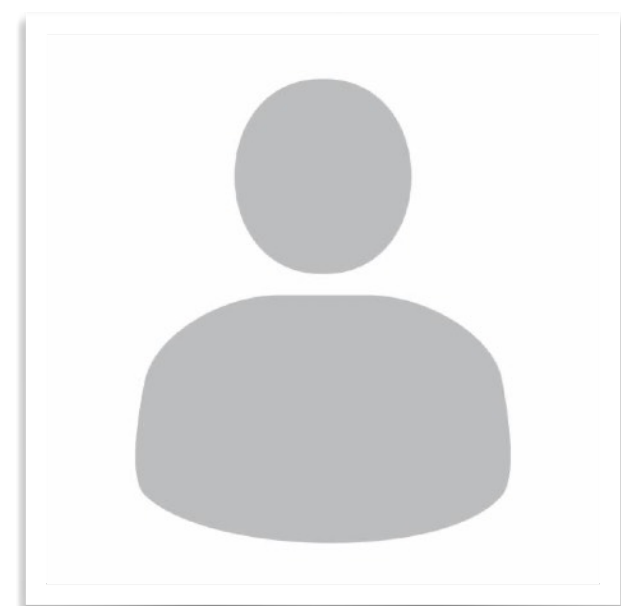


Science des données PHYS-231

Prof: Lenka Zdeborová, laboratoire SPOC, lenka.zdeborova@epfl.ch.



TAs: Cédric Koller, Odilon Duranthon, Yizhou Xu;

AEs: Julius Girardin, Raphael Faltin, Adam Chahed Ouazzani, Jules Roux, Margarita Sagitova



cedric.koller@epfl.ch, odilon.duranthon@epfl.ch, yizhou.xu@epfl.ch, julius.girardin@epfl.ch,
raphael.faltin@epfl.ch, adam.chahedouazzani@epfl.ch, jules.roux@epfl.ch, margarita.sagitova@epfl.ch

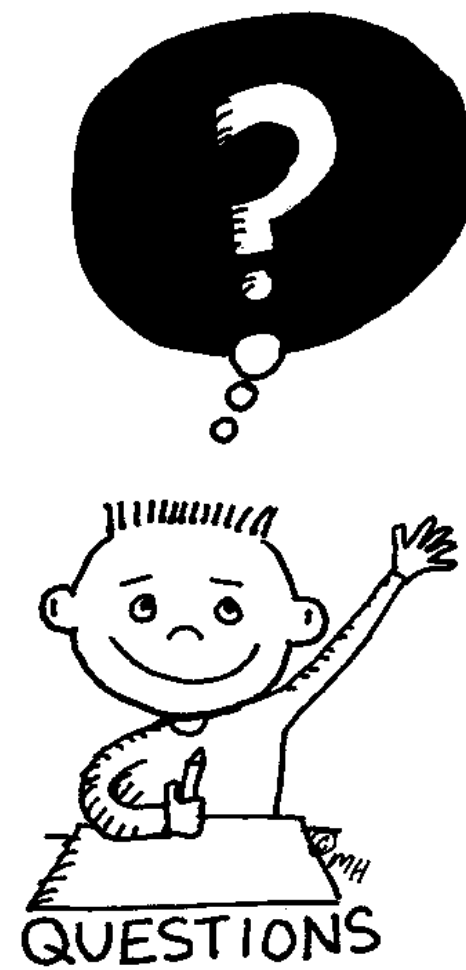
Organisation

- **Cours** (Ma 8h15-10h00, CE13).
- **Exercices** (Je 10h15-12h00, CM2, avec vos propres ordinateurs portables). Aussi important que le cours.
- **Les cours de 2023 & 2025 enregistrés** sur **Mediaspace**, pas diffusés en streaming. Lien: <https://mediaspace.epfl.ch/channel/PHYS-231%2BScience%2Bdes%2Bdonn%C3%A9es/54036>
- **Les exercices ne seront pas enregistrés**. Les solutions seront publiées après la session sur Moodle.
- Toutes les infos, et les notes de cours sur **Moodle** <https://moodle.epfl.ch/course/view.php?id=18292>
- Q&A pendant les exercices. Q&A à d'autres moments : ok par e-mail aux TA, préférable sur Moodle.
- Langue: Cours en français, Exercice: Mix de français et anglais (Python est en anglais). Question en français ou anglais comme vous préférez.

