

Gestion de la confidentialité dans les tableaux de données agrégées

Programmer avec τ -Argus



Maxime Beauté et Alexandre Awad – Département des méthodes statistiques

01 · τ -Argus : logiciel libre

02 · SAS et τ -Argus

01

τ -Argus :
logiciel libre

τ -Argus : logiciel libre

- Depuis 2014, τ -Argus est devenu un **logiciel libre**. Vous pouvez donc :
 - Exécuter librement le programme, pour tous les usages ;
 - Étudier le fonctionnement du programme et l'adapter à vos besoins ;
 - En redistribuer des copies ;
 - Améliorer le programme et publier des améliorations.
- Cela signifie donc qu'il est open source. Son **code source** est **public** et accessible sur GitHub à cette adresse :
<https://github.com/sdcTools/tauargus>
- Un groupe d'experts issus des INS européens en assure le **support**.

Le dépôt GitHub de τ -Argus

- Sur GitHub, vous pouvez :
 - (1) consulter ou télécharger le **logiciel** ou son **code** dans la version souhaitée ;
 - (2) demander directement de l'aide aux développeurs ;
 - (3) faire remonter des bugs ;
 - proposer des modifications dans le code ;
 - (4) **consulter l'aide en ligne** (FAQ, wiki, bugs connus).

(1) : <https://github.com/sdcTools/tauargus/releases>

(2) : <https://github.com/sdcTools/tauargus>

(3) : <https://github.com/sdcTools/UserSupport/issues>

(4) : <https://github.com/sdcTools/UserSupport>

02

SAS et τ -Argus

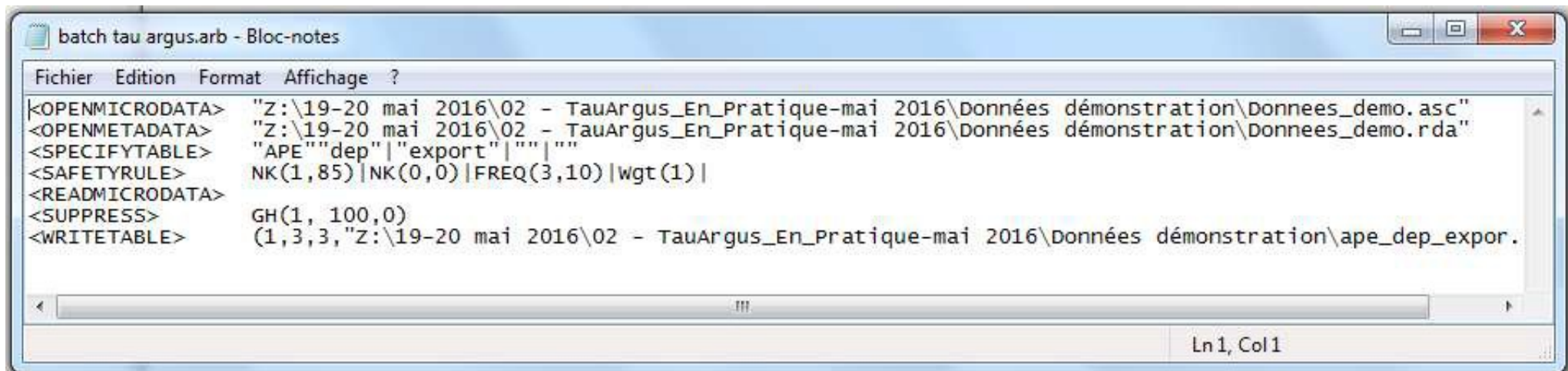
Macro SAS %Tau_Argus()

- Voici un exemple d'appel à cette macro :

```
%TAU_ARGUS (  
    tabsas           =    donnees_entreprises_usa_macro,  
    library          =    V:\Formation Tau-Argus\Mise en pratique,  
    tabulation_1     =    etat produit ventes_part,  
    tabulation_2     =    etat mois ventes_commerces,  
    tabulation_3     =    mois produit freq,  
    hierarchical_var =    mois produit,  
    hierarchy_1      =    p2 produit,  
    hierarchy_2      =    trim mois,  
    weight_var       =    poids_ent,  
    TauArgus_exe     =    Y:\Logiciels\TauArgus4.1.7b4\TauArgus\TauArgus.exe,  
    TauArgus_version =    opensource);
```

