

TP2

Visualisation de données sous Tableau

Les données Nous allons visualiser les données synthétiques de la production scientifique mondiale de 1996 à 2014.

Les données sont disponibles en format csv et nommées dataset_tableau.csv

Ce fichier comporte les variables suivantes :

- Year : année
- Rank : rang du pays en terme de production (par rapport au nombre de documents) pour l'année indiquée
- Country : pays
- Documents : nombre de documents produits
- Citable.documents : nombre de documents citables (articles, reviews et conférences)
- Citations : nombre de citations de documents produits
- Self.Citations : nombre de citations dans un document du même pays
- Citations.per.Document : nombre moyen de citations
- H.index : nombre hh d'articles ayant eu au moins hh citations (pour plus d'information : h-index sur Wikipedia)

Partie I: Imiter

Nous allons aborder les points suivants :

1. Importation des données dans Tableau
 - o Fichier texte puis choisir le fichier
2. Vérification des données (type particulièrement)
 - o Toutes sont des numériques (symbole # au dessus du nom des variables)
 - o Sauf Country, avec le symbole , qui indique que c'est une variable géographique
3. Première Feuille et premier tableau : Production scientifique
 - o Cliquer sur Feuille 1 en bas à gauche
 - o Double-cliquer sur Documents, dans les Mesures
 - o Cliquer sur le tableau dans la partie Montre moi (en haut à droite)
 - o Glisser et Déplacer Country (dans les Dimensions) dans la partie pour les lignes
 - o Faire un clic droit dans la zone au-dessus des valeurs obtenues puis choisir le tri dans l'ordre décroissant
 - o Faire un clic droit sur Country dans le tableau et choisir Filtrer
 - Cliquer sur Premiers, choisir Par champs puis OK
 - o Double-cliquer sur le titre "Feuille 1" en haut pour le changer

- o Double-cliquer sur l'onglet Feuille 1 en bas pour changer le titre de la feuille

- o On a un TOP 10 des pays sur leur production scientifique

4. Deuxième feuille et premier graphique : Evolution de la production scientifique

- o Cliquer, dans le menu, sur Feuille de calcul puis Nouvelle feuille de calcul, ou cliquer en bas de la fenêtre sur le + avec l'icône d'un graphique

- o Double-cliquer sur Documents, puis sur Year

- o Changer le titre (idem que précédemment)

- o Dans la partie à gauche, cliquer sur Analyse

- Faire un Glisser/Déplacer de Ligne de constante dans le graphique

- Choisir la case pour Year et indiquer l'année 2002

- Avec un clic droit sur la ligne, il est possible de la modifier (type de ligne, couleur, étiquette, ...)

- o Faire de même pour 2013

- o Toujours dans Analyse

- Faire un Glisser/Déplacer de Ligne de moyenne maintenant

- Sélectionner le croisement SOMME ... et Table

- Nous avons maintenant une ligne représentant la production scientifique annuelle moyenne (toujours autant modifiable)

- o Changer aussi le titre de la feuille

- o On a l'évolution de la production scientifique mondiale sur la période analysée

5. Premier Tableau de bord : Synthèse de la production scientifique mondiale

- o Cliquer Tableau de bord puis Nouveau Tableau de Bord dans le menu, ou sur l'icône en bas avec un + et un quadrillage

- o Glisser/Déplacer chacune des feuilles précédemment créées en les sélectionnant dans la partie gauche

- o Le jeu est ici de faire des Glisser/Déplacer pour avoir l'organisation que l'on souhaite

- o Il est possible d'ajouter des images, des pages web ou du texte (cf Objets dans le menu à gauche)

