

# Epi-Σûre

## Journée de formation 2018

### Manufacture des tabacs - NANTES

- **Formateur** : B. BRANGER, pédiatre, épidémiologiste, enseignant et formateur en statistiques et épidémiologie depuis 1985 (CHU Nantes, CHU Rennes, Réseau de périnatalité des Pays de la Loire.), retraité depuis 2015. Mail [branger44@gmail.com](mailto:branger44@gmail.com) et Tél 06 32 70 33 80 à Nantes.

#### - Pré-requis (à faire avant la formation ++++ pour ne pas perdre de temps)

- Amener un ordinateur portable sous Windows avec une prise multiple +++, et si possible une souris. EPIDATA est opérationnel sous Mac ou Linux pour constituer un masque de saisie, mais pas encore pour l'analyse. Les Mac devront télécharger **PlayOnMac** sur le site <https://www.playonmac.com/fr/> qui émule Windows sur Mac sans avoir besoin de télécharger Windows.
- **Télécharger Epidata** sur [www.epidata.dk](http://www.epidata.dk): page *download*: EPIDATA ENTRY 3.1 en français (*setup\_epidata\_fr.exe*) et EPIDATA ANALYSIS 2.2.2 (*setupepidatastat.exe*) sur le bureau de l'ordinateur. L'installation complète se fera en séance. Ne pas télécharger les *Epidata Entry client* ni *Epidata Entry manager*.
- **Télécharger SPSS**, version d'évaluation, valable 15 jours : <https://www.ibm.com/fr-fr/marketplace/spss-statistics> (il faut créer un compte)

#### - Programme en séance

- **Documents de référence remis sur place** – Utilisation d'un fichier-type de données réel. Pas de formule mathématique : rien que du raisonnement +++
- **Introduction à la saisie** (SPSS n'est pas adapté à la saisie)
  - Le questionnaire « papier »
    - ✓ Questions
    - ✓ Réponses : qualitatives (codées), quantitatives (unités), ordinales, date, texte....
  - Questionnaire EpiData : le masque de saisie (\*.qes) et le contrôle à la saisie (\*.chk) : penser avant d'agir +++. Création du fichier de saisie (\*.rec)
  - Comment gérer la saisie sur Excel ? *On peut exporter les fichiers \*.xls et \*.xlsx pour constituer un fichier EpiData, ou un fichier SPSS en \*.sav*
- **Analyse : description et comparaison, autres analyses**

SPSS	EpiData
<b>= Description des données</b> <ul style="list-style-type: none"><li>• Lire ou importer un fichier : Onglet <i>Fichier</i></li><li>• Etablir des pourcentages (proportions) pour les données qualitatives : <i>Analyse/ Statistiques descriptives/Fréquences</i></li><li>• Décrire des données quantitatives (moyenne, médiane, écart-type...) : idem, ou <i>Analyse/Statistiques descriptives/Effectifs</i>, ou <i>Analyse/Comparer les moyennes/ Moyenne</i></li><li>• Définir des nouvelles variables : <i>Transformer/ Calculer la variable</i>, et constituer des seuils, ou des groupes <i>Transformer/Regroupement visuel</i></li><li>• Quelques graphes : <i>Graphiques/..... Histogramme, graphe en points, barres, boîtes à moustaches, barres d'erreur</i></li><li>• Courbes de survie : <i>Analyse/survie/Kaplan-Meier</i></li></ul>	<b>* Description des données</b> <ul style="list-style-type: none"><li>• Lire ou importer un fichier (<i>read</i>) . Liste des variables : F3</li><li>• Pourcentage pour les données qualitatives ; notion d'IC à 95 % (<i>freq varqual /c</i>)</li><li>• Comment décrire des données quantitatives (moyenne, médiane...; les classes et les seuils (<i>means varquant</i>))</li><li>• Quelques graphes : <i>pie, bar, histogram, boxplot</i></li></ul>

