

# Psy1004 – Section 1:



## La statistique et les statistiques

---

### Plan du cours:

- 0.0: Bienvenue
  - 0.1: Les catégories du savoir
  - 0.2: Survol de la psychologie
  - 0.3: Le plan de cours
  - 0.4: Les assistants
- 1.0: La physique: science exacte?
  - 1.1: Science et hasard
  - 1.2: Statistiques descriptives vs. inductives
  - 1.3: Échantillonnage
    - Échelles
    - Histogrammes

Disponible sur: <http://mapageweb.umontreal.ca/cousined/home/course/PSY1004>



## 0.1: Les catégories du savoir

---

Trois catégories:

- Le savoir:

- Les connaissances théoriques

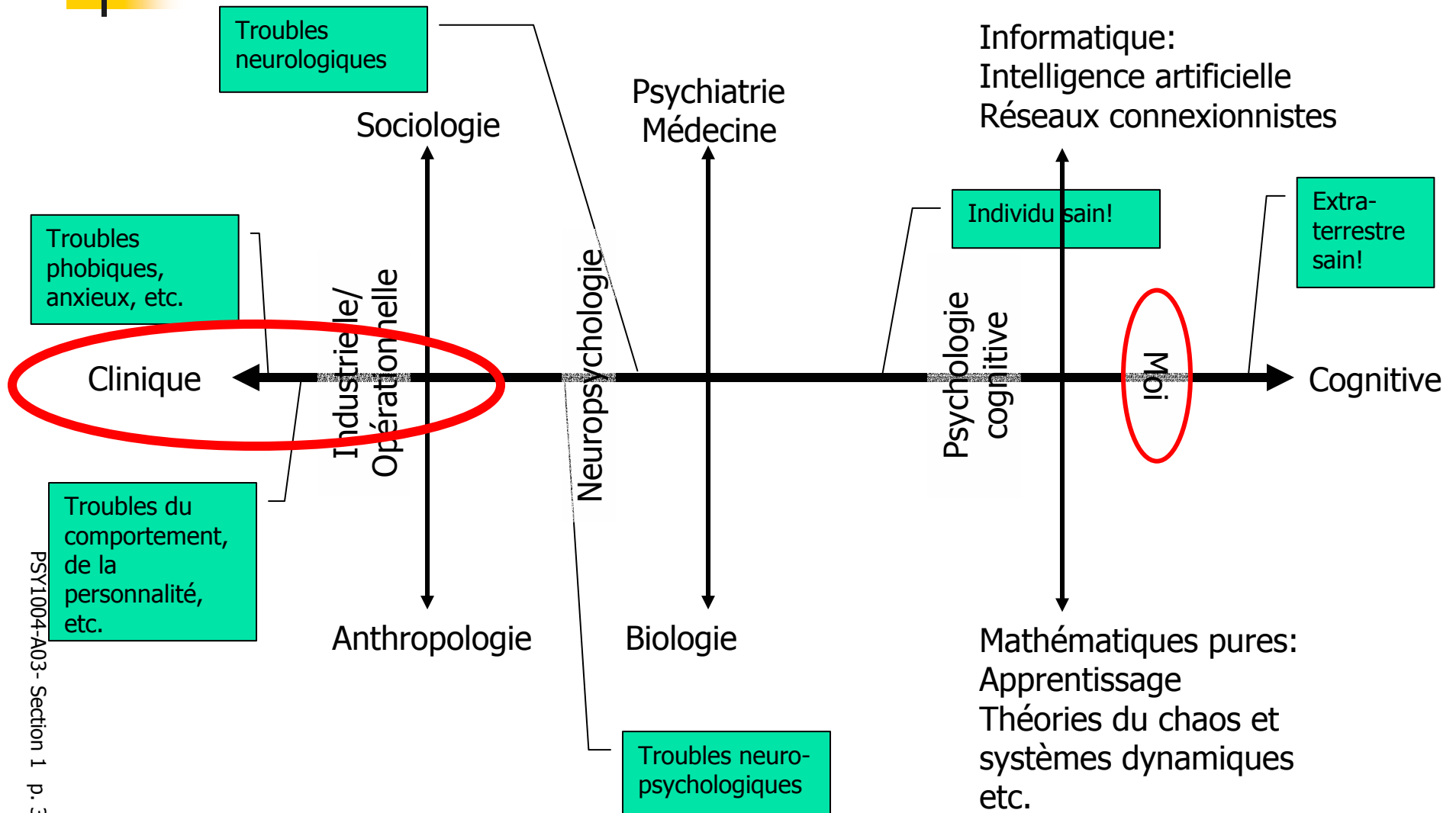
- Le savoir-faire:

- La mise en pratique, les méthodes concrètes, entre autre:
  - la méthode expérimentale
  - les méthodes quantitatives

- Le savoir-être:

- L'attitude face à la profession, au client: les règles d'éthiques.

