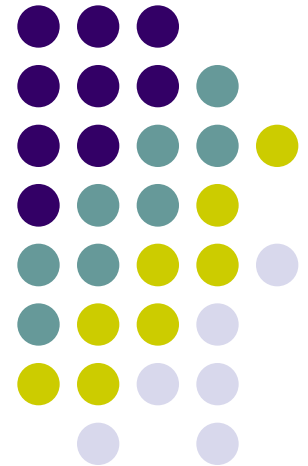


# Faciliter l'enseignement des concepts statistiques en recourant à leur histoire

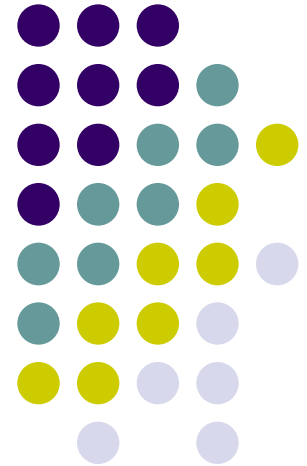
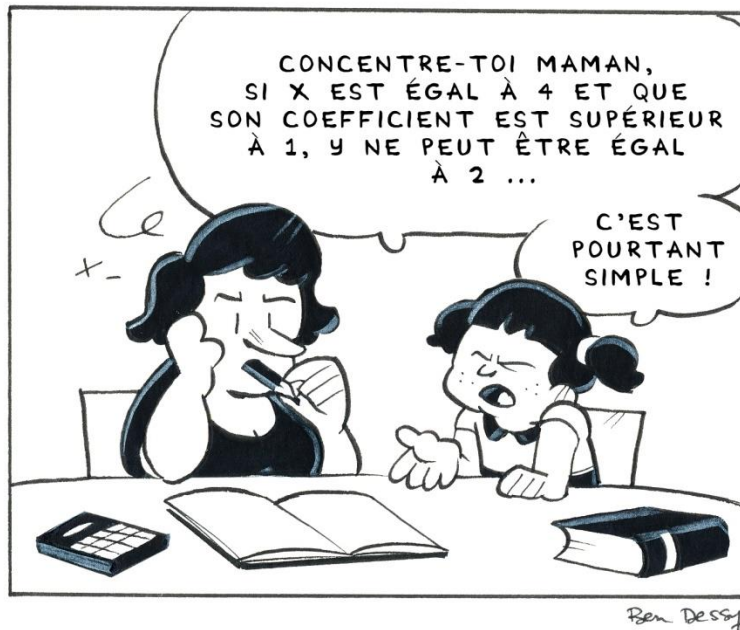
## Jean-Jacques Dreesbeke

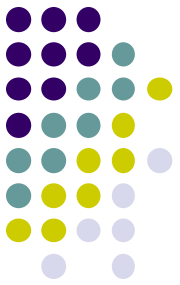
*On ne connaît pas complètement une science  
tant qu'on n'en sait pas l'histoire*

Auguste Comte, *Cours de philosophie positive*



# Enseigner la statistique à des élèves





# La méthode traditionnelle

- **Concept: la moyenne arithmétique**

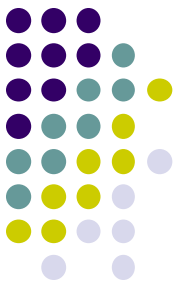
- **Définition:**

$$\{x_1, x_2, \dots, x_n\} \rightarrow \bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

→ propriétés

- **Un exemple**

- **Passons au paragraphe suivant...**



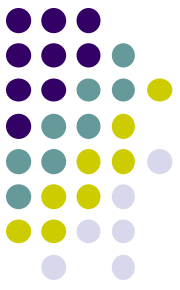
# La méthode progressive

- **Concept:** la **moyenne arithmétique**
- **Un exemple:** argent de poche hebdomadaire (€) reçu par de jeunes adolescents:  
 $\{10, 10, 10, 15, 15, 20, 50, 50, 50, 50, 60\}$   
 $\rightarrow \bar{x} = 30,9091\text{€}$
- **On exprime cela par une formule? Euh...**
- **Passons au paragraphe suivant...**



# La méthode nouvelle

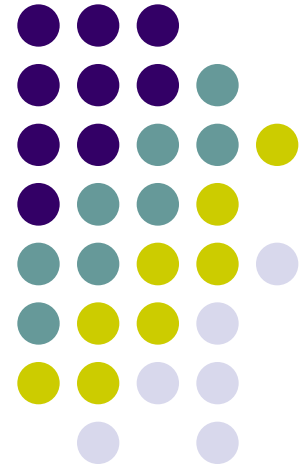
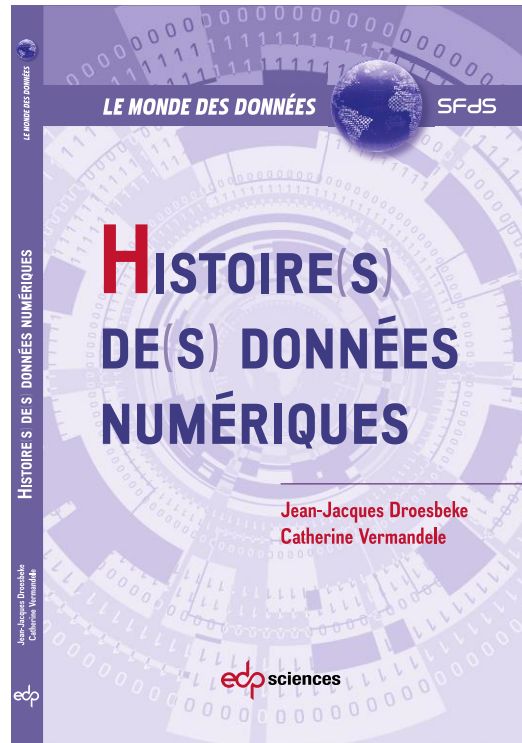
- **Concept: la moyenne arithmétique**
- **Je vous envoie un fichier de données et le logiciel statistique que j'ai appris à utiliser. À vous d'apprendre à faire de même pour obtenir la moyenne**
- **On exprime cela par une formule? Bof...**
- **Passons au paragraphe suivant...**

