

Les statistiques : info ou intox ?

Comprendre les statistiques
pour mieux s'informer



1

JECO – Atelier pédagogique Degesco – 6 novembre 2019

“Il y a trois sortes de mensonges : les petits mensonges, les gros mensonges et les statistiques”

Mark Twain (ou Benjamin Disraeli)

01 . La fabrique des statistiques

02 . Éclairer le débat public

03 . “On peut faire dire aux chiffres...”

01 La fabrique des statistiques

1.1 La statistique publique : c'est tout un système

C'est qui le chef ?

1.1 La statistique publique : c'est tout un système

C'est qui le chef ?



Statistique publique : c'est qui le chef ?



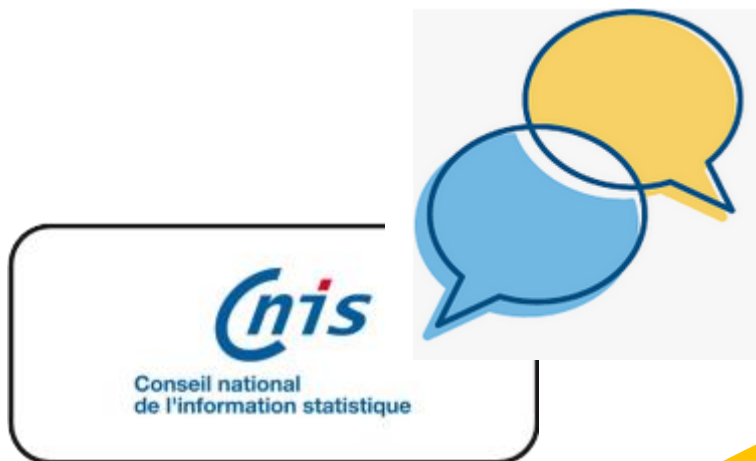
Statistique publique : c'est qui le chef ?



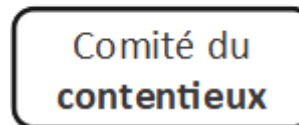
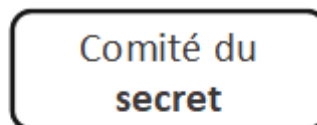
Statistique publique : c'est qui le chef ?



Statistique publique : c'est qui le chef ?

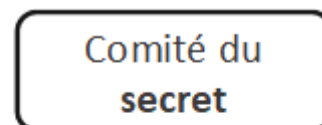


Statistique publique : c'est qui le chef ?



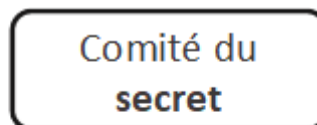
Statistique publique : c'est qui le chef ?

eurostat 



Statistique publique : c'est qui le chef ?

BANQUE CENTRALE EUROPÉENNE



1.2 La statistique publique c'est de la cuisine

... et de la bonne cuisine ! Donc :



Ça ne s'improvise pas → des recettes/méthodes rigoureuses

1.2 La statistique publique c'est de la cuisine

SONDAGE

n'est pas un gros mot. Mais il y a sondage et sondage.

1.2 La statistique publique c'est de la cuisine

Institut de statistiques	Institut de sondage
tirage aléatoire → calcul de <u>précision</u> méthodologie pour la non-réponse, etc. → https://www.insee.fr/fr/information/2838097 questions complexes → analyses détaillées économiques et sociales Statistique publique → rendre compte à la société	méthode des quotas → pas de calcul de <u>précision</u> possible questions “simples” → opinion, consommation entreprise commerciale → dépendance vis-à-vis du client, enquêtes omnibus

1.2 La statistique publique c'est de la cuisine

Les « fiches méthodologiques » présentent les principales méthodes et outils utilisés par l'Insee aux différentes phases d'un processus d'enquête, de l'échantillonnage aux méthodes d'estimation et à la rédaction de rapports qualité : cadre théorique, contexte d'application, exemples, références bibliographiques.

› **L'échantillonnage des enquêtes auprès des ménages dans le recensement annuel de la population** (pdf, 43 Ko)

› **L'échantillonnage des enquêtes auprès des ménages dans la source fiscale** (pdf, 43 Ko)

› **Stratification et calcul d'allocations dans les enquêtes auprès des entreprises** (pdf, 150 Ko)

› **La coordination d'échantillons des enquêtes auprès des entreprises** (pdf, 106 Ko)

› **Le renouvellement partiel d'échantillons d'entreprises** (pdf, 101 Ko)

› **L'échantillonnage équilibré** (pdf, 162 Ko)

› **La méthode du partage des poids** (pdf, 455 Ko)

› **La correction de la non-réponse par repondération** (pdf, 140 Ko)

› **La correction de la non-réponse par imputation** (pdf, 134 Ko)

› **Traitement des valeurs influentes dans les enquêtes** (pdf, 142 Ko)

› **Les méthodes de calage** (pdf, 76 Ko)

estimateur de Sen-Horvitz-Thompson

$$\hat{Y}_S = \sum_{i \in S} \frac{y_i}{\pi_i}$$

$$\hat{Y}_R = \sum_{i \in R} \frac{y_i}{\pi_i} + \sum_{i \in S-R} \frac{y_i^*}{\pi_i}$$

$$\forall j=1 \dots J \quad \sum_{k \in S} w_k x_{jk} = X_j \quad (1)$$

1.2 La statistique publique c'est de la cuisine

... et de la **bonne cuisine**! Donc :

Ça ne s'improvise pas → des **recettes/méthodes rigoureuses**



Ça prend du temps → le **temps long** des enquêtes
ou des données administratives

1.2 La statistique publique c'est de la cuisine

Des questionnaires au plus près du quotidien des citoyens

Exemples :

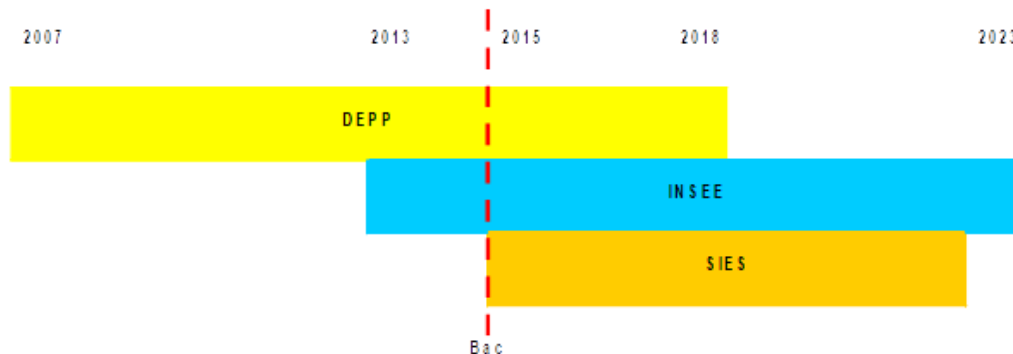
L'enquête Entrée dans la vie adulte

L'enquête Cadre de vie et sécurité

1.2 La statistique publique c'est de la cuisine

L'enquête EVA, entrée dans la vie adulte

- Panel de jeunes, entrés en 6^e en 2007
- Une enquête annuelle
- Pilotage
 - Ministère chargé de l'éducation nationale, direction de l'évaluation de la prospective et de la performance (Depp)
 - Ministère chargé de l'enseignement supérieur, sous-direction des systèmes d'information et des études statistiques
 - Insee, département de l'emploi et des revenus d'activité
- Taille de l'échantillon : 35 000 jeunes en 2007, 34 343 encore suivis en 2015 et 81,2 % de répondants à l'enquête
- Suivi de l'échantillon
- Réponses possibles sur questionnaire papier ou par internet, relance par téléphone



1.2 La statistique publique c'est de la cuisine

L'enquête Cadre de vie et sécurité

L'enquête a lieu chaque année de février à avril. Le questionnaire est administré par les enquêteurs de l'Insee sous un mode de collecte assistée par ordinateur (Capi) et, **pour les questions sensibles**, selon un **protocole adapté** permettant de **garantir une complète confidentialité des réponses** de la personne enquêtée.

