



PROJET CLASSICMODELS

Développement d'application pour le Cloud

Cours	Développement d'applications Cloud	
Professeur	TRAVERS Nicolas	
Majeure - Année	IBO - 5A	
Membres de l'équipe	LEVRET Alexandre TAZI Maxime RAMOS Miguel	

TABLE DES MATIÈRES

RESUME DU PROJET	2
ELABORATION DU PROJET	3
SPECIFICATION DES BESOINS	3
DENORMALISATION	4
DEVELOPPEMENT DE L'APPLICATION	9
LES VUES ET LES REQUETES ASSOCIEES	11
BONUS : DEPLOIEMENT DE L'APPLICATION DANS AWS	16
CONCLUSION	17

RESUME DU PROJET

Notre projet consistait à créer une application, sur le langage de notre choix, sur une base de données NoSQL étudiée pendant les cours, à savoir MongoDB.

Tout d'abord, une étape de dénormalisation des données SQL a été effectuée. En effet, nous sommes partis d'une base de données relationnelle qui représente les données d'un magasin de voitures miniaturisées, et que nous avons transformé en JSON pour ensuite l'intégrer de manière dénormalisé dans MongoDB sous différentes collections.

Nous avons donc créé une application web codée en full JS avec la stack MERN (Mongoose, Express, ReactJS, NodeJS). Cette application sert d'interface de visualisation de données pour les requêtes qui seront décrites dans la suite du rapport.

L'application comprend 3 vues :

- Une vue utilisateur standard : il s'agit de 4 requêtes les plus demandées sur la base de données sans intervention de l'utilisateur.
- Une vue Analyste/Décisionnaire : il s'agit de 2 requêtes complexes paramétrables par l'utilisateur.
- Une vue Administrateur : Il fournit des statistiques sur les données.

ELABORATION DU PROJET

SPECIFICATION DES BESOINS

La base de données SQL a été choisie à partir de ce lien :

https://relational.fit.cvut.cz/dataset/ClassicModels

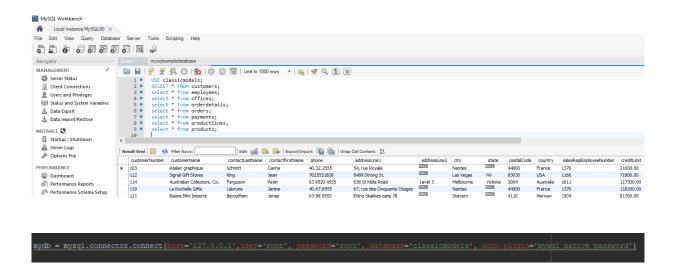
Elle contient les données de magasins de voitures miniaturisées commercialisées internationalement dans les années 2003, 2004 et 2005

Nous avons défini les besoins suivants pour l'interrogation de cette base de données :

- Pour les 4 interrogations les plus demandées nous avons :
 - La liste des clients: pour avoir une vue globale sur les différentes informations d'un client (noms, coordonnées, montant des achats)
 - Les 10 voitures les plus commandées: pour savoir la tendance des produits les plus achetés et ainsi mieux gérer les stocks.
 - Les 5 clients les plus fidèles : pour savoir avec quel client on génère le plus de revenus
 - Chiffres d'affaires mensuel pour l'année 2004 : pour comprendre le taux des ventes au cours de l'année
- Pour les 2 interrogations complexes :
 - Les profits mensuels par année et par pays des magasins : pour connaître les régions du monde où on a généré le plus de profit et ainsi faire de la publicité ciblée.
 - o Les profits mensuels moyens par année et par pays des clients :

Les étapes de la dénormalisation ont été les suivantes :

1- Connecter la base de données sous MySQL vers Python à l'aide de MySQLConnector