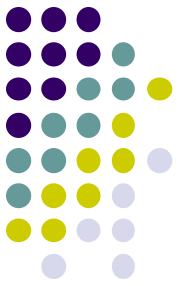


# Des questions?

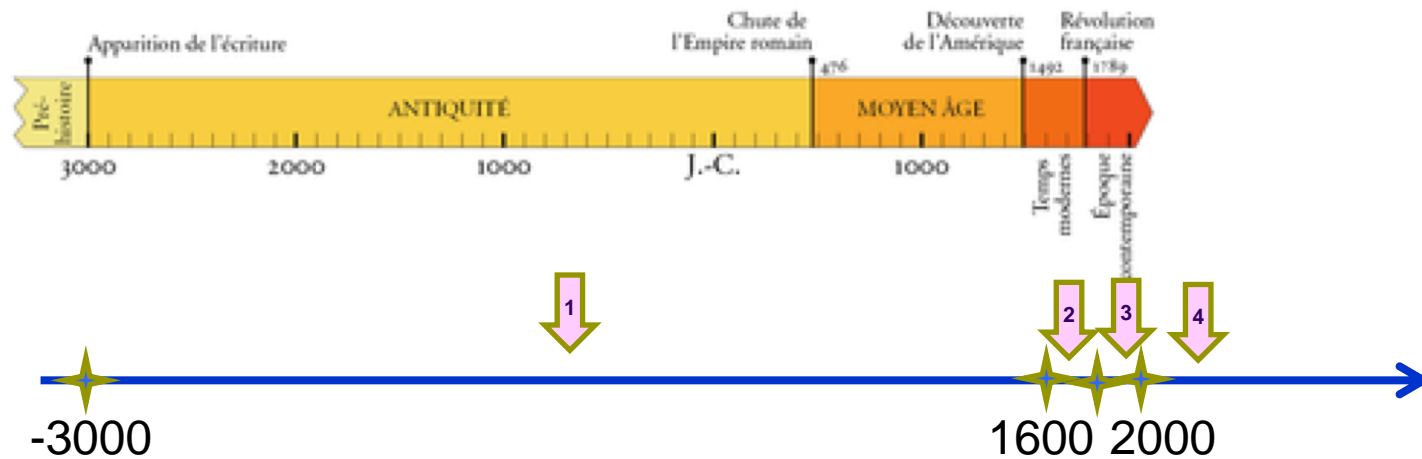
- C'est quoi le «  $i$  » dans la formule?
- Qu'est-ce que je fais avec la moyenne? **De toute façon, ma mère me donne jamais d'argent de poche...**
- Comment on fait pour sélectionner les données dans le fichier?
- **À quoi ça sert? Comment on l'interprète? Est-ce toujours utile? On r'fait un exercice, M'sieur?**

# Pourquoi ne pas utiliser l'Histoire des données?

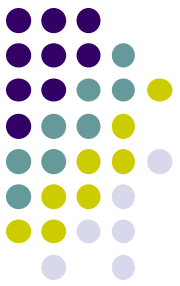




# Un voyage dans le temps







# Avant le 17<sup>e</sup> siècle

- **Données démographiques: catégories sociales**  
→ **On dénombre**



- **Données économiques et commerciales**
- **Données astronomiques, géodésiques**

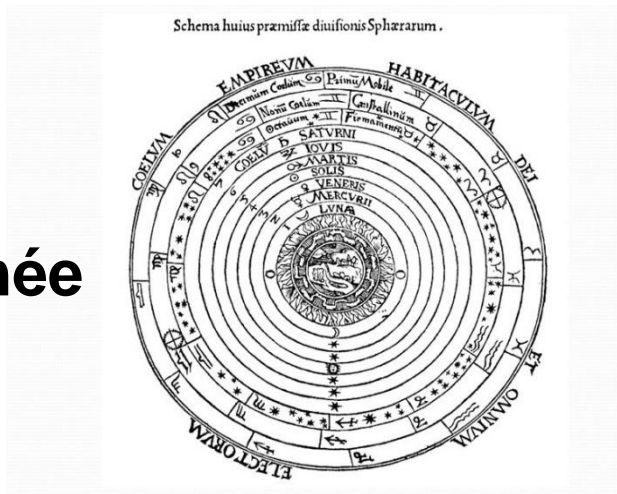
**On (tente de) mesure(r)!**



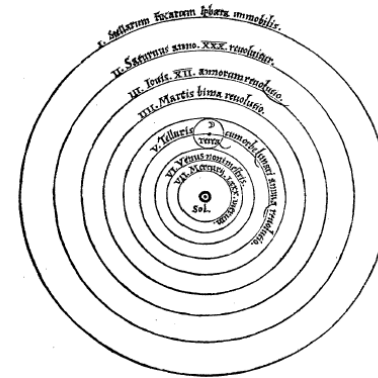
# Sans données

- On modélise!

