



UNIVERSITÉ
DE REIMS
CHAMPAGNE-ARDENNE

Web Analytics

Nils Schaefer

[nils.schaefer@sn-i.fr](mailto:nil.schaefer@sn-i.fr)

❑ Réseau mondial

- Nombreux services

❑ Adresse IP : identification sur Internet

- 4 octets avec notation pointée (233.78.45.225)

❑ Port logique : point d'entrée applicatif

- 2 octets

❑ IP + Port = Adresse applicative

❑ DNS : mécanisme de noms

- Association entre un nom et une IP
- Facultatif

Serveur Web

- ❑ Hébergement des sites Web
- ❑ Utilisation du port 80
- ❑ Protocole HTTP
- ❑ Réception et traitement des requêtes
- ❑ Production de logs
- ❑ Exemples
 - Microsoft IIS (Windows)
 - Apache (Windows et Linux)
 - Nginx (Windows et Linux)

Mesure d'audience d'un site Web

- ❑ Quantification des accès à un site Web
- ❑ Thématique aussi vieille que le Web
- ❑ Quantité de données importante
 - D'où l'importance de savoir quoi utiliser
- ❑ Plusieurs utilisations
 - Consultation
 - Prévisionnel
 - Référencement
 - Webmarketing

Aspects techniques

- ❑ Informations connues sans intervention de l'utilisateur
 - Contrairement aux informations saisies par l'utilisateur
- ❑ A la base d'indicateurs plus avancés
- ❑ Possibilité de manipuler les différentes informations
 - Peu important par rapport à la masse des utilisateurs standards
 - Toujours une faible proportion de données erronées
- ❑ D'une manière générale très fiable pour avoir une bonne idée de l'audience d'un site Web

Date et heure

- ❑ Aspect assez évidemment mais fondamental
- ❑ Utilisation de l'heure côté serveur
 - Valeur fiable
- ❑ Aucune utilisation de l'heure côté client
 - Valeur non fiable
- ❑ Régionalisation
 - Attention à la localisation des serveurs
 - Importance de gérer le décalage horaire

Adresse IP

❑ Adresse IP publique

- Reconnue sur Internet

❑ Identification d'un point de connexion à Internet

- Toujours connue d'un point de vue technique (Protocoles IPv4 et IPv6)
- Détournement avec utilisation de proxies comme TOR

❑ Différents aspects

- Juridique
- Géolocalisation

Navigateur et système d'exploitation

- ❑ Récupération du système d'exploitation et de sa version
- ❑ Récupération du navigateur et de sa version
- ❑ Infos communiquées par le navigateur
 - Terme *user-agent*
 - En-tête dans la requête HTTP
- ❑ Exemple
 - Pour Firefox 50 sur Windows 10 64 bits
 - Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:50.0) Gecko/20100101 Firefox/50.0

Taille d'écran, fonctionnalités actives

☐ Infos communiquées par le navigateur

- Détection directement par le navigateur
- En rapport avec la configuration du navigateur

☐ Important pour la mise en place des versions *Responsive*

☐ Exemple

- Résolution : 1920x1080 pixels
- Javascript : activé

Langues utilisées

- ❑ Récupération des langues utilisées
 - Classement par ordre de préférence
- ❑ Utile pour orienter l'utilisateur vers une traduction adéquate
- ❑ Info communiquée par le navigateur
 - Terme *Accept-Language*
 - En-tête dans la requête HTTP
- ❑ Exemple
 - fr,fr-fr;q=0.8,en-us;q=0.5,en;q=0.3

Cookies

❑ Fichier texte déposé sur le client

- Contenu limité à 4 Ko
- Date d'expiration limitée
- Valeur non fiable car modifiable

❑ Récupérable uniquement par le serveur à l'origine du dépôt

- Possibilité de partage des cookies entre sites (publicités, suivi)

❑ Exemple

- Dernière recherche effectuée sur le site
- Enregistrement de préférences (« Se souvenir de moi », ...)

- ❑ Données stockées côté serveur
 - Permet de suivre un utilisateur côté serveur
 - Durée courte
 - Données fiables car non accessibles
- ❑ Basé sur l'utilisation d'un cookie
 - Contient l'identifiant de session (unique)
- ❑ Souvent associé à un mécanisme d'authentification mais pas obligatoirement

Exemples de possibilités techniques...