2- Exécuter les requêtes en python et les transformer en JSON

```
tables = ["customers", "employees", "offices", "orderdetails", "orders", "payments", "productlines", "products"]
print(len(tables))

for i in range(len(tables)):

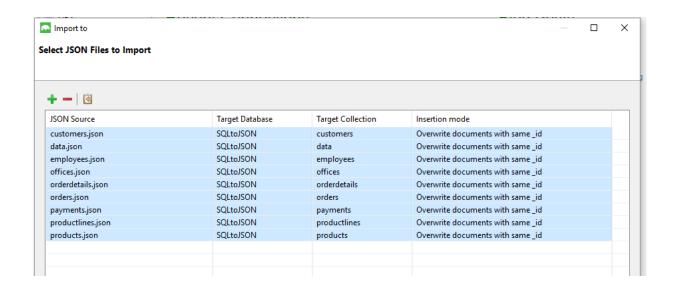
   mycursor.execute("SELECT * FROM " + tables[i])
   row_headers=[x[0] for x in mycursor.description] #this will extract row headers
   myresult = mycursor.fetchall()

   json_data=[]
   for result in myresult:
        json_data.append(dict(zip(row_headers_result)))
   print(json_data)
   print(json_data)
   print(json_data, default=decimal_datetime_default))

with open(tables[i] + '.json', 'w') as outfile:
        json.dump(json_data, outfile, default=decimal_datetime_default)
```

customers.json	28/11/2018 19:39	Fichier JSON	42 Ko
data.json	27/11/2018 23:13	Fichier JSON	42 Ko
employees.json	28/11/2018 19:39	Fichier JSON	5 Ko
offices.json	28/11/2018 19:39	Fichier JSON	2 Ko
orderdetails.json	28/11/2018 19:39	Fichier JSON	342 Ko
orders.json	28/11/2018 19:39	Fichier JSON	61 Ko
payments.json	28/11/2018 19:39	Fichier JSON	27 Ko
productlines.json	28/11/2018 19:39	Fichier JSON	4 Ko
products.json	28/11/2018 19:39	Fichier JSON	46 Ko
SQLtoJSONconverter.py	28/11/2018 19:39	JetBrains PyChar	2 Ko

3- Importer les JSON dans MongoDB sous différentes collections



4- Créer les collections dénormalisées avec des requêtes en utilisant des \$Lookup :

Pour la 1^{ère} requête, la liste des clients, il n'y avait pas de dénormalisation à effectuer, il suffisait d'afficher le contenu de la collection customers.

Pour la 2^{ème} requête, les 10 voitures les plus commandées, on devait dénormaliser les tables orders, orders_details et products pour créer une collection nommée orders.

Pour la 3^{ème} requête et 6^{ème} requête, les 5 clients les plus fidèles et les profits mensuels par année et par pays des clients, on devait dénormaliser les tables customers et payments pour créer une collection nommée customerspayments.

Pour la 4^{ème} requête, chiffres d'affaires mensuel pour l'année 2004, il n'y avait pas de dénormalisation à effectuer, il suffisait d'exploiter les données de la collection payments.

Pour la 5^{ème} requête, les profits mensuels par année et par pays des magasins, on devait dénormaliser les tables customers, employees, offices et payments pour créer une collection customersoffciespayments.

```
idb.getCollection("customersoff(cespsysments").find()

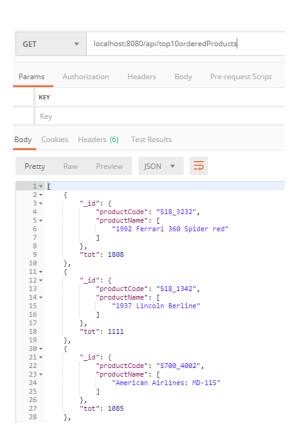
Find 33

"...ia": ObjectId("Sbfc1c57lbebbf50444deab6"),
"...ia": ObjectId("Sbfc1c57lbebbf50444deab6"),
"...ia": ObjectId("Sbfc1c57lbebbf50444deab6"),
"...ia": ObjectId("Sbfc1c57lbebbf50444deab6"),
"...ia": ObjectId("Sbfc1c57lbebbf50444deab6"),
"...ia": ObjectId("Sbfc1c57lbebbf50444deb6"),
"...ia": ObjectId("Sbfc1c57lbebbf5044deb6"),
"...ia": ObjectId("Sbfc1c57lbebbf50444deb6"),
"...ia": ObjectId("Sbfc1c57lbebbf5044deb6"),
"...ia": ObjectId("Sbfc1c57lbebbf5044deb6"),
"...ia": ObjectId("Sbfc1c57lbebbf5044deb6"),
"...ia": ObjectId("Sbfc1c57lbebbf5044dec57"),
"..
```

Au niveau du back-end:

Pour créer le serveur web, où l'application pourrait se lancer, nous avons utilisé le module Express sous NodeJS. Avec ce module nous avons pu créer une API que le client pourrait requêter pour avoir les données voulues à partir de différents endpoints (routing).