Des **associations d'aide aux victimes** ont apporté leur expérience quant à la formulation de ces questions sensibles et au protocole d'enquête. Afin que des personnes d'origine étrangère puissent répondre, les questions sont traduites en arabe, en portugais, en turc ou en anglais, les quatre langues les plus parlées en France par les non-francophones.

Plan de sondage

L'échantillon est composé de **22 800 logements (soit environ 15 500 ménages répondants)**. Il comprend en 2018 une extension sur les quartiers prioritaires.

Sont interrogées les personnes de 14 ans ou plus au 1er janvier de l'année d'enquête (une seule personne par ménage).

1.2 La statistique publique c'est de la cuisine

... et de la **bonne cuisine**! Donc :

Ça ne s'improvise pas → des **recettes/méthodes rigoureuses**

Ça prend du temps → le **temps long** des enquêtes ou des données administratives

Ça plaît toujours et à tous → **séries longues**
et **comparaisons internationales**



1.2 La statistique publique c'est de la cuisine

... et de la **bonne cuisine**! Donc :

Ça ne s'improvise pas → des **recettes/méthodes rigoureuses**

Ça prend du temps → le **temps long** des enquêtes ou des données administratives

Ça plaît toujours et à tous → **séries longues** et **comparaisons internationales**

(et c'est meilleur le lendemain :)

02 Éclairer le débat public

2.1 Panorama

- **les chiffres** qu'il faut connaître,
- **les réponses** aux questions que vous vous posez... ou pas

2.1 Panorama

- Immigrés et étrangers :

- les chiffres qu'il faut connaître,
- les réponses aux questions que vous vous posez... ou pas

2.2 Insee en Bref

- Comment l'Insee mesure :

- la pauvreté,
- le pouvoir d'achat,
- le PIB,
- etc.

03 “On peut faire dire aux chiffres...”

Préambule : les statistiques, construction sociale

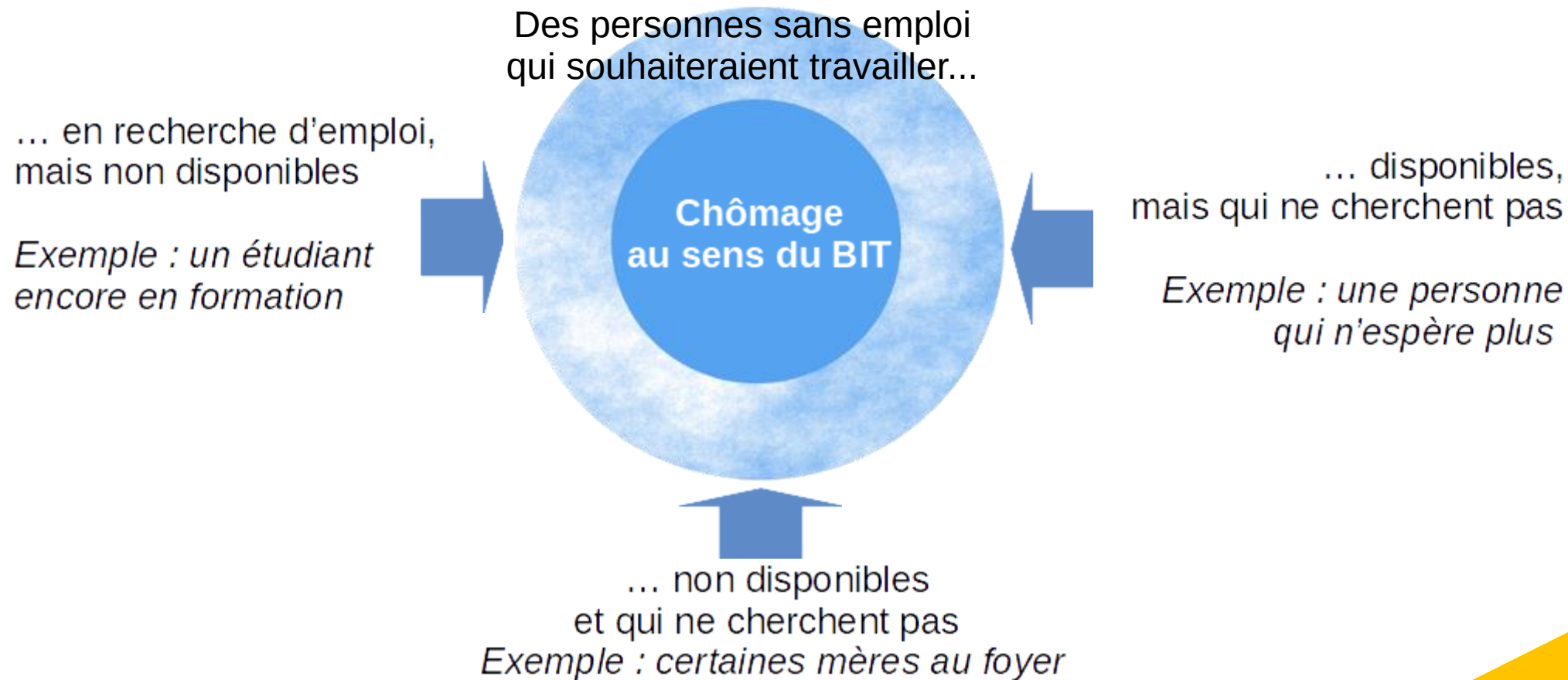
- **Les données** ne sont pas données, elles **sont construites**

Quantifier c'est d'abord **convenir** puis **mesurer...**

- les **conventions** sont **discutables** (par ex. pauvreté, chômage)
- les **mesures** doivent répondre à des critères **interrogeables** (normes internationales connues)

... pour savoir « ***De quoi on parle*** »

Un exemple de convention : le halo du chômage



03 “On peut faire dire aux chiffres...”



3.1 L'effet questionnaire

La formulation des questions

La formulation des questions a une influence (volontaire ou involontaire) sur les réponses des interrogés

- Tournure des questions
 - **Conformité sociale** : « Êtes-vous d'accord... »
 - **Verbes introducteurs** « Pensez-vous que... » / « Souhaitez-vous ... » / Ø
 - **Formulation des questions positives/négatives/neutres**
 - Préférence des questionnés pour le oui
 - **Vocabulaire utilisé : sécurité ≠ insécurité, interdire ≠ autoriser**
 - « Vous sentez-vous en sécurité dans votre vie quotidienne ? »
19 % des personnes interrogées répondaient ne pas se sentir en sécurité.
 - « Vous sentez-vous en insécurité dans votre vie quotidienne ? »
28 % des enquêtés ont déclaré ressentir une certaine insécurité



La formulation des questions

Il faut favoriser les questions précises et simples avec des termes accessibles à tous et non ambigus.

- **Questions vagues**

- Ex : « Pour chacune des personnalités politiques suivantes, dites-moi si vous en avez plutôt une image positive ou plutôt une image négative ? »

Bonne image d'une personnalité $\neq$ adhésion aux idées

- **Questions trop compliquées**

- Négation : « Ne pensez-vous pas qu'il faudrait interdire... »
- Mots compliqués (imminent, administré...) sauf s'il s'agit d'un milieu professionn

- **Questions gênantes**

- Thème « sensible » (sexualité, argent...) → réponse honnête ?



