

**Laboratoire de Technologies de e-commerce et mobile
(Java, mobiles & R)
Partie "Technologies Java et exploration des données dans
le contexte des big Datas " :**

Compléments pour le projet "Maraicher en ligne "

**3^{ème} Informatique orientation développement d'application (gestion)
2023-2024**



1. UE concernée

L'Unité d'Enseignement "**Programmation réseaux, web et mobiles**" (10 ECTS - 135h) se structure en Informatique de gestion en trois Activités d'apprentissage de la manière suivante :

- AA: Réseaux et technologies Internet (60h - 45%)
- AA: Programmation.Net (30h - 22%)
- AA: [Technologie de l'e-commerce et mobiles](#) (45h – 33%)

2. Le contexte

Le contexte de ce laboratoire de " AA : Technologie de l'e-commerce et mobiles " est le même que celui du laboratoire de "Réseaux et technologies Internet", à savoir celui de "Le maraicher".

3. Organisation et répartition du cours de « Technologie de l'e-commerce et mobiles » (théorie + laboratoire)

Le cours de « Technologie de l'e-commerce et mobiles » (45h) est scindé comme suit :

Points de matière	Théorie (50% des points)	Laboratoire (50% des points)	Titulaire
Partie 1. Technologie JAVA Mobile (Android) - SSL/TLS – SMTP/POP3/IMAP	12h (représente 25% des 50% attribués à la théorie)	10h (représente 25% des 50% attribués au laboratoire)	Mr Charlet
Partie 2. Exploration des données dans le contexte des big datas R / corrélation, régressions multiples / ANOVA2 / ACP / ACM / CAH / HCPC	12h (représente 25% des 50% attribués à la théorie)	10h (représente 25% des 50% attribués au laboratoire)	Mr Quettier

Total théorie = moyenne géométrique des parties 1 et 2 de théorie

Total labo = moyenne géométrique des parties 1 et 2 de laboratoire

4. Règles d'évaluation appliquées aux deux parties

Comme on sait, la note finale pour l'UE considérée se calcule par une moyenne géométrique des notes des AA constitutives.

Pour ce qui concerne l'évaluation de l'AA " Technologie de l'e-commerce et mobiles " pour les deux parties définies ci-dessus, voici les règles de cotation utilisées les enseignants de l'équipe responsable de cette AA.

a. Première session (Janvier 2024):

L'évaluation établissant la note de l'AA "Compléments Programmation Réseaux" est réalisée de la manière suivante :

- examen de théorie: un examen **écrit (les deux parties en même temps durant une seule séance de 2h)** en janvier 2024 (sur base d'une liste de points de théorie à développer durant le cours théorique) et coté sur 20;
- laboratoire en évaluation continue: correspond l' évaluation 1 (voir ci-dessous) cotée sur 20 qui constitue la note d'évaluation continue (**non remédiable en 2^{ème} session**);
- examen de laboratoire: correspond à l'évaluation 2 (voir ci-dessous) sous forme d'un examen oral en janvier 2024 et coté sur 20;
- **note finale total: moyenne géométrique de la note de l'examen de théorie (poids de 50%), de la note d'évaluation continue (poids de 15%) et de la note de l'examen de laboratoire (poids de 35%).**

Dans ces conditions, ***une note beaucoup trop basse parmi les trois ne peut que conduire à l'échec de l'AA considérée.***

b. Deuxième session,

Un **report de note** est possible pour **des notes supérieures ou égales à 10/20** en ce qui concerne :

- la note de théorie;
- la note de laboratoire de l'évaluation 2 (examen de laboratoire).

Les évaluations de théorie et du laboratoire d'examen ayant des **notes inférieures à 10/20** sont donc **à représenter dans leur intégralité** (le refus de représenter une évaluation complète de laboratoire entraîne automatiquement la cote de 0).

La note de laboratoire de l'évaluation 1 n'est pas remédiable et n'intervient plus dans le calcul de la note finale (principe général de l'évaluation continue).

La note finale pour la **seconde session**: **moyenne géométrique de la note de l'examen de théorie (poids de 50%), et de la note de l'examen de laboratoire (poids de 50%).**

c. Travaux par équipe

Dans le cas où les travaux sont présentés par une équipe de deux étudiants, chacun d'entre eux doit être capable d'expliquer et de justifier l'intégralité du travail sans de longues recherches dans le code de l'application proposée (pas seulement les parties du travail sur lesquelles il aurait plus particulièrement travaillé).

d. Explications supplémentaires demandés lors des évaluations

Dans tous les cas, tout étudiant doit être capable d'expliquer de manière générale (donc, sans entrer dans les détails) les notions et concepts théoriques qu'il manipule dans ses travaux (par exemple: keystore SSL, structure et traitement multipart d'un mail, ..., ANOVA2, ACP, ACM, ...).