❑ Capture à la volée d'une saisie

- Conservation des saisies intermédiaires

❑ Capture d'un formulaire non validé

- Conservation des données déjà saisies (email...)

❑ Analyse des mouvements de la souris

- Comprendre le comportement des clients

❑ Adaptation des prix au client

- Prix « optimal » par produit et par client

Les indicateurs

- ❑ Informations construites à partir des éléments techniques
- ❑ Similaires à toutes les solutions Web Analytics
 - Quelques variations sémantiques
 - Disponibilité variable en fonction des éléments techniques
- ❑ Création d'informations de haut-niveau
 - Plus proche des utilisateurs
 - Plus grande valeur ajoutée
 - Plus lisible pour des personnes non informaticiennes
 - Accessibles aux personnes du marketing, de la communication...

- ❑ Chaque demande du navigateur génère un *hit*
 - Chaque requête HTTP
- ❑ Une page Web génère plusieurs *hits*
 - Un *hit* pour la page HTML
 - Un *hit* par image
 - Un *hit* par feuille de style
 - ...
- ❑ Information peu représentative
- ❑ Intéressant d'un point de vu technique

Bande passante

- ❑ Mesure de la bande passante sortante
 - Quantité de données émises par le serveur Web
- ❑ En corrélation avec le rôle du serveur
- ❑ Intéressant d'un point de vu technique
 - Dimensionnement du serveur

Pages vues

- ❑ Chaque chargement de page est pris en compte
- ❑ Un utilisateur peut visiter plusieurs fois une même page
- ❑ Proportionnel aux *hits*
 - En fonction du nombre de fichiers associés aux pages HTML
- ❑ A interpréter en fonction du nombre de pages du site

Nombre de visites

- ❑ Une visite correspond à un utilisateur qui visite le site « en continu »
 - Visiteur qui se déplace de page en page
- ❑ Evaluation basée sur un délai maximum entre deux pages
 - Si le délai est dépassé, il s'agit alors d'une nouvelle visite
- ❑ Utilisation de la notion de session
 - Terme session parfois utilisé à la place de visite
- ❑ Indicateur populaire
 - Simple à comprendre
 - Permet de suivre le développement d'un site

Nombres de visiteurs uniques

- ❑ La notion de visiteur unique permet d'identifier un même visiteur lors de ses différentes visites du site
- ❑ Utilisation d'une période de temps
 - Au-delà de cette période on considère de nouveau le visiteur comme unique
- ❑ Utilisation d'un cookie
 - Date de dernière visite

Taux de rebond

- ❑ Pourcentage de visites ne comportant qu'une seule page vue
- ❑ On ne peut pas mesurer le temps réel passé sur la page !
 - Que veut dire « passer du temps » sur une page ?
- ❑ Indicateur assez populaire
- ❑ Interprétation délicate
 - Souvent vu comme une valeur à minimiser
 - Est-ce que le fait de ne consulter qu'une page est négatif ?
 - Dépendant du site
 - Dépendant de chaque page

Taux de sortie

- ❑ Pourcentage de visite se terminant sur une page donnée
- ❑ A mettre en association avec le taux de rebond d'une page
- ❑ Permet de remarquer de manière plus fiable une page qui pose problème
 - Manque d'invitation à l'action ?
 - Contenu qui incite à la méfiance ?
 - ...

Durée de visite

- Notion approximative
- Basée sur la date de première page vue et de dernière page vue
 - Au cours d'une visite
- Comme pour le taux de rebond on ne dispose pas de la durée de consultation de la dernière page
- Est-ce qu'entre deux pages consultées la consultation est continue ?
 - Aucune garantie

Taux de conversion

□ Définition d'objectifs

- Inscription à une newsletter
- Création d'un compte
- Commande d'un produit
- ...

