

# ICT for Smart Grids & Smart Cities

## Data Science

---

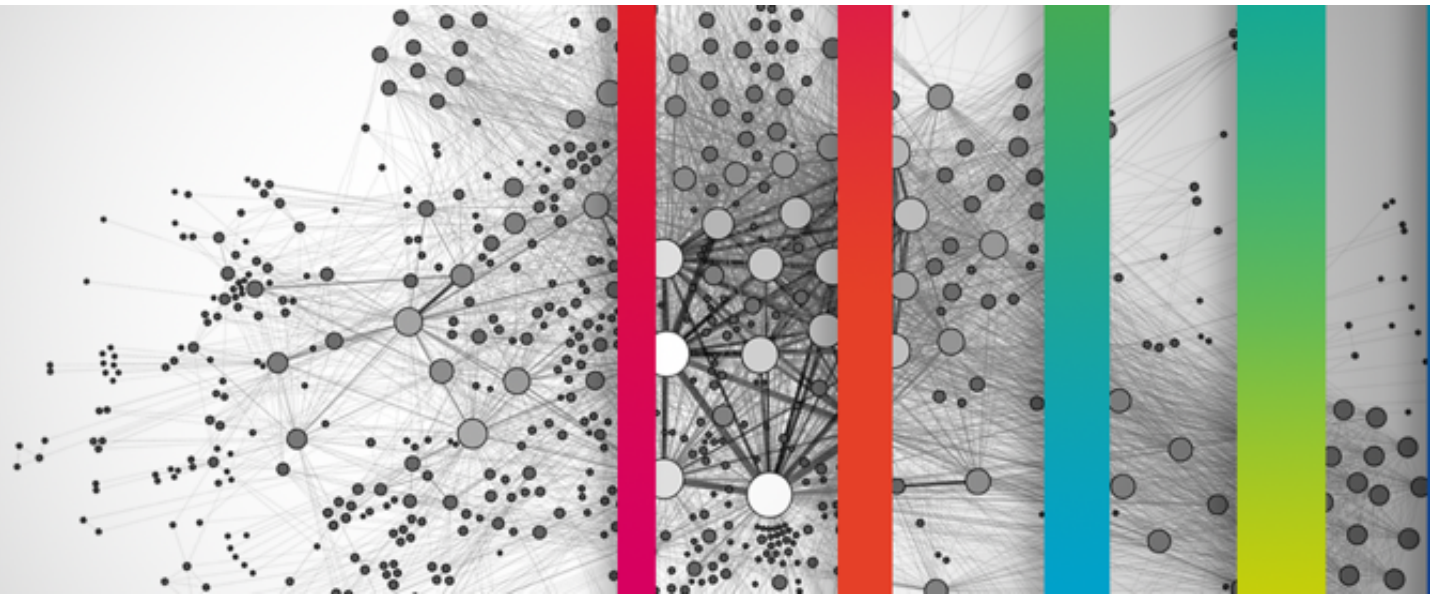
**Geneveva Vargas-Solar**  
**French Council of Scientific Research, LIG**  
**[geneveva.vargas@imag.fr](mailto:geneveva.vargas@imag.fr)**

<http://vargas-solar.com/data-centric-smart-everything/>

<https://classroom.google.com/c/MTQ4MzcwMjY1MDEz?cjc=5bz2tk6>

Slack channel: [https://join.slack.com/t/colenationale-5jr8199/shared\\_invite/zt-hhf9euv7-bmp7Kn9LL68RyzdhJnbKxA](https://join.slack.com/t/colenationale-5jr8199/shared_invite/zt-hhf9euv7-bmp7Kn9LL68RyzdhJnbKxA)





# THIS LECTURE

# OBJECTIVE & LEARNING OUTCOMES

## **GENERAL OBJECTIVE**

- Describe the principles of the new emerging data science
- Teach how to apply its method and techniques for analysing data collections through experiments running on top of Big Data analytics environments willing to analyse, model and predict the behaviour of complex systems.

## **LEARNING OUTCOMES**

Design and implement data science experiments devoted to the analysis, modelling and predicting the behaviour of complex systems.

- Understand theoretically and technically the steps of a general data science process.
- Apply tools for collecting and preparing data collections from datasets produced under different conditions.
- Understand and know how to use existing data analytics techniques: aggregation, consolidation, statistics, machine learning methods.
- Make decisions on the data analytics techniques to apply according to the data properties & the analytics objective.

# CONTENT

## **Introduction**

- Data centric sciences: Principles and common aspects
- Digital data collections: Characteristics and properties
- Data science: Big data, data analytics algorithms & tools

## **Designing experiments**

- Data collections life cycle
- Exploring and preparing data collections for building corpora
- Architectural settings: From in house to large scale experiments

## **Descriptive statistics**

- Preparing data sets
- Explanatory data analysis
- Estimation

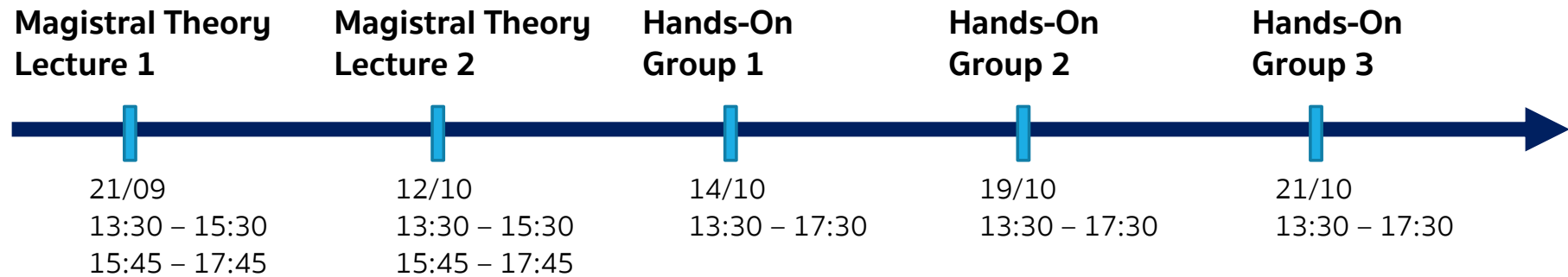
## **Statistical inference**

- The frequentist approach
- Measuring the variability in estimates
- Testing hypothesis

## **Unsupervised learning**

- What is learning?
- Clustering

# ORGANISATION & EVALUATION



**Hands on reports**

**20%** (according to the official planning)

**Final exam/Project**

**80%** (according to the official planning)

# CONTACT & ADVISING

Dr. Genoveva Vargas-Solar  
French Council of Scientific Research, LIG lab

## **ADVISING & QUESTIONS:**

Slack channel: [https://join.slack.com/t/colenationale-5jr8199/shared\\_invite/zt-hhf9euv7-bmp7Kn9LL68RyzdhJnbKxA](https://join.slack.com/t/colenationale-5jr8199/shared_invite/zt-hhf9euv7-bmp7Kn9LL68RyzdhJnbKxA)

<https://classroom.google.com/c/MTQ4MzcwMjY1MDEz?cjc=5bz2tk6>

<http://vargas-solar.com/data-centric-smart-everything/>



Genoveva [VARGAS-SOLAR.COM](https://www.vargas-solar.com)

Senior Scientist, CNRS, France