

Atelier découverte : R, données et RGPD

Informations du groupe

Nom des élèves :

Etape 1 - Les tiques

Prenez le temps de lire les différents documents faisant référence aux tiques avant de répondre aux questions.

Une tique est...

- Un insecte
- Un arachnide
- Un crustacé

Valider

✔ Bravo ! Les tiques appartiennent au groupe des arachnides, comme les araignées et les scorpions. Les adultes possèdent 8 pattes, alors que les insectes en ont 6.

Pourquoi les scientifiques s'intéressent-ils particulièrement aux tiques ?

- Parce qu'elles peuvent transmettre des agents pathogènes
- Parce qu'elles détruisent les arbres
- Parce qu'elles produisent du venin

Valider

✔ Exact ! Les tiques peuvent transmettre certaines bactéries, virus ou parasites. On parle alors de vecteurs biologiques et les maladies associées sont des maladies vectorielles.

Quelle maladie est la plus souvent associée aux piqûres de tiques en France ?

- La grippe
- La maladie de Lyme
- La varicelle

Valider

✔ Très bien ! La maladie de Lyme est l'une des maladies vectorielles les plus connues en Europe. Elle est

causée par une bactérie transmise par certaines tiques.

Quelle méthode permet aux scientifiques de mieux comprendre les risques liés aux tiques ?

- Observer uniquement les tiques au microscope
- Collecter et analyser des données sur les piqûres et les tiques observées
- Mesurer la vitesse de déplacement des tiques

Valider

Exact ! Les scientifiques utilisent de grandes quantités de données pour étudier la répartition des tiques et identifier les facteurs qui influencent les risques de piqûres.

Après une promenade en forêt, quelle est la meilleure action pour limiter les risques liés aux tiques ?

- Se laver les mains uniquement
- Vérifier soigneusement sa peau et ses vêtements
- Boire beaucoup d'eau

Valider

Parfait ! Une détection rapide permet de retirer une tique avant qu'elle ne reste fixée longtemps, ce qui contribue à réduire certains risques.

Etape 2 - Comprendre le rôle des données

Pourquoi les chercheurs ont-ils besoin de partager des données ?

- Pour permettre à d'autres chercheurs de vérifier et compléter les résultats
- Pour vendre les données à des entreprises
- Pour rendre les études plus longues à réaliser

Valider

Exact ! Le partage des données favorise la transparence scientifique. Il permet à d'autres équipes de reproduire les analyses, de vérifier les résultats et parfois de faire de nouvelles découvertes à partir des mêmes données.

Dans un projet de science participative comme CiTIQUE, quel est le rôle des citoyens ?

- Ils remplacent complètement les chercheurs
- Ils contribuent à la collecte de données utiles à la recherche
- Ils décident des résultats scientifiques

Valider

Bravo ! Les citoyens participent à la collecte d'observations ou de données. Grâce à leur contribution, les chercheurs peuvent étudier des phénomènes sur de très grandes zones géographiques.

Qu'est-ce qu'une donnée personnelle ?

- Une information qui permet d'identifier directement ou indirectement une personne
- Une information concernant uniquement la météo
- Une information qui appartient à un laboratoire

Valider

Exact ! Un nom, une adresse e-mail ou même certaines combinaisons d'informations peuvent permettre d'identifier une personne. Ces données doivent être protégées.

Quel est l'objectif principal du RGPD ?

- Faciliter la publicité sur internet
- Protéger les données personnelles des citoyens
- Empêcher les chercheurs de travailler avec des données

Valider

Parfait ! Le RGPD (Règlement Général sur la Protection des Données) vise à protéger les personnes et à encadrer l'utilisation de leurs données personnelles.

Que signifie 'anonymiser' une donnée ?

- La supprimer définitivement
- La traduire dans une autre langue
- Retirer les informations permettant d'identifier une personne

Valider

Très bien ! L'anonymisation permet d'utiliser certaines données à des fins de recherche tout en protégeant l'identité des participants.

