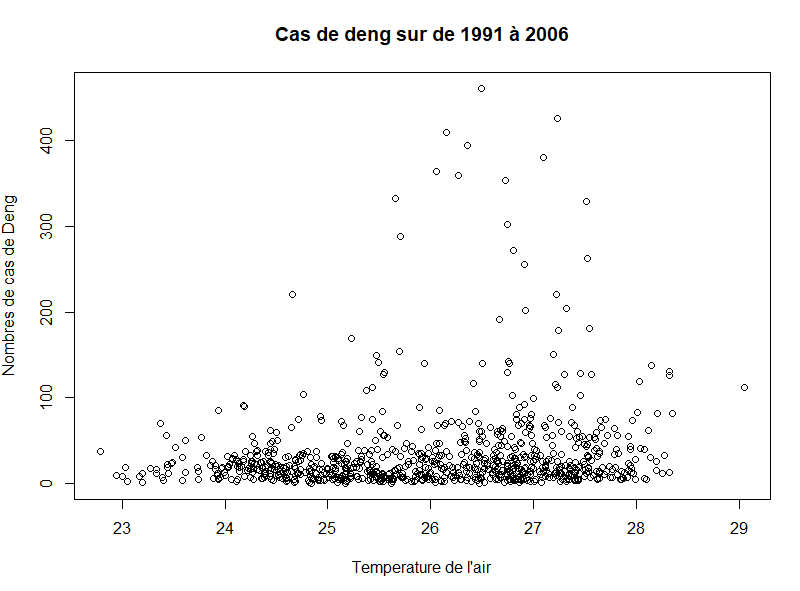
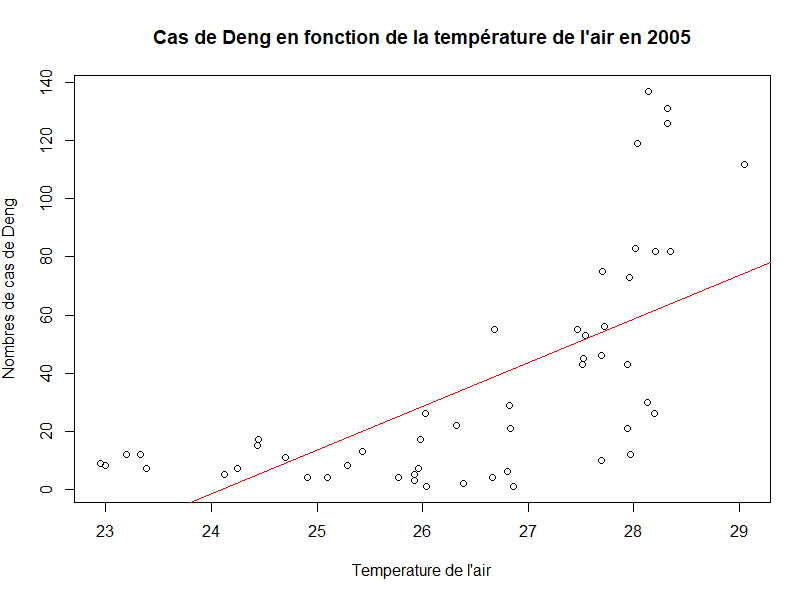
**Recherches sur la SAE – Régression linéaire sur les cas de Deng**

Pour commencer, nous avons importer les fichiers csv afin de pouvoir croiser quelques variables. En première variable, nous avons sélectionné la variable de la température maximum de l’air mais cela n’a rien donné à par un graphique chaotique. Ensuite, nous avons créer un dataframe afin de pourvoir mettre uniquement une variable avec les différents cas de Deng pour voir s’il y a un certain lien avec la variable ‘reanalysis\_air\_temp\_k’ ou plutôt la température de l’air.

Ensuite, nous avons préféré nous concentrer sur un **modèle linéaire**. Avant toute recherche, nous avons créé un dataframe avec uniquement la variable ‘reanalysis\_air\_temp\_k’ sur la période de 1991 jusqu’à 2006. Voici-ci-dessous, le nuage de points correspondant à cette période.

Par la suite, nous avons choisi de nous concentrer uniquement sur 1 période. Nous avons commencé nos analyses en partant de l’année 2006 jusqu’à descendre progressivement.

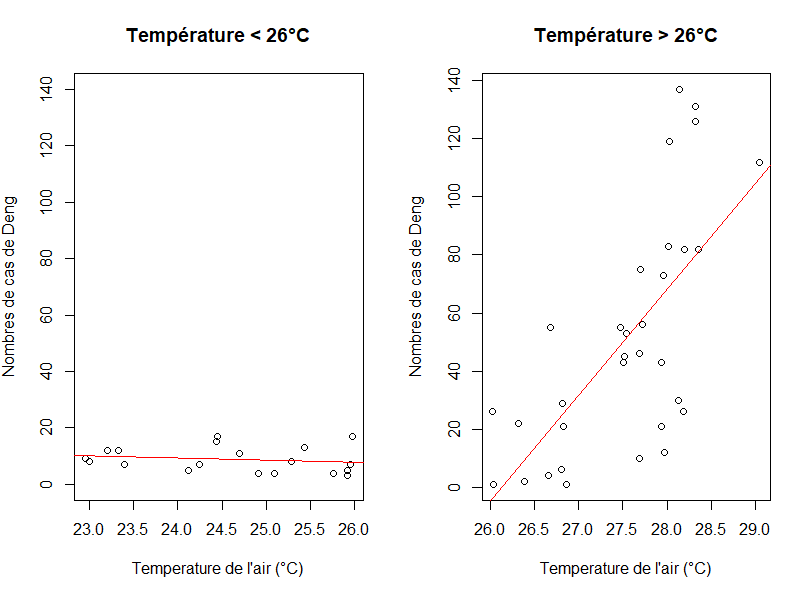
**Mais l’année 2005 a retenu notre attention de part son nuage de points :**



