

## Compétences

L'objectif de l'enseignement est de former les étudiants au **traitement et à l'analyse des données biologiques** acquises au laboratoire, sur le terrain ou en entreprise et à leur **représentation graphique**.

Il s'agira d'apprendre à exploiter au mieux les données biologiques dans le **contexte du *Big Data***,

Il sera axé sur les **statistiques descriptives**, l'**analyse factorielle** et la **régression linéaire**.

## Description

- Introduction sur le Big Data en Biologie
- La représentation graphique
- Gestion des erreurs
- Comment résumer une distribution expérimentale à l'aide de quelques paramètres (position, dispersion, forme)
- Réalisation de documents de synthèse de qualité professionnelle
- Initiation à l'analyse factorielle
- Le modèle linéaire : régression linéaire et analyse de variance

- ⇒ **Cours, Exercices et Travail Pratique à chaque séance**
- ⇒ **Salles informatiques du Bât 336**
- ⇒ **Acquisition d'une compétence très demandée**
- ⇒ **Un plus pour votre CV** (recherche / ingénierie / entreprise)
- ⇒ **L'accent est mis sur la pratique**
- ⇒ **Un site web** où vous retrouverez cours, exercices corrigés, auto-évaluations, forum, liaison avec l'enseignant  
**(vidéos des cours et de présentation de l'enseignement)**

Travail sur des jeux de données réels

Thématiques : Environnement, Epidémiologie, Génomique, Pollution, Biomédical

Travail organisé en groupes de 2 à 4 en salle

Utilisation de logiciels dédiés

Préparation des projets en classe

Apprentissage d'une démarche

Vous servira dans vos TP et vos stages ainsi que dans la suite de votre cursus

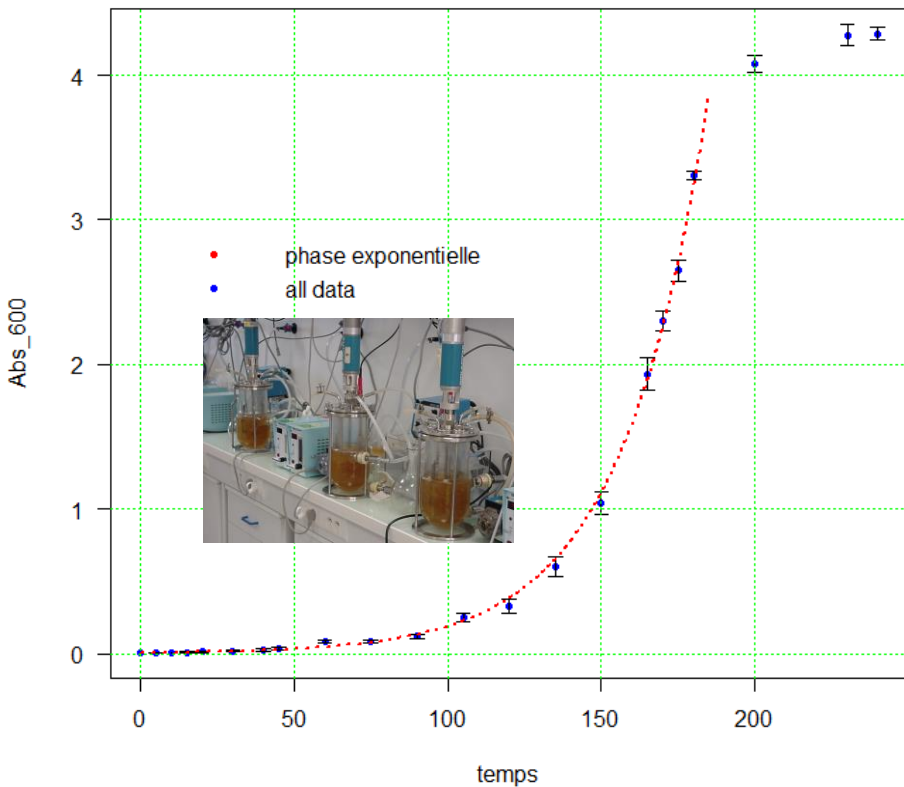
Les étudiants apprendront à raisonner avant d'appliquer le modèle statistique adapté à leur problématique.

