**SPECIFICATIONS TECHNIQUES**

1. Extraction des données brutes – Base de données

Les données utilisées sont issues de l’*infobox* de la page du site Wikipédia dédiée à chaque pays. Pour éviter une requête systématique du client au(x) serveur(s) Wikipédia qui, en cas de fort afflux, produirait l’effet d’une DDOS, le choix a été fait de télécharger une fois pour toutes les *infobox* dans l’application côté serveur, sous forme de fichiers *.json*. La structure naturelle pour le stockage et la lecture de ces fichiers est une base de données, établie au format SQL : les requêtes pour la produire et y accéder sont effectuées sous Python3, grâce au module *Sqlite3*. Bien entendu, la mise à jour des données est possible : il suffit de télécharger les *infobox* Wikipédia les plus récentes, et de réexécuter le script de création de la BDD.

La création de la base de données requiert le lancement du script Python « creation\_BDD.py ». Pour éviter que la lecture et le débogage ne soient indigestes, les tâches à réaliser sont décomposées en fonctions et sous-fonctions, chacune prenant en charge une tâche bien spécifique précisée par son nom :

* Récupération d’une donnée d’un pays dans les fichiers *.json* : préfixe « **get\_** »
* Sauvegarde d’une donnée pour tous les pays dans la base de données : préfixe « **save\_** »
* Destruction d’une donnée pour tous les pays dans la base de données : préfixe « **delete\_** »

Suivant ces considérations, le programme se déroule en trois grandes phases :

1. Définition des fonctions

* De récupération des données :

Paramètres d’entrée (excepté cv\_coords) :

🡪 *wp\_info* : dictionnaire (*dict*) contenant les informations du pays. Il est directement issu de la lecture du fichier *.json*. Les fonctions sont décrites ci-dessous par table de la BDD.

🡪 *pays*  : chaîne de caractères (*str*). Paramètre parfois non nécessaire.

* ‘GLOBAL’ :
* « get\_names » : nom commun et nom conventionnel d’un pays (sous réserve d’existence).

Format de renvoi : liste de deux chaînes de caractères *str*.

**NB** : si l’un des deux noms n’existe pas, on affecte la valeur de l’autre nom, ou bien de l’intitulé du fichier *.json* du pays.

* « get\_capitale » : capitale du pays.

Format de renvoi : chaîne de caractères *str*.

* « get\_flag » : intitulé de l’image PNG du drapeau renfermée dans le dossier compressé *flag.zip* contenu dans *Serveur*.

Format de renvoi : chaîne de caractères *str*.

**NB** : l’affichage du drapeau est effectué par le client via une requête HTML.

* « get\_coords » : coordonnées de la capitale d’un pays (sous réserve d’existence).

Format de renvoi : dictionnaire avec deux clés en chaînes de caractères *str*, deux valeurs flottantes *flott*. Exemple : *{‘lat’=10.4, ’lon’=250.6669}*.

**NB** : la fonction appelle « *cv\_coords*» en interne pour convertir au format numérique.

* « cv\_coords » : conversion d’une chaîne de caractères décrivant la position d’une capitale en dictionnaire Python de latitude et longitude.

Format d’entrée : chaîne de caractère *str*  à mettre en forme.

Format de renvoi : dictionnaire avec deux clés en chaînes de caractères *str*, deux valeurs flottantes *flott*. Exemple : *{‘lat’=10.4, ’lon’=250.6669}*.

* « get\_adresse\_wiki » : adresse de la page Wikipédia de la capitale.

Format de sortie : chaîne de caractère *str* de l’adresse de la page.

**NB** : fonction inusitée pour le moment. Renseignée dans une version future de l’application. Pour le moment, l’adresse de la capitale est récupérée manuellement en 3).

* ‘ECONOMIE’ :
* « get\_devise » : devise utilisée dans le pays.

Format de renvoi : chaîne de caractères *str*.

* « get\_PIB\_nominal » : PIB nominal courant d’un pays en dollar américain.

Format de renvoi : flottant *float*.

* « get\_PIB\_par\_tete » : PIB nominal par tête d’un pays en dollar américain.

Format de renvoi : flottant *float*.

* ‘MISCELLANEOUS’ :
* « get\_sens\_circulation » : sens de circulation pour les voitures (deux valeurs possibles, *right* ou *left*).

Format de renvoi : chaîne de caractères *str*.

* « get\_code\_appel » : code d’appel à l’international du pays.

Format de renvoi : chaîne de caractères *str*.

* « get\_domaine\_internet » : domaine des sites internet du pays.

Format de renvoi : chaîne de caractères *str*.

* ‘POLITIQUE’ :
* « get\_nom\_chef\_etat » : nom du chef d’état, indépendamment de son titre.

Format de renvoi : chaîne de caractères *str*.

* « get\_type\_chef\_etat » : titre du chef d’état, indépendamment de son nom.

Format de renvoi : chaîne de caractères *str*.

