Pour chaque question posée ci-dessous, suivre les étapes suivantes :

1. Importer le tableau de données
2. Faire une ou plusieurs représentations graphiques judicieuses afin d’explorer les données en fonction de la question posée
3. Déterminer de quels types sont les variables dépendantes et indépendantes
4. Suivre les étapes de la démarche statistique afin de répondre à la question posée (hypothèses, choix du test justifié, validation du modèle ou des conditions du test, réponse à la question biologique)
5. Retrouver le résultat du test grâce à la table statistique appropriée

Exercice 1

Tableau de données « meteo.csv ».

Question 1 :

* A) existe-il une relation entre la température annuelle moyenne (*Temperature\_an*, variable dépendante) et la latitude (*Latitude*, variable indépendante) ?
* B) Tracer sur le graphique la droite ajustée par la régression linéaire (si la relation est significative)
* C) Que vaut le coefficient R² ?

Question 2 :

* A) existe-il une relation entre la précipitation annuelle moyenne (*Precipitations*, variable dépendante) et la latitude (*Latitude*, variable indépendante) ?
* B) Tracer sur le graphique la droite ajustée par la régression linéaire (si la relation est significative)
* C) comparer la valeur du coefficient R² calculé à la question 1 et à la question 2, qu’en pensez-vous ?

Question 3 :

* A) existe-il une différence de température moyenne (*Temperature\_an*, variable dépendante) selon la zone géographique des villes étudiées (*Geographie*, variable indépendante) ?
* B) Si besoin, déterminez quelles modalités sont différentes les unes des autres grâce à un test post-hoc

Exercice 2

Tableau de données « crime.csv ».

Explication du tableau dans « crime\_data\_summary.docx »

Question 1 :

* A) existe-il une relation entre le taux de criminalité (*CrimeRate*, variable dépendante) et le taux d’homme (nombre d’homme pour 1000 femmes latitude) (*Males*, variable indépendante) ?
* B) Tracer sur le graphique la droite ajustée par la régression linéaire (si la relation est significative)
* C) Que vaut le coefficient R² ?

Question 2 :

* A) existe-il une relation entre le taux de criminalité (*CrimeRate*, variable dépendante) et le salaire médian hebdomadaire (*Wage*, variable indépendante) ?
* B) Tracer sur le graphique la droite ajustée par la régression linéaire (si la relation est significative)
* C) Que vaut le coefficient R² ?

Question 3 :

* A) Transformer en catégorie la colonne état du sud
* B) extraire les valeurs de taux de criminalité des états du sud (Southern == 1) et des états du nord dans 2 vecteurs différents
* C) Existe-t-il une différence du taux de criminalité moyen entre les états du nord et du sud ?

Question 4 :

* A) le taux de criminalité moyen est-il différent à la fin des 10 ans de l’étude (*CrimeRate* et *CrimeRate10*), attention les données sont-elles indépendantes ?