêtes « linux » par rapport à toutes les requêtes Google faites depuis la Franche-Comté sur la période (qui doit être bien faible de toutes les façons). Toujours est-il que c'est la Franche-Comté qui donne le la (le 100) et les autres régions qui suivent en conséquence.

Je me trompe peut-être mais du coup je lis qu'il y a environ un tiers de chercheurs Google « linux » en moins en Midi-Pyrénées (65) et quasiment moitié moins en Alsace (52). Au risque de me répéter on peut donc affirmer que selon Google il y aurait deux fois plus de chercheurs « linux » en Franche-Comté qu'en Alsace.

Ce qui ne signifie nullement qu'il y aurait deux fois plus de postes sous GNU/Linux en Franche-Comté qu'en Alsace parce qu'on peut très bien imaginer qu'un jour une région soit tellement à *la pointe Linux* qu'elle n'ait plus besoin de faire des recherches sur le terme générique « linux » (mais plutôt sur « bash », « emacs » ou que sais-je encore).

Sinon, on a aussi les villes :



Dans la mesure où Google ne prend pas uniquement le mot-clé « linux » en considération mais référence également les requêtes connexes qui contiennent ou non ce mot-clé, il nous est proposé deux tableaux.

Le premier tableau est constitué du top ten des requêtes associées au mot-clé « linux » (là encore si vous avez un compte Google vous aurez la possibilité d'avoir plus de dix résultats en téléchargeant le fichier CSV associé à votre recherche) :



Le second tableau est constitué des requêtes associées au mot-clé « linux » ayant connu la plus forte hausse sur la période étudiée (*Breakout* signifie que la hausse dépasse 5000% !) :

| Rising searches | | |
|-----------------|-------------------------------|----------|
| 1. | play on linux | Breakout |
| 2. | eeepc | Breakout |
| 3. | xandros | +1460% |
| 4. | linux xandros | +1050% |
| 5. | toutou linux | +320% |
| 6. | virtualbox | +240% |
| 7. | linux sur ps3 | +220% |
| 8. | linux mint | +100% |
| 9. | puppy linux | +80% |
| 10. | asus | +80% |

À la lecture de ces tableaux on peut supputer par exemple qu'Ubuntu est bien la reine des distributions en France (mais que Mandriva n'est pas morte), que les linuxiens ont de plus en plus envie de jouer sur leur ordi et que l'Eeepc (sous Xandros) a bien été le phénomène annoncé.

Digne de confiance ?

Bon, c'est bien gentil tout ça mais l'outil est-il fiable ? Telle est la question qui en fait n'aura pas de réponse. D'abord parce que la localisation via notre adresse IP n'est pas d'une rigueur absolue. Et ensuite parce que Google garde jalousement ses données et ses algorithmes de calculs.

La seule chose qu'il soit possible de faire c'est de taper des requêtes dont vous les résultats sont attendus afin de vérifier si l'outil est conforme à vos prévisions. Ainsi si l'on tape « marée » (en France) ce sont bien les régions de la côte atlantique qui apparaissent en premier.

Ce n'est pas très scientifique tout ça mais plusieurs exemples du même acabit m'ont fait supposer que globalement Google Insights Search donne des tendances plausibles.

Je pense néanmoins que des pays encore peu connectés peuvent venir fausser certains résultats car on se retrouve alors proportionnellement avec des internautes locaux pionniers et donc à mon avis plus technophiles que la moyenne. Et puis il y a aussi la question des langues, des alphabets différents, etc. N'hésitez pas à apporter d'autres bémols dans les commentaires.

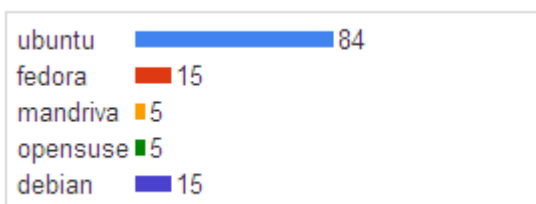
En tout cas fiable ou pas, je suis sûr que de nombreux *marketeux* vont prendre d'assaut l'outil pour affiner leurs prochaines campagnes commerciales !

