

Données ouvertes et orientation :

Ce que les chiffres disent vraiment de nos choix

Classe de seconde — Thème : Les données structurées et leur traitement

Durée : 1 h 30 — En binôme

Nom – Prénom : _____
Nom – Prénom : _____
Classe : _____ Date : _____

Contexte

Vous êtes en classe de seconde et vous devrez bientôt choisir vos trois enseignements de spécialité pour la première. Ce choix vous engage : il influence vos études supérieures, donc votre métier, donc votre niveau de vie.

Tout choix s'appuie sur des informations. Mais ces informations peuvent être fausses, partielles ou biaisées. Aujourd'hui, vous allez apprendre à manipuler des données ouvertes officielles pour répondre vous-mêmes à des questions concrètes sur l'orientation et l'emploi.

Partie I — Les choix de spécialités

Ouvrez le fichier **1.choix_spe_options.xlsx** . Il provient de la Note d'Information 2026 de la DEPP (Direction de l'évaluation, de la prospective et de la performance — Ministère de l'Éducation nationale).

Vocabulaire

- Donnée structurée : information organisée en tableau, avec des lignes et des colonnes.
- Descripteur : titre d'une colonne
- Valeur : contenu d'une cellule (ce qui est affiché dans la case)
- Enregistrement : une ligne entière du tableau (un objet décrit par ses descripteurs).

1. Combien de feuilles (onglets) différentes contient le fichier ? Citez leurs noms.

2. Sur la feuille « Choix spe en premiere », citez les quatre **descripteurs** du tableau.

3. Quelle est la valeur de la cellule B4 ? Que signifie cette valeur ?

La colonne « Part de filles » est vide. C'est à vous de la calculer pour chaque spécialité. La formule est : **Part de filles = Nombre de filles ayant choisi l'enseignement / Nombre d'élèves ayant choisi l'enseignement**. Dans la cellule de la colonne « Part de filles » correspondant à la ligne Mathématiques, écrivez la formule de calcul. Au tableur, pour écrire une formule il faut commencer par « = ». Ici votre formule sera de la forme " $= X / Y$ " où X et Y représentent les coordonnées des cases utiles au calcul.

Étirez la formule pour calculer la part de filles pour toutes les spécialités (jusqu'à la ligne **Total en première général**).

4. Si vous avez correctement fait la manipulation, la part des filles à **Total en première générale** est 56. Si le choix de spécialité n'était pas lié au genre, quelle proportion de filles devrait-on trouver dans chaque enseignement de spécialité ?

Cliquez sur la case D3 et triez le tableau par ordre croissant de « Part de filles ».

5. a) Quelles sont les trois spécialités où la part de filles est la plus élevée ? Et celles où elles sont les moins représentées ?

5. b) À quelles catégories appartiennent les spécialités où les filles sont les plus représentées ? Et celles où elles sont les moins représentées ?

Ouvrez maintenant la feuille « Choix de spécialités ». On y retrouve les données sur une cohorte (élèves qui étaient en première en 2024 et qui sont passés en terminale en 2025). A l'aide des manipulations vues précédemment, complétez les cases colorées en bleu.

En vous inspirant des formules dans les cases **B25**, **C25** et **D25**, complétez les cases **G25**, **H25** et **I25**. (il faut adapter certaines formules car en terminale les élèves ne conservent que deux enseignements de spécialité).

6. Quelle valeur obtenez-vous dans la case **H25** ?

A l'aide d'une formule bien choisie, remplissez la colonne « Ecart » qui calcule la différence entre la part des filles en 2025 et la part des filles en 2024 pour chaque spécialité.

7. Dans quelle spécialité, la part de filles a-t-elle diminué le plus ?

Ouvrez maintenant la feuille « Options en terminale ».

En terminale générale, les élèves qui ont gardé la spécialité Mathématiques peuvent prendre une option supplémentaire de maths :

- Mathématiques complémentaires : option plus facile, pour ceux qui ne veulent pas pousser les maths.
- Mathématiques expertes : option plus difficile, qui ouvre les portes des études supérieures les plus sélectives (ingénierie, prépas scientifiques exigeantes, classes prépas commerce ECG, médecine ...).

