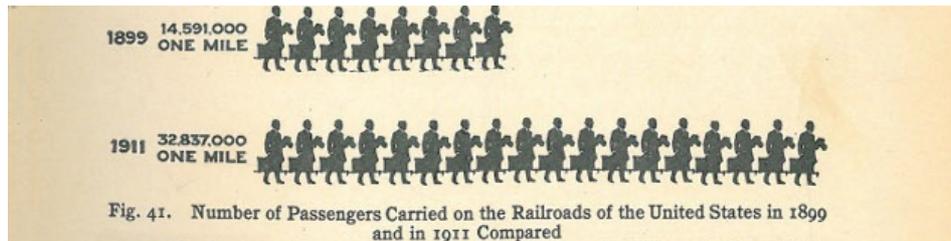


LES DATA EN FORME

LE 6 FÉVRIER 2012 PAULE D'ATHA

Tournée hebdomadaire de veille "ddj" pour l'équipe des journalistes de données d'OWNI. Du lourd, de la big data, au menu de deux magnifiques applications. Et un soupçon de déception quand même.

En guise d'amuse-gueule, **un petit article** [EN] sur le design informationnel, où l'on voit bien – à travers le livre *Graphic Methods for Presenting Facts* de 1917, notamment – que nos infographies d'aujourd'hui n'ont rien de révolutionnaire. On en profite ici pour dire bonjour à **Karen**, grande veilleuse parmi les veilleurs, pour qui la plupart des liens de **journalisme graphique** que nous vous fournissons chaque semaine ne sont pas inconnus.



DataWOW

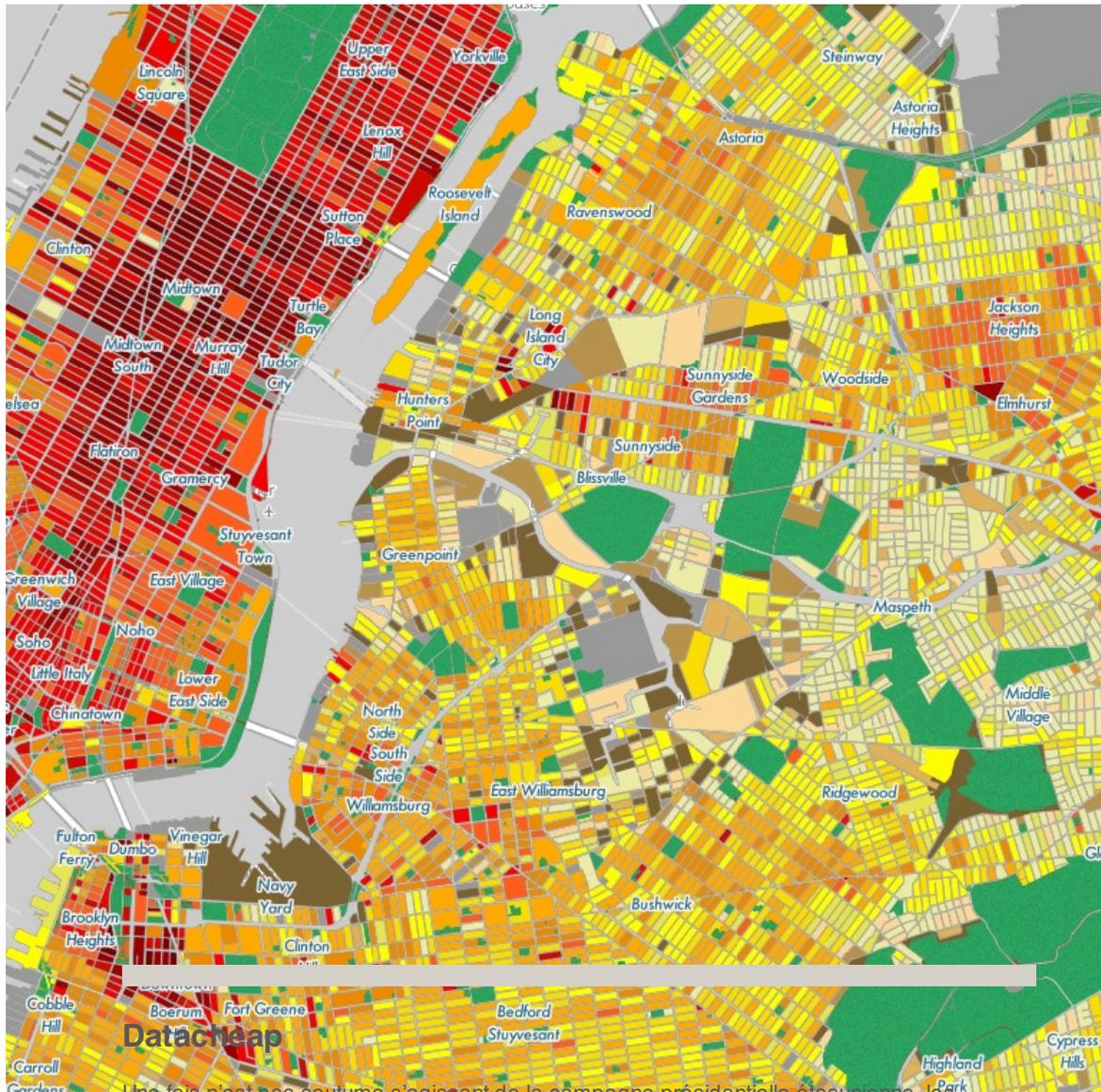
Alors que nous sommes en train de finaliser pour la présidentielle française un très beau joujou dont nous sommes très fiers et que nous avons hâte de vous montrer, nous ne pouvons pas manquer l'occasion cette semaine de faire honneur à deux applications de très haut niveau. Si vous rêviez de vous promener aux confins de la "big data" sans vraiment oser rentrer dedans, en voici deux magnifiques illustrations.