- Ajouter un petit texte explicatif du résultats éventuellement

- o Il est aussi possible de modifier la taille du tableau de bord

- o Nous avons notre premier tableau de bord

- Revenir sur le TOP pour avoir les 20 premiers pays

6. Troisième feuille et première carte : Cartographie de la production scientifique

- o Créer une nouvelle feuille

- o Double-cliquer sur Country

- une carte doit apparaître avec des points

- o Dans le menu Montre-moi, choisir la carte avec les continents en couleurs

- nous avons une carte choroplèthe
- o Double-cliquer sur Documents
 - les pays sont maintenant avec une couleur en fonction de leur production scientifique
- o Nous avons une carte choroplèthe indiquant la production scientifique des pays
- o Glisser/Déplacer Citations dans l'icône infobulle dans Repères
- o Faire de même avec H.index
 - Faire un clic droit sur la zone SOMME(H.index), et choisir Mesure (somme) puis Moyenne
- o En pointant la souris sur un pays, nous avons aussi le nombre total de citations et le H index moyen annuel
 - Cliquer sur Infobulle pour personnaliser l'affichage
- o Changer le titre aussi du graphique et de la feuille

7. Quatrième feuille : Evolution de la contribution des plus importants pays

- o Créer une nouvelle feuille
- o Glisser/Déplacer Country sur les lignes
- o Idem sur Filtres
 - Choisir les 10 premiers sur la production totale de documents
- o Glisser/Déplacer Year dans les colonnes
 - Clic sur la flèche à droite de Year dans les colonnes et choisir Discret
 - On a maintenant un tableau
- o Double-cliquer sur Documents
- o Dans Analyse (dans le menu), choisir Pourcentage de puis colonnes
- o Dans Montrez-moi, choisir les diagrammes en barres empilées
- o Changer le titre et le nom de la feuille
- o Nous avons l'évolution de la contribution des 10 plus gros pays producteurs scientifiques (un peu biaisé ici, car c'est la répartition entre les 10, mais c'est révélateur tout de même)

8. Cinquième feuille : Evolution par pays, en base 100

- o Faire un clic-droit sur la deuxième feuille créée (celle avec l'évolution de la production scientifique), puis choisir Dupliquer
- o Glisser/Déplacer la feuille (nommée Feuille 5 normalement), après les autres
- o Supprimer la ligne horizontale indiquant la moyenne
- o Glisser/Déplacer Country sur Couleurs (dans Repères)
 - Il y a trop de pays pour tous les afficher
 - Choisir Filtrer puis ajouter
 - Chercher et ajouter les pays suivants :
 - United States
 - France
 - China
 - Germany

- United Kingdom
- o Cliquer sur la fenêtre à droite de SOMME(Documents) dans Lignes
 - Choisir Ajouter un calcul de table rapide
 - Dans le Type de calcul, sélectionner Pourcentage de
 - Dans la partie En rapport avec, choisir Premier, puis fermer la fenêtre
- o Nous avons les évolutions en base 100 des 5 pays sur la période analysée
- o Cliquer sur la flèche à droite de Country dans Filtres
 - Choisir Afficher le filtre
- o Il est maintenant possible de choisir les pays à afficher dans le graphique
 - Regarder l'évolution impressionnante de certains pays tels que :
 - Serbia
 - Macao
 - Montenegro
 - Afghanistan
 - Iran
 - Qatar
 - ...

9. Sixième feuille : TOP par année, à taille variable

- o Dupliquer le premier tableau et placer le à la fin
- o Faire un clic-droit dans le vide du panneau de gauche et choisir Créer un paramètre...
 - Donner lui un nom (par exemple pChoixTaille)
 - Choisir Entier comme type et cliquer sur Liste pour définir les valeurs possibles
 - Cliquer pour ajouter les valeurs (ici 3, 5, 10 et 50 par exemple)
 - Noter que la valeur par défaut est la première indiquée mais que c'est modifiable si besoin
 - Cliquer sur OK
- o Cliquer sur la flèche à droite de Country dans le panneau Filtres et cliquer sur Modifier le filtre...
 - o Retourner dans la partie Premiers
 - Si vous cliquer sur la flèche à droite du 10, vous pouvez sélectionner votre paramètre (pChoixTaille donc normalement)
 - Vous avez maintenant un TOP 3
 - o Pour afficher le contrôle du paramètre, cliquer sur la flèche à droite du paramètre dans le panneau de gauche et choisir Afficher le contrôle de paramètre
 - o Vous pouvez maintenant sélectionner le nombre de pays affichés dans votre TOP
 - o Recréer un paramètre comme précédemment

