

# LES DATA EN FORME

LE 18 JUILLET 2012 PAULE D'ATHA

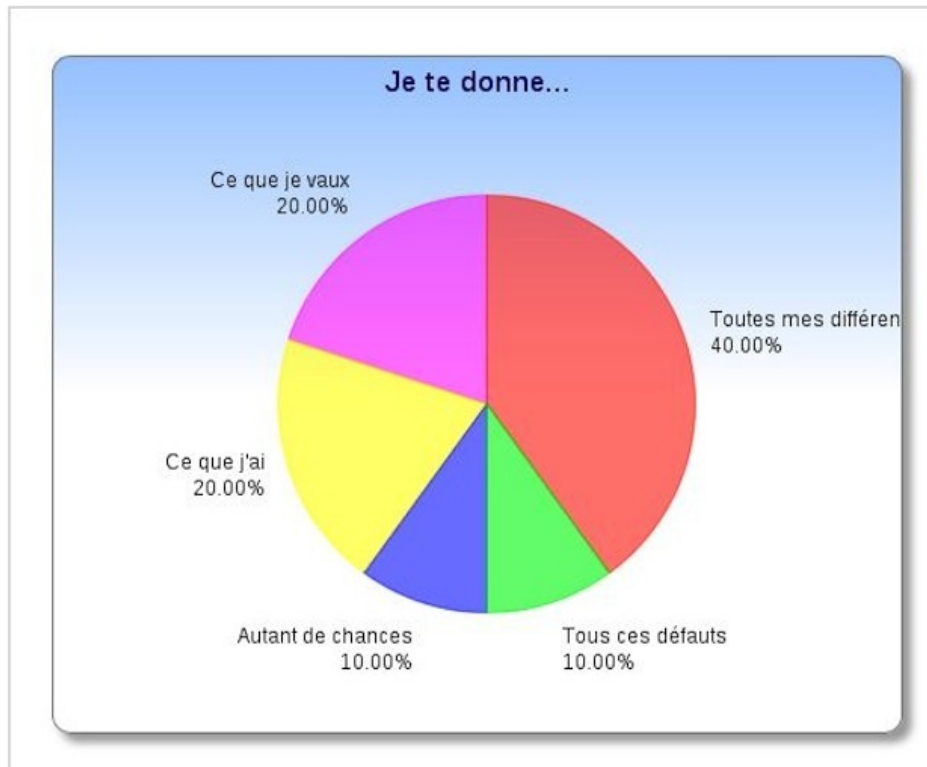
La veille des journalistes de données d'*Owni* prend de la légèreté en cette période d'été : on chante, on joue au Lego, on se balade dans les idées du monde... Et bien sûr, on se prépare aux Jeux Olympiques.

Une fois n'est pas coutume, nous allons cette semaine partir du plus LOL pour arriver au plus sérieux. Parce que c'est juillet, les vacances et que le soleil est plus garanti sur nos écrans que dans le ciel parisien.

## Je chante, tu chantes, il dataviz

Lancé par **Florent Maurin** (par ailleurs chroniqueur sur la chouette émission estivale de France Inter dédiée sur les questions numériques : **Antibuzz**)<sup>1</sup>, ce tumblr sobrement intitulé « **Dalalalataviz** » résume les chansons des plus grands artistes (Abba, Jackson 5... Bratisla Boys) en une dataviz. Le résultat est très drôle et les auteurs débordent d'imagination. Si l'idée vous inspire, lancez-vous. Chacun est libre de soumettre ses traductions visuelles : Florent Maurin précise qu'il s'agit d'un projet « **open-soupe** ».

#goldman



## Mémo Lego

Les briques de couleurs et tailles diverses produites par la société Lego s'adaptent parfaitement à la data visualization et à l'infographie, comme l'avait montré l'événement Expoviz et **ses ateliers**. La société Lego s'en est également servi pour expliquer simplement **le principe d'une infographie** : passer d'un ensemble fouilli de data à des données représentées visuellement en les ayant au préalable triés et arrangées.



