[ST] AudienceLab - Export Excel

# Description

Nous souhaitons mettre en place un processus d’export de rapport d’audience pour chaque annonceur. Les données sont présentes dans une base de données MySQL. Il faut générer un fichier Excel en respectant.

# Problématique technique

Les contraintes de ce développement sont les suivantes :

* utiliser la base de données MySQL en entrée
* respecter le format Excel de sortie (voir exemple)
* développer une méthode PHP pour générer ce fichier excel
* ne pas utiliser de framework spécifique mais utiliser la bibliothèque [phpexcel](https://phpexcel.codeplex.com/)
* le fichier Excel généré sera placé dans un répertoire précis configurable (variable globale)
* le fichier Excel généré devra respecter un nom spécifique : Tradelab-advertiser\_id-pixel\_name-date\_generation.xls
* le code devra tourner sur la version PHP suivante : PHP 5.5.13-2
* ajouter une table *exelate\_data\_output\_version* pour historiser chaque fichier généré avec les colonnes : ID, PIXEL\_ID, FILENAME, TIMESTAMP

# Solution technique détaillée

Mettre en place un script PHP qui permet de générer un fichier Excel en processus synchrone à partir d’un pixel\_id.

Chaque fichier Excel (=rapport) est lié à un pixel ID (table : *exelate\_pixel*). Chaque annonceur peut posséder un ou plusieurs pixels. Un annonceur est défini par un usr\_id (table *exelate\_pixel*).

Chaque rapport prend les données de la table *exelate\_data* pour un pixel donné. Les données sont liées à des segments prédéfinis qui sont présents en table *exelate\_segment*. Chaque donnée dépend d’une sous-catégorie définie dans *exelate\_segment\_subcategory* qui elle-même dépend d’une catégorie définie dans la table *exelate\_segment\_category*.

Les catégories qui apparaissent sont traduites en français dans le fichier de sortie. La traduction de chaque catégorie se trouve dans la colonne *fr* de la table *exelate\_segment*.

Le format Excel est défini selon l’exemple ci-joint.

* La première feuille de ce fichier “**SUMMARY**” présente des données générales :
  + **moyenne** : la valeur d’index moyenne de toutes les données de tous les segments d’un pixel
  + **minimum** : la valeur d’index minimum de toutes les données de tous les segments d’un pixel
  + **maximum** : la valeur d’index maximum de toutes les données de tous les segments d’un pixel

Pour le reste, il faut classer l’ensemble des valeurs des index par ordre croissant. Voir [Wikipédia](http://fr.wikipedia.org/wiki/Quartile) pour plus d’explication sur les termes statistiques.

* + **1er quartile** : correspond à la valeur d’*index* du segment \_id de la partie des 25 % inférieures des données (valeur max de ce 1er quartile).
  + **médiane** : correspond à la valeur d’*index* du segment \_id de la partie des 50 % inférieures des données (valeur max de ce 2ème quartile). C’est aussi la “valeur du milieu”.
  + **3ème quartile** : correspond à la valeur d’*index* du segment \_id de la partie des 75% inférieures des données (valeur max de ce 3ème quartile).
  + **centile 10** : correspond à la valeur d’*index* du segment \_id de la partie des 10 % inférieures des données (valeur max de ce 1er centile).
  + **centile 90** : correspond à la valeur d’*index* du segment \_id de la partie des 90 % inférieures des données (valeur max de ce 90ème centile).
  + **centile 95** : correspond à la valeur d’*index* du segment \_id de la partie des 95 % inférieures des données (valeur max de ce 95ème centile).
* La seconde feuille “**DEMOGRAPHIC**” présente les données de la catégorie socio-démographique. Toutes les données sont classées par sous-catégorie (Âge, Sexe, Carrère, Revenu, Foyer, Urbanité…). Les valeurs présentées sont les *index* présentes dans table *exelate\_data*.
* La troisième feuille “**INTERETS**” présente les données de la catégorie intérêt. Toutes les données sont classées par sous-catégorie (Beauté et mode, Aménagement intérieur, High tech,…). Les valeurs présentées sont les *index* présentes en table *exelate\_data*.
* La dernière feuille “**INTENTIONS**” présente les données de la catégorie intention. Toutes les données sont classées par sous-catégorie (automobile, shopping, voyage, finance et assurance, grande consommation, Services…). Les valeurs présentées sont les *index* présentes en table *exelate\_data*.

Merci de respecter la forme de chaque feuille (logo et présentation). Les sous-catégories peuvent être différentes dans le fichier Excel d’exemple. Il faut prendre en compte les sous-catégories présentes en base de données.

Les gradients de couleurs correspondent à la formule suivante (visible dans la feuille Excel : mise en forme conditionnelle) :

* **rouge** lorsque la valeur de l’*index* vaut le centile 10
* **jaune** lorsque la valeur de l’*index* vaut la valeur médiane
* **vert** lorsque la valeur de l’*index* vaut le centile 90

Ainsi toutes les valeurs en dessous du centile 10 seront rouges (même couleur sur la tranche du centile 10). De même toutes les valeurs au-dessus du centile 90 seront vert (même couleur sur la tranche du centile 90).

Contact

Pour toute question, veuillez vous adresser au contact suivant :

Vincent Sattler

vs@tradelab.fr