Site web associé à l'enseignement

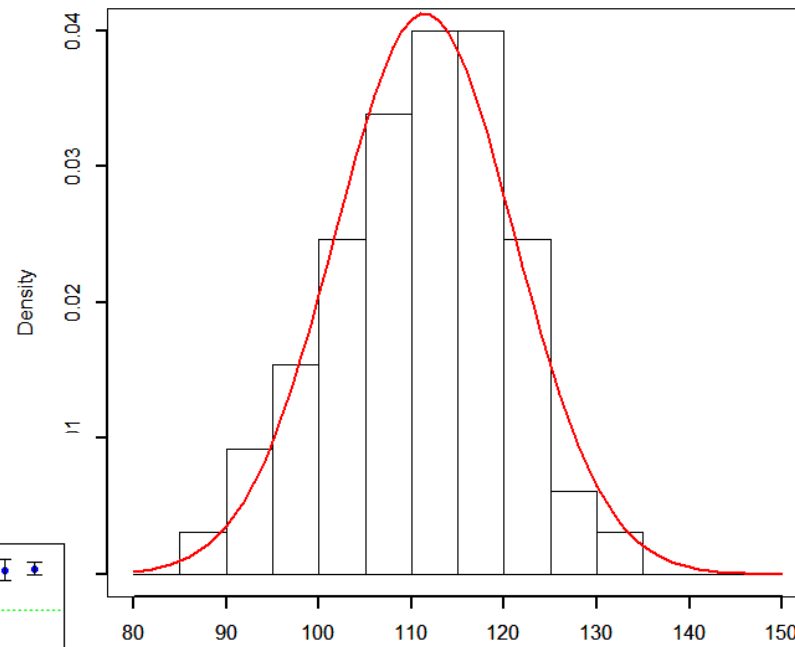
# La représentation graphique

- méthode
- gestion des barres d'erreurs
- apprentissage de logiciels

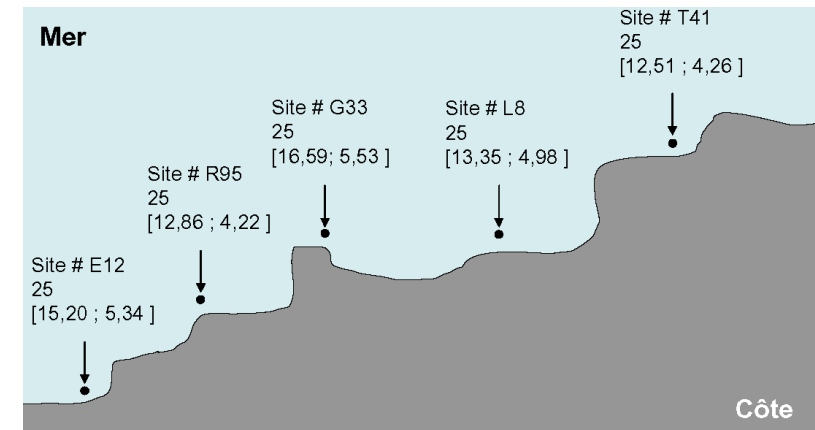
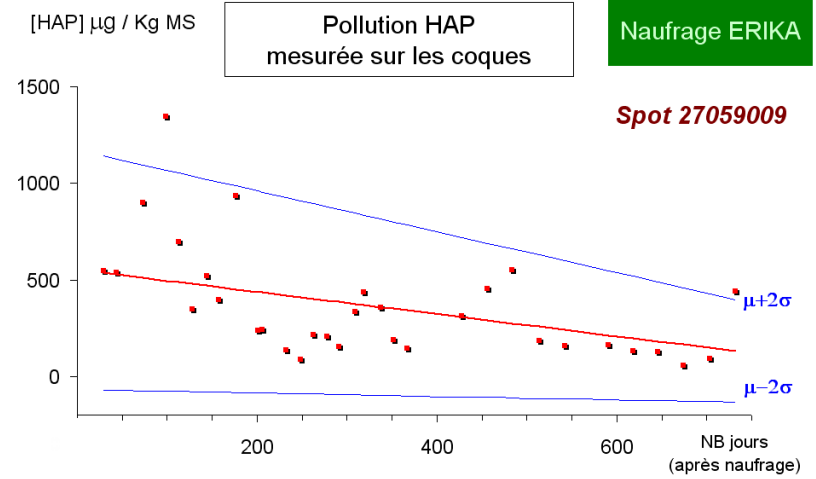
Croissance Bactérienne de la souche ORS2205