Pourquoi les chercheurs souhaitent-ils recueillir un grand nombre de données ?

- Pour rendre leurs ordinateurs plus puissants
- Pour obtenir des résultats plus fiables et représentatifs
- Pour remplir leurs bases de données

Valider

Exact ! Plus le nombre d'observations est important, plus les scientifiques peuvent identifier des tendances réelles et limiter l'influence du hasard.

Lorsqu'une personne participe à un projet scientifique, pourquoi est-il important qu'elle soit informée de l'utilisation de ses données ?

- Pour qu'elle puisse décider librement de participer ou non
- Pour augmenter le nombre de données collectées
- Pour accélérer le travail des chercheurs

Valider

Bravo ! L'information et le consentement sont des principes importants de l'éthique de la recherche et de la protection des données.

Quelle affirmation décrit le mieux une donnée ouverte (Open Data) ?

- Une donnée accessible à tous pour être consultée ou réutilisée sous certaines conditions
- Une donnée secrète réservée aux chercheurs
- Une donnée qui n'a jamais été vérifiée

Valider

Exact ! Les données ouvertes favorisent la diffusion des connaissances, l'innovation et la transparence scientifique. Elles permettent aussi de développer de nouvelles analyses à partir de données déjà collectées.

Etape 3 : manipulation de données sur R

À quoi sert principalement le logiciel R ?

- À analyser et visualiser des données
- À créer des jeux vidéo
- À réparer des ordinateurs

Valider

Exact ! R est un langage de programmation largement utilisé par les scientifiques pour analyser, représenter et comprendre des données.

Dans un tableau de données, que représente généralement une ligne ?

- Une observation ou un individu
- Une variable ou une caractéristique
- Un graphique

Valider

Bravo ! Dans un tableau, chaque ligne correspond souvent à une observation. Par exemple, dans CiTIQUE, une ligne peut représenter une tique observée ou une piqûre signalée.

Que représente généralement une colonne dans un tableau de données ?

- Une observation
- Une variable ou une information mesurée
- Un résultat d'analyse

Valider

Exact ! Une colonne contient une caractéristique commune à toutes les observations, comme une date, un lieu ou une espèce.

Pourquoi les scientifiques utilisent-ils des graphiques ?

- Pour rendre les résultats plus compliqués
- Pour mieux comprendre et communiquer les données
- Pour rendre les résultats plus faciles à interpréter

Pour remplacer complètement les données

Valider

Très bien ! Les graphiques permettent de repérer rapidement des tendances, des différences ou des évolutions dans les données.

À quoi sert principalement une application Shiny ?

- À créer des applications interactives à partir de données
- À stocker les données sur internet
- À remplacer les bases de données

Valider

Bravo ! Shiny permet de transformer des analyses réalisées avec R en applications interactives accessibles à tous via un navigateur web.

Quel est l'avantage principal d'une application interactive par rapport à un graphique statique ?

- Elle permet d'explorer les données selon différents choix
- Elle affiche plus de couleurs
- Elle nécessite moins de données

Valider

Exact ! L'utilisateur peut filtrer, sélectionner ou comparer des données selon ses besoins sans modifier le programme.

Pourquoi est-il important de vérifier la qualité des données avant de les analyser ?

- Parce que des erreurs dans les données peuvent conduire à de mauvaises conclusions
- Parce que cela rend les graphiques plus colorés
- Parce que cela augmente automatiquement le nombre de données

Valider

Parfait ! Une analyse n'est fiable que si les données utilisées sont elles-mêmes de bonne qualité.

Dans une application Shiny, que se passe-t-il généralement lorsqu'un utilisateur modifie un filtre ?

- Les résultats ou graphiques se mettent à jour automatiquement
- L'application se ferme
- Les données sont supprimées

Valider

Exact ! C'est l'un des principes de Shiny : les résultats réagissent automatiquement aux actions de l'utilisateur.