## DATA



### A quel gros ressemblez-vous ?

Où vous situez-vous sur l'échelle d'obésité mondiale ? C'est à cette étonnante question que **la nouvelle application de la BBC** vous propose une réponse. Après nous avoir permis de **savoir combien de personnes** vivaient sur la planète le jour de notre naissance ou encore d'évaluer l'étendue des pertes durant la seconde guerre mondiale en se basant sur **nos amis facebook**, la BBC poursuit dans sa série d'applications alliant graphiques, données mondiales et forte personnalisation.

Ici, le « **calculateur** » vous demande de saisir votre âge, votre sexe, votre poids et votre pays pour vous comparer ensuite à l'échelle mondiale. Le résultat est souvent étonnant : le poids moyen est ainsi plus élevé en Jamaïque qu'aux Etats-Unis.

### Your numbers

Obesity Index	National	Global
 <b>19</b> <small>bmi</small>	 <b>Below average</b>	 <b>Below average</b>
BMI is an <b>estimate</b> of how "overweight" or "obese" a person is	You have a lower BMI than <b>84%</b> of females aged 15-29 in your country	You have a lower BMI than <b>74%</b> of females aged 15-29 in the world

**Did you know?**  
If everyone in the world had the same BMI as you, it would **remove 56,484,590 tonnes** from the total weight of the world's population

**You're most like someone from Vietnam\***

Share your result    \* Compared with other females aged 15-29 in Vietn



### L'anamorphose des athlètes

Impossible d'échapper à l'actualité sportive, même dans la planète Data. A partir du 27 juillet, Londres accueille les Jeux Olympiques. Le Datastore du *Guardian* en a pris son parti en créant une **rubrique spécifique** consacrée aux données olympiques de 2012 : graphiques de comparaisons entre les JO accueillis par Londres (en 1908, 1948 et 2012), liste complète des athlètes britanniques concourant à la compétition, etc. Cette rubrique recense également les visualisations pertinentes réalisées sur ce sujet, comme cette **carte en anamorphose**, réalisée pour l'association **Join In** par **Kiln**, spécialiste de cette technique, créateurs notamment de **The Carbon Map**. Cette carte des JO permet de visualiser l'origine des athlètes selon leur sport. Là où certains faits semblent évidents (les athlètes concourant pour la navigation à voile viennent plus souvent des régions côtières), d'autres moins connus sont mis en lumière : la

prépondérance du sud comme origine des tennisman, le curieux équilibre nord/sud pour le basketball, écrasant toutes les régions.