Comment serez-vous évalués ?

- 50%: 2 ou 3 exercices notés au cours du semestre (commencés en séance, à terminer à la maison pour la semaine d'après). Les dates seront données 2 semaines à l'avance. Principalement du codage, dans un notebook Jupyter. Résolu individuellement. Tout le matériel utilisé (codes, livres, articles, chatGPT) doit être dûment cité ainsi que les noms de toutes les personnes avec lesquelles vous avez discuté. A télécharger sur Moodle.
- 50%: Examen final 3h : Questions sur des concepts théorique ou concernant des morceaux de code tirés des exercices/devoirs, quelques calculs. Pas d'ordinateur à l'examen final. Une page A4 de notes personnelles (soit manuscrit, soit avec une police d'au moins 10 points).

Des questions sur l'organisation ?



De quoi parle ce cours ?

Partie I. Semaine 1 - 5. **Introduction à l'analyse de données.** Outils numériques pour l'algèbre linéaire. Analyse en composantes principales et ses applications. Méthodes et algorithmes d'analyse de données de base, régression linéaire et réseaux de neurones.

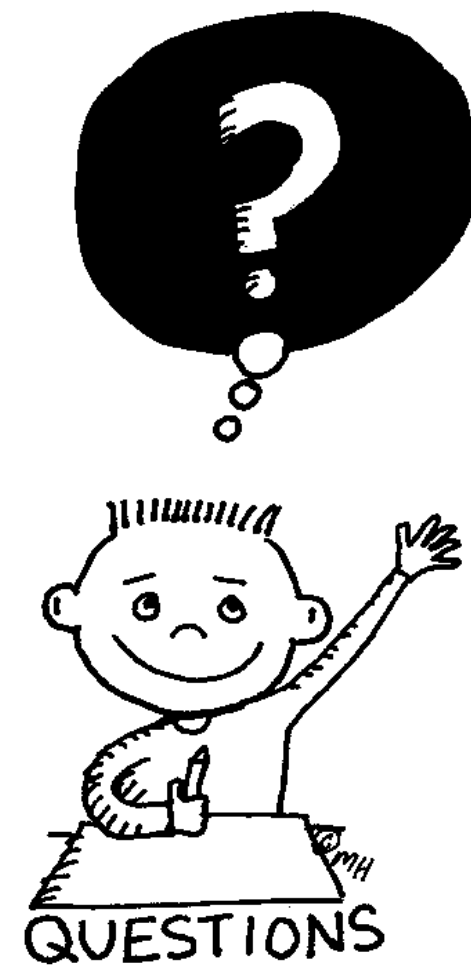
Partie II. Semaine 6 - 10. **Quantification de l'incertitude.** Propagation des erreurs. Estimation et inférence statistiques et leurs applications en physique expérimentale.

Partie III. Semaine 11 - 14. **La stochasticité dans les systèmes et modèles physiques.** Marches aléatoires et mouvement brownien. Échantillonnage de distributions de probabilité. Applications de la loi des grands nombres et du théorème central limite en physique.

Acquis de formation

- [Programmation en Python](#). Apprendre à apprendre un nouveau langage de programmation (plus facile que c++).
- Comment extraire la science [d'équations pouvant être résolues numériquement](#). Analyser les données grâce à l'algèbre linéaire et le calcul multivariable.
- Extraire des informations à partir de grands ensembles de données. Préparation à [l'apprentissage automatique](#) (machine learning).
- Fondamentaux pour comprendre comment estimer [les barres d'erreur dans les expériences de physique](#).
- Introduction à [la stochasticité en physique](#). Obtenir des informations sur le comportement de grandes ensembles d'éléments aléatoires. Préparation à la physique statistique.

Des questions sur le contenu ?



EPFL

Project Introduction



Zhenyu Cai

3rd-year PhD in CHILI Lab

zhenyu.cai@epfl.ch

My PhD project is to design and evaluate **dashboards** to support teaching and learning in **exercise sessions**

<https://youtu.be/6Eq5GUxwv2c?feature=shared>