## 0.2: Survol de la psychologie





## 0.3: Le plan de cours

---

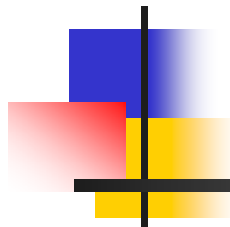
- Contenu du cours
- Formation en laboratoire sur SPSS:
  - Logiciel disponible aux laboratoires d'informatiques: A332, voir <http://www.fas.umontreal.ca/micros/mvlabo/labomv.html>
- Outils nécessaires:
  - Site web: <http://mapageweb.umontreal.ca/cousined/home/course/PSY1004>
  - Calculatrice avec fonctions statistiques recommandée
- Livre suggéré:
  - Howell, Méthodes statistiques en sciences humaines.
- Logiciel à utiliser:
  - SPSS (statistical package for the social sciences)



## 0.4: Les assistants

---

- Évaluations:
  - 4 travaux pratiques, seul ou à deux
  - 2 examens à choix multiples.
- Deux assistants:
  - Nicolas Dupuis-Roy et Jade Girard
  - Disponibilités: Jeudi de 13h00 à 16h00 au D-478, pavillon Marie-Victorin.
  - Il y aura une formation sur SPSS dans le labo à une date à préciser.
  - Ils sont là pour vous: Abusez!
- Le cours est difficile.



# Bloc 1

---

Disponible sur: <http://mapageweb.umontreal.ca/cousined/home/course/PSY1004>



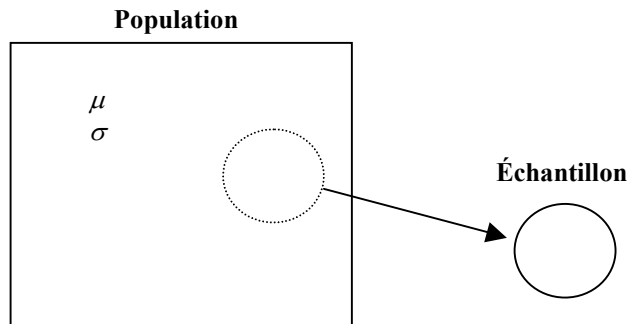
# 1.0 Survol de la démarche

---

- Les statistiques ne sont pas une fin en soi, mais un outil qui s'inscrit dans une démarche globale:
  1. Obtenir des données (« échantillonnage »)
    - Via une expérience, un sondage, etc. Cf. Méthode expérimentale
  2. Explorer les données
    - Les données semblent-elles correctes? Pas d'erreur de transcription, d'erreur dans la méthode expérimentale?
  3. Décrire les résultats
    - Réduire des milliers de chiffres à quelques « statistiques », i.e. des descripteurs qui condensent les résultats.
  4. Réaliser des tests statistiques et conclure
    - Indiquer si les différences qu'on voit à l'œil sont réelles.

# 1.1: Les données en psychologie

- Une mesure parfaite n'existe pas: Variabilité entre les individus (inter-sujet), entre les moments de mesures (intra-sujet), et liée à l'instrument de mesures (arrondissement) → Bruit dans les données.
- Pour réduire le bruit: mesurer plusieurs fois, i.e. échantillonner la population.



Échantillon inclus dans Population.

Il doit être représentatif:

- Face à la population (validité externe)
- Face à l'expérience (validité interne)  
ceci grâce à la randomisation.

et quantifié:

Variable aléatoire si mesurée (V.D.)

Facteur si imposé par le chercheur (V. I.).





## 1.1: Les données en psychologie

---

Science: Découverte de régularités (lois) et schéma d'explication.

