Présentation des modules

**Objectif :** À la fin de ces modules, vous serez capable d’importer des données, les travailler en tableaux et graphiques et sortir un rapport 100% automatisé.

**Déroulement :** Chaque séance sera composée d'un cours théorique avec des slides et des petits exemples qui seront disponibles sur le réseau \\nas-01\SBE. Puis à la fin des sessions, nous prévoyons de faire un exercice d'application sous le format "code-with-me", qui consiste à coder ensemble et en temps réel.

**Module 1 - Nusaïbah - Introduction au langage de programmation R et à l'interface RStudio**

Le but de ce module est de connaître R et RStudio, de vous familiariser avec le vocabulaire de langage de programmation et d'être autonome à la fin du module.

Introduction de l'interface RStudio et ses fonctionnalités. Puis, les bases du langage R : les packages, les opérateurs, les types de variables et les fonctions usuels.

**Module 2 - Charlotte – Manipulation de données**

Le but de ce module est de savoir importer différentes tables de données (.xlsx, .sas, .xpt, .csv, .txt …). Puis, manipuler les données : sélectionner des colonnes, filtrer des lignes, réaliser des jointures, créer des variables dérivées, manipuler des dates, faire une recherche de texte.

**Module 3 - Nusaïbah – Data visualisation avec ggplot2**

Le package ggplot2 sera l’objet du module. Initialisation, les types de représentations, personnalisation (taille, couleurs, légende, titre).

Présentation des packages complémentaires, compatibles à ggplot2 :

* plotly qui génère des graphiques intératifs ;
* ggsci qui génère des palettes de couleurs adaptés aux revues scientifiques.

**Module 4 - Dan - Analyse statistique - description et modélisation**

**4.1 analyse descriptive**

Présentation du package crosstable. Fonction du même nom que le package qui calcule des statistiques (moyenne, médiane, IQR, min, max, nombre d’observations, nombre de NA).

**4.2 analyse analytique**

Présentation des différents modèles statistiques, régression linéaire, régression logistique, la survie (KM et Cox). Présentation de fonctions qui vérifient des hypothèses de linéarité. Représentation des résultats de modèles dans R.

**Module 5 - Dan - Recherche clinique**

Présentation du package EDCimport (import des données, vérification/nettoyage des données, répresentation graphique des données).

Présentation du package grstat pour la description des tables AE.

**Module 6 - Charlotte - Reporting : Officer**

Programmer un rapport automatisé sur R avec le package officer. Détails des paramètres du package : ajout de paragraphe, table, figure, personnalisation des titres, légendes, notes, la mise en page, l’ajout des références.