Nouvelle variété céréalière - Culture BIO



Col	1	2	3	4	5	6	7
Line	temps	DO1	DO2	DO3	DO4	DO5	DO6
1	0	0.005	0.005	0.003	0.005	0.005	0.005
2	5	0.004	0.008	0.006	0.007	0.005	0.006
3	10	0.008	0.007	0.007	0.006	0.006	0.007
4	15	0.007	0.005	0.005	0.005	0.013	0.012
5	20	0.005	0.013	0.010	0.014	0.010	0.016
6	30	0.012	0.022	0.016	0.021	0.018	0.017
7	40	0.024	0.036	0.031	0.018	0.018	0.012
8	45	0.042	0.021	0.028	0.038	0.038	0.033
9	60	0.082	0.080	0.067	0.090	0.091	0.073
10	75	0.086	0.084	0.067	0.096	0.099	0.080
11	90	0.097	0.132	0.139	0.112	0.118	0.121
12	105	0.258	0.217	0.251	0.214	0.254	0.303
13	120	0.371	0.327	0.273	0.276	0.326	0.380
14	135	0.583	0.666	0.580	0.624	0.499	0.657
15	150	1.000	0.920	1.104	1.072	1.135	1.014
16	165	1.917	1.892	1.825	2.137	1.894	1.937
17	170	2.410	2.267	2.250	2.245	2.286	2.361
18	175	2.599	2.748	2.585	2.662	2.593	2.708
19	180	3.331	3.275	3.305	3.334	3.316	3.267
20	200	3.998	4.085	4.075	4.047	4.123	4.147
21	230	4.360	4.194	4.320	4.318	4.268	4.206
22	240	4.291	4.241	4.365	4.261	4.272	4.284