Les plates-formes eXplorer de la société suédoise **NComVA** manipulent de très gros volumes de données pour créer de remarquables visualisations statistiques. Par exemple, **Europe Explorer** [EN] permet de naviguer dans un large *dataset* (téléchargeable) d'informations économiques et démographiques sur les pays européens et de personnaliser (c'est le point fort) l'affichage de ces données – telles que la croissance du PIB, le taux de fertilité, l'espérance de vie à la naissance, ou encore la population par tranches d'âge.



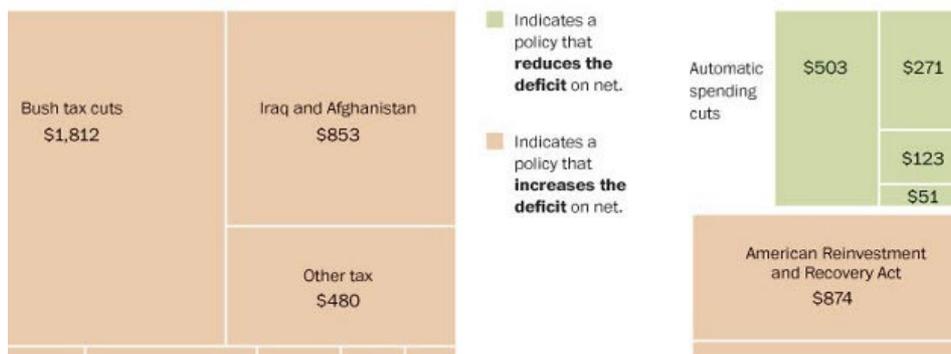
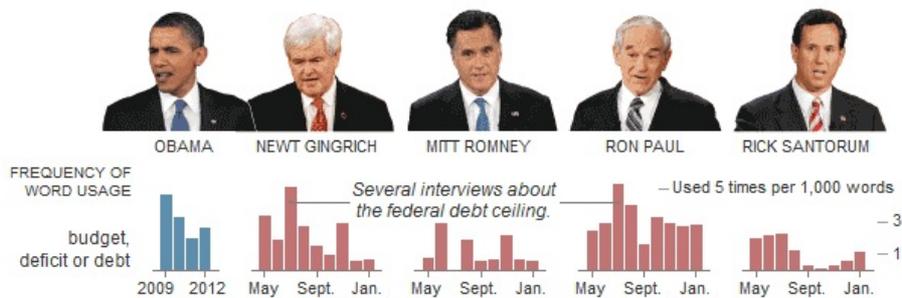
Total Annual Building Energy Consumption for New York City [EN] est une application un poil titanesque réalisée par l'équipe du groupe de recherche du Professeur Viraj Modi, spécialiste de sources d'énergie et de mécanique des fluides à l'université de Columbia. Comme son nom l'indique, cette application interactive veut pouvoir fournir une estimation