Ptolémée



Copernic

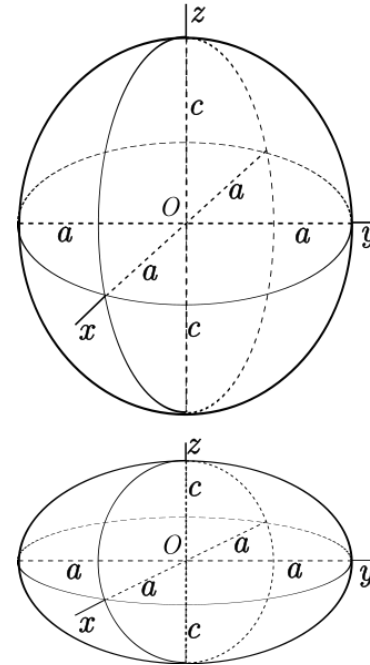
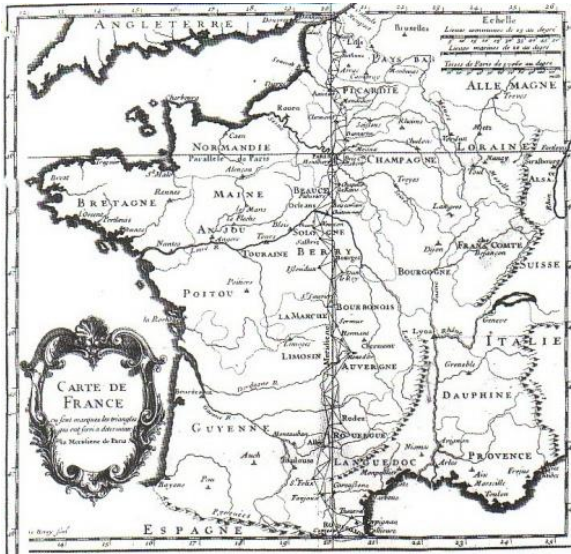


Où est le centre de l'univers?



# On a besoin de données!

- **Exemples:**



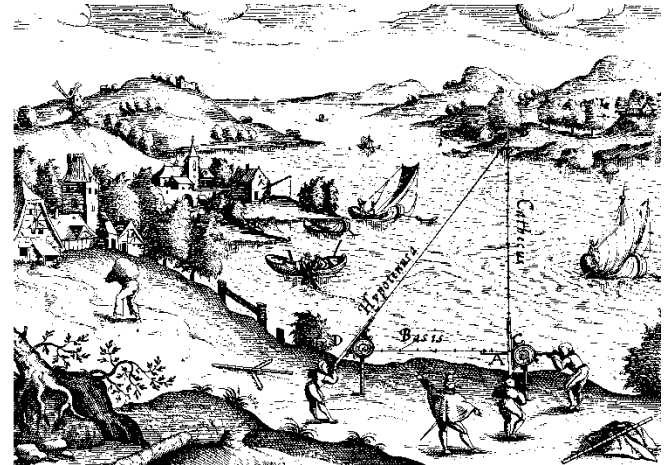
**La figure de la Terre**

- **Et quand on a plusieurs données?**



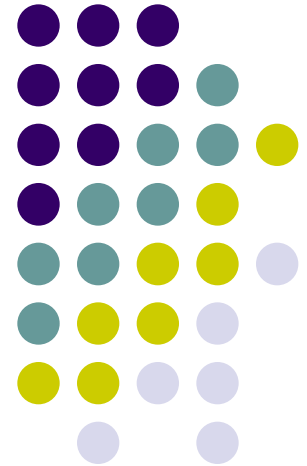
# Aux 17<sup>e</sup> et 18<sup>e</sup> siècles

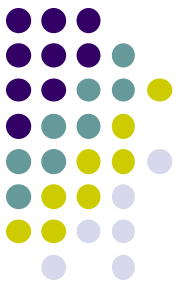
- On dénombre à nouveau
- On mesure de plus en plus: des distances, des angles... (on quantifie aussi)



**Que faut-il choisir comme « bonne mesure »?**

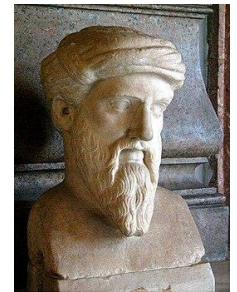
# Prendre la moyenne arithmétique?





# Petit retour en arrière

- - **550**: **Pythagoriciens** (portée musicale et proportion) → **moyenne arithmétique**



- **a** et **b** positifs:

**moyenne arithmétique:**  $\frac{a+b}{2}$

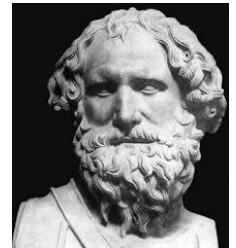
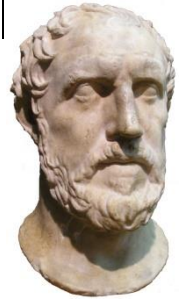
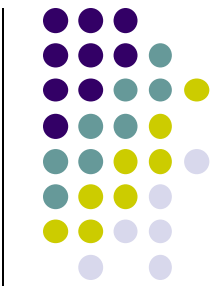
**moyenne harmonique:**  $\frac{2}{\frac{1}{a} + \frac{1}{b}}$  ; **moyenne géométrique:**  $\sqrt{ab}$

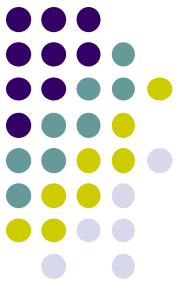
- **La Musique** d'Aristide Quintilien

# Le traitement des données

- **- 428:** Thucydide (la guerre du Péloponnèse)  
→ Usage du **mode**
- **- 162 à - 127:**  
**Hipparque:** mesures astronomiques  
→ **moyenne des deux valeurs extrêmes**  
(midrange)

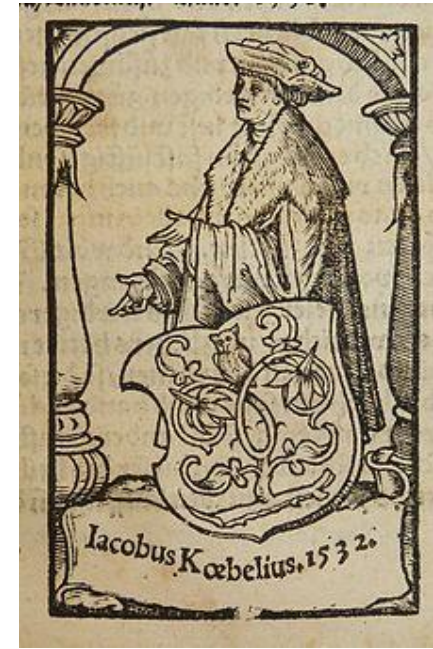
Et la moyenne?





