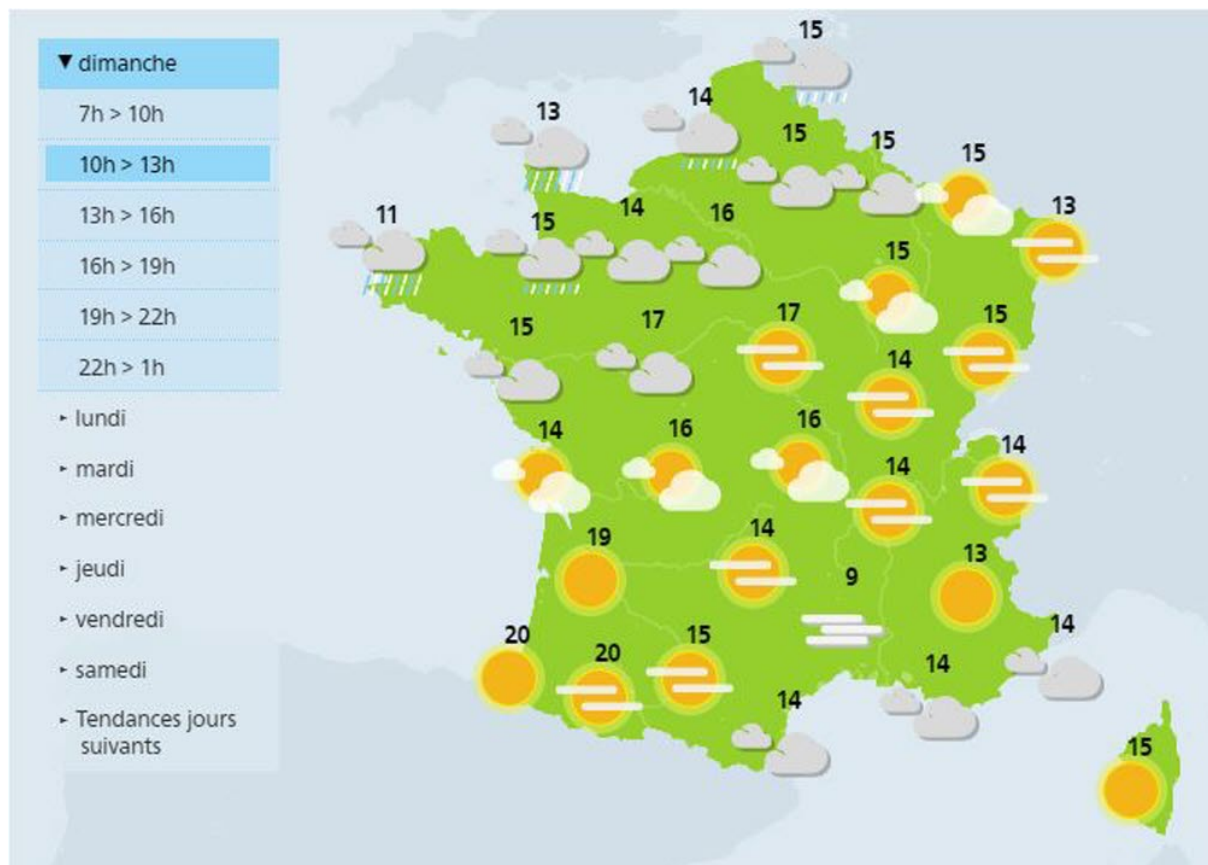


RAPPORT PROJET D'INFORMATIQUE : PROJET CYMETEO



Sommaire

1. Difficultés rencontrées	3
2. Répartitions des tâches au sein du groupe	4
3. Options manquantes	5

1. Difficultés rencontrées

La première difficulté rencontrée était l'organisation de nos structures afin de pouvoir gérer autant de données en même temps. Le choix a été fait de se pencher sur des arbres, dont chaque noeud est une structure possédant toute les informations inchangeables d'une station (ID, code postal, coordonnées etc..) et d'une liste qui elle contient toute les valeurs fluctuant en fonction des données (température, date etc...). C'est la structure avec laquelle nous pensions qu'il était le plus intéressant de travailler par le fait déjà de son organisation croissante. Cette idée nous est venue lorsque nous nous sommes rendu compte que l'on devait trier un nombre très important de données avec une faible complexité. De plus, nous nous sommes aussi rendu compte qu'il était impossible de travailler qu'avec des arbres, car la complexité du programme devenait trop importante, ou qu'avec des listes, car il était trop compliqué de trier toutes les données comme cela. L'utilisation des deux est donc un bon compromis. Cette organisation structurée nous a permis de faire face à la deuxième grosse difficulté du projet qui est tout simplement de savoir comment lire de manière efficace un fichier de plusieurs millions de lignes. Pour pouvoir stocker plusieurs millions de données dans un arbre à liste, il est nécessaire de pouvoir compter sur une fonction lire irréprochable. Mais celle-ci fut très certainement la plus complexe à réaliser. Le but d'une fonction lire de csv classique est de récupérer chaque colonne dans une variable pour pouvoir réutiliser la donnée. Pour ce faire, la fonction en c la plus connue est strtok qui permet de lire une partie précise d'une ligne d'un fichier. Mais malheureusement cette fonction n'a pas été très utile car elle possède un défaut majeur qui est qu'elle ne prend pas en compte les cases vides. En effet le fichier qui avait pour but d'être lu possède de nombreuses cases vides et était donc complètement incompatible avec cette fonction strtok. Nous avons donc dû utiliser une autre fonction qui a été créée justement dans le but de résoudre ce problème. Mais là encore nous avons dû nous heurter à un autre problème. Cette même fonction possède elle-même un autre défaut qui n'est autre qu'une erreur de segmentation pour de trop gros fichiers. Ce problème était directement lié à l'utilisation de la fonction elle-même, mais il n'y avait pas de réelle explication la concernant. Finalement après de nombreuses recherches nous avons fini par réussir à résoudre le problème en utilisant une copie de la ligne dans la fonction.

