

A

Directives pédagogiques :

- Aucun document n'est autorisé.
- Il sera tenu compte de la clarté des raisonnements.

Toute tentative de fraude sera sanctionnée par la note zéro.

Lisez attentivement les recommandations données à la fin de l'épreuve.

Nom :

Prénom :

Partie I : Statistique Descriptive

La compagnie INTER-WEB compile le nombre de minutes de connexion de chacun de ses 200 clients au cours du son premier mois d'existence. On regroupe les observations de la façon suivante :

Nombre de minutes	Nombre de clients
[0 , 60[10
[60, 120[30
[120, 180[100
[180, 240[30
[240, 300[20
[300, 360]	10

a) Identifier la population d'analyse :

.....

b) Quel est le caractère étudié ici? Notons par X ce caractère.

c) De quel type de caractère s'agit-il?

.....

d) Peut-on calculer la moyenne de ce caractère ?

e) Si Oui, donner la formule : ainsi que la valeur

.....

f) Donnez la valeur du mode de ce caractère :

.....

g) Donnez la valeur de la médiane de ce caractère : ainsi que la l'étendu :

h) Quel est le nb de clients dont le nb de connexion est au moins 270 minutes ?

.....

i) Quelle devrait être la tarification (en Dh/heure de connexion) par client si la compagnie estime le revenu moyen des ventes de ce premier mois à 30 000 Dh ?

.....

j) Quelle serait la nouvelle estimation du revenu moyen si la compagnie introduit des frais de maintenance de 4Dh par mois pour chaque client ?

.....

.....

.....

Partie II Régression linéaire (cf. sorties correspondantes dans l'annexe) les retourner avec vos feuilles de réponses commentées et avec une légende.

On soupçonne l'existence d'une relation entre la taille d'un fichier (en kilo octets) et la durée de son téléchargement (en s). La table suivante contient les observations relatives à 10 fichiers téléchargés dans des conditions identiques:

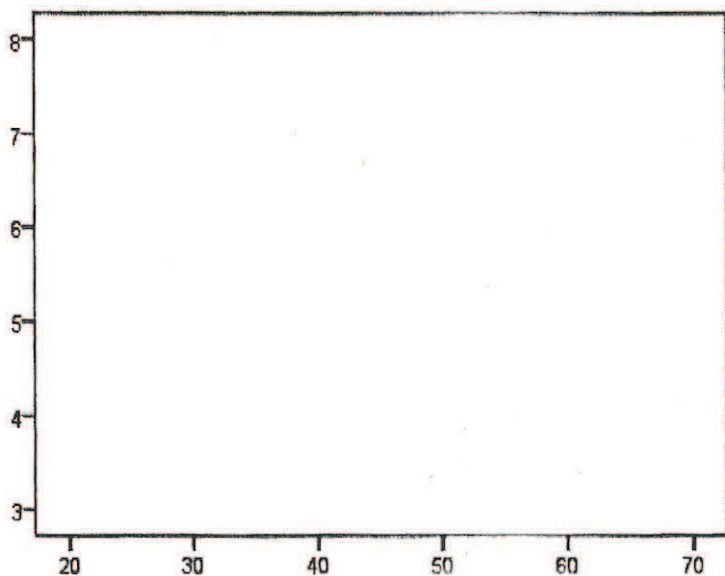
Fichier N°	1	2	3	4	5	6	7	9	9	10
Taille du Fichier (en Kb)	21	40	31	62	28	50	55	43	61	30
Durée du téléchargement (en s)	7	5	6	1	5	3	2	6	3	5

2.1 Définir la variable dépendante et la variable indépendante.

.....

.....

2.2 Faire une représentation graphique du nuage du point (dans le cadran ci-dessous) et justifier la linéarité de la relation entre la taille d'un fichier et la durée de son téléchargement.



.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

2.3 Écrire l'équation de la droite de régression

2.4 Donner une interprétation au coefficient de la variable indépendante.

2.5 Énumérer les hypothèses nécessaires pour une analyse de régression linéaire

2.6 Est-ce que ces hypothèses semblent être satisfaites dans ce cas ? Justifier votre réponse.

2.7 Au seuil de signification $\alpha=0,05$, peut-on conclure que la constante du modèle est nulle ?

2.8 Donner le temps moyen de téléchargement pour un fichier de taille égale à 35 Kb.

Partie 3 : Analyse factorielle

(cf. sorties correspondantes dans l'annexe) les retourner avec vos feuilles de réponses commentées et avec une légende.

Des données ont été collectées pour une étude sur la pollution de l'air dans les villes des États-Unis. Pour 41 villes on a pris des mesures concernant les variables suivantes:

temp: température annuelle moyenne en degrés Fahrenheit. **manu**: nombre de manufactures de plus de 20 employés. **popul**: taille de la population de la ville en milliers. **wind**: vitesse annuelle moyenne du vent en miles par heure. **precip**: précipitation annuelle moyenne en pouces. **predays**: nombre annuel moyen de jours avec précipitation. (Voir un extrait du fichier de ces données dans l'annexe).

3.1 Pour une Analyse en Composantes Principales (ACP) de ces données, est-il plus approprié d'analyser la matrice des covariances des variables ou leur matrice des corrélations ? Justifier.

3.2 En étudiant la matrice de corrélation du tableau de données (voir Annexe), identifier les variables "fortement corrélées". Est-il possible de réduire le nombre de variables ? Justifier.

3.3 Utilisez la table intitulée "Indice KMO et le test de Bartlett" pour justifier la pertinence d'une Analyse en Composantes Principales

Les tableaux donnés dans l'annexe résultent d'un traitement, par SPSS, en vue de réduire le nombre de facteurs. Consultez-les pour répondre aux questions qui suivent :

3.4 Donner le pourcentage de variation expliquée par les 2 axes ayant les plus grandes valeurs propres.

3.5 Justifier la nécessité de recourir à une rotation

Pour la suite, on retiendra les 3 premiers facteurs.

3.6 Utilisant les résultats obtenus par une rotation Varimax. Donner une interprétation des 3 facteurs.