Mode batch

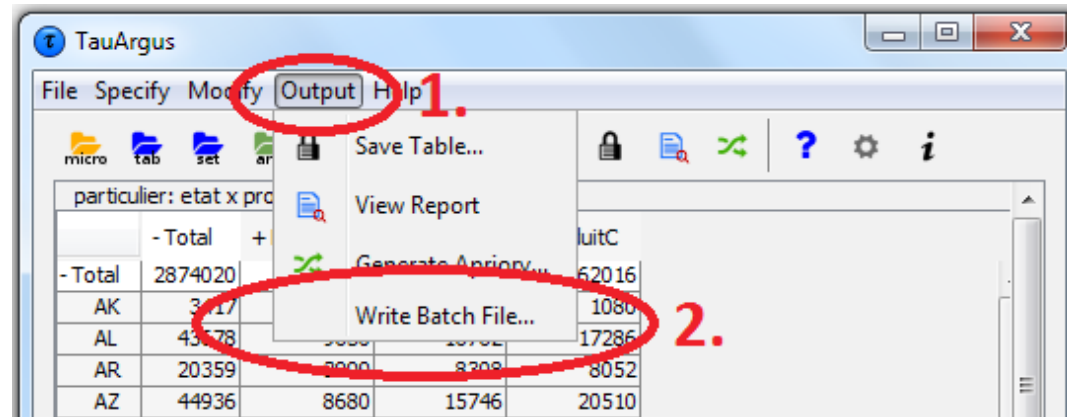
Définition : Un **batch** est un fichier .arb contenant du texte en ligne de commande permettant de **faire appel à distance** à τ -Argus.



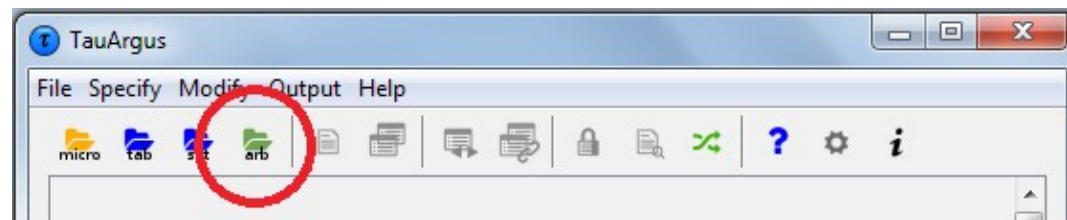
```
batch tau argus.arb - Bloc-notes
Fichier  Edition  Format  Affichage  ?
<OPENMICRODATA>  "Z:\19-20 mai 2016\02 - TauArgus_En_Pratique-mai 2016\Données démonstration\Donnees_demo.asc"
<OPENMETADATA>   "Z:\19-20 mai 2016\02 - TauArgus_En_Pratique-mai 2016\Données démonstration\Donnees_demo.rda"
<SPECIFYTABLE>   "APE""dep"|"export"|" ""|" ""
<SAFETYRULE>     NK(1,85)|NK(0,0)|FREQ(3,10)|wgt(1)|
<READMICRODATA>  GH(1, 100,0)
<SUPPRESS>       (1,3,3,"Z:\19-20 mai 2016\02 - TauArgus_En_Pratique-mai 2016\Données démonstration\ape_dep_expor.
<WRITETABLE>
```

Ln1, Col1

- Après avoir effectué vos traitements, vous pouvez les **enregistrer** dans un fichier batch .arb (*Output* → *Write Batch File...*) :



- ... qui peut plus tard être ouvert dans t-Argus pour **rejouer automatiquement** les traitements effectués :



Macro SAS %Tau_Argus()

- - À partir d'une table SAS, cette macro :
 - génère les fichiers de données individuelles et de métadonnées associées en entrée de τ -Argus ;
 - génère un fichier batch en fonction des paramètres de la macro ;
 - produit les masques de secret et génère des statistiques descriptives sur ces masques.
- Elle permet d'**industrialiser entièrement** les traitements liés à la confidentialité au sein d'un programme SAS.

Macro SAS %Tau_Argus()

- - La macro est disponible en open source sur le dépôt GitHub InseeFrLab.
→ <https://github.com/InseeFrLab/SASTauArgus>
 - Vous n'avez pas à toucher au code SAS de la macro. En revanche, en cas de doute, n'hésitez pas à consulter **les commentaires au début du code** : ils décrivent précisément chacun des paramètres.