- Nommer le pChoixAnnee
 - Choisir Entier pour le type et Plage pour les valeurs
 - Plutôt que de chercher les années min et max, nous pouvons les récupérer directement. Cliquer sur Définir à partir du champ... et sélectionner Year
 - Choisir 2014 comme valeur actuelle puis cliquer sur OK
 - o Faire un clic-droit dans le vide du panneau de gauche et choisir Créer un champs calculé... et nommer-le cChoixAnnee
 - Entrer la formule suivante dans le cadre dédié
IF ([Year] = [pChoixAnnee]) THEN 1 ELSE 0 END
 - Faire un clic sur la flèche à droite du champ Choix année et cliquer sur Convertir en dimension
 - Glisser-déplacer cette dimension dans le panneau Filtres
 - Sélectionner la valeur 1 et cliquer sur OK
 - Afficher le contrôle du paramètre pChoixAnnee pour pouvoir choisir l'année présentée
 - o Nous avons un TOP à taille variable et sur une année choisie
 - Il est possible de modifier les titres des contrôles de paramètres éventuellement
10. Première Histoire : Analyse de la production scientifique
- o Cliquer sur Histoire, puis Nouvelle Histoire dans le menu, ou sur l'icône avec un + sur un livre en bas
 - o Changer le titre et le nom de l'histoire
 - o Glisser/Déplacer le tableau de bord dans le cadre prévu
 - o Double-cliquer sur Ajouter une légende pour ajouter un texte présentant le tableau de bord
 - o Cliquer sur Nouveau plan vierge
 - Glisser la carte
 - Ajouter une légende
 - Si on souhaite cacher la légende, qui n'a que peu d'intérêt ici, retourner sur la feuille correspondant à la carte
 - Cliquer dans le menu sur Analyse, puis Légendes et décocher Légende des couleurs
 - retourner sur l'histoire pour voir que la légende a bien été enlevée
 - Idem, on peut recadrer la carte si on le souhaite (toujours dans la feuille)
 - o Ajouter la répartition de la production scientifique des 10 pays les plus productifs
 - o Idem pour l'évolution en base 100 et pour le TOP par année (dans deux plans différents)
11. Vous devriez avoir à la fin ce tableau!

Partie II: Adapter

Réaliser les feuilles suivantes :

- Le nuage de points entre le nombre moyen de documents produits par an et le nombre moyen de citations par an, pour chaque pays, avec
 - o le nom de chaque pays indiqué sur le graphique
 - o une couleur (entre du vert pour le 1er et du rouge pour le dernier) par pays en fonction de son rang médian
 - o une taille en fonction du H-index moyen
 - o deux lignes de références au niveau des moyennes de chaque variable
 - o des axes logarithmiques
- Une représentation graphique de l'évolution des rangs des pays, uniquement sur les 10 premiers rangs, avec le pays en 1er en haut, avec
 - o une couleur par pays
 - o le nom de chaque pays sur le graphique
- Une carte des pays dont le H-index moyen est supérieur à une valeur choisie par l'utilisateur, avec
 - o valeurs dans le domaine des valeurs possibles pour cet indice,
 - o la couleur des pays dépendante de la valeur de l'indice, avec une échelle de couleur fixe
- Une carte des pays représentant le TOP réalisé plus haut avec
 - o une sélection des pays identique à celle du TOP
 - o idem pour l'année
 - o une couleur dépendante du nombre de documents produits

De plus, réaliser :

- Un dashboard intégrant le TOP et la dernière carte, avec
 - o le contrôle du nombre de pays, sous forme de radio button (bouton rond avec sélection unique)
 - o le contrôle de l'année, sous forme de slider
- Une story regroupant tous ces éléments de manière cohérente