5. En conclusion

Le laboratoire à la fois pour les parties 1 et 2 de " Technologie de l'e-commerce et mobiles " comporteront donc deux évaluations. (avec rentrée d'un dossier papier tel que décrit dans l'énoncé). La deuxième (Android, SSL/TLS) sera évaluée lors de l'examen de laboratoire en **janvier 2024** (un dossier papier ne sera plus nécessaire).

	Partie 1 (Mr Charlet)		Partie 2 (Mr Quettier)	
Evaluation 1 Non remédiable	Messagerie électronique SMTP / POP3 + dossier papier	Semaine du 16/10/2023	ANOVA1 – Régression/corrélation multiples	Semaine du 16/10/2023
Evaluation 2	Janvier 2024 le même jour pour les deux parties : Partie 1 : Android, SSL/TLS - partie 2 : ANOVA2, ACP, ACM, CAH, HCPC + dossier papier			

Pour chaque évaluation, le délai est à respecter impérativement.

Remarque importante :

Pour rappel, lors de chaque évaluation, chaque étudiant est sensé connaître les bases théoriques qui lui ont permis de réaliser les développements proposés. Dans le cas contraire, on sera amené à considérer qu'il a développé sans comprendre ce qu'il faisait ...

PARTIE 1

Les travaux de l'évaluation 1: messagerie électronique

Compétences développées : Maîtriser les concepts et techniques de la messagerie électronique

Dossier attendu : Exemples de réponses obtenues pour ces requêtes dans l'application.

SMTP/POP3 en Java

6. Un mail user agent en Java

- Développer une petite application **Application_Mail_Java** qui est un mail user agent, c'est-à-dire permettant de traiter le courrier électronique échangées entre le maraicher et ses fournisseurs.
- Les fonctionnalités implémentées dans un GUI **Java** de gestion classique d'e-mail (type "**mini-outlook**"), sont d'envoyer et de recevoir un mail composé de deux manières différentes :
 - ✓ un mail simple de type texte;
 - ✓ un mail composite avec des pièces attachées qui sont des **images** gif ou jpg, ou un **digest** de contrôle d'intégrité ou encore un simple fichier texte (.txt / .pdf).
- L'utilisateur devrait idéalement être prévenu dans un délai de 5 minutes (10 secondes lors de l'évaluation) de l'arrivée d'un nouveau message. A cet effet, l'application utilisera un thread de polling qui "interrogera" périodiquement la boîte aux lettres.
- Une fonctionnalité supplémentaire permettra de "tracer" un message reçu, c'est-à-dire de lister les agents mails (les MTAs) par qui ce message est passé.

2. Un plus (option): Un second mail user agent en C

Une petite application C **Application_Mail_C** leur permettant d'envoyer des mails limités au format RFC 822.

Remarque : pas de perte de point si vous ne réalisez pas cette partie (un plus car option)

Les travaux de l'évaluation 2: Android, SSL/TLS

Compétences développées :

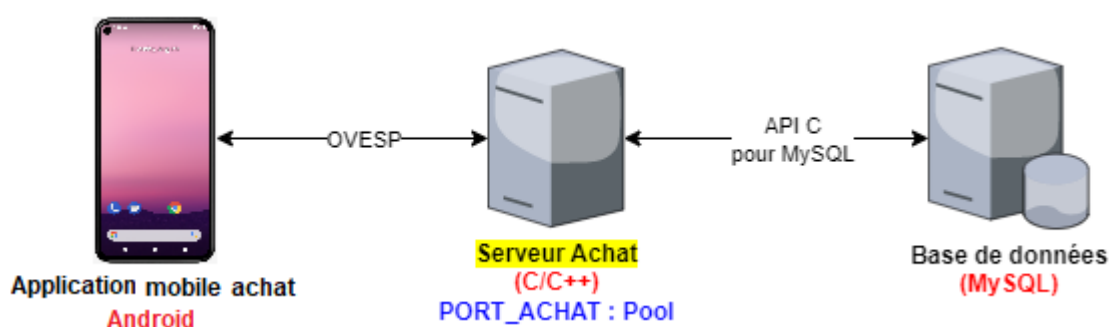
- Maîtriser les techniques de base du développement Android
- Savoir utiliser SSL/TLS

Dossier attendu :

- Schéma général de l'application Android en termes d'activités, d'intents, d'events, d'éventuels fragments (formalisme libre)
- Capture des échanges d'initialisation et de « datas » lors d'une session SSL/TLS

3. L'application mobile Android pour les réservations

Dans le contexte du maraicher, développer une application mobile afin de passer une commande au moyen de son smartphone.