Voici quelques exemples des résultats des requêtes http sous Postman :



```
localhost:8080/api/monthlyOfficesPaymentsYears?year=2004
       KEY
  ✓ year
       Key
Body Cookies Headers (6) Test Results
   Pretty Raw Preview JSON ▼
                "results": [
                        "_id": {
    "country": "Australia",
    "month": "02"
                          value": 22474.17
     10 -
                        "_id": {
    "country": "Australia",
    "month": "03"
     11 🕶
     12
                        },
"value": 44894.74
     15
                            "country": "Australia",
"month": "04"
     19
     20
     21
                          ,
'value": 31670.37
```

Certains endpoint contenaient des paramètres.

Le principe était que l'on puisse récupérer les paramètres par l'utilisateur et de traiter ces paramètres avec le module mongoose qui permet d'exécuter des commandes de Mongo DB à partir du code JS.

```
app.get('/api/customers', (req, res, next) => {
    var query = "{}"
    var object_query = JSON.parse(query);
    customers.find(object_query)
    .exec()
    .then((query_result) => res.json(query_result))
    .catch((err) => next(err));
});

app.get('/api/top10orderedProducts', (req, res, next) => {
    opUnwind = { $unwind: "$orderDetails" };
    opGroup = { $group: { _id: { productCode: "$orderDetails.productCode", productName: opSort = { $sort: { tot: -1 } };
    opLimit = { $limit: 10 }
    orders.aggregate([opUnwind, opGroup, opSort, opLimit])
    .exec()
    .then((query_result) => res.json(query_result))
    .catch((err) => next(err));
});
```

Au niveau client:

Pour afficher les résultats des requêtes sous une forme de visualisation de données, nous avons utilisé le framework ReactJS avec les librairies react-table et react-charts.

ReactJS est basé sur des états de données qui changent en fonction d'évènements. Nous avons donc créé des états de données correspondant à chaque requête, et au lancement de l'application (évènement du chargement de DOM) les requêtes sont exécutées.

Les requêtes sont exécutées par l'utilisation de fonctions de « fetch » qui font des requêtes HTTP sur les endpoints que nous avons créé au niveau du serveur.

Voici par exemple la fonction qui fetch les données de l'endpoint customers et qui met à jour l'état des données.

LES VUES ET LES REQUETES ASSOCIES



La première vue, montre les 4 premières requêtes avec find et aggregates.

La liste des clients :

Les 10 voitures les plus commandées :

```
4 // 2 - Les 10 voitures les plus commandés
5 opthwind = {$\smidth{\text{Shift}} \text{"sorderDetails"};}
6 opfroup = {$\script{\text{group} : {\left[disproductCode: "$orderDetails.productName:"}$orderDetails.productName"}, "tot":{$\script{\script{shift}} \text{"sorderDetails.quantityOrdered"}}}};
8 optimit = {\script{\text{shift}} : {\script{\
```

Les 5 clients les plus fidèles :

```
12 opProject = { $project: { paymentTotal: { $sum: "$payments.amount"}, "customerNumber": 1, "customerName":1, "employee":1}}
13 opSort = {$sort:{"paymentTotal":-1}}
14 opLimit = {$limit: 5}
15 db.customerspayments.aggregate([opProject, opSort, opLimit])
                                                              Document Aggregate Document
                                                                                                                                                       Aggregate 🖂
          Document
                          Document Document
                                                                                                                   Document
                                                                                                                                    Document
50
                                 ∨ Documents 1 to 5
          "CustomerNumber": NumberInt(141),
"customerName": "Euro+ Shopping Channel",
"paymentTotal": 715738.98
4
5
6
7
7
8
9
         "customerNumber": NumberInt(124),
"customerName": "Mini Gifts Distributors Ltd.",
11
12 }
13 {
14
15
         "_id" : ObjectId("5bfc1c571bebbf50444deab8"),
"customerNumber" : NumberInt(114),
"customerName" : "Australian Collectors, Co.",
```