L'ordre des questions

L'enchaînement des questions

- Conditionnement négatif d'une question-amorce (effet de priming ou d'amorçage)

Sondage 1 : affirmations préparatoires à effet négatif

A. Parmi ces deux affirmations, laquelle est selon vous la plus vraie ?

● Les politiques environnementales pour réduire les émissions de gaz à effet de serre dans le secteur de l'automobile vont avoir de lourdes conséquences sur l'emploi

● La création d'une taxe écologique prélevée sur les revenus contribuera à réduire le pouvoir d'achat des Français

Affirmations préparatoires à effet négatif

B. Après avoir bien réfléchi, que pensez-vous des mesures visant à réduire les émissions de gaz à effet de serre ?

Contre	20 %
Hésite	11 %
Pour	69 %

Sondage 2 : contrôle

Question directe sans influence (pas d'affirmations préparatoires).

Que pensez-vous des mesures visant à réduire les émissions de gaz à effet de serre ?

Contre	8 %
Hésite	6 %
Pour	86 %

Le type de réponses

Choix des réponses

- Propositions définies ou libres
- Présence du « ne se prononce pas » et de l'intervalle moyen ou non pour forcer l'enquêté à se positionner
- Autant de modalités positives que négatives dans le choix des réponses

Influence sur les réponses

- Ordre de présentation des réponses
- Nombre de modalités



L'ordre des réponses

Exemple (ordre des modalités)

Le problème le plus grave aujourd'hui	Ordre	Ordre inverse
Le chômage	34,2	19,2
Le terrorisme	8,5	7
La faim dans le monde	21,3	17,7
La guerre	15,3	18,3
La surpopulation du monde	1,5	1,7
Le racisme	5,0	8,5
Le non respect des droits de l'homme	8,0	10,4
L'insuffisante formation professionnelle des jeunes	1,3	6,2
La délinquance	1,2	8,2

Les modalités des réponses

Exemple (nombre de modalités)

- « Est-ce que votre vie de travail vient en conflit avec votre vie personnelle ou familiale ? »

	Début 1988
Oui	38
Non	62
<i>Ne sait pas</i>	-
Ensemble	100

(en %)

	Début 1990
Très souvent	6
Souvent	10
Quelquefois	30
Jamais	53
<i>Ne sait pas</i>	0
Ensemble	100

46

Source : CREDOC, Enquête « Conditions de vie et Aspirations des Français » début 1988.

Source : CREDOC, Enquête « Conditions de vie et Aspirations des Français », début 1990.

03 “On peut faire dire aux chiffres...”



3.2 Manipulations



Regroupement des résultats

Exemple d'un sondage sur le travail dominical



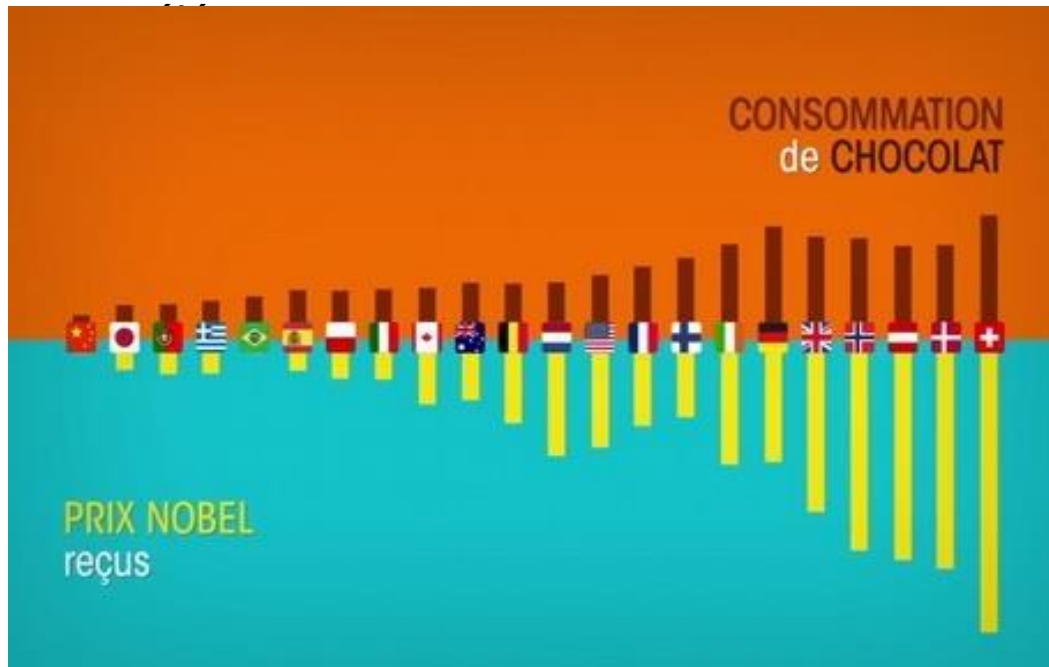
- « Travailler le dimanche et être payé davantage qu'en semaine. Si votre employeur vous proposait de travailler le dimanche, accepteriez-vous ? »
 - A : non jamais 33 %
 - B : peut-être de temps en temps 50 %
 - C : oui toujours 17 %
- Résultat du sondage : 67 % des Français sont favorables au travail le dimanche (regroupement catégories B et C)





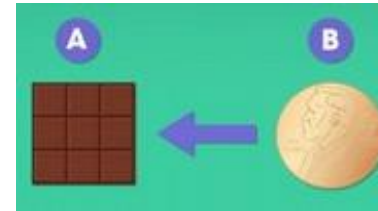
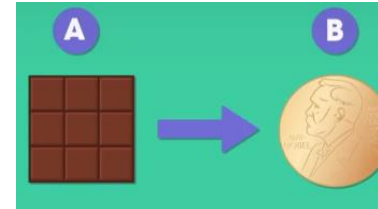
Corrélation et causalités

Exemple : La consommation de chocolat et le nombre de prix Nobel reçus dans un pays sont



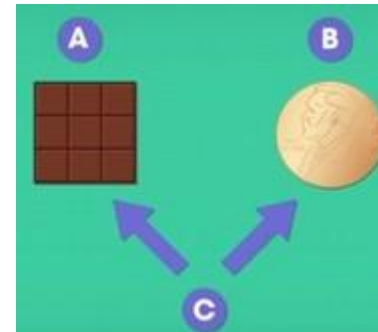
Causalité

- Le chocolat rend intelligent ?
- On célèbre le prix Nobel en mangeant du chocolat ?