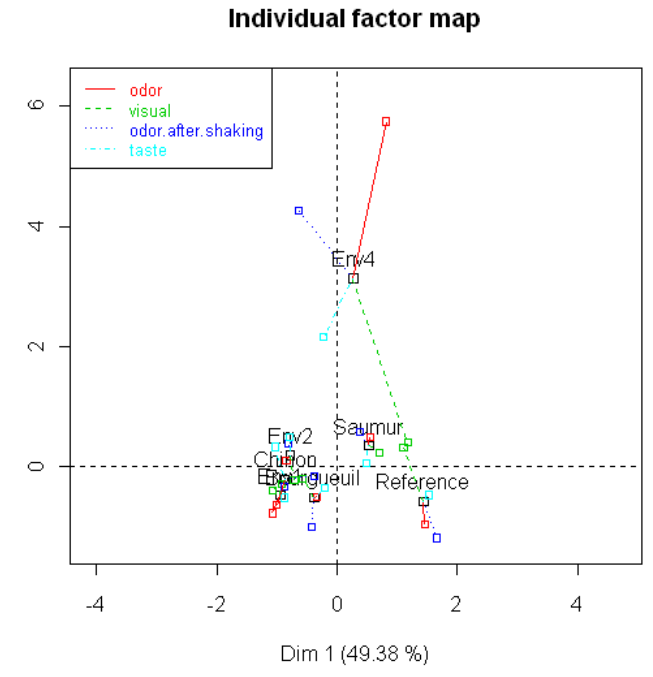
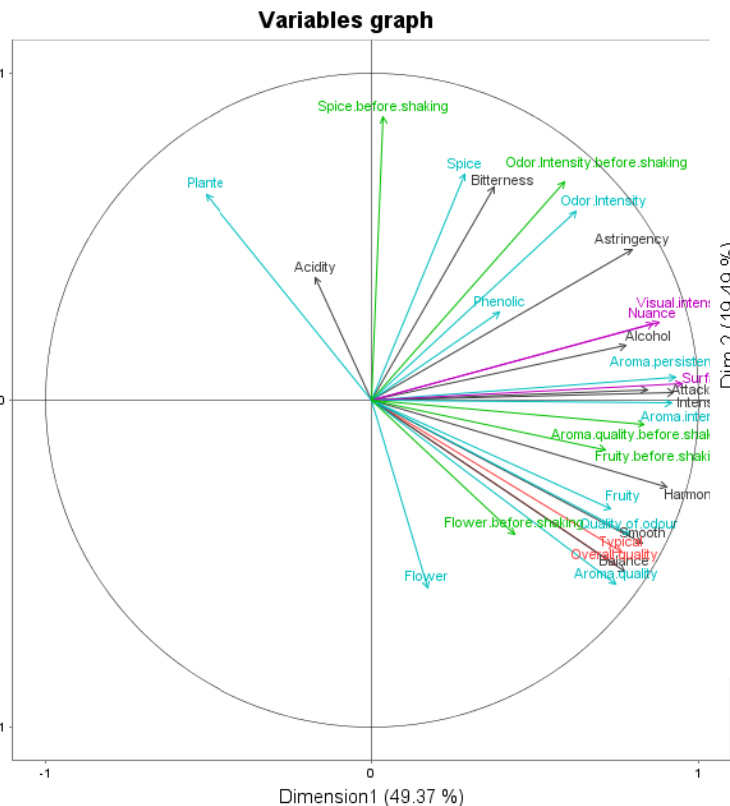
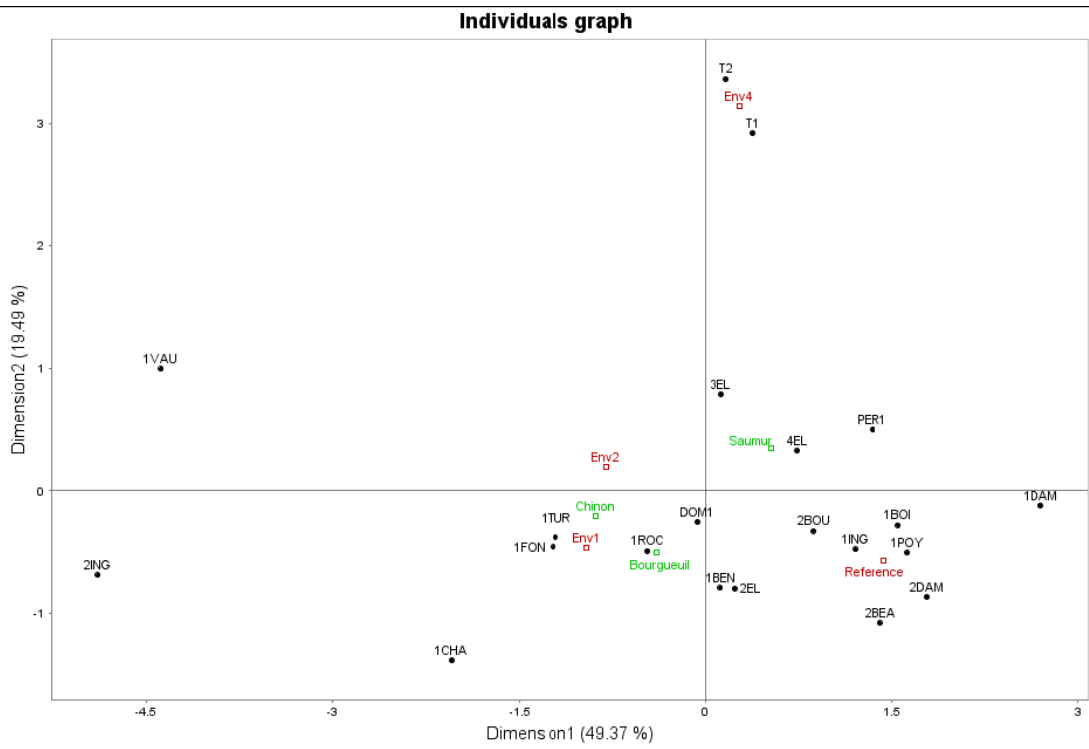


Activité enzymatique	
Individus Témoins	Individus exposés pendant 2 jours à une dose d'insecticide
27.0	26.0
28.8	19.0
26.0	27.1
34.5	19.6
28.8	22.0
23.0	15.5
32.2	24.2
25.8	17.0

## Analyse Factorielle Multiple : Evaluation sensorielle de vins rouges

Analyse gustative de vins (individus) à l'aide de descripteurs sensoriels (variables). Le tableau de données comporte 21 lignes (vins) et 31 colonnes. Les deux premières colonnes sont des variables qualitatives : le label (Saumur, Bourgueuil ou Chinon) et le type de sol (Référence, Env1, Env2 ou Env4). Les 29 colonnes suivantes sont des variables sensorielles continues. La moyenne sur les juges est donnée pour chaque vin. Le tableau de données est structuré en différents groupes de variables.

