

Module : Analyse statistique				Code	
				ING-4-SDIA-S7-P3	
Période	S7	Volume horaire	42	ECTS	4

<i>Responsable</i>	Ahmed Dhouibi	<i>email</i>	Ahmed.dhouibi@yahoo.fr
<i>Equipe pédagogique</i>	Ahmed Dhouibi		

1. Objectifs de Module (*Savoirs, aptitudes et compétences*)

Ce module porte sur les méthodes d'analyse statistique dans les cas uni varié et bi varié.

Acquis d'apprentissage :

A la fin de cet enseignement, l'élève sera capable de :

- Maîtriser les notions de base d'une série statistique. (**C1.2**)
- Caractériser les spécificités d'une série de données. (**C1.2**)
- Simuler et tester des hypothèses émises sur un jeu de données. (**C1.3**)
- Concevoir la méthode adéquate d'analyser une série donnée. (**C1.1**)
- Communiquer les principaux résultats et les valider (**C3.3**)
- Visualiser graphiquement une série de données et titrer les principaux résultats à partir de cette visualisation (**C2.3**)

2. Pré-requis (*autres UE et compétences indispensables pour suivre l'UE concernée*)

- Mathématiques de l'ingénieur
- Théorie de la probabilité

3. Répartition d'Horaire de Module

<i>Intitulé de l'élément d'enseignement</i>	<i>Total</i>	<i>Cours</i>	<i>TD</i>	<i>Atelier</i>	<i>PR</i>
Module : Analyse statistique	42	24	12	6	

4. Méthodes pédagogiques et moyens spécifiques au Module

(*pédagogie d'enseignement, ouvrages de références, outils matériels et logiciels*)

- Supports de Cours
- Projecteur et Tableau
- Travaux dirigés
- Logiciels: langage R

Bibliographie

<i>Titre</i>	<i>Auteur(s)</i>	<i>Edition</i>

Statistique avec Excel	Argentine VIDAL	2015
Statistique Inférentielle Avancée	Olivier Gaudoin	2012
Statistics controversies and concepts	David Moore and Co	2019
Introduction to the Introduction to the Practice of Statistics Practice of Statistics	David Moore and co	2014

5. Contenu (<i>Descriptifs et plans des cours / Déroulement / Détail de l'évaluation de l'activité pratique</i>)	Durée allouée	
Module 1 : Analyse statistique.		
.....		
Séance 1	Cours	3H
<ul style="list-style-type: none"> ● Introduction-Concepts de base : - L'importance de la statistique pour un data scientiste - Pourquoi faire de l'analyse statistique - Compétences nécessaires attendues du cours - Analyse conceptuelle : population, échantillon, individu (observation), variable statistique, quantitatif, qualitatif - Nécessité du traitement statistique (organisation, regroupement, ...) à partir d'une série brute 		
Séance 2	Cours	3H
<ul style="list-style-type: none"> ● Tableaux statistiques-Représentations graphiques : - Comment établir un tableau statistique à partir d'un jeu de données - Elaboration du tableau compte tenu de la nature du caractère - Comment lire et interpréter un tableau statistique - Visualisation graphique d'un tableau selon le type de caractère - Tableau VS Graphique : que faut-il choisir, pourquoi et utilité 		
Séance 3	TD	3H
<ul style="list-style-type: none"> ● Activités de base (TD1) 		
Séance 4	Cours	3H
<ul style="list-style-type: none"> ● Paramètres de tendance centrale : - Mesures d'un caractère quantitatif : - Moyennes, Médianes, Modes pour chaque type de caractère - Limites des paramètres de tendance centrale 		
Séance 5	Atelier	3H
<ul style="list-style-type: none"> ● Paramètres de dispersion : - Notion de variabilité, utilité et son sens - Principaux paramètres de variation - Apport des paramètres de variation par rapport à ceux de tendance centrale 		
Séance 6	TD	3H
<ul style="list-style-type: none"> ● Activités (TD2) 		
Séance 7	Cours-TD	3H
<ul style="list-style-type: none"> ● Paramètres de forme et de concentration - Activités TD3 		

<ul style="list-style-type: none"> - Courbe de concentration - Indice de Gini - Mise en évidence des inégalités dans une série 		
Séance 8 <ul style="list-style-type: none"> ● Analyse bi variée : corrélation entre deux variables quantitatives- Activités - Corrélation de pearson - Corrélation de spearman - Corrélation de Kandall 	Cours-TD	3H
Séance 9 <ul style="list-style-type: none"> ● Analyse bi variée : corrélation entre deux variables qualitatives- Activités ● Test du Khi-deux ● Analyse de la dépendance entre deux variables qualitatives ● Hypothèses prises en considération pour effectuer le test ● Activités TD 4 	Cours-TD	3H
Séance 10 <ul style="list-style-type: none"> ● Analyse bi variée : corrélation entre deux variables quantitatives qualitatives- Activités - ANOVA à un facteur - ANOVA à deux facteurs 	Cours-TD	3H
Séance 11 <ul style="list-style-type: none"> ● Tests non paramétriques à deux variables - Introduction à la statistique des rangs - Quel apport de cette statistique - Quelle est sa puissance - Cas de deux échantillons indépendants - Cas de deux échantillons dépendants <ul style="list-style-type: none"> ● Test d'ajustement ou de conformité 	Cours-TD	3H
Séance 12 <ul style="list-style-type: none"> ● Tests non paramétriques sur plusieurs échantillons : - Cas de trois ou plus échantillons indépendants - Cas de trois ou plus de trois échantillons dépendants 	Cours-TD	3H
Séance 13 <ul style="list-style-type: none"> ● Test de diagnostic des modèles - Test de normalité - Test d'homoscédasticité - Test d'autocorrélation des erreurs - Outils graphiques utilisés pour le diagnostic des modèles 	Cours-TD	3H
Séance 14 <ul style="list-style-type: none"> ● Nettoyage et transformation des données - Traitement des outliers - Algorithmes - Transformation basique des données (polynomiale, logarithmique) 	Cours-TD	3H

- Transformation de Box Cox
- Transformation de yeo jonhson

6. Mode d'évaluation de Module (*nombre, types et pondération des contrôles*)

Eléments d'enseignement	Coeff	DS	EX	TP	PR
Module - Analyse statistique	2	40%	60%		

Pour valider le module, les étudiants passeront un examen dont le coefficient est de 60%, un DS dont le coefficient est de 40%.

La durée de tous les examens (Examen, DS...) est de 1h30.

Le DS est planifié 7 semaines après le début du module et portera sur les thématiques suivantes :

- Tableaux statistiques-Représentations graphiques
- Paramètres de tendance centrale et de dispersion
- Paramètres de forme et de concentration

L'examen, il est planifié après l'écoulement des 14 semaines et portera sur toutes les thématiques enseignées tout au long des 42 heures.

Le module est validé si l'étudiant obtient une moyenne supérieure ou égal à 10 sur 20.

Dans l'ensemble, l'examen vise à évaluer les étudiants à la fois sur leurs capacités à maîtriser les notions de base d'une série statistique, caractériser les spécificités d'une série de données, et concevoir la méthode adéquate d'analyser une série donnée et communiquer les résultats.