* « get\_regime » : régime politique du pays, en lien avec le titre du chef d’état.

Format de renvoi : chaîne de caractères *str*.

* ‘DEMOGRAPHIE’ :
* « get\_habitants » : nombre d’habitants recensés ou estimés.

Format de renvoi : chaîne de caractères *str*.

**NB** : retourne ‘inconnu’ s’il n’est pas renseigné.

* « get\_habitants\_annee » : année pour laquelle le nombre d’habitants recensés ou estimés est donné. Complémentaire de « get\_habitants ».

Format de renvoi : chaîne de caractères *str*.

**NB** : retourne ‘inconnu’ s’il n’est pas renseigné.

* De sauvegarde de données :

Aucun renvoi (fonctions de type *void*).

* ‘GLOBAL’ :

Paramètres en entrée : conn, pays, continent, wp.

Respectivement : connecteur à la base de données (objet *Sqlite3.Connection*) ; nom du pays (chaîne de caractère) ; nom du continent (chaîne de caractère) ; informations du pays (dictionnaire).

« save\_global » : sauvegarde des informations globales pour le pays renseigné dans la table ‘GLOBAL’.

* ‘ECONOMIE’ :

Paramètres en entrée : conn, pays, wp.

Respectivement : connecteur à la base de données (objet *Sqlite3.Connection*) ; nom du pays (chaîne de caractère) ; informations du pays (dictionnaire).

« save\_economie » : sauvegarde des informations économiques pour le pays renseigné dans la table ‘ECONOMIE’.

* ‘MISCELLANEOUS’ :

Paramètres en entrée : conn, pays, wp.

Respectivement : connecteur à la base de données (objet *Sqlite3.Connection*) ; nom du pays (chaîne de caractère) ; informations du pays (dictionnaire).

« save\_miscellaneous » : sauvegarde des informations additionnelles (vrac) pour le pays renseigné dans la table ‘MISCELLANEOUS’.

* ‘POLITIQUE’ :

Paramètres en entrée : conn, pays, wp.

Respectivement : connecteur à la base de données (objet *Sqlite3.Connection*) ; nom du pays (chaîne de caractère) ; informations du pays (dictionnaire).

« save\_politique » : sauvegarde des informations politiques du pays renseigné dans la table ‘POLITIQUE’.

* ‘DEMOGRAPHIE’ :

Paramètres en entrée : conn, pays, wp.

Respectivement : connecteur à la base de données (objet *Sqlite3.Connection*) ; nom du pays (chaîne de caractère) ; informations du pays (dictionnaire).

« save\_demographie » : sauvegarde des informations démographiques du pays renseigné dans la table ‘DEMOGRAPHIE’.

* AFFICHAGE DE TOUTES LES DONNES :
* « print\_all\_info » : affiche les informations décommentées dans la fonction
* ENREGISTREMENT DE TOUTES LES DONNES :

Paramètre d’entrée : conn connecteur à la BDD (objet *Sqlite3.Connection*)

* « save\_all\_info » : sauvegarde des informations de tous les pays dans les tables respectives.

**NB1** : cette fonction itère sur les deux continents, puis sur chaque pays de chaque continent, les fonctions précédentes en préfixe « save ».

* De suppression des données :

Ces fonctions établissent une requête SQL permettant de vider une table spécifique des données qu’elle contenait (si existence desdites données). Aucun renvoi.

Paramètre d’entrée : conn connecteur à la BDD (objet *Sqlite3.Connection*)

* ‘GLOBAL’ :

« delete\_global » : supprime les informations pour tous les pays renseignés dans la table ‘GLOBAL’.

* ‘ECONOMIE’ :

« delete\_economie » : supprime les informations pour tous les pays renseignés dans la table ‘ECONOMIE’.

* ‘MISCELLANEOUS’ :

« delete\_miscellaneous » : supprime les informations pour tous les pays renseignés dans la table ‘MISCELLANEOUS’.

* ‘POLITIQUE’ :

« delete\_politique » : supprime les informations pour tous les pays renseignés dans la table ‘POLITIQUE’.

* ‘DEMOGRAPHIE’ :

« delete\_demographie » : supprime les informations pour tous les pays renseignés dans la table ‘DEMOGRAPHIE’.

* ANNIHILATION DE TOUTES LES DONNES :

Paramètre d’entrée : conn connecteur à la BDD (objet *Sqlite3.Connection*)

* « delete\_all\_info » : destruction des informations de tous les pays dans les tables respectives.

1. Génération de la base de données :

* Exécution d’un « nettoyage » préalable : si une table existe déjà, elle est écrasée.
* Exécution de la fonction de création de la base de donnée complète.

1. Modifications manuelles :

Comme mentionné en 1), certaines informations sont inexistantes pour des pays particuliers. Il faut alors résoudre ces difficultés au cas par cas. La fonction « modifications\_manuelles » est exécutée pour gérer tous les cas litigieux : valeurs manquantes, inexactes,… etc.