En physique, les chercheurs n'ont pas ce problème:

- "Une expérience est réussie si on élimine moins de 50% des mesures"! → La physique est une science facile, contrairement à la psychologie.

En psychologie:

- un grand nombre de facteurs rendent les mesures "bruyantes".
- Pour compenser, il faut effectuer plus de mesures:



Détails



## 1.2: Explorer les données (1/3)

---

Pour explorer les données:

1. Faire un graphique des moyennes selon vos différents groupes.
  1. Les moyennes font-elles du sens?
2. **Vérifier le graphique des fréquences**
  1. Y a-t-il des données aberrantes?
  2. Avez-vous mesuré deux populations sans le savoir (multimodalité).
3. Apporter des correctifs aux données si nécessaires, puis reprendre à 1.
4. L'expérience s'est-elle bien déroulée?
  1. Si oui, passé aux étapes suivantes: décrire les résultats et faire les tests requis.



## 1.2: Explorer les données (2/3)

---

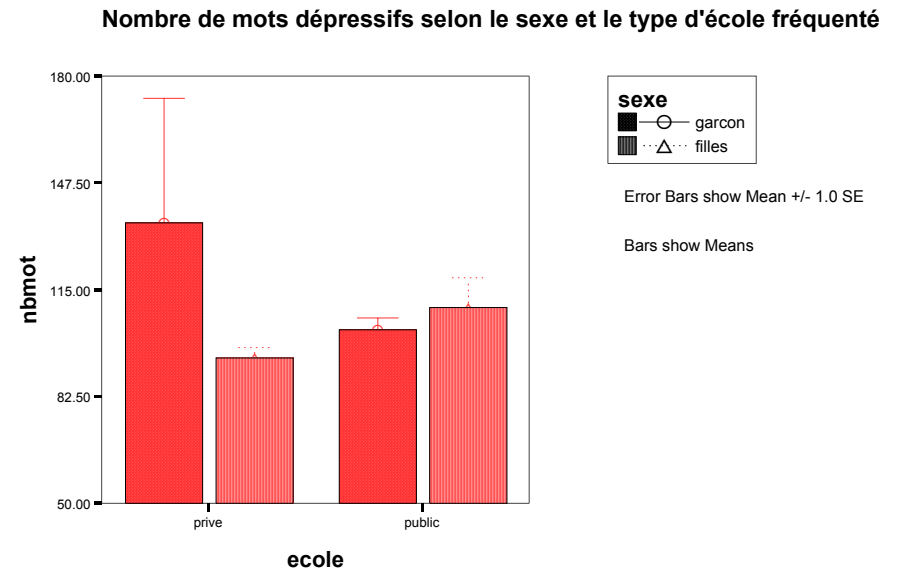
### Exemple

- Le nombre de mot à consonance dépressive exprimés par les garçons et les filles fréquentant le public ou le privé.
- 4 groupes indépendants de 20 étudiant-e-s sélectionnés au hasard parmi 5 écoles équivalentes en terme de revenu des parents.
- Les mots à consonance dépressive sont mesurés lors d'une entrevue portant sur le bien-être d'une durée d'une heure.



