|  |  |  |  |
| --- | --- | --- | --- |
|  | Traitement des données  M2 — Examen — septembre 2016 | | |
| *Prénom* | *Groupe* | / |
| *Nom* |  |

|  |
| --- |
| **Consignes**   * Mettez en avant votre raisonnement et la manière dont vous obtenez la solution ; * Faites attention au soin et à la lisibilité. |

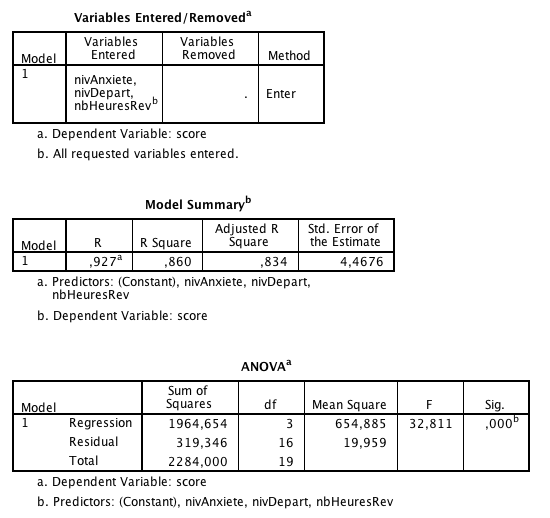
|  |  |
| --- | --- |
| **QUESTION 2 (régression linéaire et multiple)** | / |

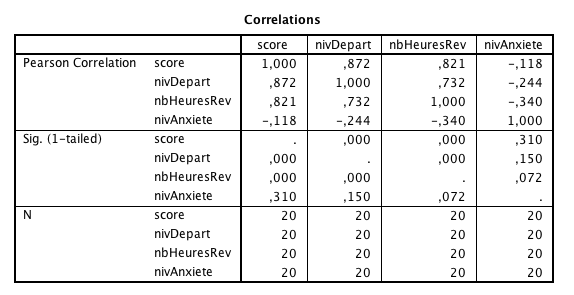
Le fichier examRevision.xlsx contient 20 observations pour un examen. Chaque observation concerne un étudiant et reprend :

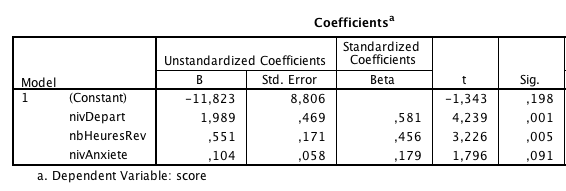
1) le score qu’il a obtenu à cet examen (cote sur 100)  
2) le nombre d’heures de révision   
3) le niveau d’anxiété  
4) des points représentant son niveau au départ.

Sur base des tableaux ci-dessous :

* 1. Déterminez le modèle de régression proposé.
  2. Sur base de trois indicateurs différents, précisez si ce modèle est judicieux.   
     Soyez complet et rigoureux dans votre argumentation.
  3. Précisez les hypothèses du test F effectué dans le tableau ANOVA.







|  |  |
| --- | --- |
| **QUESTION (Analyse en composantes principales)** | / |

1. A quoi sert une analyse en composantes principales ?
2. On s ‘intéresse au climat des différents pays d’Europe. Pour cela, on a recueilli les températures moyennes mensuelles (en degré Celsius) pour les principales capitales européennes ainsi que pour certaines grandes villes. Ces données ainsi que les résultats de l’acp sont fournis dans les tableaux et graphiques ci-dessous.  
   1. Donnez l’explication contextuelle (interprétation) des deux nouvelles composantes.   
        
      Composantes 1 : c’est lié à la moyenne des température(horizontale)   
      Composantes 2 : c’est lié à l’écart de température entre les mois froid et les mois chauds qui sont corrélées entre eux
   2. Quelle part de l’information de départ prennent-elles en compte ?  
      Composantes 1 : 87.398

Composantes 2 : 10.415

* 1. Enoncez trois indications (idées différentes) données par le vecteur octobre.  
     - Représente le plus la moyenne de l’année avec septembre

-

-

* 1. Vrai ou faux. Si faux, corrigez la phrase. Si vrai, justifiez.
     1. Pour chacune des villes, la température du mois d’avril est fortement corrélée à la température moyenne de l’année.
     2. La variable juin a joué le rôle le plus important dans la composante 2.  
          
        Faux c’est la variable juillet
     3. La proximité de deux villes sur la composante 1 implique systématiquement qu’elles ont deux profils de température moyenne semblables au cours de l’année.
     4. S’il fait froid en janvier, il fera froid en juin.  
          
        Faux car au niveau du deuxième composant, on voit bien qu’en janvier et juin ne sont pas corrélés entre eux.
  2. Ci-dessous les axes pour le graphique des individus. Interprétez la position des deux points.