Le protocole « **Online VEgetables Shopping Protocol** » (**OVESP**) développé dans le cadre du laboratoire de Réseaux et technologies Internet est réutilisé.

Rappel des différentes commandes sont décrites ci-dessous :

Serveur Achat -- Protocole OVESP – PORT_ACHAT			
Commande	Requête	Réponse	Actions / Explications
« Login »	Login, password, nouveau client ou pas	Oui ou non, message (+ idClient) ou raison	Vérification de l'existence et du mot de passe du client / Création d'un nouveau client dans la table clients

« Consult »	idArticle	idArticle ou -1, intitulé, stock, prix, image	Consultation d'un article en BD → si article non trouvé, retour -1 au client
« Achat »	idArticle, quantité	idArticle ou -1, quantité ou 0, prix	Si article non trouvé, retour -1. Si trouvé mais que stock insuffisant, retour d'une quantité 0 → Si ok, le stock est mis à jour en BD et le contenu du caddie est mémorisé au niveau du serveur → actuellement aucune action sur tables factures et ventes
« Caddie »		Contenu du panier : (idArticle, intitulé, quantité, prix) × nombre d'articles du panier	Retourne l'entièreté du contenu du caddie au client
« Cancel »	idArticle	Oui ou non	Supprime un article du caddie et met à jour à la BD
« Cancel All »			Supprime tous les articles du caddie et met à jour la BD
« Confirmer »		Numéro de facture générée	Création d'une facture et BD et ajout des éléments du caddie dans la BD
« Logout »			Si Caddie en cours, vide le caddie et met à jour la BD

Consignes :

L'application est découpée en plusieurs activités (3 au minimum) , la technique des intents, éventuels fragments.

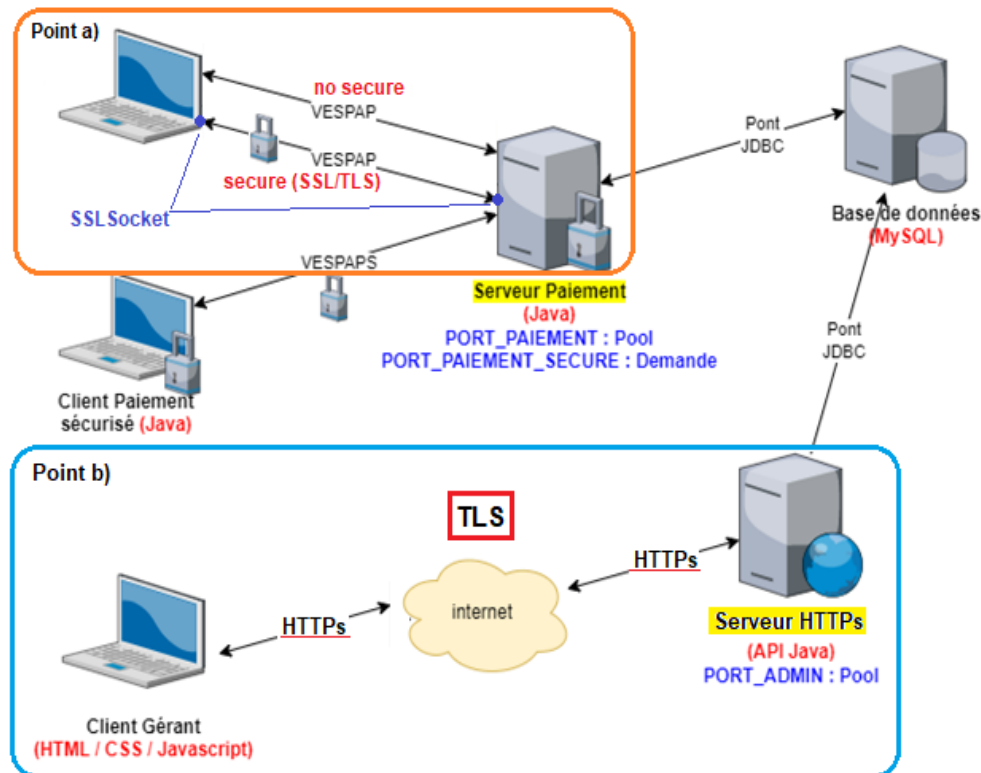
Possibilité du choix de la langue (internationalisation) au moins 2 langues.