Lien indirect

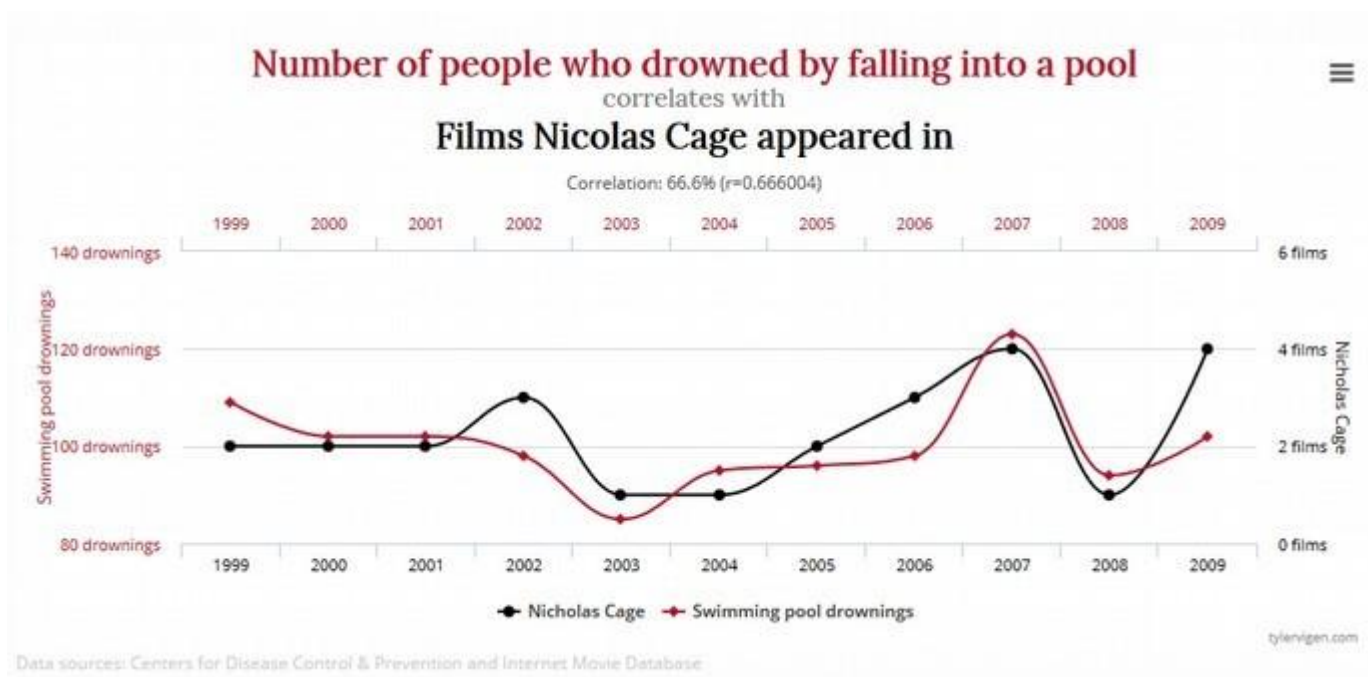
- Hypothèse : Un pays riche dispose de plus d'agent pour l'éducation et la recherche. Parallèlement sa population plus riche peut consommer plus de chocolat.





Corrélation illusoire

Exemple : plus **Nicolas Cage** apparaît dans des films,
plus le nombre de **morts noyés dans une piscine** augmente



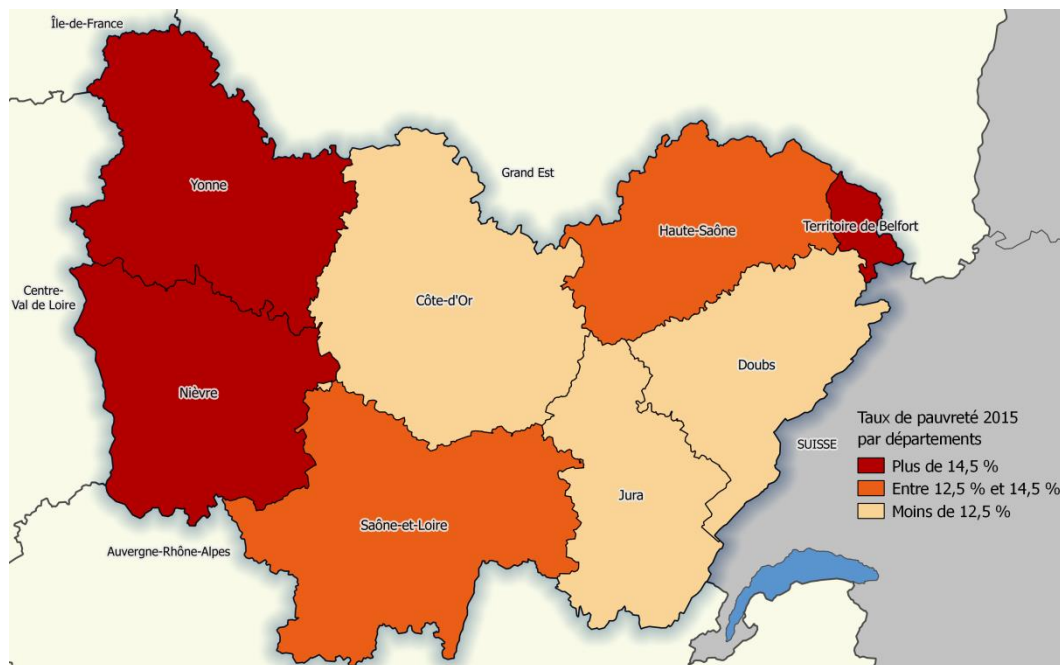
→ La corrélation n'implique pas la causalité !

03 “On peut faire dire aux chiffres...”