	Label	Soil	Odor.Intensity. before.shaking	Aroma.quality. before.shaking	...	Visual. intensity	Nuance	...	Odor. Intensity	Quality. of.odour	...	Attack. intensity	Acidity	...	Overall. quality	Typical
2EL	Saumur	Env1	3.074	3	...	4.321	4	...	3.407	3.308	...	2.963	2.107	...	3.393	3.25
1CHA	Saumur	Env1	2.964	2.821	...	3.222	3	...	3.37	3	...	3.036	2.107	...	3.214	3.036
1FON	Bourgueuil	Env1	2.857	2.929	...	3.536	3.393	...	3.25	2.929	...	3.222	2.179	...	3.536	3.179
1VAU	Chinon	Env2	2.808	2.593	...	2.893	2.786	...	3.16	2.88	...	2.704	3.179	...	2.464	2.25
1DAM	Saumur	Reference	3.607	3.429	...	4.393	4.036	...	3.536	3.36	...	3.464	2.571	...	3.741	3.444
2BOU	Bourgueuil	Reference	2.857	3.111	...	4.464	4.259	...	3.179	3.385	...	3.286	2.393	...	3.643	3.393
1BOI	Bourgueuil	Reference	3.214	3.222	...	4.143	3.929	...	3.429	3.5	...	3.393	2.607	...	3.714	3.357
3EL	Saumur	Env1	3.12	2.852	...	4.214	3.857	...	3.654	3.077	...	3.25	2.179	...	3.393	3.071
DOM1	Chinon	Env1	2.857	2.815	...	4.037	3.893	...	3.357	3.346	...	3.286	2.286	...	3.2	3.5
1TUR	Saumur	Env2	2.893	3	...	3.704	3.407	...	3.222	3.259	...	2.893	2.357	...	3.179	2.964
4EL	Saumur	Env2	3.25	3.286	...	3.857	3.643	...	3.607	3.385	...	3.321	2.429	...	3.571	3.5
PER1	Saumur	Env2	3.393	3.179	...	4.714	4.5	...	3.481	3.385	...	3.357	2.429	...	3.148	3.556
2DAM	Saumur	Reference	3.179	3.286	...	4.222	4.071	...	3.481	3.423	...	3.393	2.286	...	3.571	3.929
1POY	Saumur	Reference	3.071	3.107	...	4.714	4.536	...	3.357	3.444	...	3.519	2.111	...	3.929	3.481
1ING	Bourgueuil	Env1	3.107	3.143	...	4.071	3.893	...	3.357	3.37	...	3.185	2.286	...	3.643	3.296
1BEN	Bourgueuil	Reference	2.929	3.179	...	3.889	3.429	...	3.286	3.308	...	3.393	2.393	...	3.75	3.571
2BEA	Chinon	Reference	3.036	3.179	...	3.786	3.607	...	3.444	3.5	...	3.071	2.571	...	3.536	3.269
1ROC	Chinon	Env2	3.071	2.926	...	3.679	3.393	...	3.37	3.36	...	3.071	2.393	...	3.464	3.444
2ING	Bourgueuil	Env1	2.643	2.786	...	2.607	2.536	...	2.889	2.8	...	2.179	2.25	...	2.37	2.321
T1	Saumur	Env4	3.696	3.192	...	4.321	4	...	3.737	3.08	...	2.963	2.407	...	2.643	2.571
T2	Saumur	Env4	3.708	2.926	...	4.321	4.107	...	3.727	2.885	...	3.333	2.571	...	2.852	2.75