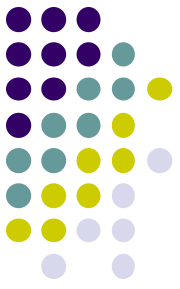
# Pourquoi ne pas utiliser une moyenne?

- **Début du 16<sup>e</sup> siècle: Jakob Köbel (1462-1533)**  
→ usage implicite de la moyenne





# Une moyenne, oui ou non?



- **Fin du 16<sup>e</sup> siècle: Tycho Brahé (1546-1601)**  
→ moyenne arithmétique, mais...
- **Johannes Kepler (1571-1630): idem**
- **17<sup>e</sup> siècle: Henry Gellibrand (1597-1637)**



**Robert Boyle (1627-1691)**





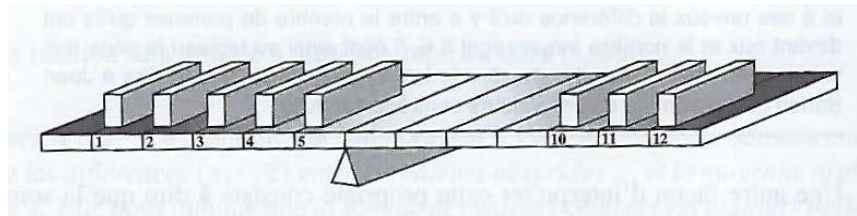
# Y'a-t-il quelqu'un qui ose?

- **18<sup>e</sup> siècle: Vauban (1633-1707):**

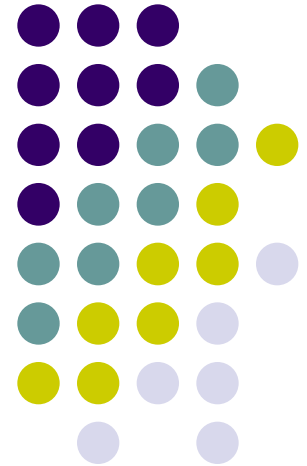
**moyenne proportionnelle  
valeur commune**

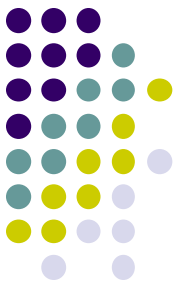


- **Roger Cotes (1682-1716)**



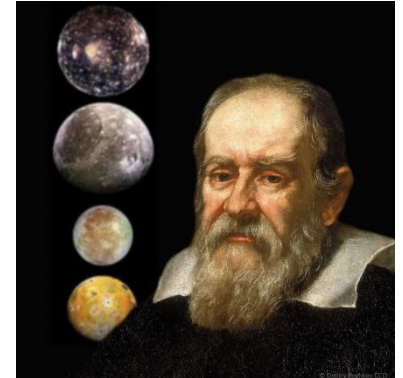
# La solution: s'intéresser à l'erreur





# Galileo Galilei (1564-1642)

- *Dialogo sopra i due massimi sistemi del mondo* (1632)
- $f(x) = |x - x_1| + |x - x_2| + \dots + |x - x_n|$
- $x = \text{moyenne?}$  **Non: une médiane!**

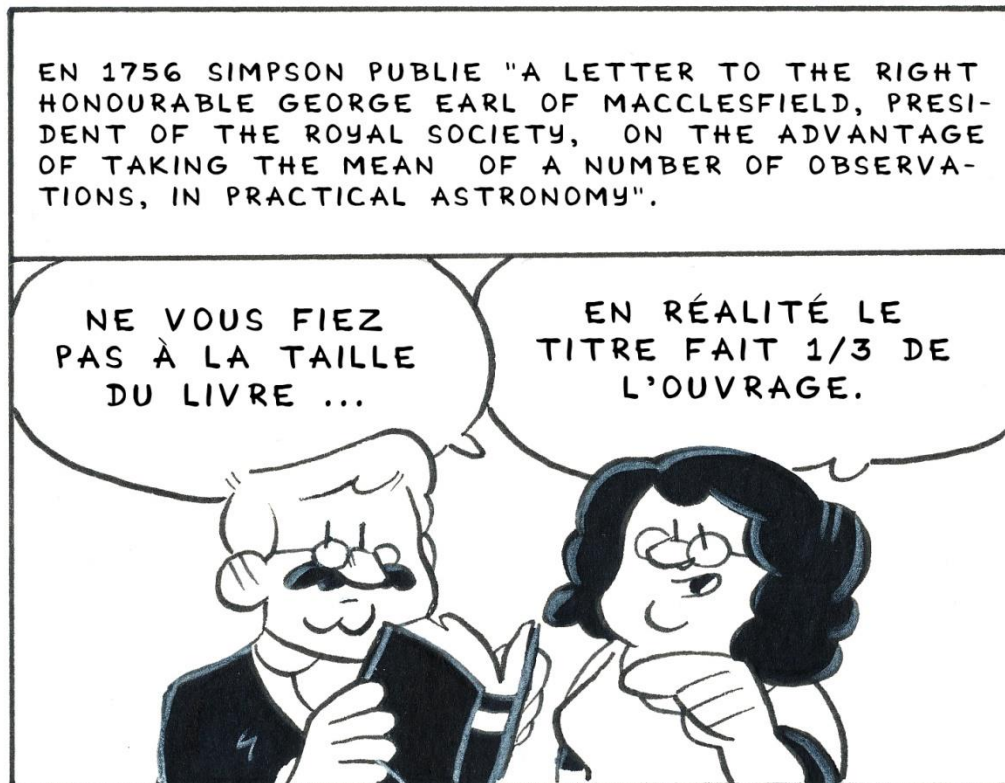
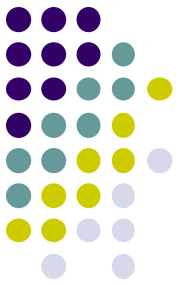