Quelques exemple d'un monde libre qui se cherche

Ces précautions d'usage étant posées, on peut faire mumuse avec l'outil, d'autant qu'il est permis de comparer plusieurs requêtes entre elles.

Quelques distributions GNU/Linux

Comparons mondialement Ubuntu, Fedora, Mandriva, OpenSuse et Debian. On retrouve bien la suprématie actuelle d'Ubuntu (aucune idée de la signification du 84 derrière Ubuntu, par contre je pense là encore qu'on peut supposer qu'Ubuntu est environ dix-sept fois plus recherchée que Mandriva).

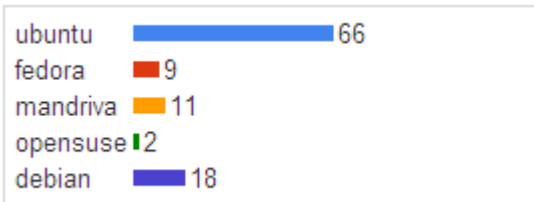


Au niveau des trois pays de tête ça donne :

- Cuba, Italie, Indonésie pour Ubuntu

- Sri-Lanka, Inde, Taïwan pour Fedora
- Russie, République Tchèque, Biélorussie (et France cinquième) pour Mandriva
- République Tchèque, Russie, Allemagne pour OpenSuse
- Cuba, Biélorussie, République Tchèque pour Debian.

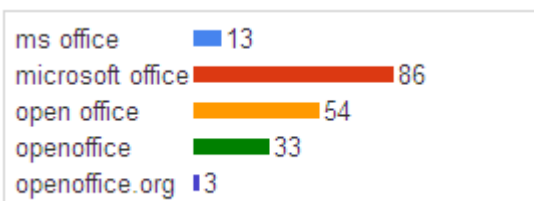
Si l'on se restreint à la France, ça bouge un peu puisque par exemple Mandriva s'en sort mieux face à Ubuntu.



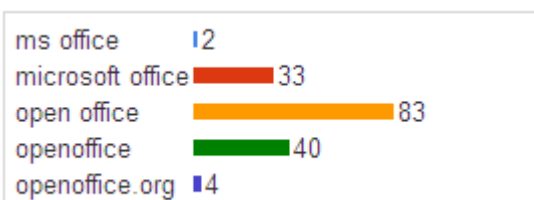
La petite guerre des suites bureautiques

J'ai tenté de comparer les suites bureautiques MS Office et OpenOffice.org (comparaison difficile en raison des différentes appellations)

Dans le monde (avec l'Allemagne et la France qui se classent premières pour « open office ») :



Et en France :



On s'en sort bien sur ce coup là ☐

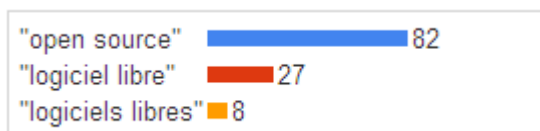
Open Source vs Free Software

Comparons les requêtes « Open Source » et « Free Software ». On constate tout d'abord que « Open Source » est plus demandé que « Free Software ». Va-t-il finir par s'imposer ?



Ce sont plusieurs villes indiennes (Delhi en tête), Djakarta et Singapour qui remportent la palme des recherches « Free Software ». Alors que les villes « Open Source » sont San Francisco, encore Delhi, puis Milan, Amsterdam, Washington, Sydney, Londres. Il y a clairement là une différence géographique (et éventuellement sociale).

En France cela ne sera à rien de comparer car « Free Software » est remplacé par « Logiciel Libre ». Du coup on obtient ceci :



Et tant qu'on y est pourquoi ne pas comparer « Richard Stallman » et « Linus Torvalds ». De notoriété mondiale équivalente, vous constaterez avec moi que Stallman est très populaire dans les pays d'Amérique Latine tandis que Torvalds a une forte côte dans les contrées occidentales.

À vous de jouer

Fin de mon petit tour de manège avec Google Insights Search. J'en appelle aux commentaires pour nous signaler vos propres requêtes que vous aurez jugées intéressantes à communiquer.