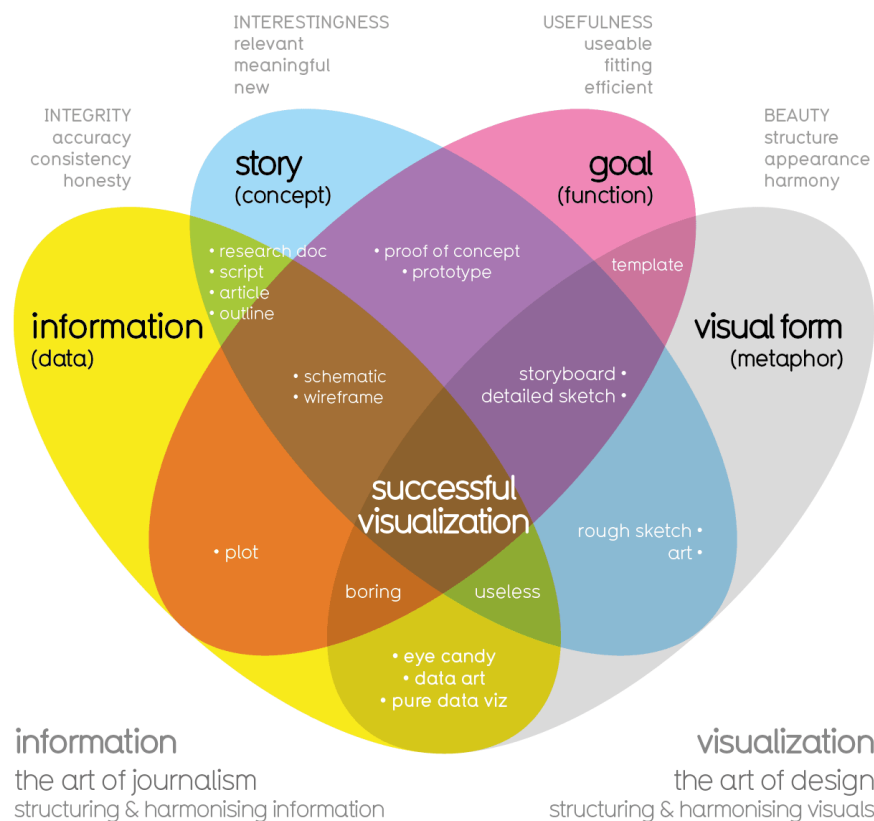




## DATAVIZ part 2:

Quel graph pour quelle information ?

### What Makes a Good Visualization?



David McCandless  
[InformationisBeautiful.net](http://InformationisBeautiful.net)

find out more  
[bit.ly/KIB\\_Books](http://bit.ly/KIB_Books)

# Deuxième partie

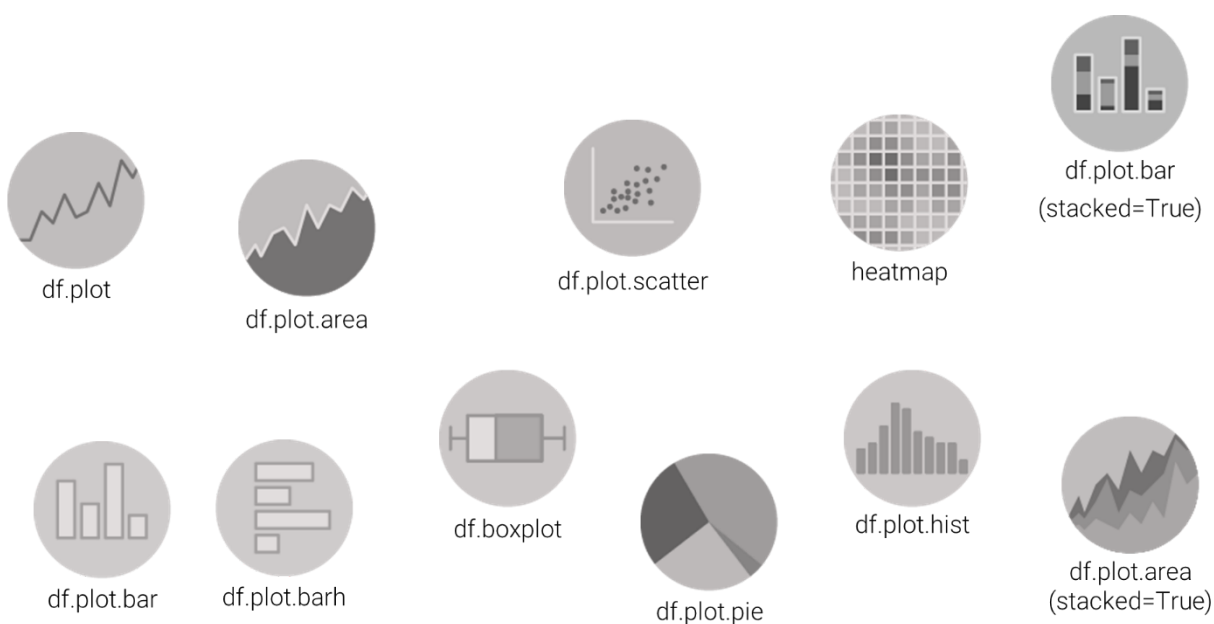
## Cette partie est découpée en 4 étapes :

