

Fichier valpineta_log

Ce fichier contient les données des log du site valpineta récupérés dans trois csv qui correspondent ici à trois feuilles : log_visit ; log_action ; log_link_action

Feuille log_visit

Les « individus » sont ici les visites : une visite est une succession d'actions faites par un visiteur. Nous avons pour chaque visite les informations suivantes :

idvisit	Identifiant de la visite
visitorID	Identifiant du visiteur
visit_last_action_time	la date-heure de la dernière action de la visite
visit_first_action_time	la date-heure de la première action de la visite
visitor_returning	si la visite est la première visite pour ce visiteur ou non. Normalement il n'y a que deux valeurs, 0 et 1. Ici, il y a la valeur 2, on ne sait pas trop à quoi elle correspond.
visitor_count_visits	le nombre de visites que le visiteur a effectué jusqu'à celle-ci
visit_entry_idaction_name	Identifiant idaction du nom de la page de la première action de cette visite
visit_entry_idaction_url	Identifiant idaction de l'URL de la première action de la visite
visit_exit_idaction_name	Identifiant idaction du nom de la page de la dernière action de cette visite
visit_exit_idaction_url	Identifiant idaction de l'URL de la dernière action de la visite
visit_total_actions	Nombre d'actions effectuées lors de cette visite
visit_total_interactions	Nombre d'interactions effectuées lors de cette visite : il y a interaction quand il y a changement de page. Un téléchargement est considéré comme action mais pas comme interaction, car on reste sur la même page.
visit_total_searches	Nombre de recherches sur le site effectuées lors de cette visite
referer_keyword	Mot clé saisi dans le moteur de recherche quand la visite provient d'un moteur (referer_type=2)
referer_name	Moteur de recherche quand la visite provient d'un moteur (referer_type=2)
referer_type	Type de "provenance" : 1-entrée directe ; 2-moteur de recherche ; 3-autre site web ; 7-Réseau social
referer_url	url de provenance
location_browser_lang	Langue utilisée dans le navigateur du visiteur
config_device_model	Appareil
config_device_type	Type de l'appareil : 0-ordinateur de bureau ; 2-tablette ; 1-portable ; 5-TV connectée ; 10-phablette (NULL-generic mobile)
config_os	Système d'exploitation
config_os_version	Version du système d'exploitation
visitor_localtime	Date/heure de la visite dans l'heure du visiteur
config_resolution	Résolution d'écran utilisée par le visiteur pour effectuer cette visite
config_cookie	Indique si le navigateur du visiteur a activé les cookies ou non
config_flash	Indique si le navigateur du visiteur peut visualiser ou non les fichiers flash
config_java	Indique si le navigateur du visiteur peut exécuter Java ou non
config_pdf	Indique si le navigateur du visiteur peut visualiser ou non les fichiers PDF
config_quicktime	Indique si le navigateur du visiteur utilise quicktime pour lire les fichiers multimédias ou non

config_realplayer	Indique si le navigateur du visiteur peut lire les fichiers multimédias Realplayer ou non
config_silverlight	Indique si le navigateur du visiteur peut exécuter des programmes silverlight ou non
config_windowsmedia	Indique si le navigateur du visiteur utilise Windows Media Player pour lire les fichiers multimédias
visit_total_time	Durée de la visite
location_city	Ville dans laquelle se trouvait le visiteur lors de sa visite sur le site
location_country	Chaîne de deux caractères décrivant le pays dans lequel se trouvait le visiteur lors de sa visite sur le site
location_latitude	Latitude du visiteur lors de sa visite sur le site
location_longitude	Longitude du visiteur lors de sa visite sur le site
location_region	Région du pays dans laquelle se trouvait le visiteur
visitor_seconds_since_first	Nombre de secondes depuis la dernière visite de ce visiteur (le cas échéant)
visitor_seconds_since_last	Nombre de secondes depuis la dernière visite de ce visiteur (le cas échéant)

Feuille log_link_visit_action

Les « individus » sont ici les « actions dans les visites », c'est-à-dire les « interactions », qui correspondent aux changements de pages (dans le site ou vers un lien externe), et les autres actions comme les téléchargements ou les recherches sur le site.

Pour chaque action nous disposons des informations suivantes :

idlink_va	Identifiant de l'action dans la visite
idvisit	Identifiant de la visite
idaction_url_ref	Identifiant de l'Url de la page correspondant à l'action précédente (type d'action 1)
idaction_name_ref	Identifiant du Nom de la page correspondant à l'action précédente (type d'action 4)
custom_float	Stocke le temps qu'il a fallu au serveur pour exécuter cette action
server_time	Date-heure convertie dans le fuseau horaire UTC
pageview_position	Rang de la page vue dans la visite, en commençant à 1. Si l'action enregistrée n'est pas une nouvelle page (téléchargement par exemple), le rang ne change pas
idaction_name	Identifiant du nom de la page correspondant à l'action (type d'action 4)
idaction_url	Identifiant de l'URL de la page correspondant à l'action (type d'action 4)
time_spent_ref_action	Durée de l'action précédente en secondes

On dispose donc des Url et des noms des pages qui ont précédés l'action et des pages qui correspondent à l'action. Ces Url et ces noms sont identifiées par un numéro et l'URL ou le nom correspondant de la page est récupéré dans la feuille log_action

Feuille log_action

Le tableau log_action permet de faire correspondre les identifiants utilisés dans la feuille précédente pour décrire les actions dans les visites, aux URL, noms de pages, types d'actions...

Nous disposons des informations suivantes :

idaction	Identifiant de l'action
name	Nom de l'action
type	Type de l'action : 1-URL vers une page du site suivi ; 2-URL d'un lien, sur le site, vers une page externe ; 3-URL d'un fichier qui a été téléchargé depuis le site Web suivi ; 4-l'action est le titre de la page d'une page du site en cours de suivi ; 8-l'action est une action de recherche de site
url_prefix	si le nom est une URL, il s'agit du préfixe de l'URL : 0: 'http://' ; 1: 'http://www.' ; 2: 'https://' ; 3: 'https://www.'

Fichier sitemap

Le fichier sitemap contient deux feuilles : sitemap et log_action

- Sitemap : liste des pages du site avec le type et la langue (Français ou Espagnol)
Les types des pages :
 - Les pages des menus :
 - ✓ Menu principal : PageAccueil, PageRefuge, PageVallée, PageActivité, PageMedias
 - ✓ Familles d'activités : PageEscalade, PageExcursion, PageGlace, PageSkiRaquette, PageSkiFondPiste, PageVélo
 - Les activités : Alpinisme, Canyoning, catégorie, Escalade, Excursion, Glace
 - PageMentionsLegales et PagePolitiqueConfidentialité
- Log_action : mêmes informations que dans la feuille log_action du fichier précédent, valpineta_log. Enrichies des informations de sitemap.

Fichier origine ?

Dans la feuille log_visit, les données sur l'origine géographique des visiteurs ne sont pas très faciles à traiter en l'état. Il pourrait être intéressant d'avoir une table sur les origines géographiques, qui permettrait de les compléter, nettoyer, catégoriser...