Avec un coefficient de détermination de 0,42, nous avons choisi de garder la variable ‘reanalysis\_air\_temp\_k’ uniquement sur l’année 2005.

On voit que ce graphique a 2 parties intéressantes : les cas de Deng pour la température inférieure à 26°C et les cas de Deng pour la température supérieure à 26°C. (A noter que l’on a converti la température en °C qui était au tout début en Kelvin pour une lecture plus parlante)

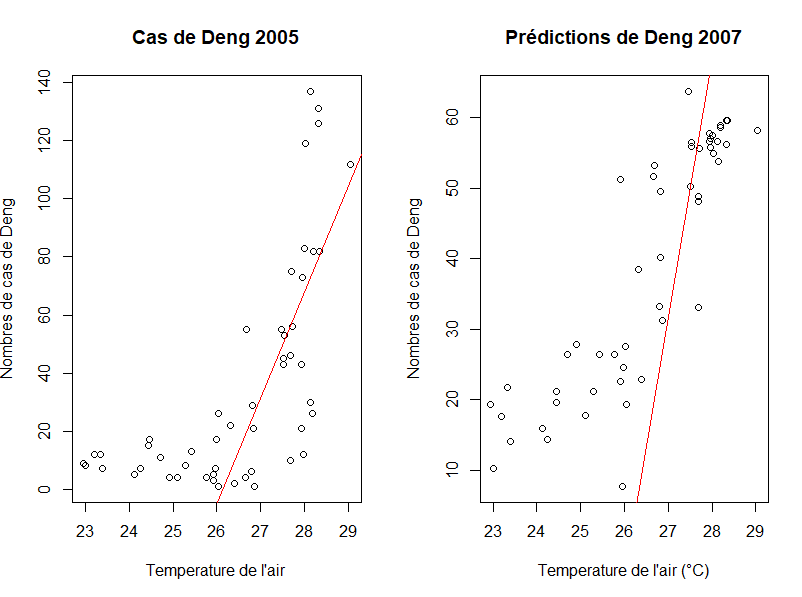
**Nous avons donc coupé le graphique en 2 parties :**



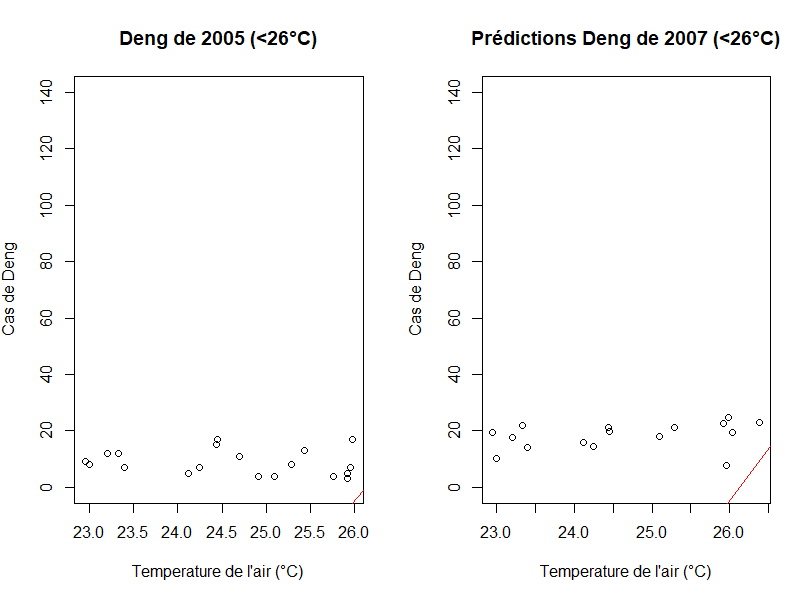
Après avoir choisi notre variable, nous avons commencé les prédictions de cas de Deng pour 2007. Nous avons noté que la semaine 52 de l’année 2005 est corrompue car elle s’appelle semaine 53 et la température de l’air n’existe pas. Donc nous avons choisi de la supprimer.

Après avoir prédit les cas de Deng, nous avons importer les prédictions dans 2 dataframe différents. Un dataframe pour les températures inférieures à 26°C et un dataframe pour les températures supérieures à 26°C.

A la fin, nous avons choisi de comparer les différents nuages de points.

 **Ci-dessous, c’est le nuage de points des cas de 2005 et des prédictions de 2007 :**

**Ensuite, voici le nuage de points pour les températures inférieure à 26°C des cas de 2005 et des cas de 2007 :**



**Pour finalement, comparer les nuages de points pour les températures supérieure à 26°C des cas de 2005 et des cas de 2007 :**

