

06. Fichier VCard

Durée 1h30
(+ 1h30 pour
l'approfondi

Thématique

**DONNEES
STRUCTUREES**

Activité
Débranchée
non

Description de l'activité :

Dans cette activité, l'enseignant.e explique la manipulation des données ouvertes provenant du site data.gouv.fr afin d'en extraire de l'information pertinente.

Durant ce cours l'enseignant.e fait un suivi direct des élèves en situation de manipulation des informations à traiter.

Objectifs pédagogiques :

Objectifs généraux	Objectifs intermédiaires	Compétences
Connaître la notion de données ouvertes	Savoir traiter des données à l'aide d'un tableur (filtrer, trier, croiser)	

Tags :

Serveur DNS #Nom de domaine #TLD (Top-Level Domain) #IP (Internet Protocol)

Matériel :

- Ordinateur connecté à internet

Déroulé de l'activité

Étape 1 : Recherche (10 minutes)

Les élèves téléchargent sur le site [Accueil - data.gouv.fr](https://www.data.gouv.fr) des fichiers sur les émissions CO2 de véhicules.

Étape 2 : Traitement de l'information (10 minutes)

Les élèves décompressent le fichier et l'ouvrent, dans un premier temps, avec un éditeur de texte de type Notepad++ afin de regarder sa structure et notamment les descripteurs et les délimiteurs (L'enseignant.e peut aussi en amont donner le fichier à jour déjà décompressé).

Étape 3 : Utilisation du tableur (10 minutes)

Les élèves ouvrent ensuite le fichier avec un tableur (s'il y a des problèmes à l'ouverture, il faut faire fichier -> importer et préciser les délimiteurs utilisés)

Étape 4 : Extraction des informations (45 minutes)

Les élèves trient le fichier de façon à pouvoir en extraire les informations qui leur sont demandées selon les questions directes de l'enseignant.e.

1. Le tri

Nous souhaitons trier les véhicules suivant leur taux d'émission de CO2 et leur type de carburant.

- Quel est le descripteur de la collection des voitures commercialisées en France indiquant le type de carburant ? **cod-cbr (fichier carlab-annuaire-variable)**
- Quelle est la donnée indiquant que le véhicule fonctionne à l'essence sans plomb ? **ES**

2. Le filtre

2.1. Dans le menu **Données** du tableur, des outils de filtres permettent de ne faire afficher que certaines lignes d'une feuille de calcul suivant certains critères. La fonction **Filtrer** (également appelée **AutoFiltre** dans certains tableurs) insère, au niveau d'une ou de plusieurs colonnes de données, une zone combinée permettant de sélectionner les enregistrements (lignes) à afficher. Cette fonction s'utilise de la manière suivante :

- Sélectionner (= mettre en surbrillance) les colonnes auxquelles on désire appliquer le filtrage (ne rien sélectionner revient à sélectionner toutes les colonnes)
- Activer la fonction **Filtrer** : des boîtes combinées sont visibles dans la première ligne de la plage sélectionnée.
- Cliquer sur la flèche de déroulement située dans l'en-tête de la colonne et choisissez un ou plusieurs éléments.
- Une fois la fonction de filtrage exécutée, seules les lignes dont le contenu correspond aux critères de filtre sont affichées. La colonne utilisée pour le filtre est identifiée par une icône différente.

2.2. Si on applique un filtre supplémentaire sur une autre colonne de plage de données filtrées, alors les autres zones combinées listent seulement les données filtrées.

- a) Pour afficher à nouveau tous les enregistrements, sélectionner tout dans toutes les colonnes ayant servi à filtrer les données.
- b) Pour cesser d'utiliser la fonction Filtrer, sélectionner toutes les cellules sélectionnées initialement et désactiver la fonction Filtrer.

3. Les données

REMARQUE, EN FONCTION DE LA MISE A JOUR DU SITE DATA GOUV LA LISTE EST EVOLUTIVE, IL EST DONC CONSEILLER A L'ENSEIGNANT D'EFFECTUER L'EXERCICE ET DE CREER SES PROPRE COPIES D'ECRAN POUR ALIMENTER ET CORRIGER L'ACTIVITE AVEC LES BONNES DONNEES.