de la consommation d'énergie de chaque bâtiment new-yorkais à partir d'une **étude ad hoc** [EN] et d'une **méthodologie assez épurée** [EN]. Rien qu'imaginer le temps qu'il a fallu pour rassembler les données et les harmoniser... donne un peu mal à la tête. Nous avons là,



Une fois n'est pas coutume s'agissant de la campagne présidentielle étasunienne, les éléments incontournables de notre veille cette semaine sont moyennement excitants. Petite déception, vu que le *New York Times* et le *Washington Post* nous ont habitués à beaucoup mieux. Ou alors on devient plus exigeants avec le temps, possible aussi.

Pour accompagner un (bon) **article d'analyse politique** [EN] du discours de l'état de l'Union, le NY T s'est donc fendu d'une **infographie** [EN] comparant la fréquence d'usage des mots du Président Obama avec ceux prononcés par ses adversaires républicains putatifs. La mise en forme en histogramme est ici assez originale dans le contexte de l'analyse lexicale – pour nous être également **prêtés au jeu**, on vous confirme que le rendu d'un tel corpus est loin d'être évident à modéliser. En revanche pour le reste, deux écueils majeurs (l'échelle du temps et l'échelle de numération) rendent particulièrement décevante cette expérience : primo, Obama est scruté sur ses quatre adresses à la Nation (2009-2012) tandis que les candidats républicains sont étudiés sur les neufs derniers mois au cours de "*plusieurs entretiens*". Secundo, le nombre d'occurrences observées ne dépasse 7 qu'une seule fois – la plupart du temps, on compare des fréquences situées entre 1 et 5 fois ; rendant le diagnostic particulièrement biaisé au vu de la taille du corpus étudié.



La deuxième “déception” de la semaine (les guillemets parce qu’il nous reste un peu d’humilité quand même) provient de la **visualisation en treemap** [EN] qui accompagne un **papier économique** plutôt technique [EN, paywall] qui déchiffre les “déficits” Obama et qui établit une comparaison avec ceux provoqués par Bush fils. Le format **treemap**[EN] est toujours performant pour comparer des budgets de **celui, brillant, de Jean Abbiateci**) mais on attend – forcément – du *WashPo* autre chose qu’une pauvre image compressée, des cubes à géométrie variable et des projections jusqu’à 2017.

Datacoq

D’autant que l’économie est un sacré terrain de jeu pour faire de l’infographie et de la bonne datavisualisation. Pour preuve, le blog “*Echosdataviz*” de **Frédéric Vuillod**, qui a pondu **une vidéo très simple et très didactique** à l’occasion de l’introduction en bourse de Facebook. Deux regrets : 18 secondes de pub en pré-roll (désolé) et une vidéo non-embeddable. On va dire que c’est pour faire du business à la française.