Les paramètres basiques de la macro %Tau_Argus()

- *tauargus_exe* → Chemin d'accès au logiciel
- *tauargus_version* → (rien)/opensource (à partir de la version 4.0.0, toujours mettre opensource)
- *tabsas* → Nom de la table SAS
- *library* → Répertoire où se trouve tabsas
- *tabulation_x* (x = 1 ... 9) → Liste des variables d'une tabulation, séparées par un espace, en commençant par les variables de ventilation et en finissant par la variable de réponse (ou FREQ pour un tableau d'effectifs)

Les paramètres usuels de la macro %Tau_Argus()

- *primary_secret_rules* → DOM et/ou P et/ou FREQ (DOM FREQ par défaut)
 - *dom_k* → seuil pour la règle de dominance (85 par défaut)
 - *p_p* → seuil pour la règle du p % (10 par défaut)
 - *frequency* → seuil pour la règle de fréquence (3 par défaut)
- *outputtype* → Choix du format de sortie (SBS par défaut)
 - 1 pour CSV-file
 - 2 pour CSV file for pivot table
 - 3 pour Code-value file
 - 4 pour SBS format (par défaut)
 - 5 pour Intermediate file

Les paramètres usuels de la macro %*Tau_Argus*()

- *solver* → hypercube/modular/optimal/(rien) (hypercube par défaut)
- *weight_var* → Nom de la variable de poids
- *holding_var* → Nom de la variable identifiant (siren...)
- *hierarchical_var* → Liste des variables hiérarchiques (niveau le plus fin)
- *hierarchy_x* (x = 1 ... 9) → Description d'une hiérarchie (du moins fin au plus fin)
- *linked_tables* → Les tabulations_x sont-elles liées ? (yes par défaut)
- *synthesis* → Un fichier de synthèse est-il créé ? (no par défaut)
- *apriori_x* (x = 1 ... 9) → Un fichier d'a priori doit-il être appliqué à la tabulation x ? (no par défaut)

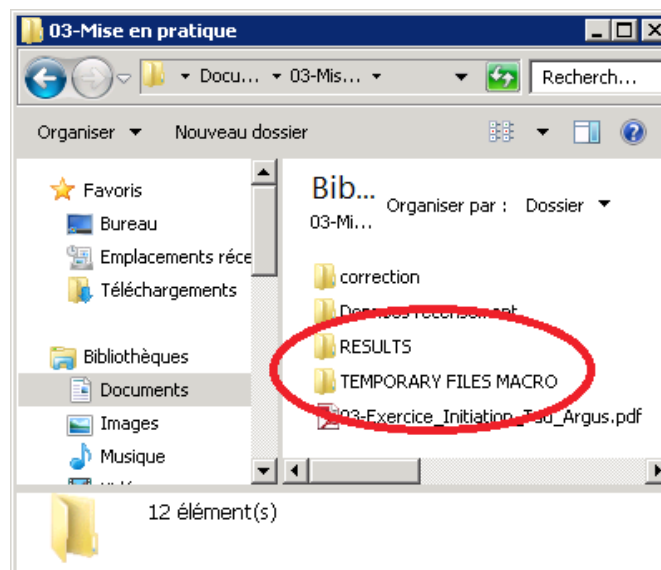
Retour sur l'exemple

```
%TAU_ARGUS (  
    tabsas           =    donnees_entreprises_usa_macro,  
    library          =    V:\Formation Tau-Argus\Mise en pratique,  
    tabulation_1     =    etat produit ventes_part,  
    tabulation_2     =    etat mois ventes_commerces,  
    tabulation_3     =    mois produit freq,  
    hierarchical_var =    mois produit,  
    hierarchy_1      =    p2 produit,  
    hierarchy_2      =    trim mois,  
    weight_var       =    poids_ent,  
    TauArgus_exe     =    Y:\Logiciels\TauArgus4.1.7b4\TauArgus\TauArgus.exe,  
    TauArgus_version =    opensource);
```

Retour sur l'exemple

batch_sas.arb
etat_mois_ventespart.html
etat_mois_ventespart.sbs
etat_produit_ventespart.html
etat_produit_ventespart.sbs
LogBook.txt
metadata.rda
microdata.asc
mois_produit_freq.html
mois_produit_freq.sbs

*TEMPORARY
FILES MACRO*



RESULTS

etat_mois_ventespart.sas7bdat
etat_mois_ventespart.xls
etat_produit_ventespart.sas7bdat
etat_produit_ventespart.xls
mois_produit_freq.sas7bdat
mois_produit_freq.xls

Et merci pour votre attention ! 😊

Maxime Beauté
maxime.beaute@insee.fr
01 87 69 55 43

Alexandre Awad
alexandre.awad@insee.fr
01 87 69 55 14