SPSS	EpiData
<p><b>= Comparaison (analyse univariée) : les tests statistiques sans formules et avec l'interprétation</b></p> <ul style="list-style-type: none"> <li>• Comparaison de deux (ou plus) pourcentages, Pourcentages en lignes ou en colonnes, test de <math>\chi^2</math> (ou Fisher), <i>Analyse/Statistiques descriptives/tableaux croisés</i></li> <li>• Risque (risque relatif et odds ratio), test de kappa : <i>idem</i></li> <li>• Comparaison de deux moyennes : test t de Student et Wilcoxon (Mann-Whitney) : <i>Analyse/Comparer les moyennes/Tests t</i> pour échantillons indépendants</li> <li>• Comparaison de plusieurs moyennes : ANOVA et Kruskal-Wallis : <i>Analyse/Comparer les moyennes/ANOVA à 1 facteur</i>, ou <i>Analyse/Tests non paramétriques/ Echantillons indépendants</i></li> <li>• Tests non paramétriques si la taille de l'échantillon est faible ou si la répartition n'est pas "normale" (voir <i>Analyse/Statistiques descriptives/Fréquences/ Kurtosis ou Skewness</i>, et <i>Graphiques/histogramme/Afficher courbe gaussienne</i>), puis <i>Analyse/Tests non paramétriques/Echantillons indépendants/Paramètres = choix des tests</i></li> </ul> <p><b>= Comparaison (analyse multivariée) : tenir compte de tous les facteurs, lever les facteurs de confusion</b></p> <ul style="list-style-type: none"> <li>• Le critère de jugement est qualitatif binaire (0 et 1) : régression logistique : <i>Analyse/Régression/Logistique binaire</i>. Mode complet ou ascendant ou descendant. Cocher <i>Options/CI pour exp(B)</i> (c'est l'OR avec son intervalle de confiance)</li> <li>• Le critère de jugement est quantitatif (ex : durée accouchement), et les variables explicatives sont quantitatives aussi : régression linéaire multiple : <i>Analyse/Régression linéaire</i>. Le critère de jugement est quantitatif et les variables explicatives sont soit quantitatives soit qualitatives : ANOVA : <i>Analyse/Comparer les moyennes/ANOVA à 1 facteur</i>, ou <i>Analyse/Comparer les moyennes/Modèle linéaire général/Univarié</i> (mauvaise appellation...)</li> <li>• Le critère de jugement est une durée de survie : modèle de Cox : <i>Analyse/Survie/modèle de Cox</i>. Idem <i>Options/CI pour exp(B)</i> pour <i>Hazard Ratio</i></li> </ul> <p><b>* Autres commandes fréquentes</b></p> <ul style="list-style-type: none"> <li>• Coefficient de corrélation : <i>Analyse/Corrélation bivariée</i></li> <li>• Valeur diagnostique : Courbes ROC : <i>Analyse/Courbes ROC</i></li> <li>•</li> </ul> <p><b>* Utilitaires</b></p> <ul style="list-style-type: none"> <li>• Afficher le fichier en format tableur ou avec les noms des variables. <i>Définir un label</i> pour expliciter les codages des variables (1 → masculin)</li> <li>• Sélection d'une partie d'un fichier : <i>Données/ Sélectionner des observations</i></li> <li>• Sauvegarde fichiers, données : <i>Fichier/ Enregistrer sous...</i></li> <li>• Sauver les sorties de l'analyse : fichier en *.spv : <i>Fichier/Enregistrer</i>. Possibilité d'export en format Word : <i>Fichier/Exporter</i></li> </ul>	<p><b>* Comparaison : les tests statistiques sans formules : lesquels, comment les interpréter ?</b> Commande suivie de /t</p> <ul style="list-style-type: none"> <li>• Comparaison de deux pourcentages ou plus : <math>\chi^2</math> (et Fisher pour deux pourcentages) : <code>tables varqual1 varqual2 /c /t /ex</code></li> <li>• Comparaison de deux moyennes : test t de Student et Wilcoxon (Mann-Whitney) : <code>means varquant1 varqual /t ou kwallis varquant varqua</code></li> <li>• Comparaison de plusieurs moyennes : ANOVA et Kruskal-Wallis : <code>means varquant varqual /t ou kwallis varquant varqual</code></li> </ul> <p><b>* Autres commandes fréquentes</b></p> <ul style="list-style-type: none"> <li>• Risque relatif et odds ratio : <code>tables varqual1 varqual2 /rr /sa ou /o /sa</code></li> <li>• Coefficient de corrélation : <code>correlate var1 car2</code></li> <li>• Régression linéaire : <code>regress vary var</code>, graphe en points : <code>scatter varx vary</code></li> <li>• Commandes fréquentes : <code>select, define newvar #, if... then..., savedata</code></li> <li>• <code>Savedata</code></li> <li>• Retrouver fichiers de sortie : <code>EAoutput.htm</code> et <code>temp.pgm</code></li> </ul>