Remarques importantes :

Afin de ne pas impacter le fonctionnement de l'application mobile achat, si le Serveur achat n'est pas disponible (problème(s)), vous pouvez implémenter un serveur basique mono-connexion (il peut être en java) afin d'assurer les échanges (voir protocole) entre le client mobile et ce dernier. Dans ce cas de figure, l'évaluation de cette partie de laboratoire n'est pas conditionnée par la présence ou pas / problèmes éventuels avec le serveur achat. Seul impératif, le contenu des différents messages du protocole devra être affiché coté serveur. Pas d'impact sur l'évaluation de l'application mobile.

Réutilisation de librairies bien réfléchies au départ (analyse) par rapport aux deux cours : RTI/compléments

4. Les communications sécurisées par SSL



a. Client paiement avec tunnel sécurisé

Le but est de reprendre le client paiement non sécurisé du laboratoire de RTI et d'assurer le chiffrement et l'authentification à l'aide du protocole SSL/TLS. Vous avez donc à ajouter à ce client la possibilité de choisir entre un mode non secure (socket normale) et mode secure (SSLSocket). Du point de vue du serveur, un port supplémentaire (port pour la SSLSocket) est en « écoute ».

Le protocole reste identique que ça soit en secure ou en non secure (rien à changer). De cette manière, on met bien en évidence la notion vue au cours, de sous-couche propre à SSL/TLS. Vous avez, bien entendu, la nécessité de créer deux certificats (un pour le client et un pour le serveur) qui seront signés par votre CA personnel (oui il faudra un CA « non officiel » => pas de certificat auto-signé).

Remarques importantes :

Plan B1, Si l'intégration de la possibilité de choisir entre la version secure et non secure au sein de la même application cliente, vous pouvez implémenter un nouveau client exclusivement réservé pour une connexion de type SSL/TLS.

Plan B2, Si l'intégration au sein du serveur existant, de la SSLsocket pose problème(s), vous pouvez concevoir un serveur basique mono-connexion réservé exclusivement à cette partie. Seul impératif, le contenu des différents messages du protocole devra être affiché coté serveur. Pas d'impact sur l'évaluation.

b. Le service WEB de gestion du stock du magasin

Le service WEB de gestion de stock du magasin du laboratoire de RTI est actuellement accessible via des requêtes en http. Il va de soi que de nos jours, ce service doit s'appuyer sur des échanges sécurisés un minimum. Comme l'avez certainement compris, le but de cette partie est de passer en HTTPS (port supplémentaire) autrement dit d'ajouter (un complément, ce n'est pas le nom du cours d'ailleurs) une sous-couche SSL/TLS au service WEB demandé dans le laboratoire de RTI.

C'est bien « un ajout/un complément » rien ne change du point de vue l'implémentation / fonctionnement / protocole du serveur de web défini dans le laboratoire de RTI. Seul une sous-couche supplémentaire est ajoutée (voir définition du protocole SSL/TLS : sous-couche intercalée entre la couche d'application et de transport du modèle TCP/IP). Il

Les échanges seront dès lors **intégralement chiffrée et authentifiée en utilisant le sous-protocole SSL/TLS**. Bien entendu, vous prendrez soin de vérifier à l'aide par exemple de « Wireshark » ou autre outil similaire si les échanges sont bien chiffrés une fois le protocole SSL/TLS initialisé. Il est inutile de vous rappeler que vous devrez confectionner au minimum un certificat pour le serveur.

Remarques importantes :

- Comme préciser dans le cadre de l'application mobile Android, afin de ne pas conditionner le fonctionnement de cette partie à l'existence du serveur WEB du laboratoire de RTI, vous pouvez implémenter un serveur WEB basique HTTPS afin d'assurer les échanges (voir protocole) entre le client WEB et ce dernier. Dans ce cas de figure, l'évaluation de cette partie de laboratoire n'est pas conditionnée par la présence ou pas / problèmes éventuels avec le serveur. Seul impératif, le contenu des différents messages du protocole devra être affiché côté serveur. Pas d'impact sur l'évaluation.
- Lors de l'utilisation de votre Browser habituel pour communiquer avec le serveur, une alerte pourrait être émise par ce dernier ou par votre protection anti-virus/web/... . En effet, le certificat utilisé par le serveur n'est pas un certificat officiel (c'est vous qui l'avez fabriqué). Ce dernier n'est donc pas un certificat validé/certifié/signé par un CA de confiance. Cependant, vous avez la possibilité d'accepter de prendre le risque et d'outrepasser cette page d'alerte. Et le magasin de certificats alors ?
- Réutilisation de bibliothèques bien réfléchies au départ (analyse) par rapport aux deux cours : RTI/compléments