1. Comprendre les intérêts de chacun des graphes les plus classiques
2. Un peu de lecture
3. Revue de pièges courants
4. Étude de cas

## Modalités

- Durée du projet : 1 jour
- Ce projet sera réalisé en équipe pour la plus grande partie
- Des slides serviront de support aux discussions tout au long de la journée, ces slides seront mis à disposition sur le google drive.

## Étape 1 : Les graph classiques



## Objectifs de l'activité :

Connaître les graphes les plus classiques, savoir quand les utiliser :

- Selon le type de données disponibles :  
*Variables qualitatives ou quantitatives*
- Selon l'information que l'on souhaite faire passer :  
*distribution, corrélation, classement, proportion, évolution ...*

## Consignes

Travail en équipes de 3 ou 4 personnes.

1. Pour chacun des deux jeux de données, proposer des visualisations (parmi les graphs classiques listés ci-dessus) qui permettent de répondre aux questions posées (seul le type de graph importe ici).
  - Dataset 1 : Titanic (dataset très classique fourni par Kaggle)
    - Quelle est la répartition des passagers par classe ?
    - Quelles différences selon le sexe ?
    - Distribution des passagers selon âge et sexe
    - Quel est le taux de survie selon le sexe ? l'âge ? la classe ?
    - Y a-t-il un lien entre prix payé pour le billet et l'âge des passagers ?
  - Dataset 2 : Consommation et productions nationales
    - Comment varient les différentes grandeurs présentées dans ce jeu de données ?
    - Quel est l'évolution de la consommation ?
    - Quand a-t-on un excédant / un déficit de production ?
    - Y a-t-il une corrélation entre les différents types de production ?
2. Créez un tableau avec les différentes figures produites, une capture d'écran de la figure suffit. Spécifiez le type de graph (scatter, line plot, ...) ainsi que le type de variables représentées (quantitatives ou qualitatives). Entamez une réflexion sur la meilleure manière de faire passer le message (c'est-à-dire de répondre à la question posée): que manque t'il à la figure, quelles sont les simplifications possibles, l'usage de couleurs etc...
3. Téléchargez le tableau sur le drive afin de pouvoir comparer les différents tableaux créés par chaque groupe. On en discute ensemble en fin de matinée!

## Compétences :

Je sais choisir le type de graph le plus pertinent permettant de répondre à une question posée.

## Étape 2 : Un peu de lecture

### Objectifs de l'activité :

- Comment créer un graph : <https://blog.datawrapper.de/better-charts/>
- Comment utiliser les couleurs : <https://blog.datawrapper.de/colors/>

## Étape 3 : Pièges courants

### Objectifs de l'activité :

Bar graph, line graph, pie graph, et boxplot : quels sont les problèmes les plus courants ?

### Data-to-viz, une mine d'informations:

Concernant les pièges à éviter : <https://www.data-to-viz.com/caveats.html>

Pour savoir quel graph choisir : <https://www.data-to-viz.com/>

### Consignes :

- Réfléchir ensemble aux « pièges » présentés
- Passer un peu de temps à explorer l'arbre proposé par <https://www.data-to-viz.com/>  
-> est-ce que vous auriez répondu différemment aux questions posées dans la première partie?

## Étape 4 : Études de cas

### Objectifs de l'activité :

Pour plusieurs cas d'étude, analyse et correction de graph.

### Ces études de cas proviennent du livre :

[Storytelling with data](#), de Cole Nussbaumer Knaflic

### Consignes :

Pour chacun des graphs projetés :

- Par équipe de 4, lister les défauts des graphs, et proposer des améliorations
- Comparaison des graphs créés par chaque équipe
- Analyse du graph proposé par l'auteur

**Note:** les propositions de graphs améliorés seront réalisées avec du papier/crayon, afin de se concentrer sur l'aspect visualisation, oubliez le code! Par contre n'oubliez pas les couleurs! 😊