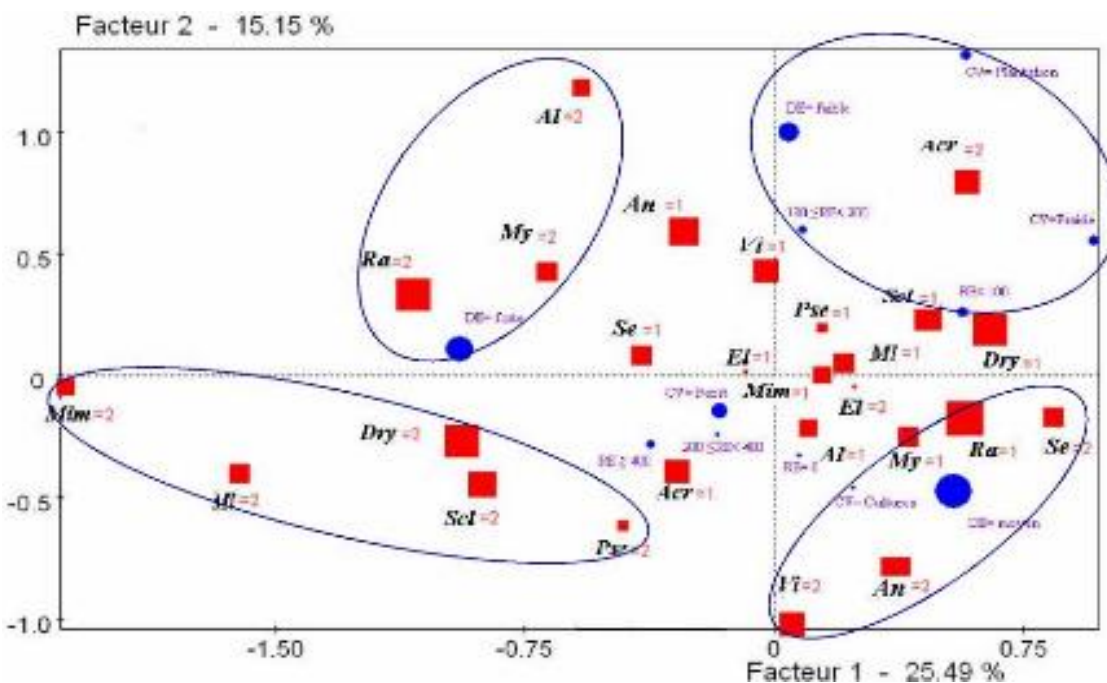
# Analyse Factorielle (AFC)

Facteurs influençant la **production piscicole des étangs** (Ouest-Cameroun)



Mise en place d'un étang d'inondation

HISTOGRAMME DES 13 PREMIERES VALEURS PROPRES			
	NUMERO	VALEUR	POURCENTAGE
PROPRE	CUMULE		
1	0.2549	25.49	25.49
2	0.1515	15.15	40.64
3	0.1087	10.87	51.50
4	0.0926	9.26	60.76
5	0.0798	7.98	68.74
6	0.0661	6.61	75.35
7	0.0521	5.21	80.57
8	0.0503	5.03	85.60
9	0.0430	4.30	89.90
10	0.0365	3.65	93.55
11	0.0280	2.80	96.35
12	0.0208	2.08	98.43
13	0.0157	1.57	100.00



> La **production** piscicole d'un étang **dépend** de l'**environnement végétal** des sites et des **conditions hydrologiques** du cours d'eau qui l'inonde

- réhabiliter d'autres étangs afin d'augmenter la surface exploitable ;
- organisation des pisciculteurs autour d'un groupe (partage d'expérience);
- protéger les couverts forestiers car ils assurent les meilleurs productions des étangs



## Contrôle des connaissances

- **Contrôle continu [1/3]**
- **Examen écrit [1/3]**
- **Mini Projet [1/3]** (thématique choisie par l'étudiant)

première session :  $NF = 0,4*EE + 0,3*CC + 0,3*CC\ TP$

deuxième session :  $NF = 0,5*EE + 0,25*CC(R) + 0,25*CC\ TP(R)$

**Le contrôle des connaissances portera sur un examen écrit qui mettra l'accent non sur des calculs mais sur des stratégies de traitement des données et sur l'analyse de résultats présentés sous forme de graphiques et de tableaux. Un projet individuel sera rendu que nous démarrerons en classe.**

**Vous ne verrez plus les stats comme une matière absconde mais bien comme une composante de l'analyse indispensable de vos données sur le terrain comme au laboratoire ou au bureau d'étude.**