8. Calculez la part de filles dans Mathématiques complémentaires et dans Mathématiques expertes. Quels pourcentages obtenez-vous ? Comment pourriez-vous expliquer cela ?

À retenir

Le problème n'est pas que les filles soient moins compétentes en mathématiques : en première, même si cela n'est pas proportionnel au nombre de filles au lycée général (56%) elles sont presque aussi nombreuses que les garçons à choisir la spécialité (48 %).

Mais beaucoup arrêtent ou se dirigent vers l'option la plus facile. Ce phénomène a un nom : **l'autocensure**. Les filles, à compétences égales, s'autorisent moins à viser les filières mathématiques exigeantes.

Partie 2 — Et après le bac ?

Vous savez maintenant quelles spécialités sont les plus choisies par les filles et les garçons au lycée. Mais qu'est-ce que cela change après le bac ? Vous allez le découvrir en explorant les données Parcoursup.

Télécharger un fichier de données ouvertes

Un fichier de données ouvertes (open data en anglais) est un fichier de données mis à disposition de tous, librement et gratuitement par une institution (ici le ministère de l'enseignement supérieur).

Allez sur le site de données ouvertes de l'enseignement supérieur : <https://data.enseignementsup-recherche.gouv.fr/>

Dans le moteur de recherche interne, cherchez les données « Parcoursup Proposition d'admission » :

The screenshot shows the 'Plateforme open data' website for the French Ministry of Higher Education and Research. The header includes the ministry's logo, the platform name, the URL 'data.enseignementsup-recherche.gouv.fr', and buttons for 'Connexion' and 'Inscription'. A navigation bar contains links like 'Accueil', 'Nos thématiques', 'Nos points d'intérêts', 'Explorez tous les jeux de données', 'Créez vos cartes', and 'Utilisez les API'. The main content area features the text '212 jeux de données sur l'enseignement supérieur et la recherche' and 'Jeux de données populaires'. A search bar contains the text 'Rechercher parmi les jeux de données' and 'parcoursup propositions admission'. Below the search bar, a section titled 'MEILLEURS RÉSULTATS' lists two results: 'Parcoursup : propositions d'admission ...' and 'Parcoursup 2018 - vœux de poursuite d...'. To the right, a snippet of the first result is visible, titled 'Parcoursup : propositions d'admission dans l'enseignement supérieur des élèves de terminale diplômés du'.

Sélectionnez

Utiliser les outils pour...

Explorer les données

Parcourir le jeu de données et accéder au schéma

Puis

Exporter les données

Enfin choisissez de télécharger le fichier au format csv

Formats de fichiers plats

- CSV [Jeu de données entier](#)
Le CSV utilise le point-virgule (;) comme séparateur.
- JSON [Jeu de données entier](#)
- Excel [Jeu de données entier](#)

Vous avez ainsi téléchargé le fichier de données ouvertes qui nous intéresse.

Comprendre la structure

CSV signifie Comma-Separated Values, ce qui veut dire en français « valeurs séparées par des virgules ». C'est un format de fichier très simple pour stocker des données structurées en tableau.

Un fichier CSV est en réalité un fichier TEXTE : il ne contient que des caractères, pas de mise en forme (pas de couleurs, pas de polices, pas de formules). Cela le rend léger et lisible par n'importe quel logiciel — un tableur, un éditeur de texte, ou même un programme Python.

Ouvrez le dossier dans lequel vous avez téléchargé le fichier csv → CLIC DROIT sur le fichier → Ouvrir avec... → Bloc-notes (ou autre logiciel de lecture de fichier texte comme notepad++ ou gedit)

9. Recopiez ici la deuxième ligne telle qu'elle apparaît :

La **structure d'un fichier CSV** est la suivante :

- Chaque LIGNE du fichier correspond à un **enregistrement**.
- Sur chaque ligne, les **valeurs** sont séparées par un séparateur. En France, on utilise le plus souvent le point-virgule « ; » car la virgule est déjà utilisée comme séparateur des nombres décimaux (3,14 plutôt que 3.14).
- La première ligne contient en général les **descripteurs**.