3.3 Un graphique peut en cacher un autre

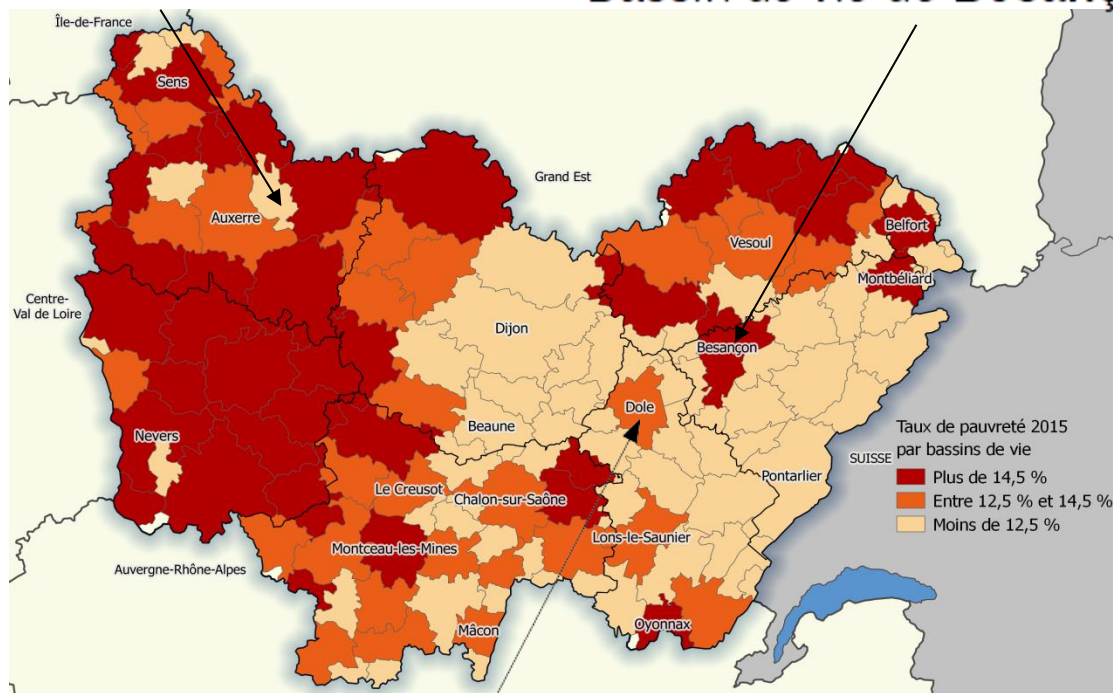
Pauvreté départementale



Pauvreté par bassin de vie

Bassin de vie de **Chablis** : 11,3 %

Bassin de vie de **Besançon** : 15,2 %

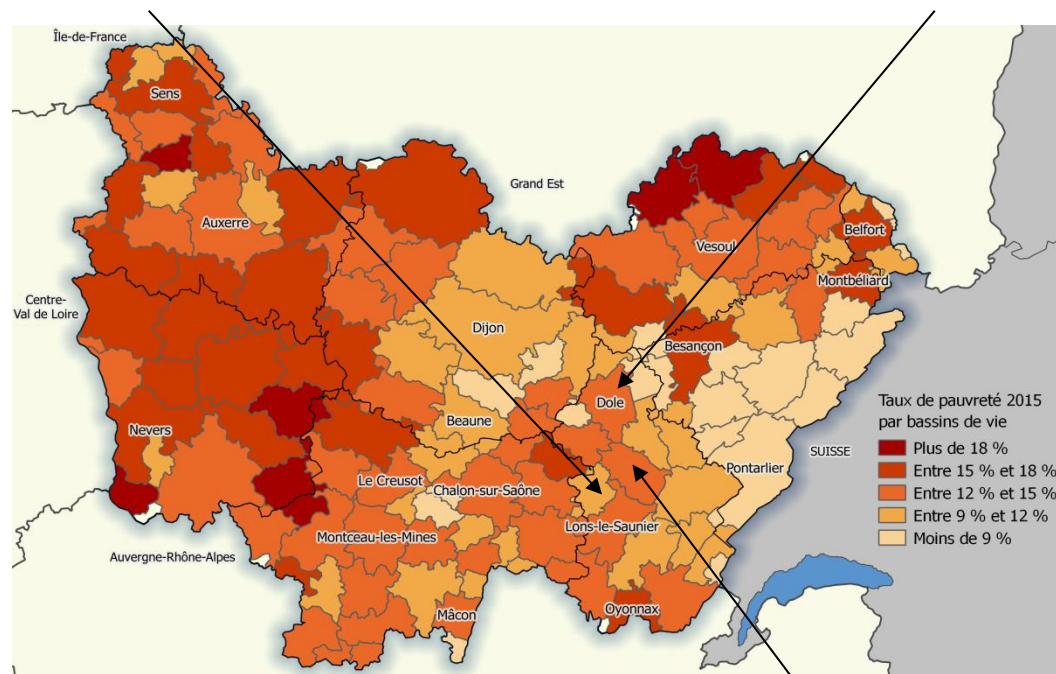


Bassin de vie de **Dole** : 14,5 %

Pauvreté par bassin de vie (variante)

Bassin de vie de **Bletterans** : 11,6 %

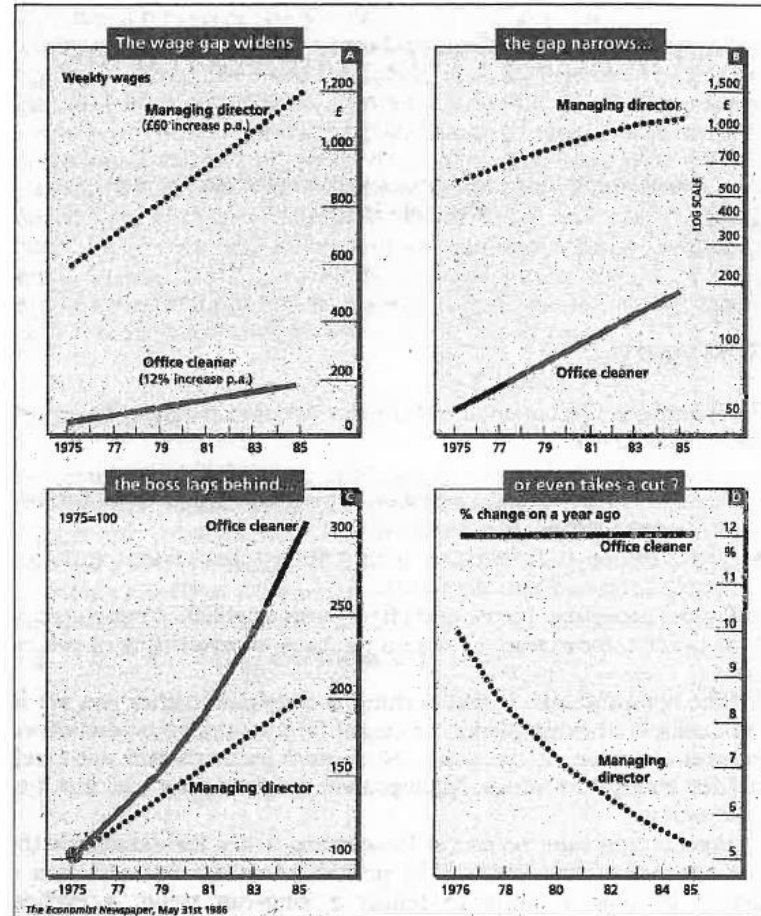
Bassin de vie de **Dole** : 14,5 %



Bassin de vie de **Poligny** : 12,2 %

Qui gagne ?

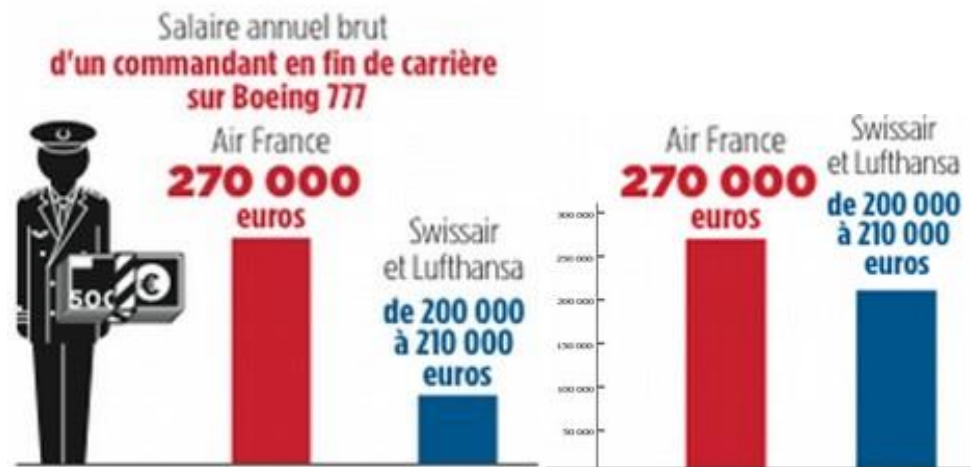
Figure 4.17 Four ways of watching wages



Influence des axes

Choix de l'échelle :

- Axe ordonnées (axe Y)
 - Échelle complète (commence à 0) ou incomplète



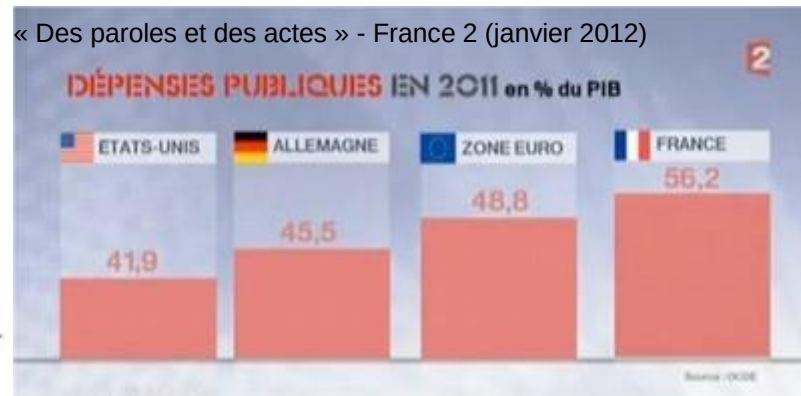
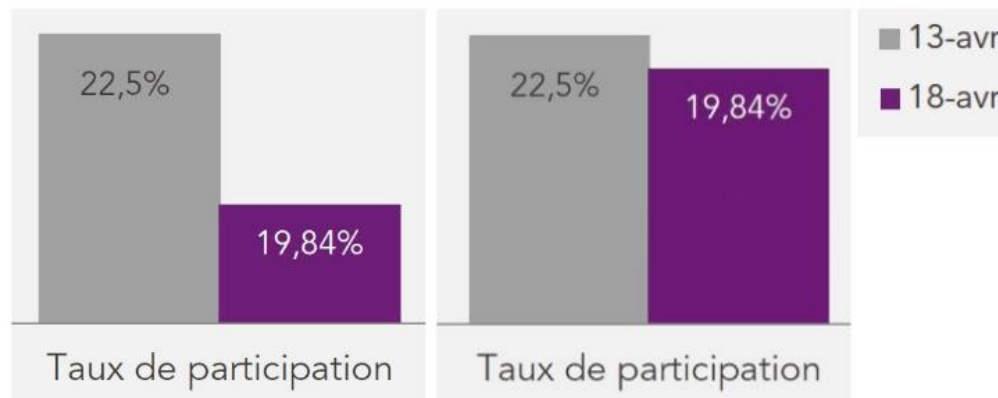
Graphique publié en oct. 2015 par

