

Mini-projet 2023

Base de données – Science des données

Objectif de la partie base de données : vous devez récupérer des données, les insérer dans une base de données relationnelle puis interroger cette base afin de construire plusieurs tableaux que vous utiliserez pour produire des analyses statistiques dans le cadre du cours Science des données 2.

Données : Ce travail utilisera *a minima 2 bases de données* mais vous êtes encouragés à en ajouter d'autres.

Démarche attendue

1. Analyser le descriptif des données pour comprendre comment elles sont structurées.
2. Trouver une utilisation, une question à laquelle répondre. C'est à vous de définir votre problématique.
3. Quel champ ces deux sources de données ont-elles en commun ? Ce champ sera important pour croiser les données (jointure).
4. Quelles données sont importantes pour répondre à la question définie en 2.
5. Proposer un MCD et un MOD.
6. Créer les tables puis pré-traiter les données pour pouvoir les importer.
7. Lister les requêtes qui vous paraissent pertinentes en langage naturel (l'objectif est de varier le type de requêtes).
8. Implémenter les requêtes.
9. Interpréter les résultats.
10. Rédiger un rapport décrivant votre démarche.

Calendrier

- 06 mars remise du MCD
- 07 mai (rapport(*)+vidéo);
- 09 mai (notation croisée)

Rendu :

1. Un export de la base en sql
2. Un rapport de maximum 15 pages décrivant :
 - une introduction contenant des liens vers l'ensemble des données utilisées
 - les objectifs, la question définie qui aura guidé vos analyses
 - le MCD et MOD
 - le listing des requêtes (en langage naturel et en SQL) et leur interprétation
 - les analyses statistiques
 - les difficultés rencontrées
 - les perspectives
 - une conclusion
3. Des diapositives reprenant le plan précédent utilisées pour produire une vidéo (max 12m). Sur chaque diapositive, vous mentionnerez en bas de page le nom de la personne en train de parler. Tous les membres du groupe doivent parler.

Choix des sujets par groupe de 4 étudiants (exactement) :

1. Indiquer comme nom les membres de votre groupe : NOM1/NOM2/NOM3/NOM4
2. Premier arrivé / Premier servi

1. Ou rouler en sécurité ? Répartition géographique des accidents impliquant un vélo

<https://www.data.gouv.fr/fr/datasets/bases-de-donnees-et-fichiers-details-du-recensement-de-la-population/>

<https://www.data.gouv.fr/fr/datasets/bases-de-donnees-annuelles-des-accidents-corporels-de-la-circulation-routiere-annees-de-2005-a-2020/>

<https://www.data.gouv.fr/fr/datasets/reseau-des-itineraires-cyclables/>

2. Dans quels secteurs investir ? État des entreprises en 2022

<https://www.data.gouv.fr/fr/datasets/base-sirene-des-entreprises-et-de-leurs-etablissements-siren-siret/>

<https://www.data.gouv.fr/fr/datasets/chiffres-cles-2022/>

3. Plus on est nombreux, plus on pollue ? Inventaire de gaz à effet de serre territorialisé

<https://www.data.gouv.fr/fr/datasets/inventaire-de-gaz-a-effet-de-serre-territorialise/>

<https://www.data.gouv.fr/fr/datasets/bases-de-donnees-et-fichiers-details-du-recensement-de-la-population/>

4. Jeune diplômé : dans quelle discipline a-t-on le plus de chance de trouver du travail

https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-master_donnees_nationales/information/

https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-lp_donnees_nationales/information/

<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion-professionnelle-doctorat-par-discipline/information/>

https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-dut_donnees_nationales/information/

<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/insertion-professionnelle-des-diplomes-de-doctorat-par-ensemble/information/>

<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion-professionnelle-des-diplomes-doctorat-par-etablissement/information/>

5. Jeune diplômé : Homme/Femme les inégalités sont-elles visibles ?

https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-master_donnees_nationales/information/

https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-lp_donnees_nationales/information/

<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion-professionnelle-doctorat-par-discipline/information/>

https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-dut_donnees_nationales/information/

<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/insertion-professionnelle-des-diplomes-de-doctorat-par-ensemble/information/>