Série ordonnée:  $\{x_1, x_2, \dots, x_n\}$



$x_{1/2}$

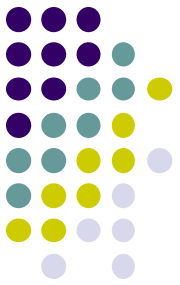
# Thomas Simpson (1710-1761)



Ben Dessy

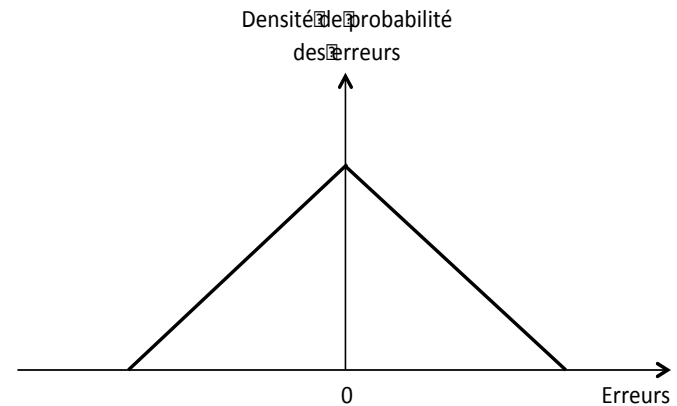
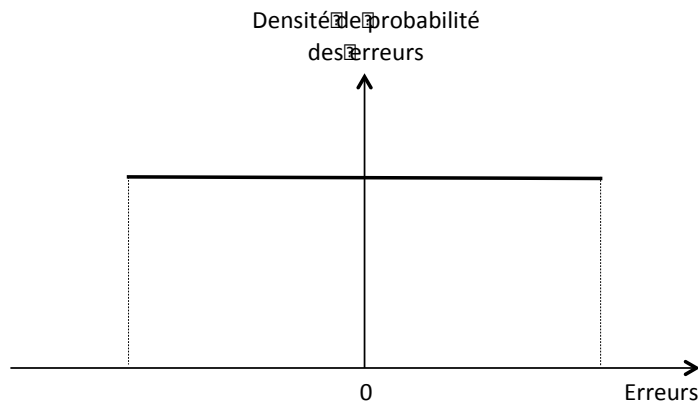


1756

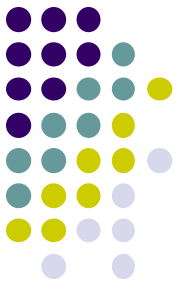


# Lois des erreurs de mesure

- Lois discrètes
- Lois continues

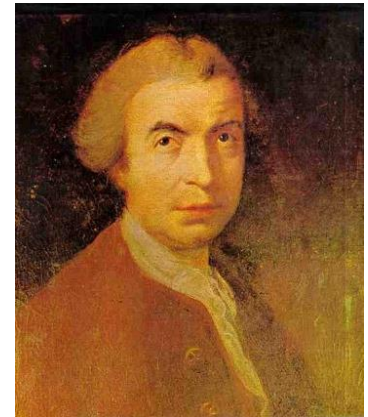


# Rogierus Josephus Boscovich (1711-1787)

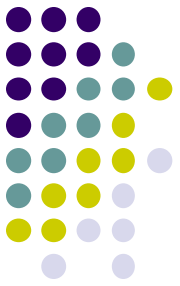


- 1740: professeur de mathématiques
- 1743: conseiller du pape Benoît XIV
- 1750-1752: mission avec Christopher Maire  
(estimer la longueur d'un arc de méridien à Rome +  
nouvelle carte géographique des états pontificaux)
- **1757: *De Literaria Expeditione  
per Pontificiam Ditionem  
et Synopsis Amplioris Operis***

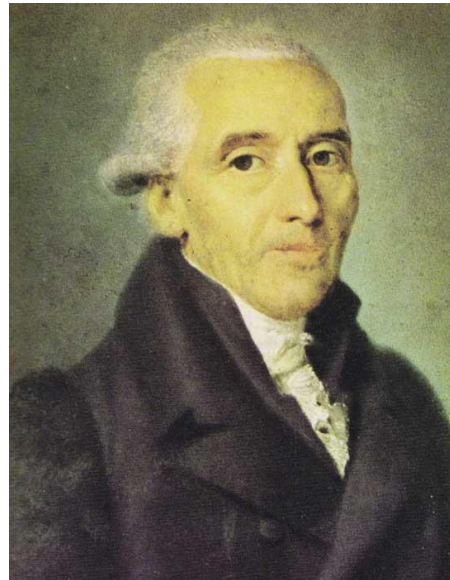
**Il privilégie l'usage de la médiane**



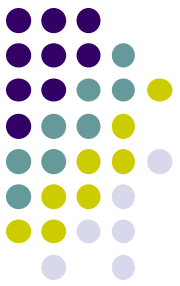
# Joseph-Louis, comte de Lagrange (1736-1813)