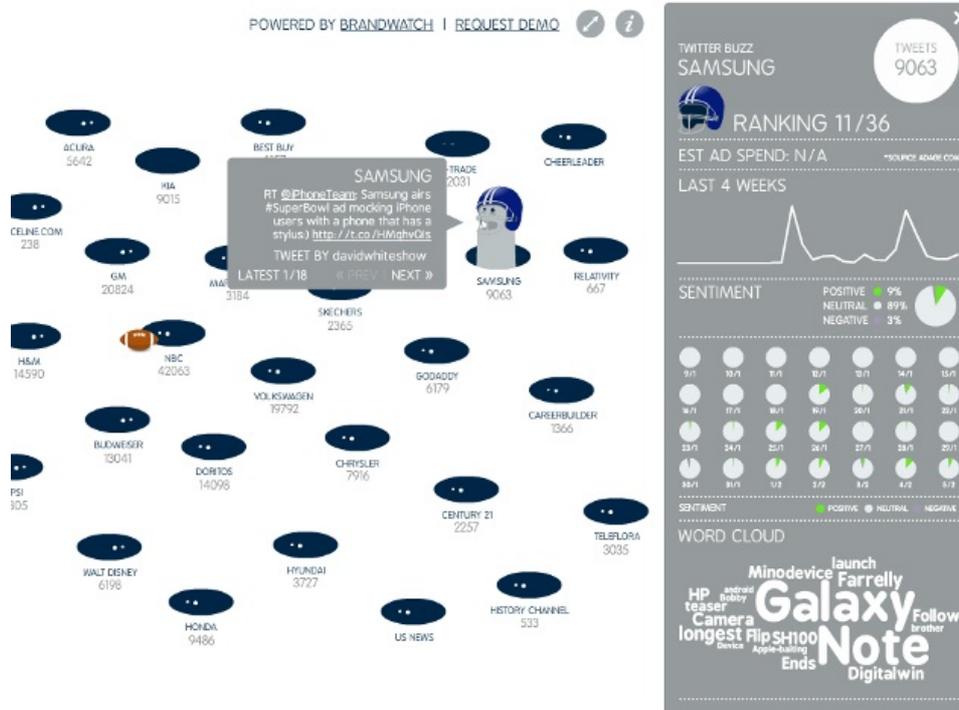
Dataworld

Et puisqu’on parle de business à la française, difficile de ne pas évoquer cette semaine le bras armé de la moralisation du capitalisme qui se réunit chaque année à Davos en Suisse.

Le célèbre Forum Economique Mondial nous gratifie d'une jolie vidéo – limite décadente – pour présenter le rapport 2012 des *“risques globaux”* qui peuvent transformer notre planète en véritable chambre de torture si les grands de ce monde ne font rien pour les détruire à la racine. Notre *“risque”* préféré est sans aucun doute *“les graines de la contre-utopie”* (*“the Seeds of Dystopia”*) – symbolisé par des drapeaux, des poings levés et un masque d'Anonymous – et leur *“côté sombre de la connectivité”* (*“the Dark Side of Connectivity”*) et son logotype de port USB, véritable empêchement d'épanouissement humain. Passée la propagande, la vidéo [EN] est objectivement bien foutue, donc pourquoi pas.

Databowl

Heureusement, pour contrer les anarchistes et les hackers, il reste le Super Bowl. On n'insultera ici ni les amateurs ni les détracteurs de ce spectacle en omettant volontiers de donner le résultat de la finale du championnat de **football américain**, qui vient de se dérouler. A ceux qui douteraient qu'il s'agit bien d'un sport, les laboratoires de Brandwatch ont pondu **une application de réputation sociale** [en] (basée sur Twitter) des annonceurs figurant dans la liste très courue de ceux qui peuvent se payer 30 secondes de pub pendant cette messe annuelle. C'est marrant, ça parle de sentiment, et donc presque d'amour.



Et en parlant d'amour, on se quitte sur la nécessaire vidéo de la semaine de Stephen Malinowski. Paix et gloire sur vos data !

Retrouvez les précédents épisodes des **Data en forme** !

SIMON PATARIN

le 6 février 2012 - 22:54 ; SIGNALER UN ABUS - PERMALINK



Vous pourriez également jeter un oeil au Public Data Explorer de Google (<http://www.google.com/publicdata>) qui offre de nombreux datasets couplés à un puissant outil de visualisation.

VOUS AIMEZ



VOUS N'AIMEZ PAS



LUI RÉPONDRE



le 9 février 2012 - 16:53 • SIGNALER UN ABUS - PERMALINK



La vidéo "Global Risks 2012", on dirait une vidéo de Patrick Clair :
<http://vimeo.com/patrickclair/videos/page:1/sort:newest>

VOUS AIMEZ



0

VOUS N'AIMEZ PAS



0

LUI RÉPONDRE