□ Quantification des cas de réussite des objectifs

□ Indicateur fondamental en e-commerce

- Comparaison par rapport aux autres sociétés du même domaine

Solutions *Site-Centric*

- ❑ Solutions centrées sur l'analyse d'un site
- ❑ Très souvent utilisées
- ❑ Par opposition aux solutions *User-Centric*
 - Utilisation de panels d'utilisateurs
- ❑ Deux types de solutions
 - Analyse des logs
 - Utilisation de marqueurs

Analyse des logs

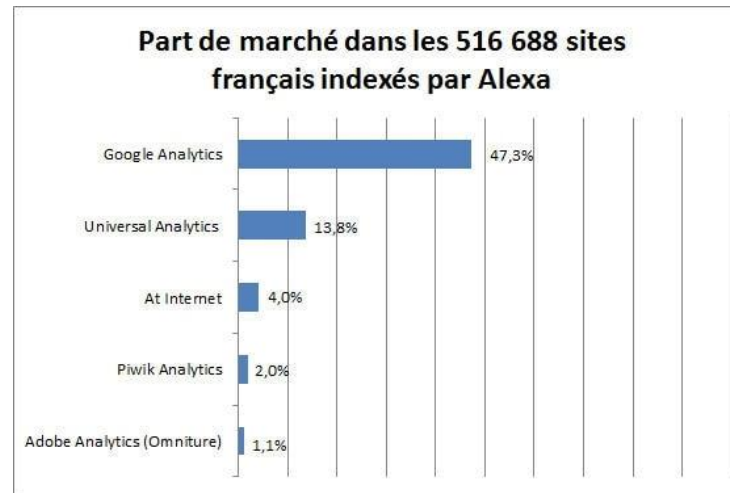
- ❑ Programme à installer sur le serveur Web
- ❑ Analyse régulière des logs du serveur Web
 - Mise à jour des statistiques
- ❑ Consomme des ressources sur le serveur
- ❑ Exemple
 - The Webalizer
 - AwStats
 - GoAccess
 - ...

Solutions avec marqueurs

- ❑ Intégration d'un code dans les pages du site
- ❑ Tout ce qui n'est pas associé au code est exclu des statistiques
- ❑ Exemples
 - Google Analytics
 - Adobe Analytics
 - AT Internet
 - IBM Digital Analytics
 - Webtrands Analytics

Google Analytics

- ❑ Solution populaire et accessible gratuitement
- ❑ Création d'un compte et mise en place simple
- ❑ Utilisation sur une grande majorité de sites (données de 2015)



Marqueur Google Analytics

- ❑ Morceau de code Javascript qui permet à Google de récupérer les informations

```
<script>
  (function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r]=i[r]||function(){
  (i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date();a=s.createElement(o),
  m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(a,m)
  })(window,document,'script','https://www.google-analytics.com/analytics.js','ga');

  ga('create', 'UA-6760939-4', 'auto');
  ga('send', 'pageview');
</script>
```

- ❑ Intégration sur toutes les pages

- Structuration en PHP avec intégration d'un fichier (*require*)

- ❑ Exclu de fait l'analyse des documents non-HTML

- Prévoir éventuellement une gestion par événement

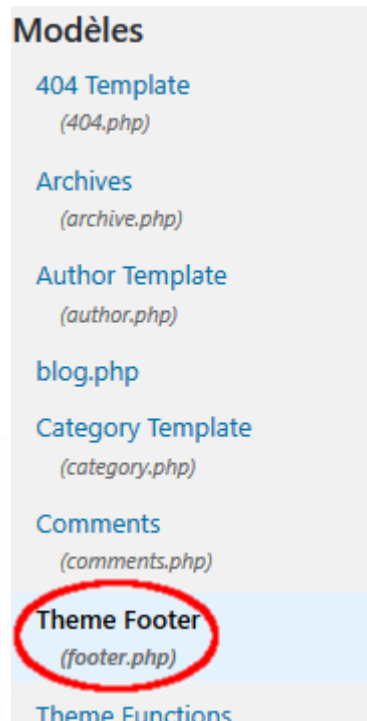
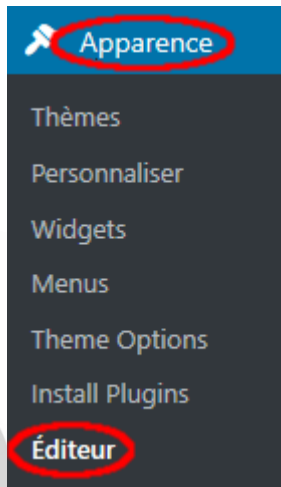
Marqueur avec Wordpress