- **1776**: *Utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations...*

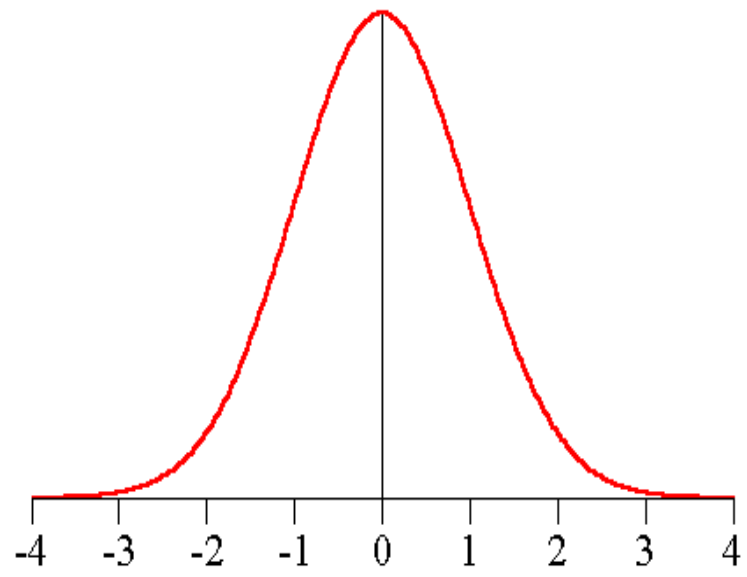
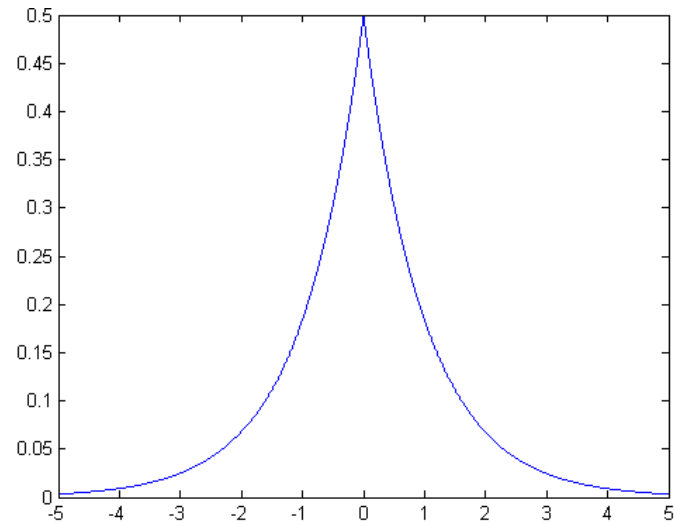


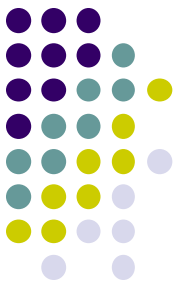




# Lois des erreurs de mesure

- Lois de Laplace et Gauss



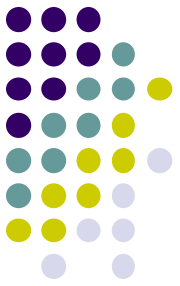


# L'attraction du milieu

- Jean Bernoulli (1744-1807):  
**Encyclopédie méthodique (1789): *Milieu***
- « *Quand on a fait plusieurs observations d'un même phénomène, et que les résultats ne sont pas tout à fait d'accord entre eux, on a coutume alors de prendre le milieu entre tous les résultats, parce que de cette manière les différentes erreurs se répartissent également dans toutes les observations, l'erreur qui peut se trouver dans le résultat moyen devient aussi moyenne entre toutes les erreurs* ».



# Carl Friedrich Gauss (1777-1855)



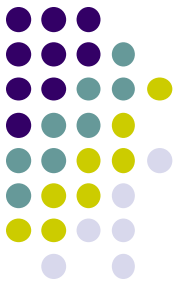
## *Découverte de Cérès (1801 →...)*

Recherche de l'orbite par un ajustement de données observées au moyen du *critère des moindres carrés*

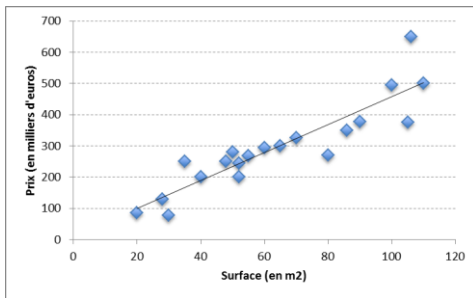


## *Trois hypothèses:*

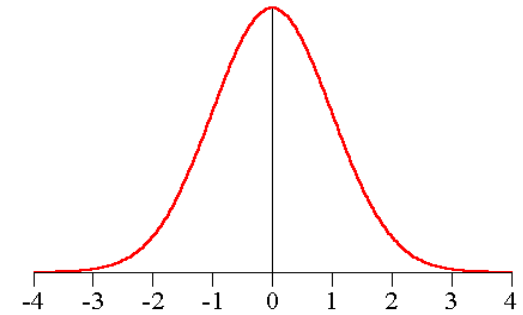
- Les petites erreurs plus vraisemblables
- La vraisemblance d'une erreur  $\varepsilon$  égale celle de  $-\varepsilon$
- En présence de plusieurs mesures d'une même quantité, la valeur la plus vraisemblable est la moyenne arithmétique