PARTIE 2

Les travaux de l'évaluation 1: ANOVA 1 – Régression/corrélation multiple

Compétences développées : Maîtriser les concepts et techniques de l'ANOVA 1 et de la régression/corrélation multiple

Dossier attendu : Solution, résultats, interprétations et justifications des résultats

ANOVA 1

1. Les civilisations précolombiennes

Le Ministère de la Culture de Batracie a étudié le nombre de récipients contenant de la bière fermentée sur divers sites archéologiques correspondants chacun à l'un des 4 types de civilisations précolombiennes suivantes: Cuacuacomeqiqi, Oxomatl, Tlaloc et Tenochtitlan. Les résultats de ces comptages (par are) sont repris ci-dessous :

Cuacuacomeqiqi	Oxomatl	Tlaloc	Tenochtitlan
93	85	100	96
120	45	75	58
65	80	65	95
105	28	40	90
115	75	73	65
82	70	65	80
99	65	50	85
87	55	30	95
100	50	45	82
90	40	50	
78		45	
95		55	
93			
88			
110			

Observe-t-on des différences significatives entre les quatre traitements et quels sont ceux qui sont, s'ils existent, à résultats similaires ?

Régression et corrélation multiple

5. Les accidents sur les routes du Minnesota

Le fichier accidents2.csv contient des données sur le taux d'accidents de voiture sur les autoroutes du Minnesota (1973). Ces données, collectées sur 39 grands segments d'autoroute, ont évidemment été collectées pour essayer de déterminer les raisons de ces accidents.

- **rate : taux d'accidents par million de véhicules**
- len : longueur du segment de collecte
- adt: densité moyenne du trafic (en milliers)
- trks: pourcentage de camions
- **sigs1: nombre de signaux routiers par mile**
- slim: vitesse maximale autorisée
- **shld: largeur (en pieds) de la bande d'urgence latérale**
- lane: nombre de bandes
- **acpt: nombre d'entrées par mile**
- itg: nombre d'échangeurs par mile
- lwid: largeur d'une bande (en pieds)
- htype: type de chaussée, soit MC ("major collector"), FAI ("Federal interstate highways"), PA ("principal arterial highway") or MA ("major arterial highways").
- rate: nombre d'accidents par million de véhicules en 1973

On demande d'étudier l'éventuelle relation entre ce taux d'accidents et nombre de signaux routiers par mile associé à la largeur de la bande d'urgence latérale. Dans un second temps, on demande d'ajouter comme 3^{ème} variable explicative le nombre d'entrées par mile d'autoroute.

Les travaux de l'évaluation 2: ANOVA 2, ACP, ACM, CAH, HCPC

Compétences développées :

Maîtriser les techniques de l'ANOVA 2 et d'exploration des données (ACP, ACM, CAH et HCPC)

Dossier attendu : Solution, résultats, interprétations et justifications des résultats

ANOVA 2

6. Bière et petits maux

L'Administration de la Santé Publique de Bidendumie a recensé le nombre de patients atteints de l'une des 4 maladies bénignes les plus fréquentes et ayant consommé l'une des 3 bières locales les plus répandues. Elle a mesuré un coefficient biochimique représentatif sur 6 patients (si possible) choisis aléatoirement et a obtenu (chaque ligne correspond à une maladie) :

	bière : La tueuse						bière : La vaisselle						bière : Le nectar royal					
gastro entérite	42	44	36	13	19	22	33	26	33	21	29	31	31	-3	25	25	24	17
mal de tête	28	23	34	42	13	27	34	33	31	36	34	38	3	26	28	32	4	16
nausées	1	29	19	29	18	22	11	9	7	1	-6	5	21	1	9	3	12	17
sinusite	24	9	22	-2	15	12	27	12	12	-5	16	15	22	7	25	5	12	19

Est-il possible d'interpréter de tels résultats ? Ils se trouvent dans le fichier bieres_petits_maux.csv.