❑ Utilisation d'une extension

- Eventuellement payante

❑ Configuration manuelle

- Intégration du code



```
Economy: Theme Footer (footer.php)

<?php if(!economy_set($options, 'hide_bottom_footer')):?>
<div class="footer-bottom">
<div class="auto-container clearfix">
<!--Copyright-->
<div class="copyright text-center">
<?php echo balanceTags(economy_set($options, 'copyrightx'));?>
</div>
</div>
<?php endif;?>
</footer>
<?php endif;?>
</div>
<!--End pagewrapper-->

<!--Scroll to top-->
<div class="scroll-to-top scroll-to-target" data-target=".main-header" style="display: block;"><span class="icon flaticon-transport"></span></div>

<?php wp_footer(); ?>
<script type="text/javascript">
jQuery("#menu-item-302 a").attr('title','Accueil');
jQuery("#menu-item-302 a").html('<img src=http://www.idx-conseil.com/wp-content/uploads/2016/12/icone_accueil.png');
jQuery("#search-2 input").attr('placeholder','Rechercher');
jQuery(".pull-right div:nth-child(1) ul").html('<li><strong style="color:#979e58;">09 72 57 64 79</strong><li style="font-size:0.95em;">1 rue d
100 Reims');
child(2) ul').html('<li><strong style="color:#979e58;">Votre espace</strong><li style="font-size:0.95em;"><a href=h
rget_blank>Accès PC/Mac</a><li style="font-size:0.95em;"><a href=https://acd.coaxis.com/ISuiteExpertMobile/?CNX=CN
a');
</script>
</body>
</html>

<script>
(function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r]=i[r]||function(){
(i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date(),a=s.createElement(o),
m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(a,m)
})(window,document,'script','https://www.google-analytics.com/analytics.js','ga');

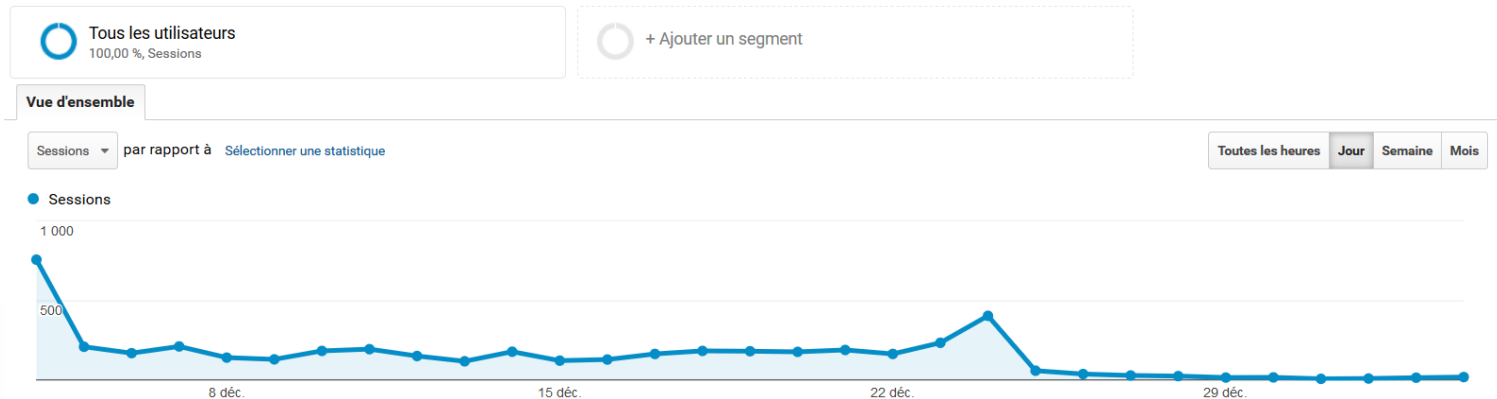
ga('create','UA-6760939-4','auto');
ga('send','pageview');
</script>
```

Rubrique « Audience »

- ❑ Regroupe l'ensemble des informations de mesure d'audience
- ❑ Vue d'ensemble
 - Principaux indicateurs
 - Résumé de l'activité du site
- ❑ Graphique
 - Plage de temps
 - Granularité (heures/jour/semaine/mois)
 - Différents segments
 - Deux statistiques