# Le sacre de la moyenne

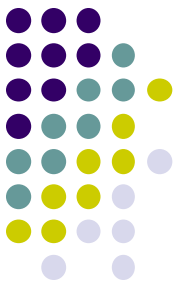


**Loi des erreurs:  
Loi « normale »**



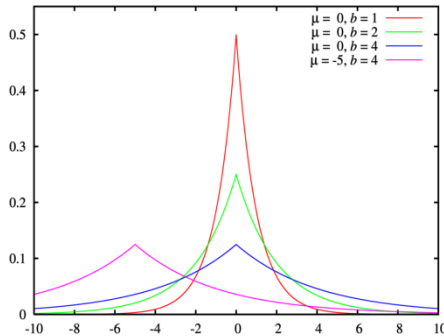
**Critère  $L_2$ :  
Moindres carrés**

**Milieu:  
Moyenne**

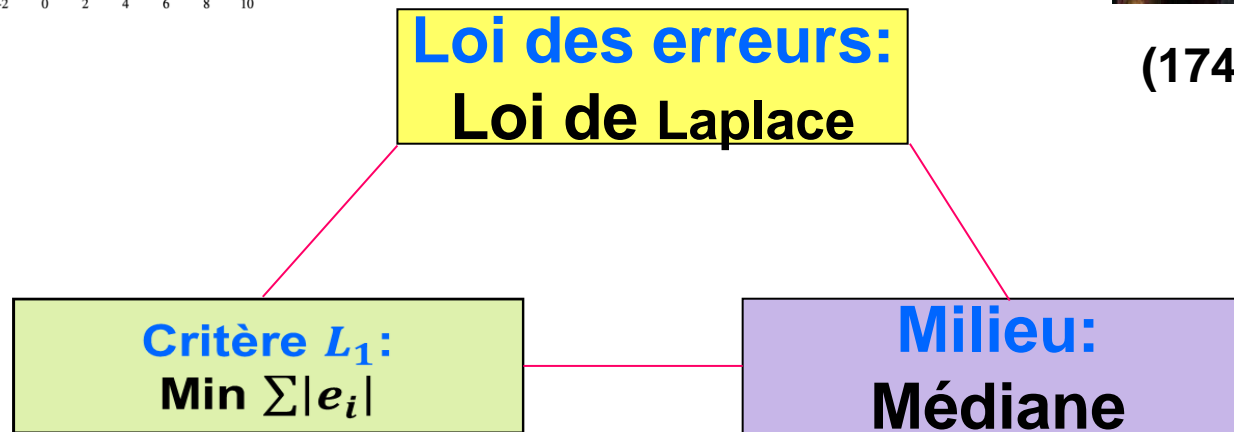


# Le perdant (momentané)

- Un autre triangle est possible

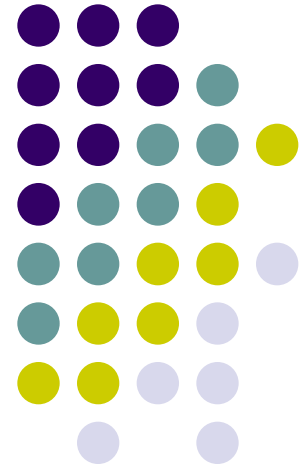


(1749-1827)

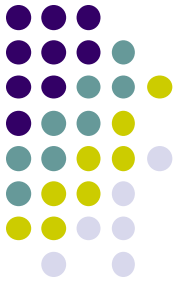


# Le règne de la moyenne

**Adolphe Quetelet**  
**(1796-1874)**



# La théorie des moyennes



- **Lettres à SAR le Duc...(1846)**

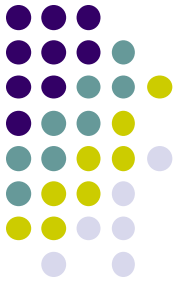
- Mille mesures de la poitrine d'une **statue**
- Mille mesures de la poitrine d'une **personne**
- Mesures de la poitrine de **mille copies d'une statue**
- Mesures de la poitrine de **mille personnes différentes**



- **Homme moyen**



Ben Dessy



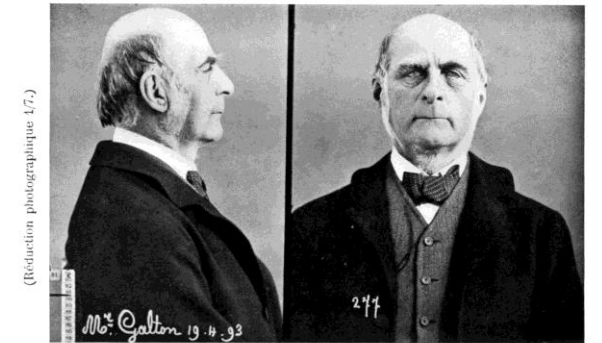
# Biométrie et eugénisme

- Darwin et l'évolution
- Francis Galton (1822-1911)

Echelons d'aptitude naturelle		Nombres d'hommes	
Sous la moyenne	Au-dessus de la moyenne	Par tranche (un million de même âge)	Proportion (un sur...)
		a	A
b	B	161 279	6
c	C	63 563	16
d	D	15 696	64
e	E	2 423	413
f	F	233	4 300
g	G	14	79 000
x	X	1	1 000 000

- Les éminents et les imbéciles
- **corrélation** et **régression**

Taille 1<sup>m</sup> \_\_\_\_\_ Long. \_\_\_\_\_ Pied g. \_\_\_\_\_ N° de cl. \_\_\_\_\_ Age de \_\_\_\_\_  
 Voite \_\_\_\_\_ Larg. \_\_\_\_\_ Médus g. \_\_\_\_\_ Au<sup>te</sup> \_\_\_\_\_ né le \_\_\_\_\_  
 Enverg 1<sup>m</sup> \_\_\_\_\_ Long. \_\_\_\_\_ Auric<sup>le</sup> g. \_\_\_\_\_ Cou<sup>l</sup> de l'iris g. \_\_\_\_\_ Par<sup>te</sup> \_\_\_\_\_ a \_\_\_\_\_  
 Buste 0. \_\_\_\_\_ Larg. \_\_\_\_\_ Cou<sup>l</sup>ée g. \_\_\_\_\_ Cou<sup>l</sup> de l'iris g. \_\_\_\_\_ Par<sup>te</sup> \_\_\_\_\_ dep<sup>l</sup> \_\_\_\_\_  
 \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ Age app. \_\_\_\_\_