Le Point

Graphique avec échelle débutant à 0

Taux de participation aux grèves SNCF

Communiqué de presse (avril 2018) Graphique « corrigé »

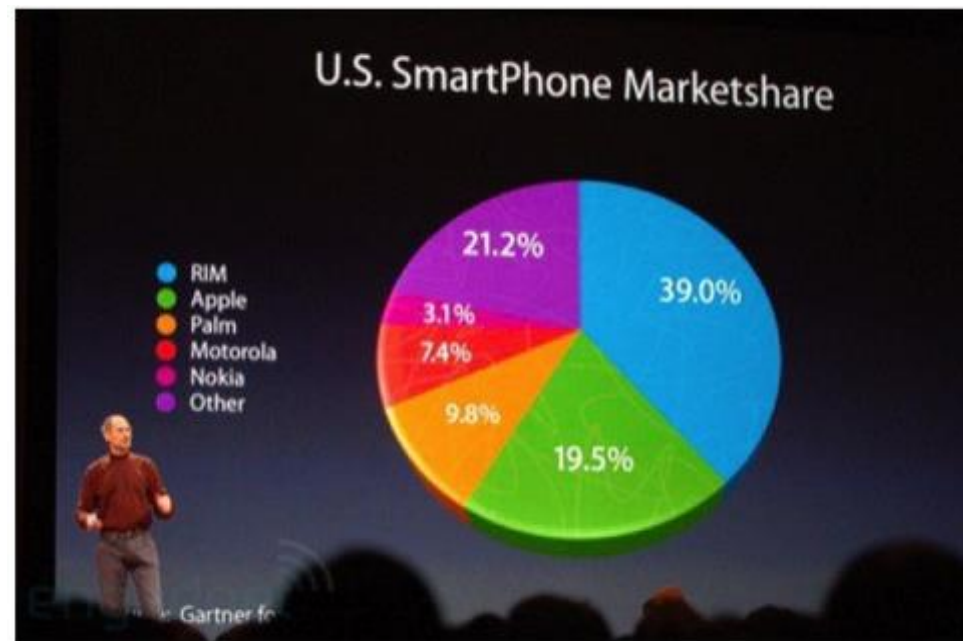


Graphiques trompeurs

Limites de la perception visuelle

- Graphique 3D

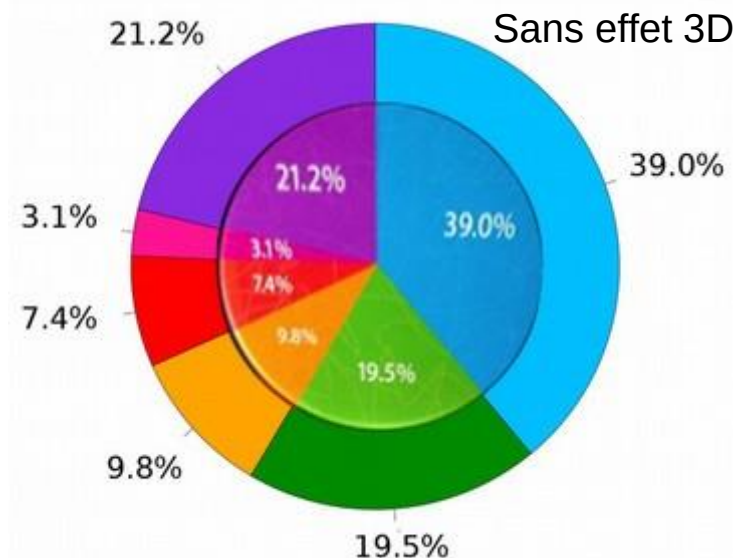
Répartition du marché américain de smartphone



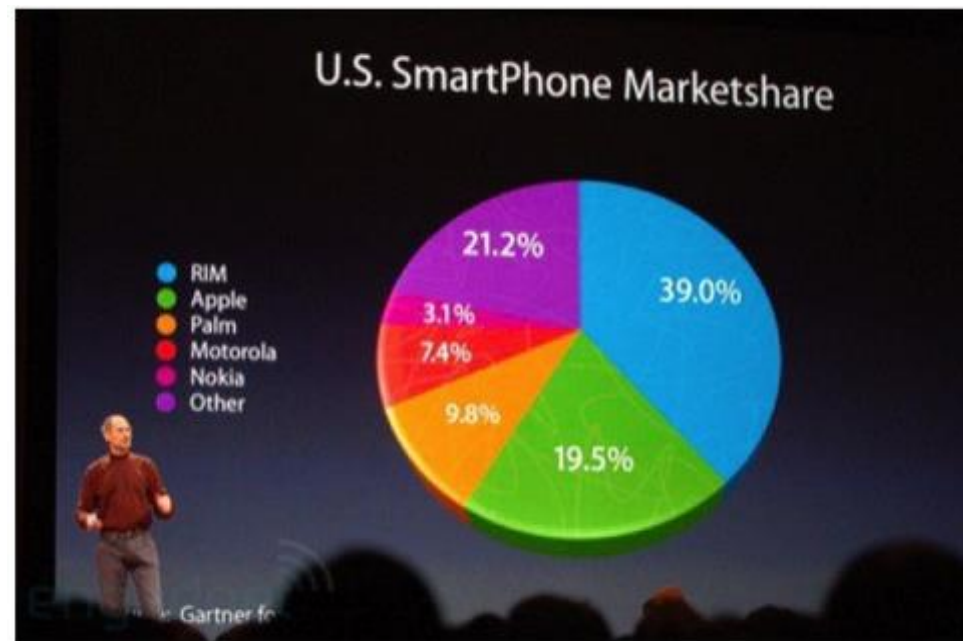
Graphiques trompeurs

Limites de la perception visuelle

- Graphique 3D



Répartition du marché américain de smartphone



Graphiques trompeurs

Influence du choix de la représentation

Nombre d'iPhone vendus



Présentation Apple - 2013 – Tim Cook

- « Les ventes cumulées d'iPhone sont en train d'exploser ! »



