

## Tronc commun de statistiques de 2ème année

### Rédaction des travaux personnels

Les travaux personnels ont pour objectif de mettre en œuvre les méthodes statistiques présentées dans le cours sur un exemple concret. Les exemples sont constitués de jeux de données réelles issues de divers domaines et disciplines. Il est recommandé d'explorer largement le problème (y compris par une recherche bibliographique). Ces travaux seront réalisés en binôme.

### Suivi

Il existe une vingtaine de projets différents. À l'issue de votre choix de sujet, on vous communiquera courant septembre le nom de votre enseignant référent de votre projet (potentiellement différent de votre enseignant de Cours/TD) pouvant être sur le site de Claude Bernard ou de Massy.

**Pour vous aider à démarrer votre projet**, une séance d'encadrement obligatoire (toute absence sera sanctionnée) sera organisée par votre enseignant de cours, par demi groupe de TDs. Cette séance a pour objectif de vous aider à charger correctement vos données (si ce n'est pas déjà fait) ainsi qu'à faire les premiers pas dans l'analyse descriptive.

**Lorsque le projet est bien avancé**, vous devez prendre rendez-vous avec votre enseignant référent afin de faire le point sur l'avancement de vos projets (ce qui a été fait, ce qui mérite d'être investigué davantage, les questions que vous avez et pistes que vous envisagez et les points importants pour la rédaction). Ce rendez-vous est obligatoire. Afin de rendre cette séance de suivi rentable pour vous, il est important de l'avoir préparée (d'avoir bien exploré les premières pistes, d'avoir préparé toutes vos questions, etc ...). Ces rencontres se feront par visioconférence, skype ou en direct selon les possibilités. Pour des rencontres en visioconférence ou skype, il est préférable d'envoyer des documents ainsi que le code R si nécessaire. Attention, il est nécessaire de programmer ce rendez-vous le plus tôt possible pour trouver un créneau dans des emplois du temps assez contraints de part et d'autre.

### Rapport

Le rapport doit prendre la forme d'un document rédigé de 15 pages maximum, hors annexes et doit comporter les éléments suivants à organiser de la manière qui vous semble la plus pertinente :

- une **introduction** courte et ciblée présentant la cadre général du projet et les questions que vous voulez étudier.
- une **partie descriptive** présentant les données et leurs caractéristiques par exemple sous forme de graphes ou de résumés statistiques. Cette partie, souvent négligée, a pour but de bien appréhender vos données (bien les connaître, détecter des anomalies comme des valeurs aberrantes, données manquantes, etc ....) et surtout elle permet de faire émerger les questions qui seront traitées dans la suite.
- une **partie d'analyse** : pour chaque question identifiée, vous
  - écrierez sous forme mathématique propre le modèle statistique qui vous permettra d'y répondre,
  - présenterez les résultats pertinents et conclurez sur la question posée.Il ne s'agit pas de chercher le modèle le plus complet possible mais plutôt des modèles permettant de répondre aux questions posées.
- une **conclusion** résumant les principaux résultats obtenus et leurs conséquences.

Certains sujets possèdent un nombre important de variables. Votre étude ne nécessite pas obligatoirement l'utilisation de toutes.

Les sorties présentées doivent permettre d'illustrer la démarche et de justifier les conclusions. Les graphiques et analyses statistiques pertinents doivent figurer dans le corps du rapport. Les analyses intermédiaires peuvent être simplement citées. Si la même démarche est appliquée plusieurs fois sur des jeux de données différents, il est inutile de la détailler systématiquement.

Les programmes R doivent être joints en annexe et référencés dans le corps du document. Les annexes ne doivent contenir aucun autre élément (ni sortie R, ni graphique).

## Source et références

Il est conseillé de rechercher des informations additionnelles sur le sujet dans diverses sources pertinentes. Toute source utilisée doit être citée et référencée de façon précise.

## Rendu du rapport

Les rapports doivent être rendus pour le lundi 22 Janvier 2018 à la direction des études du site sous forme papier et déposés dans la rubrique travaux du site [https://tice.agroparistech.fr/coursenligne/courses/24769/?id\\_session=0](https://tice.agroparistech.fr/coursenligne/courses/24769/?id_session=0) sous forme d'un fichier pdf (indiquer le numéro du projet et vos noms explicitement). Chaque jour de retard sera sanctionné de 0.5 point.

## Évaluation

L'évaluation du projet, fondée sur le rapport, se fait selon les cinq critères suivants :

- Présentation, rédaction et intérêt de l'étude ;
- Articulation de la modélisation par rapport au raisonnement biologique ;
- Précision statistique (écriture des modèles, formulation des hypothèses, etc.) ;
- Choix des sorties listings et des graphiques ;
- Interprétation biologique des résultats (cette partie est très souvent négligée or c'est elle qui donne le sens à l'analyse statistique proposée).

Toute fraude caractérisée (comme par exemple la copie totale ou partielle d'un autre rapport, passé ou présent) sera sanctionnée par une note éliminatoire.