Fermez le fichier que vous venez d'ouvrir sans enregistrer.

Ouvrez le dossier dans lequel vous avez téléchargé le fichier csv → CLIC DROIT sur le fichier → Ouvrir avec... → Open office calc (ou autre logiciel de tableur comme Excel). Si besoin sélectionnez bien « ; » comme séparateur.

Sur ces données il y a une petite erreur typologique, il faut donc cliquer sur la colonne avec les nombres décimaux et sélectionner Anglais US sinon le « . » fait que le tableur reconnait le nombre comme du texte.

Importer

Jeu de caractères : Unicode (UTF-8)

Langue : Français (France)

À partir de la ligne : 1

Options de séparateur

☐ Largeur fixe ☒ Séparé par

☐ Tabulation ☐ Virgule ☒ Point-virgule ☐ Espace ☐ Autre

☐ Fusionner les séparateurs ☐ Espaces superflus Séparateur de chaîne de caractères : "

Autres options

☐ Formater les champs entre guillemets comme texte ☐ Détecter les nombres spéciaux

Champs

Type de colonne : Anglais US

	Standard	Anglais US
1	Nombre de candidats bacheliers ayant confirmé au moins un vœu	Nombre de cand.
2	0	0.0
3	0	0.0
4	15	6.0
5	79	35.0
6	4	1.0
7	104	20.0
8	2	0.0

Aide OK Annuler

10. Combien de lignes et de colonnes contient le fichier ? Combien de valeurs différentes prend l'attribut « **Année du Baccalauréat** » ?

Le fichier ici est très volumineux. Nous allons le modifier et appliquer des filtres pour mettre en évidence des informations qui nous intéressent.

Préparation du fichier

Dans la case **G1** écrivez : « **% élèves acceptés** »

Dans la case **G2** écrivez la formule qui permet de calculer ce pourcentage « **=E2/D2*100** »

11. Que voyez-vous dans la cellule G2 ? Expliquez pourquoi.

Copiez la formule de la case G2 jusqu'en bas de la colonne G.

Sélectionnez maintenant toute la première ligne (en cliquant sur le 1)

Enfin sélectionnez dans le menu Données → Filtrer (ou Autofiltre) suivant le logiciel.

Des menus déroulants nous permettent de filtrer les valeurs qui nous intéressent / ne nous intéressent pas et de trier.

Exemple : On veut trouver les spécialités qui optimisent les chances d'être admis à Sciences Po.

Réglages :

- En **Année du Baccalauréat** on garde uniquement **2024** (année la plus récente disponible)
- En Formation (colonne C) on garde uniquement Sciences Po
- En **Nombre de candidats bacheliers ayant reçu au moins une proposition d'admission** on garde uniquement **les valeurs supérieures à 5** (les autres sont ce qu'on appelle des résultats non représentatifs)
- On trie par **% élèves acceptés** en ordre décroissant.

Ainsi les spécialités qui ont donné le plus de choix d'acceptation apparaissent dans les premières lignes.

12. Citez les 5 doublettes qui, après application de ces filtres, donnent les meilleurs pourcentages d'acceptation. Pour chaque doublette indiquez le % d'élèves acceptés.

Réinitialisez vos filtres (on peut utiliser les touches Ctrl + Z pour annuler les dernières manipulations).

13. En vous inspirant de la méthode vue précédemment, donnez en argumentant le meilleur choix de spécialités à choisir en terminale pour être accepté(e) en médecine (PASS)

14. Même question avec deux autres orientations qui pourraient vous intéresser au sein du binôme.

Partie 3 — Etude de 2 profils contrastés

Vous allez maintenant comparer deux trajectoires types :

Profil A — un profil très majoritairement masculin : doublette Mathématiques + Numérique et Sciences Informatiques (vous avez vu en partie 1 que NSI compte seulement environ 19 % de filles).

