

IUT de Vannes, BUT Informatique, 2024-2025

M.T. Pham

S2.04 : Exploitation d'une base de données

1 Données

Data.gouv.fr est une plateforme ouverte contenant des données publiques françaises relatives à de nombreux sujets de société : transports, logements, élections, santé, emploi, environnement... La plateforme contient plus de 40 000 jeux de données différents stockés sous des formats divers.

Le jeu de données du projet “**BDVelos**” est composé de plusieurs fichiers csv. Le format csv (Comma Separated Values) structure les données sous la forme de valeurs séparées par des virgules (ou autre délimiteur). La première ligne du fichier (*header*) donne le nom des colonnes/attributs. Ce format est lisible avec n'importe quel éditeur de texte et peut être mis en forme dans n'importe quel tableur.

Description des fichiers : (disponibles sur Moodle)

- “s204_comptageVelo.csv”, donne le nombre de passages de cyclistes depuis entre le 01/01/2020 et le 24/01/2023. Ces données sont complétées par d'autres informations, en particulière une variable indiquant une probabilité d'anomalie sur le compteur pour la journée en question.
- “s204_compteurs.csv.csv”, donne la liste des compteurs avec leur noms et leur sens
- “s204_quartierCompteur.csv”, donne le quartier de chaque compteur
- “s204_quartiers.csv” donne des informations sur les quartiers
- “s204_longueurPistesVelo.csv” donne la longueur totale des aménagements cyclables par quartier
- “s204_temperature.csv” donne la température moyenne par jour.

2 Travaux à réaliser

Pour cette SAE, vous devez travailler en groupe de 3 ou 4 personnes **dans le même groupe TP**.

1. Sous la forme d'un diagramme de classes UML, proposer une base de données permettant d'intégrer toutes les données des fichiers fournis. **Livrable** : un fichier pdf avec le diagramme de classe et ses éventuelles contraintes textuelles.
2. Écrire le script de création de la base. **Livrable** : un fichier .sql contenant le schéma relationnel (dans un bloc de commentaires) et le script de création de la base.
3. Proposer des requêtes et vues sur votre base en respectant les demandes suivantes :

Question	Demande	Nombre de questions
1	jointure interne	1
2	auto-jointure	1
3-4	jointure externe	2
5-6	sous-requête (avec IN et NOT IN)	2
7-8	sous-requête (avec EXISTS et NOT EXISTS)	2
9-10	fonction de groupe sans regroupement	2
11-12	regroupement avec fonction de groupe	2
13-14	regroupement et restriction (avec HAVING)	2
15-16	division (division normale et division exacte)	2
17-18	vue (pour gérer des contraintes à proposer)	2
19-20	vue (pour gérer des informations dérivables à proposer)	2
	Total	20

Livrable : fichier .sql qui contient, pour chaque demande, la question en français (dans un bloc de commentaires) correspondant à chaque requête, la requête avec MySQL, et les résultats de la requête (dans un bloc de commentaires).

L'outil utilisé pour faire le diagramme de classe est laissé libre. Pour la base de données, vous devez utiliser MySQL.

3 Modalités de dépôt et d'évaluation

Les trois livrables devront être rassemblés dans une archive .zip bien nommé **S204_gr1xx_Nom1_Nom2_Nom3_Nom4.zip** et déposés sur l'espace Moodle de la SAE S2.04 (**un seul exemplaire** par groupe) **avant le 04 avril 2025 à 23h59** (en fin de votre semaine de SAE-P3).

IMPORTANT : En plus de ces livrables, **une évaluation orale** est prévue en semaine 17 ou 18 (selon chaque groupe TP). Chaque groupe devrait présenter leur travaux à travers leur livrables. La note finale va être **individualisée** selon les réponses aux questions pour chaque membre du groupe.