32 variables

+ Ville (7 différentes)

+ Date

+ Echéance

Gérer la variable des villes :

* 7 variables binaires
* 7 modèles

Gérer la variable de la date :

* Recoder la variable pour l’utiliser (quanti [0 ; 365] / scale)
* Ne pas la traiter (les autres variables doivent suffire à gérer la saison)
* Vérifier l’impact de la variable de la date sur le delta de correction (Y)

Gérer la variable de l’échéance :

* On ne la gère pas
* Utiliser la variable dans le modèle (attention : selon le modèle utilisé, il est parfois nécessaire que l’importance des variables utilisées soit équivalente)
* Faire un modèle de correction : premier modèle global, puis sur-modèle qui utilise la prédiction & la variable correctrice

Problème : variables manquantes !

* Imputation nécessaire (car présence de valeurs manquantes dans le jeu de test)
* Attention aux valeurs manquantes en bloc : ne pas imputer, mais supprimer

Si on scale les variables, il faut garder les paramètres utilisés, pour revenir à l’original pour la prédiction.