

Projet de Séries Temporelles :

modélisation ARIMA de l'indice de production industrielle

À rendre :

- **au plus tard le 19 mai 2020** en le déposant dans [ce lien Dropbox](#) (pas d'inscription nécessaire)
- dans un dossier compressé (.zip de préférence) :
 - nommé “*pnom1_pnom2*” où “*pnom1*” est la 1ère lettre du prénom et le nom du membre 1 du binôme. Ajoutez “*_en*” à la fin si vous rédigez le rapport en anglais.
 - contenant **le rapport au format PDF, le script .R et la série utilisée au format CSV.**

Consignes générales : Ce projet doit être réalisé sous R. La notation tiendra compte de la rigueur dans la mise en œuvre des outils économétriques, de la concision et de la clarté de la présentation des résultats. Le rapport (6 pages de contenu maximum hors annexe, en français ou en anglais) fera figurer les programmes (commentés) en annexe. Ce tutorat doit être effectué seul ou (de préférence) en binôme.

Le barème (à titre indicatif) est le suivant : Partie 1 : 30%; Partie 2 : 30%; Partie 3 : 40%.

On s'intéresse à la modélisation et la prévision de l'indice de production industrielle observé en France. Il ne vous est pas demandé de connaître le processus de construction de l'indicateur. Vous ne travaillez que sur les données observées. À partir du répertoire des séries chronologiques de l'INSEE ([lien des séries de l'IPI](#)), vous devez choisir une série agrégée corrigée des variations saisonnières et des jours ouvrés (CVS-CJO), mensuelle, correspondant à n'importe quel secteur d'activité (à votre convenance) et contenant au moins 100 observations.

• Partie I : Les données

1. Que représente la série choisie ? (secteur, périmètre, traitements éventuels, transformation logarithmique, etc.)
2. Transformer si besoin la série pour la rendre stationnaire (différentiation, suppression de la tendance déterministe, etc.). Justifier soigneusement vos choix.
3. Représenter graphiquement la série choisie avant et après transformation.

• Partie II : Modèles ARMA

4. Choisir, en le justifiant, un modèle ARMA(p, q) pour votre série corrigée X_t . Estimer les paramètres du modèle et vérifier sa validité.
5. Exprimer le modèle ARIMA(p, d, q) pour la série choisie.

• Partie III : Prévision

On note T la longueur de la série. On suppose que les résidus de la série sont gaussiens.

6. Ecrire l'équation vérifiée par la région de confiance de niveau α sur les valeurs futures (X_{T+1}, X_{T+2}).
7. Préciser les hypothèses utilisées pour obtenir cette région.
8. Représenter graphiquement cette région pour $\alpha = 95\%$. Commenter
9. Question ouverte : soit Y_t une série stationnaire disponible de $t = 1$ à T . On suppose que Y_{T+1} est disponible plus rapidement que X_{T+1} . Sous quelles conditions cette information permet-elle d'améliorer la prévision de X_{T+1} ?