Graphiques trompeurs

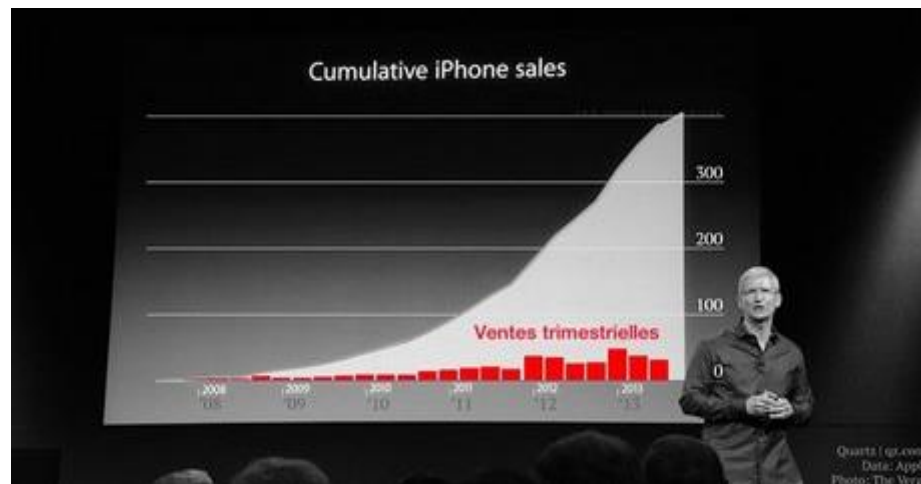
Influence du choix de la représentation

- Nombre cumulés

Nombre d'iPhone vendus



Présentation Apple - 2013 – Tim Cook



Graphiques trompeurs

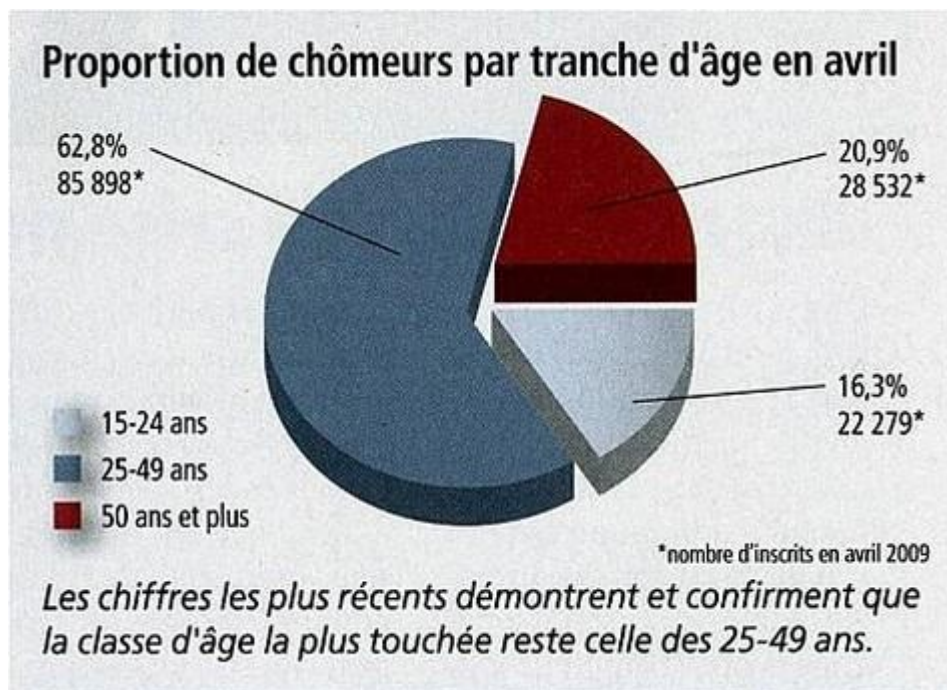
Mauvaise représentation des données

- Échelle des abscisses (horizontal) non constante

Le Matin Consommation de viande



Ne veut pas dire grand chose



Ne veut pas dire grand chose

Compare ce qui n'est pas comparable

- **Taille des 3 catégories**

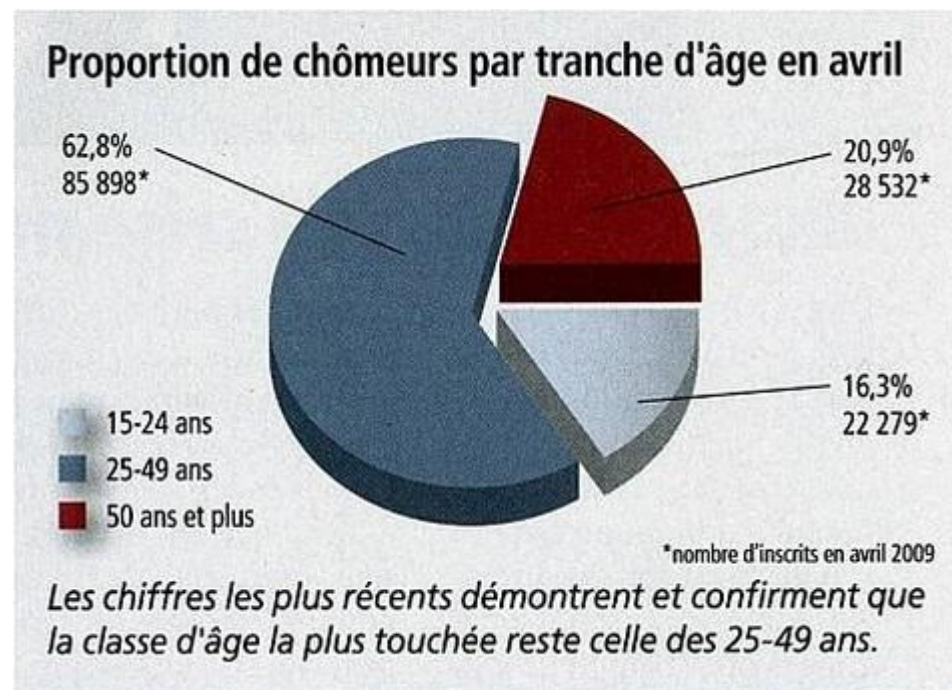
- 15-24 ans : 10 années
- 25-49 ans : 25 années
- 50 ans et plus : + de 15 années

→ La population est plus importante dans la tranche 25-49 ans.

- **Chômeurs inscrits et taux de chômage**

- | | | |
|--------------------|--------|-------|
| • 15-24 ans : | 22 279 | 4,0 % |
| • 25-49 ans : | 85 898 | 3,6 % |
| • 50 ans et plus : | 28 532 | 2,6 % |

→ Le chômage est plus élevé chez les jeunes



Source : SECO (secrétariat d'état à l'économie) 