Filtrer les entrées pour que seules les voitures fonctionnantes à l'essence sans plomb s'affichent. Les neuf premières entrées sont les suivantes :

	A	B	C	D	E
1	lib_mrq	lib_mod_doss	lib_mod	dsccom	cnit
2	ALFA-ROMEO	159	159	159 1750 Tbi (200ch)	M10ALFVP000G340
3	ALFA-ROMEO	159	159	159 1750 Tbi (200ch)	M10ALFVP000H341
14	ALFA-ROMEO	4C	4C	4C	M10ALFVP000S413
15	ALFA-ROMEO	AR8C SPIDER	8C SPIDER	8C SPIDER	M10ALFVP000S293
16	ALFA-ROMEO	AR8C SPIDER	8C SPIDER	8C SPIDER	M10ALFVP0006039
17	ALFA-ROMEO	BRERA	BRERA	BRERA 1750 TBI (200ch)	M10ALFVP0003266
18	ALFA-ROMEO	BRERA	BRERA	BRERA 1750 TBI (200ch)	M10ALFVP0004267
23	ALFA-ROMEO	GIULIETTA	GIULIETTA	GIULIETTA 1.4 MPI T-jet (120ch) S/S	M10ALFVP0006294
24	ALFA-ROMEO	GIULIETTA	GIULIETTA	GIULIETTA 1.4 MPI T-jet (120ch) S/S	M10ALFVP0007295

EXEMPLE DE TABLEUR (IMAGE INDICATIVE A MODIFIER SELON LES DONNEES VERIFIEES PAR L'ENSEIGNANT.E)

Question à poser au groupe :

Pourquoi les numéros de lignes ne se suivent-ils pas ? Les données sont filtrées, donc certaines lignes sont cachées.

- a) Dans le menu Données, choisir Trier, puis sélectionner ensuite les clefs de tri primaires et éventuellement secondaires en spécifiant si le tri est croissant ou décroissant .
- b) Trier les voitures fonctionnantes seulement à l'essence sans plomb par ordre croissant d'émissions de CO2 dans l'atmosphère.
- c) Quelles sont les deux voitures fonctionnantes seulement à l'essence sans plomb qui émettent le moins de CO2 ?

Étape 5 : Mise en pratique autonome (15 minutes)

Dans cette étape, les élèves utilisent les connaissances qu'ils ont acquises pour déterminer, quelles voitures, roulant uniquement au gazole, émettent le moins de CO2 et le plus de CO2.

Le choix est effectué selon les marques françaises Peugeot, Renault et Citroën.

Les élèves précisent : la marque, le modèle, le code national d'identification du type et le taux de CO2.

Étape 6 : Approfondissement (1h30)

Les élèves vont récupérer des données ouvertes sur le site de <https://www.insee.fr/fr/accueil> sur le découpage administratif (3 fichiers). Ils croisent ensuite les données issues de ces fichiers.

Suivre les instructions suivantes en s'appuyant de la manipulation faite en cour :

1. Aller sur <https://www.insee.fr/fr/accueil>

2. **Suivre le chemin suivant** : Accueil > Définitions, méthodes et qualités > Géographie administrative et d'étude > Téléchargement > Code officiel géographique.
3. **Cliquer sur** téléchargement des fichiers **du millésime 2019**.
4. Dans le sommaire, cliquer sur le lien Régions.
5. Télécharger le fichier csv, puis décompresser le fichier .zip dans votre espace de travail.
6. Récupérer de la même façon les fichiers csv du millésime 2019 sur les communes et les départements.
 - o De quel institut proviennent ces trois fichiers ? Quel est son rôle ?

Ces trois fichiers proviennent du site de l'INSEE (Institut National de la Statistique et des Etudes Economiques). Cet institut est chargé de la production, de l'analyse et de la publication des statistiques officielles en France : comptabilité nationale annuelle et trimestrielle, évaluation de la démographie nationale, du taux de chômage, etc. Le site web de l'INSEE fournit quantité de données ouvertes que tout citoyen peut récupérer et traiter suivant ses besoins. Il est parfois nécessaire de récupérer les données de plusieurs fichiers pour obtenir l'information désirée.

7. Importer ici les trois fichiers csv dans un même classeur sans utiliser de copier-coller.
 - a. Ouvrir les trois fichiers csv avec Notepad++ et prendre connaissance de l'encodage et du type de séparateur utilisés.