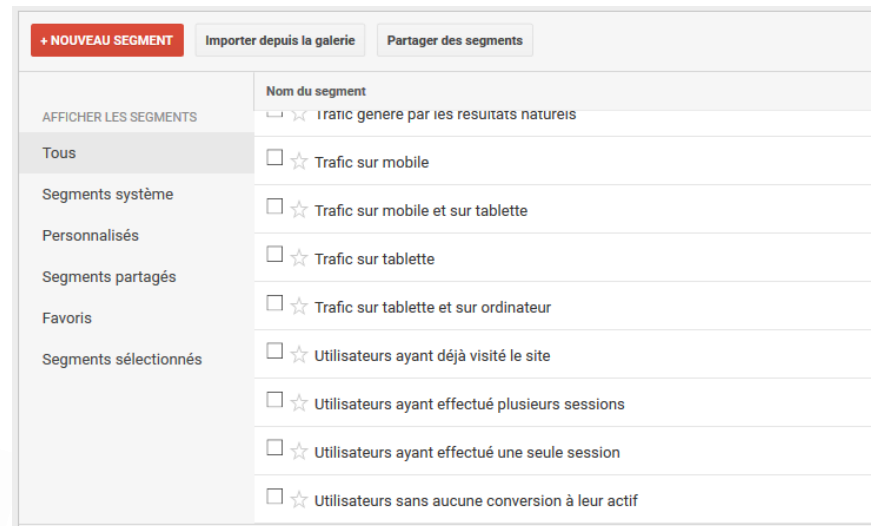
Graphique

- Plage de temps
- Granularité (heure/jour/semaine/mois)
- Différents segments
- Une ou deux statistiques



Segmentation de l'audience

- ❑ Analyse un sous-ensemble des visiteurs
- ❑ Découpage de l'audience selon différents critères
 - Segments prédéfinis
 - Segments personnalisés



Indicateurs

Données démographiques (!)

- Age et sexe

Centres d'intérêts (!)

Données géographiques

- Langue et pays

Comportement

- Nouveau/Anciens
- Fréquence et récence
- Intérêt et interaction

Indicateurs

Technologie

- Navigateur et OS
- Réseau

Mobile

Analyse comparative

- Nécessite de partager anonymement ses données
- Intéressant pour mesurer la performance de son site

Explorateurs d'utilisateurs

- ❑ Fonctionnalité pour suivre individuellement chaque utilisateur de manière anonyme afin d'étudier les comportements en se basant sur des cas précis

The screenshot displays a user analytics interface. On the left, a user profile is shown with a blue circular icon and the following details:

- Client-ID:** 1637424437.1446955981
- Date d'acquisition:** déc. 14, 2016
- Canal d'acquisition:** (represented by a blue icon)
- Catégorie d'appareil mobile:** (represented by a blue icon)

A link below reads: [Revenir au rapport "Explorateur d'utilisateurs"](#)

The main dashboard area shows session data for 4 sessions. At the top, summary statistics are provided:

- 4 Sessions**
- 00:23:56** (Durée de la session)
- 0,00 \$US** (Chiffre d'affaires)

Below these are filters and sorting options:

- Filtrer par:** Créer un segment, 4 option(s) sélectionnée(s)
- Trier par:** Ordre décroissant
- Buttons: Tout développer, Tout réduire

