

# **Master 1- Informatique**

## **Informatique Décisionnelle et Statistique**

### **Logiciel SAS**

### **Travaux dirigés**

[Afo\\_stat@yahoo.fr](mailto:Afo_stat@yahoo.fr)

#### **Exercice VI :**

Travail à rendre au plus tard le 31 décembre 20h00.

Fournir impérativement dans un fichier Zippé nommé TD6\_XXX.zip à déposer sur la plateforme Moodle où XXX représente les initiales des personnes de votre groupe.

Exemple pour Alain Follet / Jean Mouffle : Ex6\_AFO\_JMO.zip.

Les fichiers Hospi\_2019.xlsx, Hospi\_2018.xlsx et Hospi\_2017.xlsx contiennent des informations sur les structures et les activités de soins des établissements hospitaliers public et les établissements privés. Chaque fichier contient 4 onglets :

- Indicateur (description des indicateurs)
- Etablissement (contient les informations sur les établissements de santé)
  - finess : identifiant des établissements de santé
  - rs : raison sociale
  - cat : catégorie
    - CH : Centre Hospitalier
    - CHR : Centre Hospitalier Régional
    - CLCC : Centre de Lutte Contre le Cancer
    - CLI : Clinique
    - EBNL : Etablissement à But Non Lucratif
  - taille\_MCO : taille de l'établissement MCO (Médecine / Chirurgie / Obstétrique)
  - taille\_M : taille Médecine
  - taille\_C : taille Chirurgie
  - taille\_O : taille Obstétrique
- Lits et places : nombre de lits et places
- Activité globale : divers indicateurs de structure et d'activité

Il est conseillé de lire l'ensemble des questions et de regarder la structure des fichiers sources afin de constituer correctement vos bases de travail.

- 1°) Importez les données et familiarisez-vous avec les données.
- 2°) Décrivez les données (analyse, problème de qualité, outliers, ...).
- 3°) En 2019, existe-il des établissements nouveaux ou non présents par rapport à 2018 et / ou 2017 ?
- 4°) Est-ce que certains établissements ont changé de taille (taille\_MCO , taille\_M, Taille\_C et taille\_O) sur la période 2017-2019 ? Est-ce que des établissements ont changé d'activité ?
- 5°) A l'aide du nombre de lits, essayez de déterminer quels sont les seuils qui ont été utilisés pour constituer les variables taille\_M, taille\_C et taille\_O (variables catégorielles). Justifiez votre raisonnement.
- 6°) Indiquez dans un tableau croisé par catégorie d'établissement (cat) et par année, le nombre d'établissement, le nombre de lits en MCO en détaillant par Médecine / Chirurgie / Obstétrique.
- 7°) Les deux premiers caractères du N° finess correspondent au département. Indiquez le nombre d'établissement par catégorie (cat) par région.
- 8°) Indiquez dans un tableau le nombre d'accouchement, le nombre de lit d'obstétrique par catégorie d'établissement (cat) et par niveau de maternité. Quels sont les 5 établissements avec la plus forte activité d'obstétrique ?  
Ce tableau n'a de sens que pour les établissements pratiquant des actes d'obstétrique.
- 9°) Résumez les indicateurs de la question 8 par région. Indiquez également le nombre min et max d'accouchement.
- 10°) A l'aide de variables présentes dans le fichier essayez de calculer un score de « qualité » pour les accouchements (obstétrique) afin de classer les établissements. Justifiez le calcul de votre score et proposez une interprétation de votre score.  
A l'aide de votre score, indiquez quels sont les 5 meilleurs et 5 moins bons établissements.
- 11°) Etudiez la stabilité de votre score sur les années 2017, 2018 et 2019. Par exemple, constatez-vous que les établissements qui ont un score qui n'est pas stable s'accompagne de changement dans les activités et les structures ? De même, une stabilité dans le score doit se traduire par une stabilité dans les activités et les structures. Donnez des exemples concrets.
- 12°) Calculez votre score par région et par année. Précisez comment vous passez d'un score calculé par établissement à un score calculé par région / année. Existe-il des différences entre les régions ? Quelles sont les limites de votre score ? Indiquez comment pourriez-vous l'améliorer ?