



Université des sciences et de la technologie
Mohamed-Boudiaf d'Oran USTO.

Faculté des mathématiques et informatique

Département d'Informatique

M1 IAA, SID et RSID

Analyse de données AND

Fiche de TD1

Exercice 1 :

Soit le tableau suivant représente les clients et leurs achats dans un magasin

ID	Age	Nbre d'achat
1	25	3
2	30	5
3	22	2
4	35	8

- 1- Calculer la distance euclidienne entre les clients 1 et 2 sur la base de leurs caractéristiques âge et nbre d'achat.
- 2- Calculer la distance Manhattan entre client 3 et 4.
- 3- Construire le tableau de distance (euclidienne).

Exercice 2 :

On veut prédire si une boisson est thé ou café suivant les données suivantes :

Boisson	Volume (ml)	Caféine(g)
thé	250	0.025
thé	100	0.010
café	125	0.050
café	250	0.100

- Calculer la corrélation entre volume et caféine.
- Quel serait l'étiquette prédite pour une boisson avec de 125ml contenant 0.015 g de caféine.

Exercice 3 :

Ines caractérise des mélodies. Elle considère comme variables les 7 notes do, ré, mi, fa, sol, la et si. Les trois mélodies qui l'intéressent sont :

A : si si do ré ré do si la sol sol la

B : do do do ré mi re do mi ré ré do

C : sol sol si la si la sol sol si si la

Sans passer par le codage disjonctif

- 1- Quelle est la distance de Hamming entre A et B , A et C, B et C.
- 2- Quelle est la similarité de Jaccard entre A et B, A et C, B et C
- 3- Quelle sont les notes similaires ?

Exercice 4 :

Soit le tableau de données suivant :

Produit	Couleur	Taille
P1	Bleu	XL
P2	Blanc	L
P3	Bleu	M
P4	Blanc	M
P5	Bleu	L

- 1- Construire le tableau de codage disjonctif.
- 2- Construire le tableau de contingence.
- 3- Construire le tableau des fréquences, profil ligne et profil colonne.