



TP1

Analyse de des données de deux parcs éoliens en Picardie

Travail à réaliser

Cette séance a pour objectif de vous faire découvrir l'utilisation des fonctions principales de statistique descriptive du logiciel R et a aussi pour but de vous permettre de vous entraîner à personnaliser les graphiques, utiliser les aides en ligne, les forums et les sites dédiés afin de développer de l'autonomie.

Réaliser ce travail dans un script et dans un document Rmarkdown (format html)

Importation et préparation des données :

- Importer les données : ProductionP1P7.csv à mettre dans un dataframe
- Nettoyer les données : enlever les données manquantes
- Transformer la classe de la variable parc en facteur
- Transformer (dans une autre colonne) le format de la colonne Date en format date puis dans des colonnes spécifiques séparer les composantes des dates et heures
- Réaliser deux dataframes : ProductionP1 et Production P7

Représentation des données :

- Pour le parc P1 représenter les données au cours du temps (avec l'option curseur zoom)
- Pour les parcs P1 et P7 déterminer les indicateurs numériques représentant les données
- Présenter les données sous forme graphique : boîtes à moustaches et histogrammes
- Par un graphique regarder la normalité des données P1 et P7
- Faire un histogramme comparant la production totale de courant de P1 et de P7
- Faire un histogramme comparant P1 et P7 par mois, chaque année (2019,2020) sur une ligne
- Pour l'année 2020, pour chaque mois, superposer P1et P7 sur un histogramme. Tableau de graphique 2 lignes, 3 colonnes)
- Faire un tableau et un histogramme pour P1 et pour P7, de la production totale de courant par mois. Faire figurer la valeur au-dessus de la barre de l'histogramme.