

## TD 2 : Introduction au logiciel SAS

Aude Tavenard

1. Importer les fichiers `notes.txt` et `clients.txt` .

```
PROC IMPORT OUT= notes
    DATAFILE= "/folders/myfolders/Cours SAS/td2/notes.txt"
    DBMS=DLM REPLACE;
    DELIMITER=";";
    GETNAMES=YES;
    DATAROW=2;
RUN;
```

```
PROC IMPORT OUT= clients
    DATAFILE= "/folders/myfolders/Cours SAS/td2/clients.txt"
    DBMS=DLM REPLACE;
    DELIMITER=";";
    GETNAMES=YES;
    DATAROW=2;
RUN;
```

2. Faire une jointure des deux fichiers SAS obtenus.

```
proc sort data=notes ; by resto_id annee ; run ;
proc sort data=clients ; by resto_id annee ; run ;
data resto ;
    merge notes clients;
    by resto_id annee ;
run ;
```

3. Modifier la variable Quartier en supprimant le mot 'Quartier' pour chaque observation

```
data resto;
    set resto ;
    /* quand on ne sait pas combien de caractères on souhaite,
       on ne précise que la position de départ */
    quartier=substr(10) ;
run ;
```

4. Créer une nouvelle variable note au format caractère.

```
data resto;
  set resto ;
  note2= put(note,best12.);
run ;
```

5. Créer une table contenant la meilleure note pour chaque restaurant.

```
/* Option 1 */
proc sort data=resto ; by resto_id note ; run ;
data best ;
  set resto ;
  if last.resto_id ;
  by resto_id note ;
run ;

/* Option 2 */
proc sort data=resto; by resto_id descending note ; run ;
proc sort data=resto out=best2 nodupkey ; by resto_id ; run ;
```

6. Calculer la fréquence des types de restaurant en sélectionnant les observations qui correspondent à des restaurants dont le nom contient le mot "pizza" (quelle que soit sa casse).

```
proc freq data=resto (where=(index(upcase(nom), "PIZZA")>0)) ;
  table type;
run ;
```

7. Calculer la moyenne des notes pour les années 2018 et 2019 séparément en une procédure.

```
proc means data=resto (where=(annee in (2018,2019))) ;
  class annee ;
  var note;
run ;
```

8. Présenter la moyenne du nombre de clients par mois par quartier et par année dans un même tableau.

```
proc tabulate data=resto;
  class annee quartier ;
  var n_clients_moy;
  table (quartier),(annee)*(n_clients_moy)*(mean);
run ;
```