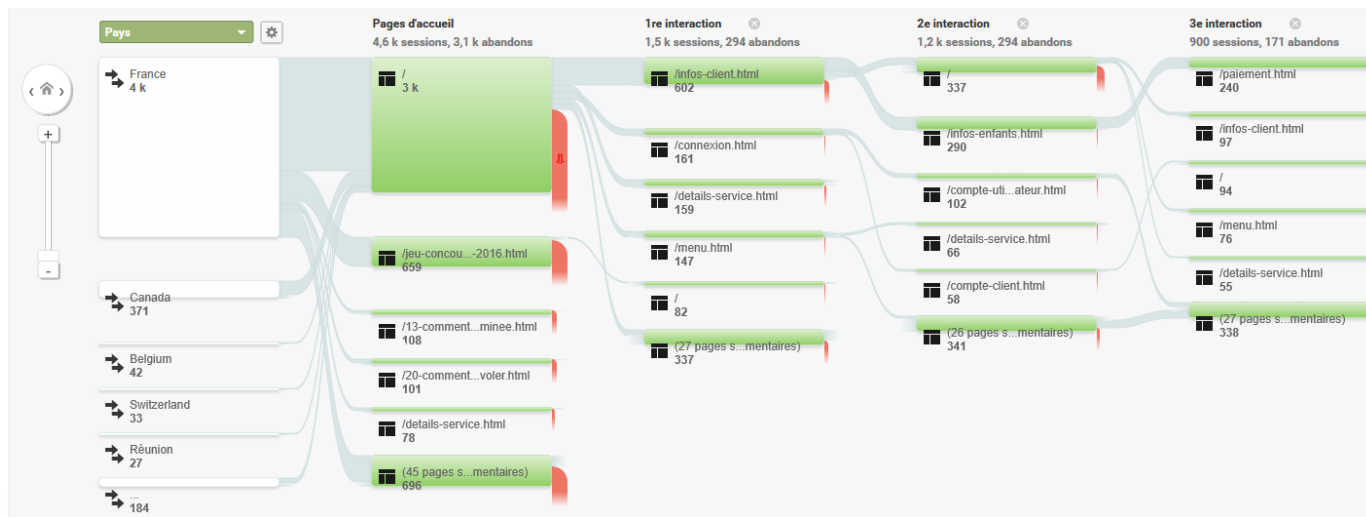
The session data is organized by date:

- déc. 21, 2016:** 1 session
 - 9:07 AM - 19:43 - Direct - 22 views
- déc. 16, 2016:** 1 session
 - 6:44 AM - 00:25 - Direct - 5 views
- déc. 14, 2016:** 2 sessions
 - 7:57 AM - 00:11 - Direct - 3 views
 - 6:20 AM - 03:36 - Direct - 9 views

At the bottom right, there are controls for displaying lines (100) and navigation (Accéder à: 1, 1 à 39 sur 39).

Flux d'utilisateurs

- ❑ Graphique spécifique original qui permet de suivre les principaux cheminements des utilisateurs sur le site



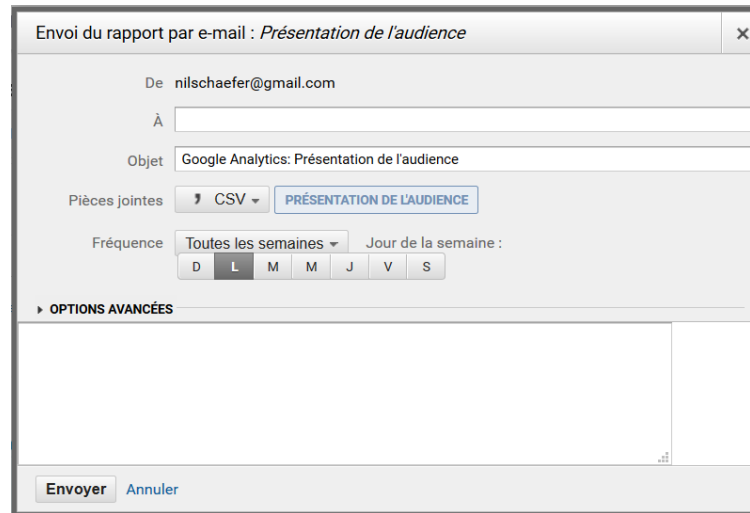
Variables et valeurs personnalisées

- ❑ Possibilité d'ajouter ses propres indicateurs !
- ❑ Il faut alors produire du code Javascript qui va permettre de définir ces variables et valeurs associées
- ❑ Attention il est interdit d'ajouter des indicateurs permettant d'identifier nominativement des utilisateurs

```
1 var _gaq = _gaq || [];  
2   _gaq.push(  
3     ['_setAccount', 'UA-XXXXXXX-1'],  
4     ['_setDomainName', 'votresite.fr'],  
5     ['_setCustomVar', 1, 'Univers', 'Electro-ménager', 3],  
6     ['_setCustomVar', 2, 'Rayon', 'Cuisine', 3],  
7     ['_setCustomVar', 3, 'Famille', 'Grille-pain', 3],  
8     ['_trackPageview']  
9   );
```

Envoi par email

- ❑ Possibilité d'automatiser les envois de rapport par email
- ❑ Très pratique pour fournir des informations spécifiques de manière régulière à un ensemble de personnes concernées