Erreurs de représentation

Surface proportionnelle à l'effectif



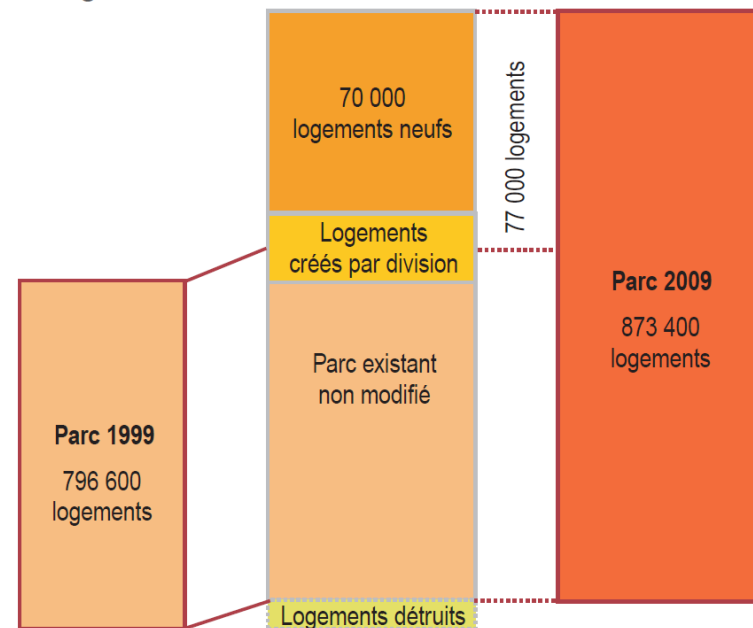
Erreurs de représentation

Surface proportionnelle à l'effectif



3 Le parc de logements évolue dans sa structure

Évolution du parc de logements entre 1999 et 2009 en Picardie



Source : Dreal, Filocom

Erreurs de représentation

Un drôle de camembert



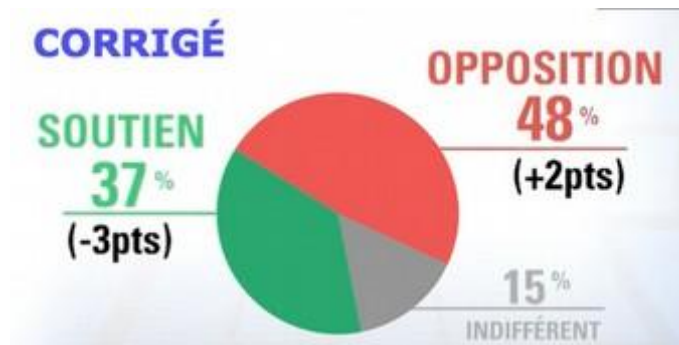
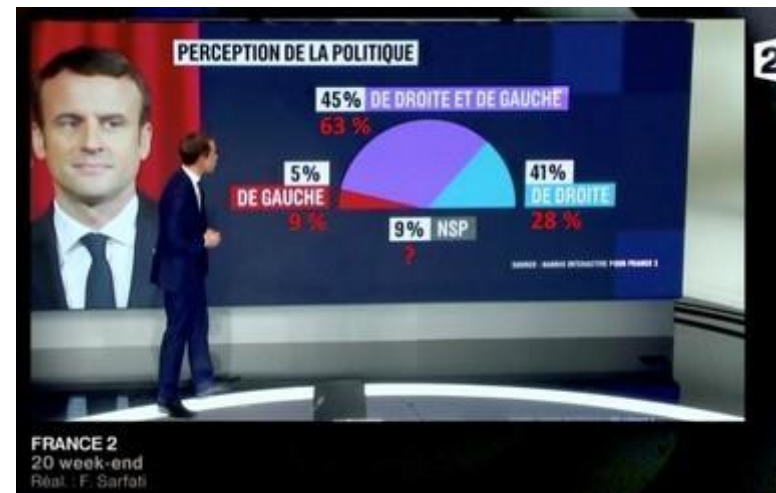
Erreurs de représentation

Un drôle de camembert



Erreurs de représentation

Un drôle de camembert



03 “On peut faire dire aux chiffres...”



3.4 On fait ce qu'on peut, mais...



Au-delà du constat brut (BFC)

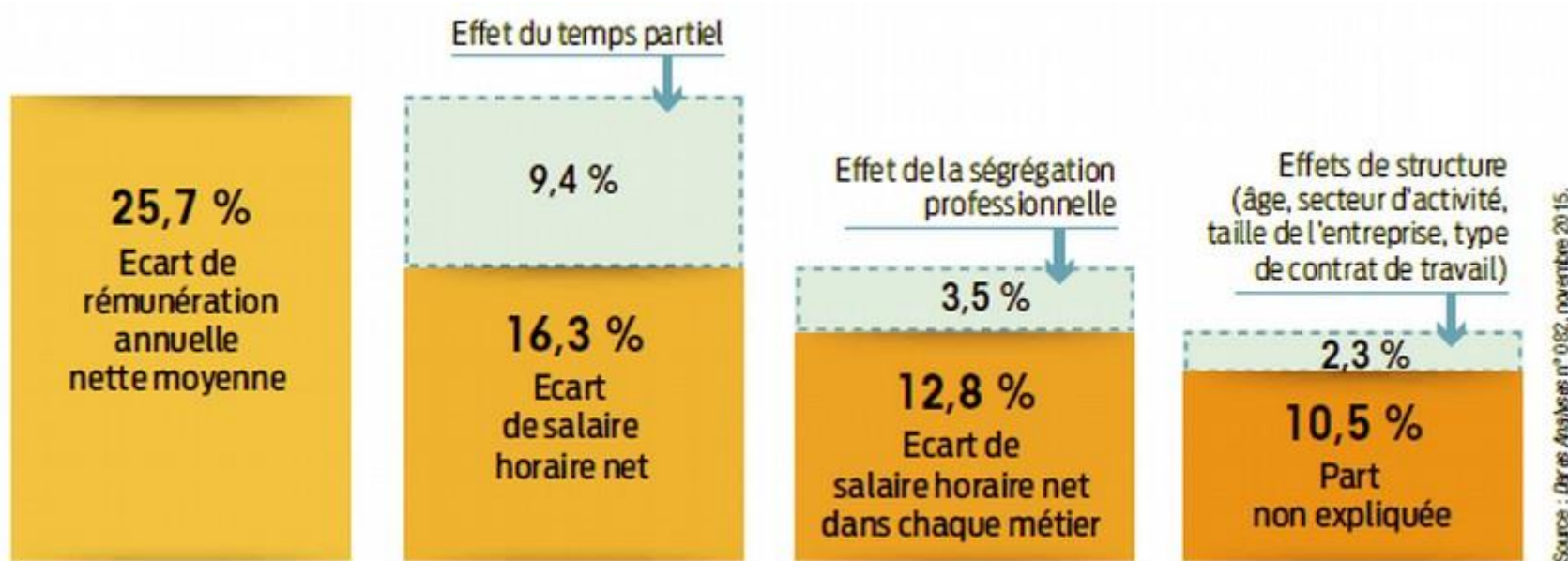
Figure_1 – Toutes choses égales par ailleurs, un écart salarial inexplicable de 6,8 %





Au-delà du constat brut (national)

- Décomposition des écarts de salaires femmes - hommes :





Mais...

Reprise dans la presse de l'étude de l'Insee :

« Les revenus et le patrimoine des ménages », *Insee Références* - Édition 2018

REDISTRIBUTION
**Le 1% le plus riche a pourtant
de quoi ruisseler**



**1 % des Français paient un
quart de l'impôt sur le revenu**



Les très hauts revenus en 2015 :
**1 % de la population perçoit 7 % des revenus
et 30 % des revenus du patrimoine déclarés**



04 Pour conclure

**Pour trouver de l'intérêt
aux statistiques publiques,
rien ne vaut la pratique !**

**Professeurs, élèves
participez à la**

**Pour trouver de l'intérêt
aux statistiques publiques,
rien ne vaut la pratique !**

**Professeurs, élèves
participez à la**

**Inscrivez-vous
avant le 6 décembre !**

**Compétition européenne
de statistiques**

ES European
Statistics
Competition

Année
scolaire
2019-2020



Professeurs et élèves de seconde, de 1ère et de terminale, PARTICIPEZ !
Inscriptions jusqu'au 6 décembre 2019 sur [insee.fr](https://www.insee.fr). Début des épreuves le 9 décembre 2019.

“Pour ma part, je reste persuadé que les statistiques, à condition de préciser ce qu’elles visent à chiffrer, demeurent un outil essentiel pour comprendre la société dans laquelle nous vivons.”

Denis Clerc (*Alternatives économiques*)

“Les chiffres ne disent n’importe quoi

**“Les chiffres ne disent n’importe quoi
qu’à ceux qui croient n’importe qui.”**

MERCI DE VOTRE ATTENTION !

Retrouvez-nous sur :

[insee.fr](https://www.insee.fr)



Jean-William Angel
Division Écoute & Communication