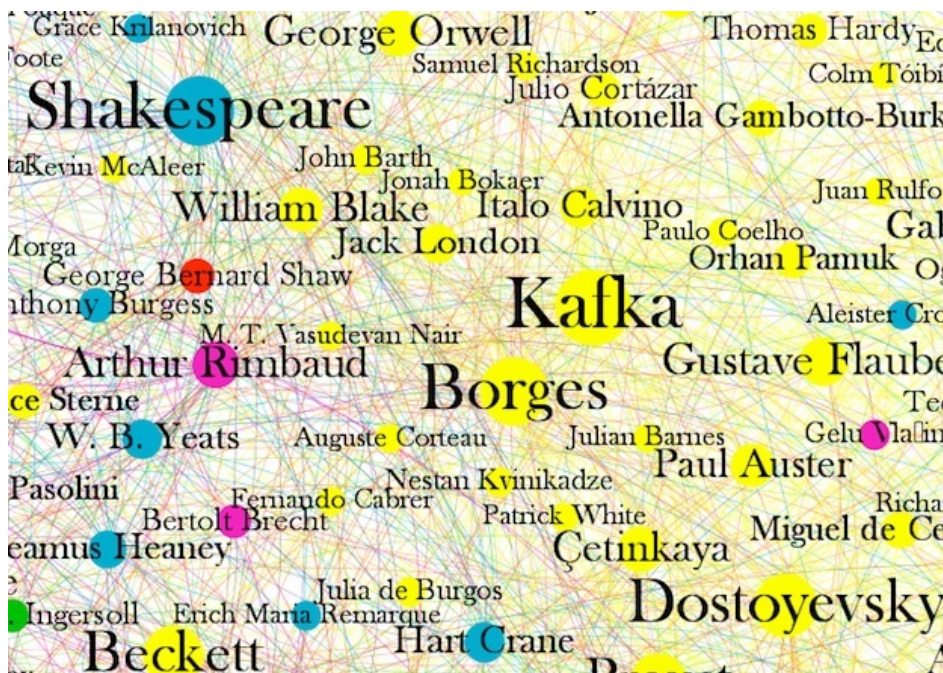


## Toutes les idées du monde

Il y a environ un mois, **Simon Rapper**, statisticien britannique, s'était essayé à la visualisation de **l'histoire de la philosophie** et des influences entre auteurs. Il a inspiré Brendan Griffen qui a reproduit son idée (scaper Wikipédia et sélectionner ceux dont il est mentionné « influencé par » ou « ayant influencé ») à l'échelle de Wikipédia tout entier : sa méthodologie est expliquée [ici](#).

**Le résultat** est impressionnant. Tant par sa taille (attention au temps de chargement) que par son contenu.

Sur ce graph, réalisé avec le logiciel de graphiques **Gephi**, la taille des nœuds est proportionnelle au nombre de connections entre deux éléments. Plus le nœud est important, plus grande est l'influence de cette personne. Nietzsche, Kant, Hegel, Hemingway, Shakespeare, Platon, Aristote et Kafka trustent sans surprise l'ensemble des liens. Cependant, d'autres connections méritent également le coup d'oeil : la forte influence des artistes comme Andy Warhol ou Marcel Duchamp par exemple.



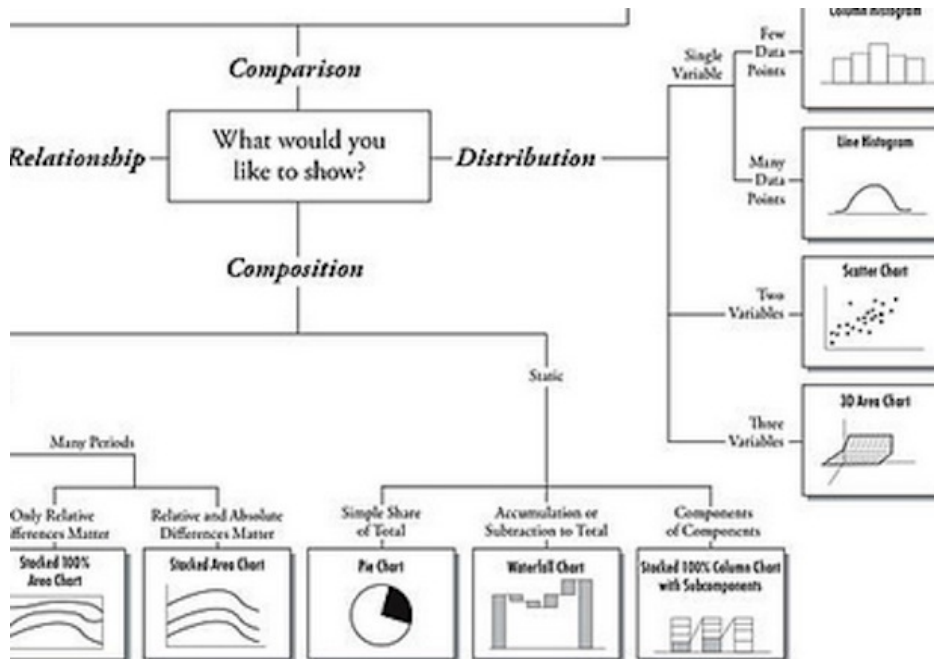
## Comment choisir son graphique

Minute pédagogie. C'est une évidence : le choix du graphique ou du type de visualisation est essentiel dans le fait de rendre lisible une infographie ou une data visualisation. Pour

simplifier cette étape, des outils ont commencé à faire leur apparition, comme le **tableau des méthodes de visualisation**, utilisant le concept du tableau de classification périodique des éléments chimiques.