# Comment faire un graphe des moyennes?

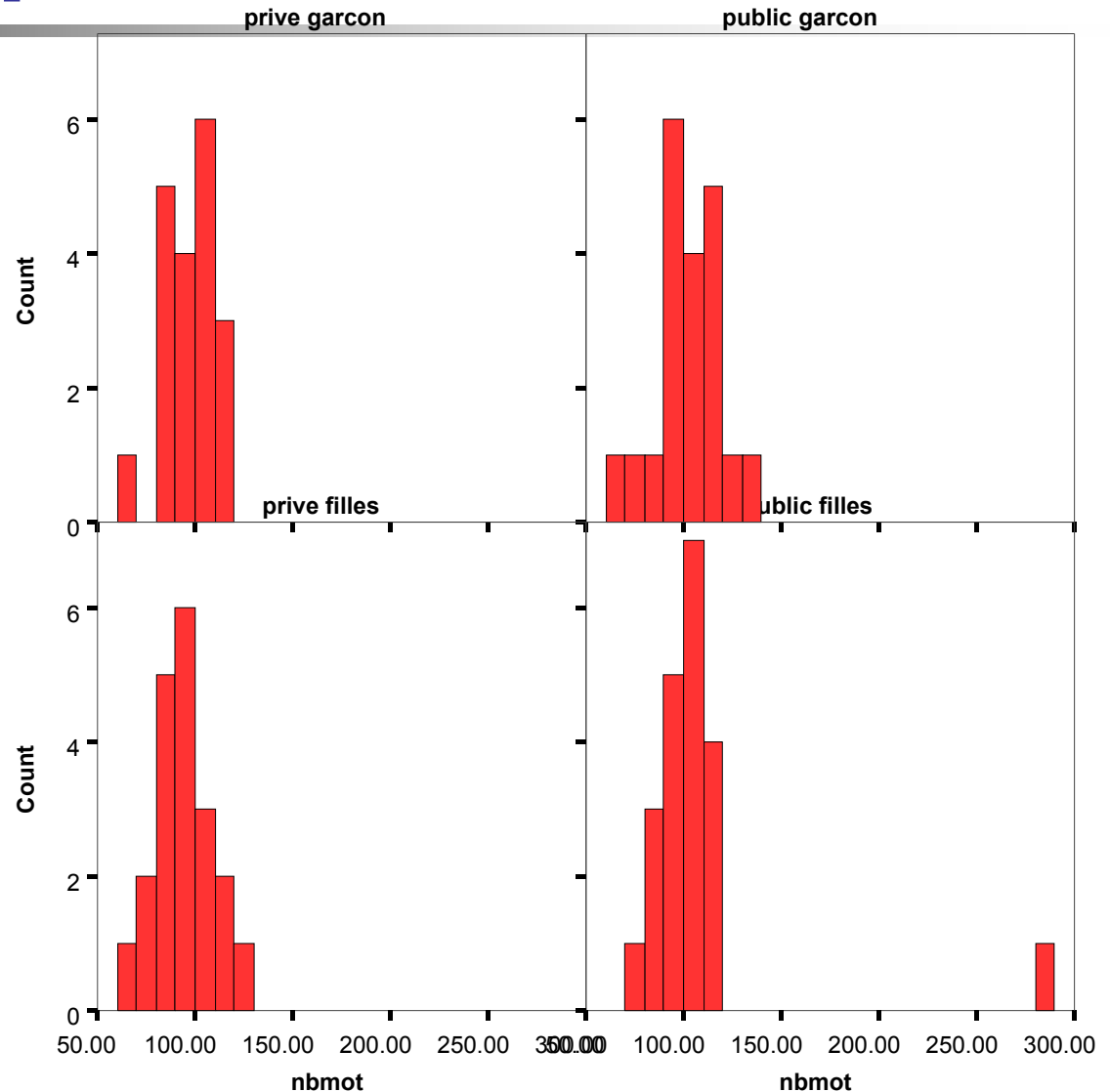
- À la main
- Avec Excel
- Avec SPSS
  - Recommandé car a) pas besoin de calculer les moyennes au préalables; b) très rapide → Démonstration.
- Autre...



# Comment faire un graphe des fréquences?

Répartition des scores selon le sexe et le type d'école

- À la main
- Avec Excel?
- Avec SPSS
  - Recommandé car très rapide → Démonstration.
- Autre...

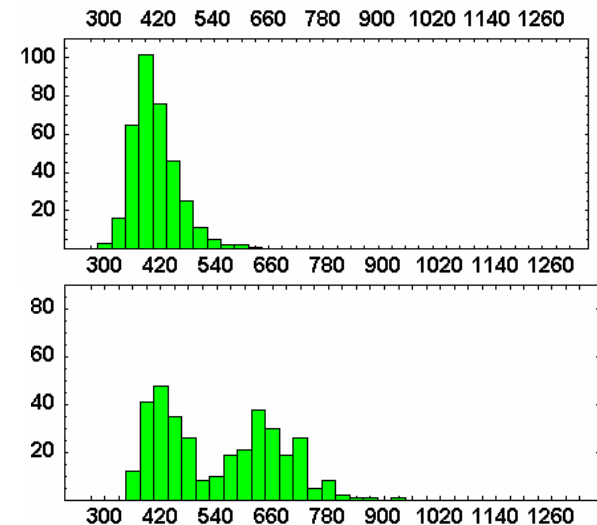


## 1.2: Explorer les données (fin)

Deux choses à vérifier dans les données, visible grâce au graphique des fréquences:

- Données extrêmes/aberrantes (« outliers »).  
Voir exemple précédent.