Pour le script, l'un des plus grands problèmes était de récupérer les paramètres des arguments : mode de température et pression, et les min et max de a d et g. Il y a d'abord eu un souci avec le fait que le getopt ne bouclait pas et cela posait donc problème car il ne cherchait pas tous les arguments. Ce problème a été réglé en ajoutant "OPTIND=0" au début de la boucle du getopt. Ensuite il fallait comme dit auparavant, récupérer les paramètres des arguments. Pour ce qui est du mode et du min cela était simple avec la \$optarg mais pour le max il y'a eu beaucoup de soucis. Avec la commande shift qui permet de décaler tous les arguments vers la gauche, il était d'abord prévu de récupérer le nombre d'arguments séparant le max du \$1 et de shift autant de fois que nécessaire pour

simplement récupérer le max. Cela créait beaucoup de décalage et de problèmes, mais nous avons ensuite trouvé qu'il suffisait de shift une seule fois pour récupérer ce que l'on souhaitait, comme fait dans le script final.

Gnuplot : Le problème avec Gnuplot c'est que l'ordre des commandes était important, des fois pour des petites différences cela ne marchait pas sans qu'on comprenne vraiment pourquoi ou même des codes erreurs "line .." qui ne nous donnaient pas la vraie ligne du problème. Il était aussi compliqué de savoir si les graphs obtenus étaient les bons, vu l'apparence de certains ou même certaines non ressemblances avec les wiki.

De plus, certains graphs sont justes et correspondent à la demande de l'énoncé mais ne ressemblent pas à grand chose à cause de l'abscisse allant parfois de 7000 à 98000 ce qui implique que les stations avec des nombres rapprochés (typiquement la France) se retrouvent les unes sur les autres, mais ce n'est pas quelque chose de changeable.

Le gnuplot n'est aussi pas commenté car les # génèrent des erreurs.

2.Répartitions des tâches au sein du groupe

Pierre: Pour la répartition des tâches, Pierre s'est d'abord questionné sur la structuration du sujet, comment structurer nos arbres en fonction des options demandées et comment s'y prendre. Il a ensuite pris en charge la réalisation du lire qui était une tâche plus qu'ardue au vu de la difficulté et de la longueur de celle-ci. Il s'est également occupé de la réalisation du main, des tirets optionnels et des fonctions liés à humidité, t2 et p2 puis a commenté le .c.

Antonin : Antonin a commencé avec la réalisation du script auquel il a rajouté au fur et à mesure les liens avec le c, il a réalisé les graph de T1,T2,P1,P2,Vent et hauteur), rédigé le readme, le rapport, et a en c réaliser les fonctions écrire lié à t1 p1, participé pour humidité et vent.

Gatien : Au début, Gatien s'est occupé des recherches concernant les stations, quel ID correspondait à quelles localisations. Il avait fait des tentatives de structuration sur le.h. Il a ensuite aidé Pierre sur quelques fonctions du c : il a fait une tentative de lire et s'est occupé du vent et de hauteur.

Mathis: Mathis a réalisé le Help.

Romain : Romain s'est renseigné sur le Gnuplot, comment les réaliser etc. Il avait réalisé le graph d'humidité. Il a également commenté le main et une partie du .h.

3. Options manquantes

Si le -o n'est pas présent alors le fichier créer n'est pas un .dat mais un .csv appelé météo2023.csv.

Nous n'avons pas traité le mode 3 de température et de pression.