7. Les médicaments contre la GCE

Une entreprise pharmaceutique s'intéresse à une maladie tropicale (la Gengivite Cephalopodique Endiablée - GCE) et a mis au point trois molécules susceptibles de soigner cette maladie : AlphaVictoire, BetaTriomphe et GammaSucces. Les tests cliniques ont été pratiqués pour mesurer un coefficient relatif d'amélioration de l'état de patients gravement atteints (plus ce coefficient d'immunité est élevé et plus l'action sera considérée comme efficace). Mais, de plus, on souhaite également tenir compte du mode d'administration des différentes molécules (par voie orale ou par injection intraveineuse). Les résultats sont :

	AlphaVictoire	BetaTriomphe	GammaSucces
voie orale	10	7	12
	12	14	9
	8	10	11
	10	11	27
	6	9	7
	13	10	8
	9	11	13
	10	7	14
	9	9	10
	8	9	11
injection	11	8	7
	18	9	6
	12	10	10
	15	9	7
	13	11	7
	8	13	5
	15	7	6
	16	14	7
	9	15	9
	13	12	6

Observe-t-on une différence significative d'efficacité soit selon la molécule, soit selon le mode d'administration ou encore selon une combinaison des deux facteurs ?

Exploration des données

8. ACP et ACM

a. Eaux minérales

Le fichier Eaux1.txt contient des données sur la teneur en divers éléments chimiques pour quelques eaux minérales commercialisées en France.

Quelles relations peut-on détecter ?

Peut-on donner une signification claire aux axes principaux ?

b. Etude de maïs

Le ministère de l'Agriculture du Malabarland a commandité une étude sur les plants de maïs afin d'optimiser les techniques de culture. Un échantillon de 100 pieds de maïs a été constitué (sur 50000 pieds possibles) et les résultats ont été compilés dans le fichier etude-agro-mais.csv.

Certaines variables s'interprètent par leur nom et pour les autres :

- ◆ Masse: masse de l'ensemble des grains du plant
- ◆ Germination.epi: le grain est-il germé sur épi ?
- ◆ Verse: le pied est-il penché ou tombé ?
- ◆ Attaque: attaqué par des insectes ?
- ◆ Hauteur.J7: hauteur 7 jours après la récolte
- ◆ Verse.Traitement: verse après traitement ?
- ◆ Nb.jours.attaque: nombre de jours entre la pousse jusqu'à l'attaque
- ◆ Censure.droite: non utilisée dans la suite

Le Ministère commande une étude globale d'exploration des données (il ne le sait pas vu sa formation, mais ceci implique donc une ACM et une ACP de l'ensemble des données). Que peut-on observer ?

c. Le retour du Titanic (ACM)

Un historien a réalisé une étude des données en rapport avec le naufrage du Titanic. Les résultats sont dans le fichier de données titanic.csv.

Que peut-on en déduire ?

La colonne 1 est l'identificateur des personnes.

La colonne 2 correspond à la classe de cabine, selon les codes

0 = équipage, 1 = première classe, 2 = seconde classe, 3 = troisième classe.

La colonne 3 est la catégorie d'âge : 0 = enfant, 1 = adulte.

La colonne 4 est le sexe de la personne : 0 = femme, 1 = homme.

La colonne 5 indique si la personne a survécu : 0 = non, 1 = oui.

9. Les classifications : CAH et HCPC

a. Histoire d'eaux (minérales)

On revient sur des données sur la teneur en divers éléments chimiques des eaux minérales de diverses provenances (plates ou gazeuses) commercialisées en France en utilisant cette fois le fichier Eaux2010.txt.

On demande de chercher à établir une classification de ces eaux minérales. On demande notamment de comparer les 4 méthodes classiques de regroupement. Une étude HCPC apporte-t-elle des informations complémentaires ?

b. Les vins italiens

Une étude internationale porte sur des vins italiens afin de déterminer si les classifications établies par les viticulteurs reposent sur des données objectives ou relèvent plutôt de traditions et autres arguments subjectifs. Cette étude dit essentiellement :

" These data are the results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The analysis determined the quantities of 13 constituents found in each of the three types of wines.

The attributes are

- 1) Alcohol
- 2) Malic acid
- 3) Ash
- 4) Alkalinity of ash
- 5) Magnesium
- 6) Total phenols
- 7) Flavanoids
- 8) Nonflavanoid phenols
- 9) Proanthocyanins
- 10) Color intensity
- 11) Hue
- 12) OD280/OD315 of diluted wines
- 13) Proline"

Le fichier "**winedata.csv**" contient ces données. On demande une étude des classifications envisageables pour ces vins.

BON TRAVAIL