Tableur type Excel :

- o Onglet Données > Fichier texte
- o Etape 1 : type délimité : Unicode utf-8
- o Etape 2 : virgule
- o Etape 3 : texte pour toutes les colonnes (sinon perte des 0 dans la première colonne)

Tableur type OpenOffice Calc :

- o Clic gauche à côté des feuilles existantes en bas
 - o Sélectionner à partir d'un fichier
 - o Unicode utf-8; séparateur : virgule; colonne en texte
- b. Lancer un tableur et ouvrir un classeur vide. Sauvegarder ce classeur avec le nom Activite-RegDepCom.
 - c. On désire importer les trois fichiers csv dans trois feuilles de calcul séparées.
 - o Rechercher sur internet comment on peut importer dans un tableur des données externes issues d'un fichier.
 - o Importer les trois fichiers puis renommer chacune des feuilles Régions, Départements et Communes suivant le type de données qu'elles contiennent.

8. Sauvegarder l'ensemble du classeur.
9. On désire ajouter pour chaque commune le nom de la région à laquelle elle appartient. Pour cela, on procède de la façon suivante :
 - a. Dans la feuille Communes du classeur Activite-RegDepCom, saisir le mot nreg (pour Nom Région) dans la cellule L1.
 - b. Dans la cellule L2 de cette même feuille, saisir l'une des deux instructions suivantes :
 - o Tableur type Excel : =RECHERCHEV(C2;Régions!\$A\$2:\$D\$19;4;FAUX)
 - o Tableur type OpenOffice Calc : =RECHERCHEV(C2;\$Régions:\$A\$2:\$D\$19;4;0)

10. Observer le résultat obtenu et expliquer le contenu de la formule saisie. On pourra utiliser l'aide du tableur.

L'instruction demande de chercher la valeur contenue dans la cellule C2 de la feuille courante (premier paramètre) dans la première colonne du tableau de la feuille Régions constituée des cellules allant de A2 (en haut à gauche) à D19 (en bas à droite) (deuxième paramètre), puis de renvoyer la valeur située sur la même ligne que la valeur trouvée et dans la colonne numéro 4 (troisième paramètre). Le dernier paramètre est un booléen (VRAI ou 1 / FAUX ou 0) qui indique si la première colonne est triée ou non. Si on indique FAUX ou 0, il cherche la valeur exacte dans tous les cas. Les symboles \$ servent à bloquer les cellules lors de la recopie de la formule dans les cellules suivantes.

11. Compléter la colonne nreg avec le nom des régions de chacune des communes.

Fiche activité élève

Suivre les instructions suivantes en s'appuyant de la manipulation faite en cour

1. Aller sur <https://www.insee.fr/fr/accueil>
2. Suivre le chemin suivant : Accueil > Définitions, méthodes et qualités > Géographie administrative et d'étude > Téléchargement > Code officiel géographique.
3. Cliquer sur téléchargement des fichiers du millésime 2019.
4. Dans le sommaire, cliquer sur le lien Régions.
5. Télécharger le fichier csv, puis décompresser le fichier .zip dans votre espace de travail.
6. Récupérer de la même façon les fichiers csv du millésime 2019 sur les communes et les départements.

De quel institut proviennent ces trois fichiers ? Quel est son rôle ?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

7. Importer ici les trois fichiers csv dans un même classeur sans utiliser de copier-coller.
 - a. Ouvrir les trois fichiers csv avec Notepad++ et prendre connaissance de l'encodage et du type de séparateur utilisés.

Tableur type Excel :

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Tableur type OpenOffice Calc :

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

- b. Lancer un tableur et ouvrir un classeur vide. Sauvegarder ce classeur avec le nom `Activite-RegDepCom`.
- c. On désire importer les trois fichiers csv dans trois feuilles de calcul séparées.
 - Rechercher sur internet comment on peut importer dans un tableur des données externes issues d'un fichier.
 - Importer les trois fichiers puis renommer chacune des feuilles Régions, Départements et Communes suivant le type de données qu'elles contiennent.

8. Sauvegarder l'ensemble du classeur.

9. On désire ajouter pour chaque commune le nom de la région à laquelle elle appartient. Pour cela, on procède de la façon suivante :

- a. Dans la feuille Communes du classeur `Activite-RegDepCom`, saisir le mot `nreg` (pour Nom Région) dans la cellule `L1`.
- b. Dans la cellule `L2` de cette même feuille, saisir l'une des deux instructions suivantes :
 - Tableur type Excel : `=RECHERCHEV(C2;Régions!A2:D19;4;FAUX)`
 - Tableur type OpenOffice Calc : `=RECHERCHEV(C2;$Régions:$A$2:$D$19;4;0)`

10. Observer le résultat obtenu et expliquer le contenu de la formule saisie. On pourra utiliser l'aide du tableur.

.....

.....

.....

.....

.....

.....

.....

.....

11. Compléter la colonne `nreg` avec le nom des régions de chacune des communes.