<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion-professionnelle-des-diplomes-doctorat-par-etablissement/information/>

6. Jeune diplômé : Est-on mieux formé à Paris

https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-master_donnees_nationales/information/
https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-lp_donnees_nationales/information/
<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion-professionnelle-doctorat-par-discipline/information/>
https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion_professionnelle-dut_donnees_nationales/information/
<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/insertion-professionnelle-des-diplomes-de-doctorat-par-ensemble/information/>
<https://data.enseignementsup-recherche.gouv.fr/explore/dataset/fr-esr-insertion-professionnelle-des-diplomes-doctorat-par-etablissement/information/>

7. Ou mange t'on le mieux ?

<https://www.data.gouv.fr/fr/datasets/resultats-des-controles-officiels-sanitaires-dispositif-dinformation-alimconfiance/>
<https://www.data.gouv.fr/fr/datasets/data-insee-sur-les-communes/>

8. Parcoursup 2020 - Quelles sont les disciplines et universités les plus attractives : vœux de poursuite d'études et de réorientation dans l'enseignement supérieur

<https://www.data.gouv.fr/fr/datasets/parcoursup-2020-voeux-de-poursuite-detudes-et-de-reorientation-dans-lenseignement-superieur-et-reponses-des-etablissements/>
<https://www.data.gouv.fr/fr/datasets/data-insee-sur-les-communes/>

9. Médicaments : Quels sont les médicaments les plus vendus et remboursés ?

<https://www.data.gouv.fr/fr/datasets/base-de-donnees-publique-des-medicaments-1/>
<https://www.data.gouv.fr/fr/datasets/open-medic-base-complete-sur-les-depenses-de-medicaments-interregimes/>
<https://www.data.gouv.fr/fr/reuses/remboursements-des-medicaments-en-france-global-et-par-medicament/>

10. Temps d'antenne et élection : A-t-on plus de chance d'être élu si on passe beaucoup à l'antenne ? Comparaison des temps d'antenne et des résultats des candidats au 1er tour

<https://www.data.gouv.fr/fr/datasets/temps-de-parole-et-dantenne-des-candidats-a-lelection-presidentielle-2022-premiere-periode-du-1er-janvier-au-7-mars-2022/>
<https://www.data.gouv.fr/fr/datasets/election-presidentielle-des-10-et-24-avril-2022-resultats-definitifs-du-1er-tour/>

11. Elections de 2022 - Quels ont été les rapports de force dans les circonscriptions ?

<https://www.data.gouv.fr/fr/datasets/candidatures-enregistrees-aux-elections-legislatives-de-2022/>
<https://www.data.gouv.fr/fr/datasets/election-presidentielle-des-10-et-24-avril-2022-resultats-du-second-tour/>
<https://www.data.gouv.fr/fr/datasets/election-presidentielle-des-10-et-24-avril-2022-resultats-definitifs-du-1er-tour/>

12. Covid : tableau de bord de l'épidémie

<https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/>

<https://www.data.gouv.fr/fr/datasets/donnees-relatives-aux-personnes-vaccinees-contre-la-covid-19-1/>

<https://www.data.gouv.fr/fr/datasets/capacite-analytique-de-tests-virologiques-dans-le-cadre-de-lepidemie-covid-19/>

<https://www.data.gouv.fr/fr/datasets/synthese-des-indicateurs-de-suivi-de-lepidemie-covid-19/>

<https://www.data.gouv.fr/fr/datasets/donnees-relatives-aux-resultats-des-tests-virologiques-covid-19/>

13. Le marché du travail en France se porte t-il mieux ?

<https://www.data.gouv.fr/fr/reuses/le-marche-du-travail-en-france/>

<https://www.data.gouv.fr/fr/datasets/data-insee-sur-les-communes/>

14. Les villes les plus sûres ?

<https://www.data.gouv.fr/fr/datasets/bases-communale-et-departementale-des-principaux-indicateurs-des-crimes-et-delits-enregistres-par-la-police-et-la-gendarmerie-nationales/>

<https://www.data.gouv.fr/fr/datasets/data-insee-sur-les-communes/>