

Cours LSTAT2170: Consignes pour la preparation des projets

Les projets sont individuels. Il est demandé d'effectuer une analyse de données réelles justifiées par les techniques étudiées au cours et aux travaux pratiques, et d'établir une prévision de longueur égale à 10% de la longueur de la série initiale (et comprenant au minimum une période complète en cas d'une saisonnalité). Chaque étudiant devra remettre un rapport écrit **individuel**.

Le rapport ne devra pas excéder quinze pages, sans compter les annexes. Le nombre de pages en annexes n'est pas limité. Veillez à ce que la taille de vos figures soit suffisamment importante, au cas nécessaire utilisez l'annexe.

Veillez aussi à imprimer un exemplaire (sans annotations) à votre intention: il sera utile lors de l'examen oral.

Vous trouverez ci-dessous **quelques conseils utiles** pour la rédaction du rapport et une liste de **FAQ** non exhaustive.

Conseils utiles :

Le rapport sera notamment apprécié sur les aspects suivants:

- Présentation préliminaire des données (avec commentaires éventuels sur sa tendance générale: y a-t-il des ruptures? des éléments surprenants dans la variance? ...).
- Justification d'une transformation préliminaire éventuelle des données.
- Motivation intuitive et justification statistique d'un modèle avec saisonnalité ou tendance éventuelles (il est par contre inutile de recopier de la théorie du cours....)
- Lors de l'analyse de la composante stationnaire, un graphe des acf et des pacf, et une interprétation (attention au modèle SARIMA !)
- Justification du choix d'un modèle, estimation de ses paramètres et écriture du modèle final, avec les coefficients estimés; à l'aide des outils suivants :
 - Test de signification des coefficients.
 - Critères automatiques de sélection de modèle : comparaison de la log-vraisemblance et de l'AIC.
 - Analyse des résidus de l'ajustement des modèles: interprétation des sorties graphiques, test de Portmanteau (Ljung-Box).
 - Analyse des résidus: y a-t-il une structure de corrélation sur le carré des résidus? Application éventuelle du module garch pour la modélisation des innovations (Attention aux potentiellement faibles significativités des (P)ACF des résidus au carré – appliquer votre compréhension intuitive pour décider si un modèle (G)ARCH est vraiment nécessaire)
- Utilisation du critère de l'erreur de prédiction sur les dernières 20% des observations pour comparer des modèles (au cas où c'est possible par la nature des données).
- Si la nature des données (ruptures, tendance non paramétrique, ...) le propose, comparaison de plusieurs modèles non paramétriques, ou de modèles non paramétriques avec un modèle paramétrique (c'est-à-dire dans lequel la tendance est estimée paramétriquement) ou, encore, comparaison entre prédiction à la Box-Jenkins et un lissage exponentiel. Cette comparaison peut être effectuée grâce à l'erreur de prévision.

- Au cas où la nature des données est essentiellement du type (G)ARCH, application de la méthodologie pour ce type des données (avec l'adaptation des analyses discutées ci-dessus à ce cas...). Attention aux problèmes des p-valeurs significatives pour des longues séries : comparer avec des tests aux niveaux plus petits que 5%)
- Prévvision de longueur égale à 10% de la longueur de la série initiale. Graphique de cette prévvision y inclus les intervalles de prévvision. Interprétation.

Rappels: les projets sont obligatoirement à soumettre par Moodle :

Ceci facilite largement le travail de correction pour les enseignants du cours. Merci de bien vouloir respecter cette consigne et ne pas soumettre votre travail par mail.

Pensez cependant de vous amener avec une copie vierge de toute annotation pour l'examen oral en juin.

Si, après un effort considérable d'analyse des données de votre projet vous avez un doute sur quel type de modèle à choisir, n'hésitez pas de demander une (breve) aide de vos enseignants, le plus tot possible.