- Multimodalité (présence de plusieurs modes). Voir exemple ci-contre → obtenues lors d'une tâche de "recherche visuelle" du type "Où est Charlie?"  
La multimodalité suggère que deux sous-populations ont été mélangées





## 1.3: Statistiques descriptives vs. inductives

---

Exemple:

- La taille des étudiants de psychologie montréalais au cours de la période 1990-2000 est de...  $\{175, 168, 184, \dots\}$ , où  $n \approx \infty$ .
- Un échantillon représentatif est  $\{165, 142, \dots\}$  où  $n =$  valeur précise.
- Une description serait: "la taille moyenne est de  $168 \pm 10$  cm".
- Statistiques descriptives: Quantités pour synthétiser de façon **utile** un échantillon.
- Une statistique: Une valeur numérique qui synthétise un aspect de l'échantillon.
- Statistiques déductives (inférence statistiques): Méthodes pour inférer les propriétés de la population entière étant donné une statistique tirée d'un échantillon particulier (i.e. généralisation).



## 1.4: Échantillonnage: Notation

---

Soit  $\mathbf{X}_i$  une mesure (une observation) obtenue d'un individu placé dans une certaine condition.

$\mathbf{X}$  est un ensemble de mesure, constituant l'échantillon:

$$\mathbf{X} = \{\mathbf{X}_i\} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\},$$

où  $n$  est le nombre de mesures.

On note:

$$\sum_{i=1}^n \mathbf{X}_i = \mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_n$$

$$\sum_{i=1}^n \mathbf{X}_i^2 = \mathbf{X}_1^2 + \mathbf{X}_2^2 + \dots + \mathbf{X}_n^2$$





# Varia

Disponible sur: <http://mapageweb.umontreal.ca/cousined/home/course/PSY1004>

## 1.4: Notation (exemple)

- Soit les données pour les deux échantillons **X** et **Y**:

<b>X</b>	<b>Y</b>	<b>X<sup>2</sup></b>	<b>X×Y</b>
8	1		
15	4		
9	2		
16	5		
11	4		
10	2		
8	3		

Calculez la valeur des formules:

$$\sum_{i=1}^7 X_i$$

$$\sum_{i=1}^7 Y_i$$

$$\sum_{i=1}^7 X_i^2 \stackrel{?}{=} \sum_{i=1}^7 X_i \times \sum_{i=1}^7 X_i$$

$$\sum_{i=1}^7 (X_i + Y_i) \stackrel{?}{=} \sum_{i=1}^7 X_i + \sum_{i=1}^7 Y_i$$

$$\sum_{i=1}^7 (X_i \times Y_i) \stackrel{?}{=} \sum_{i=1}^7 X_i \times \sum_{i=1}^7 Y_i$$

Et vérifiez si l'égalité est vraie. Pouvez-vous dire pourquoi?



## 1.4: Notation (fin)

$$\begin{aligned}\sum_{i=1}^n (\mathbf{X}_i + \mathbf{Y}_i) &= \mathbf{X}_1 + \mathbf{Y}_1 + \mathbf{X}_2 + \mathbf{Y}_2 + \dots + \mathbf{X}_n + \mathbf{Y}_n \\ &= \mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_n + \mathbf{Y}_1 + \mathbf{Y}_2 + \dots + \mathbf{Y}_n \\ &= \sum_{i=1}^n \mathbf{X}_i + \sum_{i=1}^n \mathbf{Y}_i\end{aligned}$$

$$\sum_{i=1}^n c = \underbrace{c + c + \dots + c}_{n \text{ fois}} = nc$$

$$\begin{aligned}\sum_{i=1}^n c\mathbf{X}_i &= c\mathbf{X}_1 + c\mathbf{X}_2 + \dots + c\mathbf{X}_n \\ &= c(\mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_n) \\ &= c \sum_{i=1}^n \mathbf{X}_i\end{aligned}$$



## 1.5: Bref survol de SPSS

---

- Un survol expliqué aura lieu la semaine prochaine.
- Les assistants font aussi faire une démonstration dans le laboratoire (date à venir)
- Idée aujourd'hui: comprendre l'approche utilisée par SPSS.