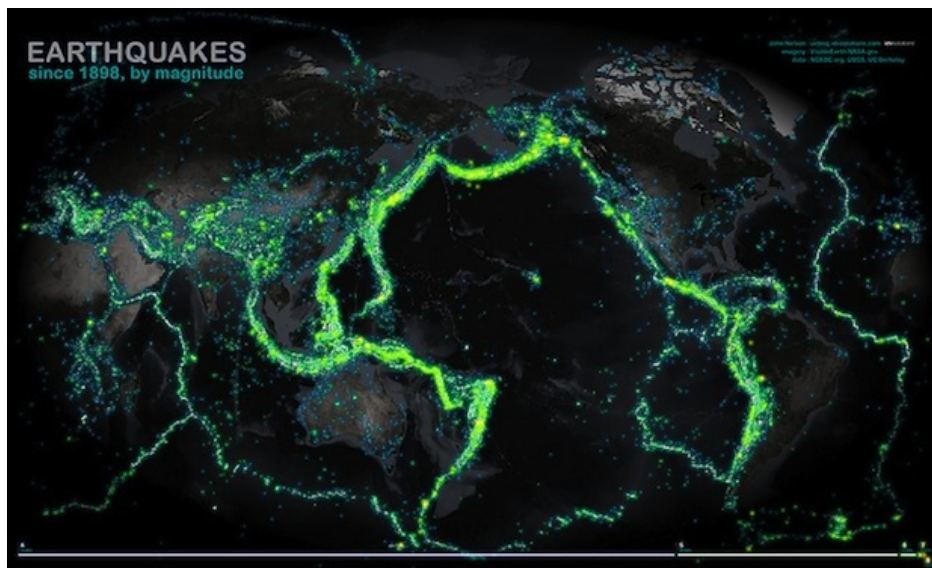
**Andrew Abela**, professeur de marketing et consultant en design, a réalisé quant à lui un **graphique de choix de graphique**, sous la forme d'un test. La question principale se situe au milieu : « Que voulez-vous montrer ? ». Puis, suivant la réponse (une comparaison, une composition, une distribution, une relation) et les caractéristiques des données (statiques, évolutives, accumulatives,...) que le lecteur souhaite, le graphique qui semble le plus approprié est affiché au bout de la ligne.

L'équipe de Juice Labs en a d'ailleurs réalisé une **version interactive**, mais presque moins lisible car elle ne centralise pas tous les éléments sur une seule page.



## Une vision d'intensité

Une belle visualisation pour terminer cette chronique : celle de **John Nelson** sur **les séismes dans le monde** depuis 1898, représentés selon leur magnitude. John Nelson a utilisé les données de l'NCEDC (Northern California Earthquake Data Center), USGS (US Geological Survey) et de l'Université de Berkeley, qu'il a replacé sur **un des fonds de visualisations fournis par la Nasa**. Outre l'esthétique permise par le fond de carte, les informations sur les zones les plus dangereuses ressortent clairement. John Nelson avait également représenté avec succès **les incendies majeurs** ayant frappé les Etats-Unis.



Bonne data-semaine à tous !

Tous les épisodes précédents des **Data en forme**.

**Paulette sur Twitter** | **Paulette sur Facebook** | **Paulette sur Pinterest**

1. Emission dans laquelle intervient également notre comparse de Paule d'Atha, **Julien Goetz** [↗]

