

Summary

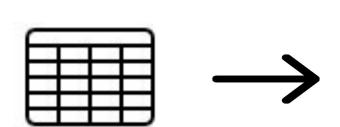
1. Objectives

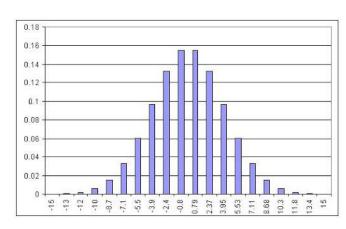
2. Organisation

1. ID Notebook

The main idea of this notebook is to automatically provide information about a dataset, i.e. by extracting meta-features.

Those information will represent the "ID card" of the dataset.

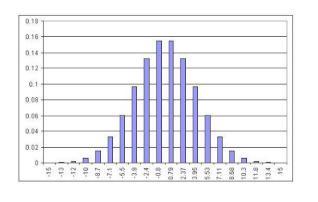


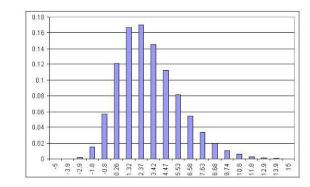


2. Comparison

Use the meta-features extracted in the notebook to compare two datasets. Useful to compare original and generated data.

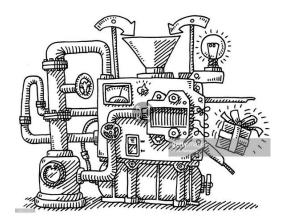
VS





3. Generative models

Generate synthetic data from real data in order to be able to release them without worrying about confidentiality, using for example medGAN.



4. Create challenges

Create challenges for students.

Generated data avoid confidentiality issues.

- Survival analysis?
- Risk factor analysis?

2. Organisation: Work Plan

	Start date	Teams	Expected end date
ID Notebook	12/03	Adrien & Thomas	26/03
Comparison	26/03	Adrien & Thomas	09/04
Generative models	09/04	Adrien & Thomas	31/06
Challenges	30/04	Ritik & Saloni supervised by Adrien & Thomas	31/06

Generative model benchmark