Profil B — un profil très majoritairement féminin : doublette Humanités, Littérature et Philosophie + Langues, Littératures et Cultures Étrangères et Régionales (vous avez vu que HLP compte plus de 80 % de filles).

Réinitialisez vos filtres et sélectionnez l'Année 2024 uniquement.

Profil A :

Réinitialisez vos filtres et sélectionnez l'Année 2024 uniquement. Sur la colonne « **Enseignements de spécialité** » sélectionnez uniquement « **Mathématiques, Numérique et Sciences Informatiques** ». Triez les résultats par ordre décroissant de la colonne « **Nombre de candidats bacheliers ayant accepté une proposition d'admission** ». N'oubliez pas d'exclure la ligne « **Ensemble des candidats bacheliers** » dans la colonne « **Formation** ».

15. a) Quels sont les 5 types de formation les plus choisis par les bacheliers ayant cette doublette ? Donnez le pourcentage d'acceptations pour chacune.

15. b) Parmi ces 5 formations, combien sont des formations sélectives (écoles, BUT, CPGE...) et combien sont des licences (peu sélectives) ?

Profil B :

Réinitialisez vos filtres et sélectionnez l'Année 2024 uniquement. Sur la colonne « **Enseignements de spécialité** » sélectionnez uniquement « **Humanités, Littérature et Philosophie, Langues, littératures et cultures étrangères et régionales** ». Triez les résultats par ordre décroissant de la colonne « **Nombre de candidats bacheliers ayant accepté une proposition d'admission** ». N'oubliez pas d'exclure la ligne « **Ensemble des candidats bacheliers** » dans la colonne « **Formation** ».

16. a) Quels sont les 5 types de formation les plus choisis par les bacheliers ayant cette doublette ? Donnez le pourcentage d'acceptations pour chacune.

16. b) Parmi ces 5 formations, combien sont des formations sélectives (écoles, BUT, CPGE...) et combien sont des licences ouvertes ?

On notera que la seule licence qui semble « sélective » est Licence **Psychologie, Sciences cognitives** **(le o a été oublié dans la base de données)* mais qu'avec d'autres spécialités les pourcentages d'acceptation sont bien meilleurs.

17. En vous inspirant de la méthode utilisée aux questions 13, 14 et 15 vous pouvez commenter les choix d'enseignements de spécialité à favoriser pour un élève qui souhaite aller en Licence Psychologie, Sciences cognitives.

18. A l'aide de ce que l'on a vu depuis le début de l'activité et de l'infographie fournie, pouvez-vous donner une (des nombreuses) raison qui peut expliquer la différence de salaire entre les hommes et les femmes ? Vous citerez des chiffres précis et des résultats de certaines questions pour justifier vos propos.

19. À partir de tout ce que vous avez vu, complétez la phrase suivante :

« Les filles ne sont pas moins bonnes que les garçons en sciences. Pourtant, elles finissent en moyenne avec des salaires inférieurs. Quels sont les choix successifs qui expliquent cette situation ? »

L'esprit critique fait aussi partie de SNT. Réfléchissez aux limites :

20. Citez deux limites des données que vous avez manipulées aujourd'hui (par exemple : ce que les chiffres ne disent pas, ce qu'ils peuvent masquer, ...).

Le constat à retenir

Ce que disent les chiffres

Les métiers les MIEUX rémunérés sont majoritairement des métiers scientifiques et techniques (informatique, ingénierie, encadrement industriel).

Ces métiers ne sont pas « réservés » aux hommes : les femmes y sont déjà présentes (entre 25 et 35 %). Mais elles y sont MINORITAIRES, ce qui leur laisse une grande place à prendre.

Le ticket d'entrée vers ces métiers, ce sont les CHOIX FAITS DÈS LA SECONDE et LA PREMIÈRE : prendre la spé Maths, garder Maths expertes en terminale, viser une école d'ingénieur·e·s, un BUT informatique, une CPGE scientifique.