

## Projet... suite et fin !

Le projet du second semestre s'inscrit dans la continuité de celui réalisé au semestre précédent. Ainsi, vous devrez repartir du sujet précédent et de votre travail précédent comme point de départ. Même si vos réalisations du premier semestre ne vous ont pas permis d'atteindre tous les objectifs initiaux, vous devrez les réutiliser. L'évaluation de cette seconde partie ne porte que sur les éléments demandés dans le sujet ci-dessous.

### 1. Présentation du projet

Comme mentionné dans le sujet précédent, votre projet vise à reprendre toutes les connaissances et compétences acquises ou renforcées dans les différentes EC traitant des bases de données sur des données réelles. En effet, ces données qui vont être utilisées sont celles du *Metropolitan Museum of Art (Met)* et concernent plusieurs centaines de milliers d'objets d'art. Il y a un peu plus de deux ans, le Met a lancé un programme d'accès ouvert visant à rendre les images et les données des œuvres du domaine public de la collection du musée disponibles dans le cadre d'une promesse de données ouvertes. Le *Met* fournit cet [ensemble de donnée](#) avec plus de 470 000 œuvres d'art, environ 64 376 **artistes** et plus de 1 000 **catégories d'art** (avec plus de 8 582 peintures et 26 969 photographies). Le programme joue un rôle important dans la mission du Met d'élargir sa portée mondiale en faisant de la collection du musée l'une des plus accessibles, découvrables et utiles sur Internet.

1. Au cours de la *première partie* de votre projet, vous avez normalement **recupéré** et **nettoyé** les données, puis vous avez proposé un **schéma global** pour manipuler ces données. Enfin, vous avez proposé quelques premières **requêtes** pour valider l'ensemble de votre travail.
2. Pour cette *deuxième partie*, vous aurez à créer un **outil** pour faciliter l'accès à ces oeuvres d'art et leur indexation via un site web (l'objectif est, bien entendu, **l'indexation**, **l'optimisation**, et **l'interfaçage** de votre base de données avec un site web. Vous n'avez pas besoin de développer « un site super cool » : l'accessibilité et la qualité de la base de données et des requêtes sont plus importantes que l'apparence :)

**Petit rappel du jeu de données** Les données sont fournies dans deux fichiers au format CSV, qui correspondent aux mêmes données dans deux langues différentes, avec les intitulés suivants :

Object Number, Is Highlight, Is Public Domain, Object ID, Department, Object Name, Title, Culture, Period, Dynasty, Reign, Portfolio, Artist Role, Artist Prefix, Artist Display Name, Artist Display Bio, Artist Suffix, Artist Alpha Sort, Artist Nationality, Artist Begin Date, Artist End Date, Object Date, Object Begin Date, Object End Date, Medium, Dimensions, Credit Line, Geography Type, City, State, County, Country, Region, Subregion, Locale, Locus, Excavation, River, Classification, Rights and Reproduction, Link Resource, Metadata Date, Repository.

## 2. Consignes

Pour cette seconde partie, vous devrez au sein de votre équipe (toujours la même) :

- Proposer des index et des possibilités d'optimisation des requêtes proposées au semestre précédent. Si vous le jugez pertinent, vous pourrez ajouter des requêtes, sachant que celles-ci devront permettre de manipuler les données sans aucune connaissance en informatique (permettre la visualisation, modification, suppression de données à partir d'une interface graphique et non de requêtes SQL)
- Proposer un tableau de bord qui permettra de tester l'ensemble des requêtes proposées et de visualiser le contenu de votre base
- Intégrer des fonctionnalités que vous devrez proposer. Voici quelques exemples :
  - Contribuez à faciliter le libre accès à l'art : (sous forme de statistiques) :
    - combien d'artistes sont, de quels pays et régions
    - recherche des photos et peintures venant des villes les plus proches
    - rechercher des artistes de pays spécifiques
    - combien de types d'art différents sont couverts par cet ensemble de données et lesquels sont les plus et les moins fréquents
    - en quelles années les artistes ont été actifs
    - *(peut-être afficher les peintures ou autres œuvres d'art d'un artiste particulier en utilisant le lien de ressource fourni ?)*
    - ...
  - Pour certaines colonnes :
    - calculer le nombre de valeurs
    - calculer la valeur min/moyenne/max
    - calculer le nombre de valeurs distinctes
    - ...
  - Proposer les modifications classiques (insertion, modification, suppression) et plusieurs comptes utilisateurs (administrateur, visualisation)

L'ensemble du travail demandé doit être réalisé en 4.5h (1CM+1TP).

## 1. Dépôts

Vous pouvez et c'est même conseillé déposer avant la date limite (détails à venir).

**Dépôt final - date limite le 20 février à 23h55 - un rapport .pdf et une archive contenant :**

- donnees.sql : un export de votre base de données
- requetes.sql : toutes les requêtes, transactions, index, etc. en lien avec le sujet
- les fichiers de code source (sous forme d'archive) vous permettant d'afficher le tableau de bord

- un rapport résumant l'ensemble de votre travail (5 pages maximum), des copies d'écran de votre tableau de bord et un manuel d'utilisation pour les utilisateurs. Ce rapport devra détailler :
  - Les index proposés en justifiant leur choix
  - Les optimisations proposées
  - Les fonctionnalités proposées

### 3. FAQ :)

- **Est-il possible de faire des groupes de plus de 4 personnes ?** Non, toujours pas !
- **Peut-on changer de groupe ?** Non, sauf cas exceptionnel (la notion d'exceptionnel sera à faire valider par les enseignants).
- **Doit-on déposer les fichiers pour chaque personne ?** Un dépôt par groupe suffit.
- **Combien de requêtes faut-il faire ?** Assez pour montrer que vous avez compris comment faire des requêtes non triviales, que les index sont pertinents et le tout en lien avec le sujet. Si en plus elles sont différentes de celles des autres groupes... Si vous en avez moins de 5 c'est probablement insuffisant, plus de 20 c'est probablement trop.
- **Faut-il faire une interface ?** Oui !
- **Peut-on utiliser autre chose que postgresql ?** Non !
- **Est-il possible d'avoir une correction du travail du premier semestre ?** Non, cela n'est pas prévu et n'aurait pas d'intérêt. L'objectif de cette seconde partie est de démontrer que vous savez interfacer votre SGBD avec une IHM (page web, application en Java, ...) et que vous avez observé vos données pour en proposer des index et des optimisations pertinentes.
- **Quel langage de programmation dois-je utiliser ?** Celui de votre choix. L'objectif de ce cours n'est pas d'apprendre à coder en Java, ou en PHP, ou en javascript, ou en ..., mais bel et bien d'apprendre à interagir avec une base de données à partir d'un langage de programmation.
- **Pour les requêtes devons-nous reprendre toutes nos requêtes du semestre précédent ou pouvons-nous en reprendre que quelques une et en ajouter d'autres que l'on se sert avec notre interface ?** Ne reprenez que quelques une de vos requêtes du semestre précédent et ajoutez-en d'autres qui sont utiles avec votre interface.