

Traitement des données sous Excel.

A chaque question doit correspondre une feuille Excel, étiqueté avec le numéro de question. Expliciter les traitements effectués et les réponses aux questions dans une zone de texte.

Aucune indication n'est donnée. Utilisez les outils à bon escient.

Charger le fichier « [credit.xlsx](#) » sous Excel.

1. Combien y a-t-il de variables dans le fichier ? Combien y a-t-il d'observations ?
2. Calculer la proportion des crédits acceptés et refusés
3. Calculer la proportion des acceptations de crédit selon "Self.Employed". Que remarque-t-on ?
4. Les variables "Age" et "Income.per.Dependent" sont elles liées ? Un graphique « nuage de points » devrait suffire pour le constater.
5. Les variables "Own.home" et "Self.Employed" sont-elles liées ? Calculer le V de Cramer.
6. Quelles les différentes valeurs possibles de "Derogatory.Reports" ? Est-elle liée d'une manière ou d'une autre à "Acceptation.Crédit" ? Une simple réponse descriptive suffit.
7. Calculer la moyenne de "Income.Per.Dependent" selon "Self.Employed".
8. Découper "Income.Per.Dependent" en 3 intervalles de largeurs égales. Donner les bornes de découpage. Créer explicitement la variable recodée et calculer les effectifs dans chaque classe.
9. Découper "Income.Per.Dependent" en 3 intervalles de fréquences égales. Donner les bornes de découpage. Créer explicitement la variable recodée et calculer les effectifs dans chaque classe.
10. Quelle approche proposez-vous pour découper "Income.Per.Dependent" en 3 intervalles, de manière à ce que dans chaque classe, les proportions de "yes" et "no" de "Acceptation.Credit" soient les plus différentes possibles ?
11. Construire la courbe de Lorentz et l'indice de concentration pour la variable « income.per.dependent ».