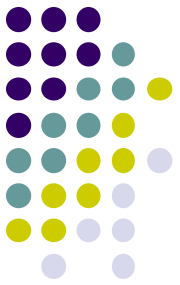


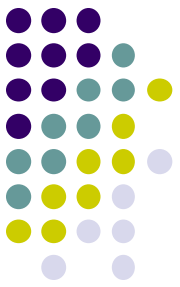
Front. \_\_\_\_\_ Racine (cavité) \_\_\_\_\_ Dent e. s. p. f. \_\_\_\_\_ Borbe \_\_\_\_\_ (pig<sup>ment</sup>)  
 Haut<sup>eur</sup> \_\_\_\_\_ Dos \_\_\_\_\_ Base \_\_\_\_\_ Lob. e. s. m. d. \_\_\_\_\_ Cheveux \_\_\_\_\_ (saig<sup>ner</sup>)  
 Larg<sup>eur</sup> \_\_\_\_\_ Haut<sup>eur</sup> \_\_\_\_\_ Saillie \_\_\_\_\_ Larg<sup>eur</sup> \_\_\_\_\_ A. trig. l. p. r. d. \_\_\_\_\_ Car. \_\_\_\_\_ Cou<sup>l</sup> \_\_\_\_\_  
 Par<sup>te</sup> \_\_\_\_\_ Par<sup>te</sup> \_\_\_\_\_ Dent<sup>ure</sup> droite \_\_\_\_\_ P. f. s. h. E. \_\_\_\_\_ Autres traits caractéristiques :  
 \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ \_\_\_\_\_ Sig<sup>net</sup> dressé par M. \_\_\_\_\_



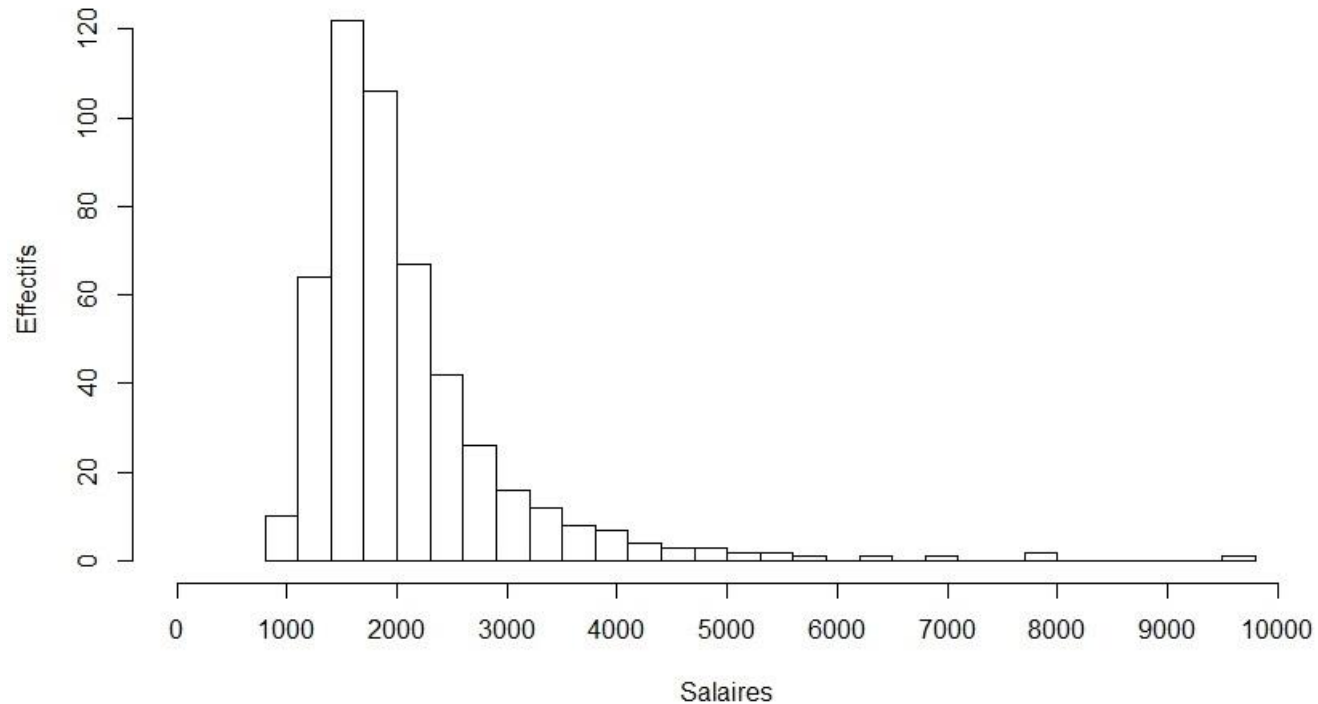
# Au 20<sup>e</sup> siècle

- **Moyennisme**
- **Standardisation**
- **Norme**



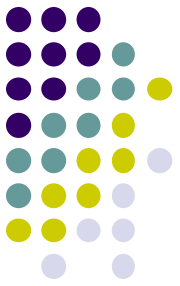


# Moyenne ou médiane?



***Moyenne*** = 2105€; ***médiane*** = 1832€

# Masse de données individuelles



Merci pour votre attention

## Conclusions

**ARMATTE M. (1995), *Histoire du modèle linéaire. Formes et usages en statistique et en économétrie jusqu'en 1945*, Thèse de doctorat, Paris, EHESS.**

**DROESBEKE J.-J. et TASSI Ph. (2015), *Histoire de la statistique, Que-sais-je?*, 2<sup>e</sup> édition rééditée, Paris, Presses Universitaires de France.**

**DROESBEKE J.-J. et VERMANDELE C. (2016), *Les nombres au quotidien. Leur histoire, leurs usages*, Collection *La statistique autrement*, Paris, Technip.**

**DROESBEKE J.-J. et VERMANDELE C. (2018), *Histoire(s) de(s) données numériques*, Collection *Le monde des données*, Paris, EDP Sciences.**