Envoi du rapport par e-mail : *Présentation de l'audience*

De nilschaef@gmail.com

À

Objet Google Analytics: Présentation de l'audience

Pièces jointes

Fréquence Jour de la semaine :

► OPTIONS AVANCÉES

Exportation

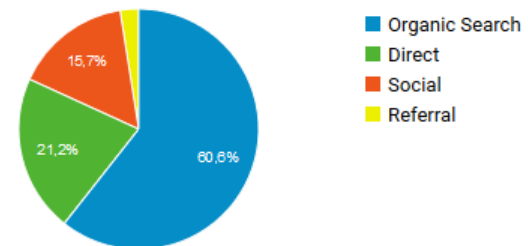
- ❑ Récupération possible des toutes les données
- ❑ Utile pour les manipuler directement et pour réaliser des présentations spécifiques
- ❑ Différents formats
 - CSV
 - TSV
 - Excel
 - PDF

Rubrique « Acquisition »

- ❑ Regroupe les informations relatives à l'origine des utilisateurs
- ❑ Fondamental de pouvoir connaître l'origine des utilisateurs
- ❑ Plusieurs canaux

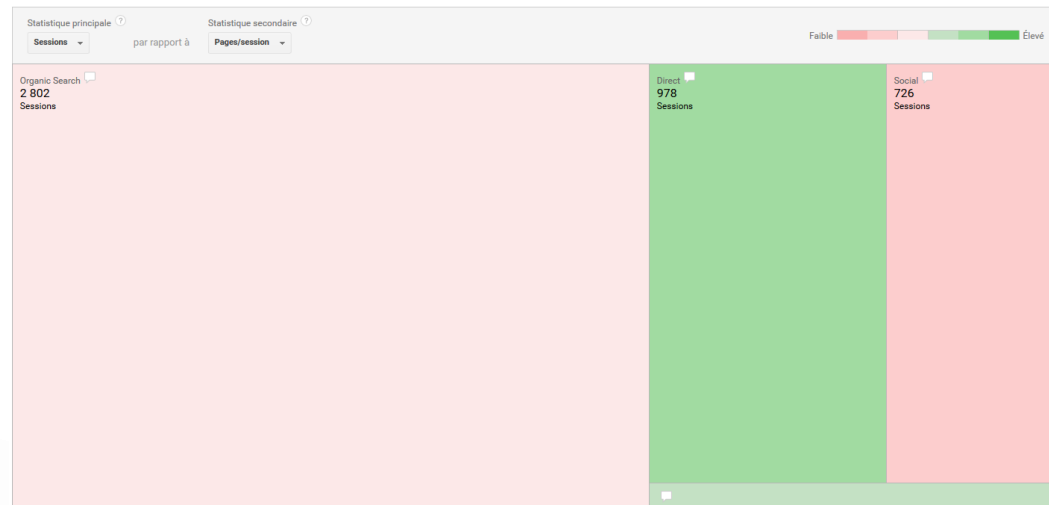
- Recherche organique
- Depuis un lien
- En direct
- ...

Principaux canaux



Treemaps

- ❑ Cartographie qui permet de faire ressortir, pour une statistique donnée, les cas où la valeur est faible ou élevée
- ❑ Notation relative



Rubrique « Comportement »

- ❑ Durée des visites
- ❑ Pages vues
- ❑ Contenu du site
 - Page de destination
 - Page de sortie
- ❑ Vitesse du site
 - Conseils d'optimisation
 - Voir aussi l'outil GTmetrix

Rubrique « Comportement »

□ Recherches sur le site

- Termes utilisés pour arriver sur le site
- Accessible uniquement avec AdWords

□ Événements

- Statistiques sur des événements personnalisés
 - Accès à des documents autres que pages Web
 - Comportement (survol, mise en pause, déplacements....)
- ``
Documentation
``

Rubrique « Conversions »

Mise en place d'objectifs

- URL à visiter
- Durée à atteindre
- ...

Chemin de navigation vers l'objectif

Entonnoir de conversion

Entonnoir multicanaux

Base de données

❑ Besoin de données complexes

- Besoin de passer par la base de données

❑ Interroger la base de données

- Langage SQL

❑ Répondre à des questions complexes

- Qu'elle est l'évolution du panier moyen entre 2012 et 2016 pour les clients inscrits avant 2010 et ayant commandé au minimum 5 fois par an ?