Chiffres d'affaires mensuel pour l'année 2004 :

```
17 // 4 - Chiffres d'affaires mensuel pour l'année 2004
18 opMatch = { $match: { "paymentDate": { $gt : "2004-01-01", $lt : "2005-01-01" }}}

19 opGroup = {$group : {_id: {$substr: ['$paymentDate', 5, 2]}, paymentsMonthTotal: {$sum: "$amount"}}};

20 opSort = {$sort: {"_id":1}}
21 db.payments.aggregate([opMatch, opGroup, opSort])
          Document
                            Document
                                            Document
                                                              Document
                                                                               Aggregate
                                                                                                Document
                                                                                                                 Document
                                                                                                                                  Documen<sup>a</sup>
50
                                ∨ Documents 1 to 12
 4 }
5 {
 6
7
8 }
9 {
         "paymentsMonthTotal" : 106652.01000000001
10
11
         "paymentsMonthTotal" : 404603.21
12 }
13 {
14
15
16 }
```



La deuxième vue, montre les 2 dernières requêtes complexes avec map/Reduce et où l'utilisateur peut choisir les années.

Les profits mensuels par année et par pays des magasins :

Les profits mensuels moyens par année et par pays des clients :

```
### Agregate Document Document Document Document Document Document Document Document Aggregate Text Text Document Docume
```

BONUS: DEPLOIMENT DU PROJET DANS LE CLOUD: AWS

Dans le but d'étendre le champ de nos connaissances nous nous sommes penchés sur un déploiement via le cloud. Nous avons choisi d'utiliser Amazon Web Services (AWS) qui, en étant gratuit pendant un an (en fonction des ressources utilisées), nous permettait également d'avoir accès à la communauté la plus importante dans le domaine du cloud, car AWS est leader dans son domaine.

Nous avons donc créé une simple instance via Terraform qui est un soft édité par la société Hashicorp qui permet de faire de l'Infrastructure as Code. Grâce à cela nous pouvons codifier notre infrastructure de A à Z (le réseau, les VM, le stockage block, le stockage objet, les règles de sécurité) ce qui permet de lancer, modifier ou supprimer en une commande son infrastructure cloud. Cela reste très pratique pour le versionning de code.

Voici un bout de code qui permet par exemple la création d'une instance en lui attribuant un nom, son OS, sa zone de disponibilité (datacenter dans une région) et son sous-réseau par exemple :

Une fois que notre infrastructure ait été mise en place il nous restait plus qu'à avoir le code source de notre application via un git clone, d'installer les composants nécessaires, d'importer les json pour MongoDB puis de laisser tourner notre applicatif.

Vous pouvez voir le résultat de notre travail à cette adresse : http://35.180.61.119/

CONCLUSION

Ce projet nous as permis de nous intéresser aux sytèmes de base de données NoSQL, leurs avantages comme leurs inconvénients. Nous avons pu appliquer les technologies NoSQL dans un cas d'utilisation simple comme une application web orienté expérience utilisateur.

Ce projet nous a aussi permis d'intégrer les connaissances acquises pendant les cours comme la dénormalisation des données, la manipulation des requêtes simples find ou aggregates mais aussi des plus complexes en utilisant des map/reduce.

Nous avons aussi été curieux dans la manière de déployer le projet dans le cloud à l'aide de AWS, ce qui facilite la visualisation en temps réel sans avoir à l'installer.

La suite de ce projet est de pouvoir intégrer une vue d'administration qui pourrait afficher les différentes statistiques sur les données.

Pour suivre l'évolution du projet :

Git: https://github.com/MiguelRamosF/classicmodels