



Dossier suivi par : Sandra CHIRAZI  
Tel. : +33 (0)2 99 05 32 42  
Sandra.chirazi@ensai.fr

**Objet PROJETS  
STATISTIQUES**

Bruz, le 30 août 2017

## Encadrer des projets statistiques de 2<sup>e</sup> année à l'Ensaï

L'Ensaï fait appel à des praticiens de la statistique pour encadrer des projets statistiques de 2<sup>e</sup> année. Cette activité pédagogique a pour objectif l'application, voire l'approfondissement, des connaissances statistiques acquises.

### Travail des élèves

Les élèves travaillent par groupes de 4 (voire 3) sur le sujet qu'ils ont choisi, sous la direction du praticien qui l'a proposé. Ils ont à circonscrire le problème soumis et à le traiter durant le temps imparti (de fin janvier à fin avril), en utilisant les concepts, méthodes et outils appropriés.

Chaque groupe rend compte de son travail sous forme d'un rapport écrit et d'une soutenance.

La quantité de travail que l'on peut attendre des élèves est de l'ordre de 4 heures hebdomadaires par élève, compte tenu de leur charge de travail extérieure au projet (ils continuent parallèlement à suivre les cours de l'école).

### Contenu statistique du projet

Pour répondre à la problématique, les élèves doivent utiliser leurs connaissances en séries temporelles, théorie des sondages, analyse de données, régression linéaire, régression sur variables catégorielles, économétrie.... D'autres outils peuvent être utilisés. Le tuteur s'engage alors à les présenter aux étudiants.

### Rôle de l'encadrant

Le tuteur amène un sujet qui se compose de :

- une problématique
- un ou plusieurs jeu(x) de données qui doit être structuré, de volume raisonnable et documenté si besoin.
- des éléments de bibliographie nécessaires aux étudiants pour contextualiser l'analyse et discuter les résultats obtenus lors de la rédaction du rapport final.

Campus de Ker Lann  
rue Blaise Pascal BP37203  
35172 BRUZ CEDEX  
Tél : 33 (0)2 99 05 32 32  
communication@ensai.fr

L'encadrement d'une telle activité nécessite de la part de l'encadrant une bonne connaissance des données qu'il fournit aux élèves et une bonne maîtrise des techniques statistiques de base utilisées pour leur traitement.

Pour des raisons d'organisation, chaque tuteur est incité à proposer plusieurs problématiques différentes, donc à encadrer les travaux de plusieurs groupes d'élèves. Les problématiques peuvent s'appuyer sur le même jeu de données.

### **Calendrier**

- D'octobre 2017 à fin novembre 2017, les futurs animateurs proposent des sujets complets (avec les jeux de données). L'Ensaï valide les sujets.
- En janvier, le recueil des sujets proposés est diffusé par l'Ensaï aux élèves, qui forment librement leur groupe. Ils émettent une liste de choix mi-janvier.
- Lors de la première séance de suivi, le tuteur rencontre les groupes qu'il encadrera au cours d'une réunion de présentation.
- Pour la suite du projet, chaque groupe d'élèves rencontre le tuteur de son projet quatre fois, dans le cadre de séances de « suivi de projet ». Deux fois au plus, ces suivis peuvent être organisés en visioconférence.
- Début mai, les élèves remettent leur rapport.
- Les soutenances des projets statistiques ont lieu en seconde quinzaine de mai. Chaque groupe présente son rapport devant un jury composé d'un président (permanent de l'Ensaï), d'un expert en communication et du tuteur.

### **Si vous souhaitez encadrer un ou plusieurs projets**

Pour chaque proposition de sujet, vous devez présenter, en une page, le thème de l'étude, la problématique, le type de données, ainsi que quelques éléments de bibliographie. Un exemple de proposition est joint en annexe.

Envoyez le tout à **Sandra Chirazi**, Enseignante en statistique ( [sandra.chirazi@ensai.fr](mailto:sandra.chirazi@ensai.fr) )

### **Aspects informatiques (données et logiciels)**

La totalité des données à traiter doit être disponible avant le début du projet. La taille des bases de données est limitée à 100 Mo. Si vous jugez cette taille insuffisante pour votre projet, vous devez me le signaler afin que l'on adapte le dispositif.

Les fichiers de données devront parvenir à l'Ensaï sur support informatique avant la fin de l'année 2017. Pour éviter tout problème de confidentialité, les données seront stockées dans un répertoire dont l'accès est limité aux seuls élèves du groupe concerné et le responsable des projets. A ce propos, tous les élèves signeront une charte de confidentialité avant le démarrage des projets.

Pour le traitement informatique des données, l'utilisation de Sas ou R est obligatoire.

### **Rémunération des intervenants**

Les projets sont rémunérés au forfait. Les éventuels frais de déplacement sont pris en charge par l'Ensaï.

## Exemple de projet de 2<sup>e</sup> année proposé

<b>Sujet - Réalisation d'une segmentation des Médecins traitant les patients atteints de cancer de la prostate</b>
--

Entreprise : Janssen
----------------------

### 1. Contexte

Janssen est la filiale pharmaceutique du groupe américain Johnson & Johnson, le 1er groupe de santé mondial. Fondée en 1953 par le docteur Paul Janssen pour créer des médicaments améliorant la qualité de vie, Janssen est aujourd'hui reconnu dans des domaines thérapeutiques hautement spécialisés comme l'oncologie, les maladies du système nerveux central, l'infectiologie, l'immunologie ou l'hématologie. La filiale française est la 1ere du groupe du groupe avec un chiffre d'affaires de 912 millions d'euros en 2012. Au sein de la Direction de la Stratégie Client, nous avons pour objectif d'améliorer la connaissance de nos clients par la mise en place de méthodes et projets permettant d'appréhender au mieux leurs besoins, de proposer une offre adaptée afin d'améliorer les prestations et les services proposés à nos clients.

### 2. Sujet proposé

L'objet de ce projet est mettre en œuvre des méthodes statistiques permettant d'obtenir une segmentation pertinente composée de groupes de clients homogènes en termes de caractéristiques, comportements, besoins et attentes, et ce pour notre clientèle d'oncologues, d'urologues et les radiothérapeutes. La segmentation doit permettre d'effectuer des actions marketing différenciées en fonction des segments, d'entrer en relation avec chacun de nos clients de la manière la plus appropriée qu'il soit en fonction de ses appétences et de ses attentes.

### 3. Méthodologie envisagée

Méthodologies d'études et de traitement statistiques relatives aux statistiques exploratoires multivariées. Outils de datamining : SAS, SPAD

### 4. Résultats attendus

Le rapport du projet devra comporter, les statistiques descriptives préalables, la description de la méthodologie d'étude (choix de la méthode pour l'imputation des données manquantes si nécessaire, choix du type de classification...) et le résultat de la segmentation réalisée par une description qualitative des particularités de chacune des classes.

La Sensibilité client devra être un axe majeur dans la construction de cette segmentation. Une attention particulière doit être portée sur la mise en perspective des data afin d'apporter des éléments de réflexion qui orienteront la stratégie client.

### 5. Bibliographie

Probabilités, analyses des données et statistiques, Gilles Saporta  
Data mining et statistique décisionnelle, Stéphane Tuffery  
Etude de cas en statistique décisionnelle, Stéphane Tuffery