**Question 1 :**

Parmi les propositions suivantes, cochez celles qui correspondent à une variable qualitative :

* l’âge quinquennal (0-4 ans, 5-9 ans, etc.)
* le statut tabagique (fumeur, non-fumeur, ancien fumeur)
* tranches de revenu du foyer
* niveau de sévérité d’une maladie coté 1 (« peu sévère ») à 4 (« très sévère »)
* température en Celsius

**Question 2 :**

Voici la distribution des durées d’hospitalisation dans un service hospitalier.

Quelles statistiques descriptives sont les plus adaptées pour la décrire ?



* ] moyenne
* médiane
* écart-type
* variance
* espace inter-quartile

**Question 3 :**

Vous réalisez une étude sur la consommation des médias sociaux par les étudiants à l’Université. Vous vous intéressez notamment au nombre d’heures passées quotidiennement par les étudiants sur ces applications ou sites. Voici les premières données collectées :

[2; 4; 1; 2; 3; 1]

Parmi les propositions suivantes concernant ces données, lesquelles sont vraies ?

* la moyenne est de 2
* la médiane est de 3
* la proportion d’étudiant déclarant passer plus de deux heures par jour sur les médias sociaux est de 1/3
* le ratio « >= 2h » / « < 2h » est de 2
* l’étendue (max – min) est pour le moment de 2,5

**Question 4 :**

Vous vous intéressez au suivi épidémiologique du COVID-19 en France. Quel indicateur de morbidité (prévalence/incidence) vous semble le plus adapté à cette maladie ?

(Justifiez votre réponse)

**Question 5 :**

La figure ci-dessous représente le suivi et le devenir de personnes incluses dans une étude de cohorte. Quel est le taux d’incidence des décès ?



**Question 6 :**

Parmi les propositions sur la figure ci-dessous, laquelle ou lesquelles sont vraies ?



* il s’agit d’un diagramme de flux
* il s’agit d’un DAG
* il n’y a pas de collision
* il faut ajuster sur l’estime de soi
* il n’est pas nécessaire d’ajuster

**Question 7 :**

Dans le cadre d’une étude cas-témoins en cours, il est nécessaire de retourner dans le dossier des cas. Ces cas étant anciens, les dossiers médicaux sont en format papier et stocké dans une archive mal rangée.

Pour récupérer les dossiers, vous décidez de tirer au sort un nombre X entre 1 et 6, de prendre le Xème dossier sur l’étagère de gauche et de prendre ensuite un dossier sur 6 dans l’ordre jusqu’à en avoir 200.

Parmi les propositions suivantes concernant votre méthode d’échantillonnage, laquelle ou lesquelles sont vraies ? Répondez en fonction des méthodes vues en cours.

* il s’agit d’un échantillonnage non aléatoire
* il s’agit d’un sondage systématique
* il y a environ 1200 dossiers dans la salle d’archive
* la fréquence d’échantillonnage est de 6
* il s’agit d’un sondage en grappes

**Question 8 :**

Parmi les propositions suivantes, lesquelles correspondent à des méthodes prises en compte des biais de confusion ?

* la randomisation
* l’échantillonnage aléatoire
* l’analyse multivariable (ajustement)
* l’évaluation en aveugle
* la puissance statistique

**Question 9 :**

On s’intéresse à la survie des personnes atteintes de cancer du rein. Pour optimiser la puissance de l’étude, cette dernière portera à la fois sur des patients prévalents (qui ont déjà la maladie au début de l’étude) et incidents (qui rejoignent l’étude au fur et à mesure quand le diagnostic est posé). Le problème posé par cette approche est que les patients prévalents sont forcément des patients qui ont survécu jusqu’à la date de début d’étude et ont donc une présentation moins sévère de la maladie que certains patients incidents. **Il s’agit d’un biais de survie sélective.**

De quel type de biais s’agit-il ?

* un biais de confusion
* un biais d’information
* un biais de sélection
* un biais de mémoire

**Question 10 :**

À quoi correspond l’acronyme POOLS pour la définition des critères d’éligibilité ?

* Puissant
* Objectif
* Ordonné
* Logique
* Scientifique