Soyhuce challenge :

Contexte: Nous sommes une société de taxi à New York et nous sommes confrontés à un problème. En effet, nos chauffeurs perdent énormément de temps dans les lieux de prises des clients. Connaître en temps réel la concentration de taxi dans New York aurait pour nous une valeur ajoutée significative. De plus, nous aimerions pouvoir analyser à posteriori les trajets des taxis afin de connaître les points de concentration majeurs sur une période de temps donnée, les heures de pointe habituelles... Aujourd’hui, nous avons collecté des données que nous stockons dans une base de données SQLite et avons une application java pour la visualisation. SQLite : faire une recherche sur ce que c’est/ Java : proposer une API Flask permettant la visualisation avec Python Le nombre de taxis étant en forte augmentation, notre solution actuelle ne permet plus d’effectuer les traitements en temps réel et les analyses sont trop longues pour être pertinentes.

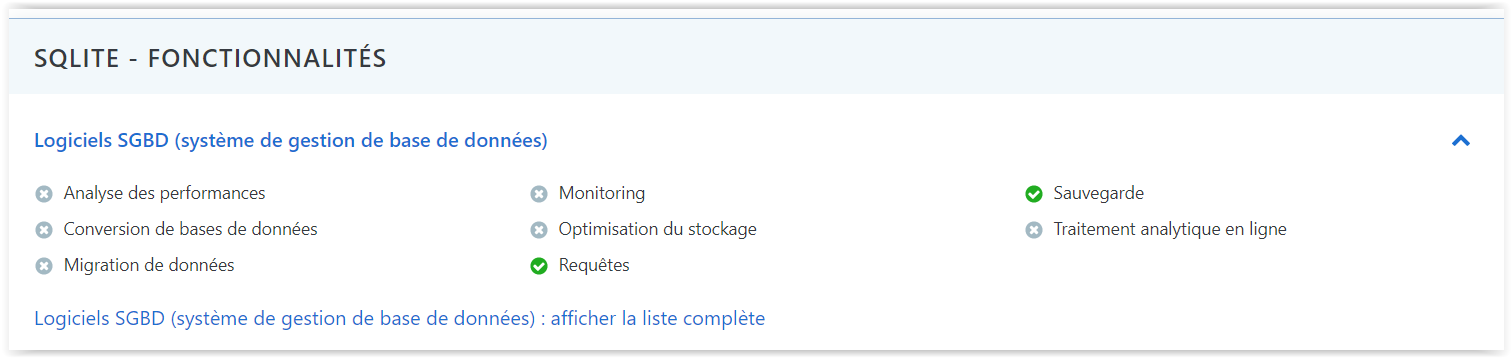
Objectifs: Le challenge se décompose en deux parties:

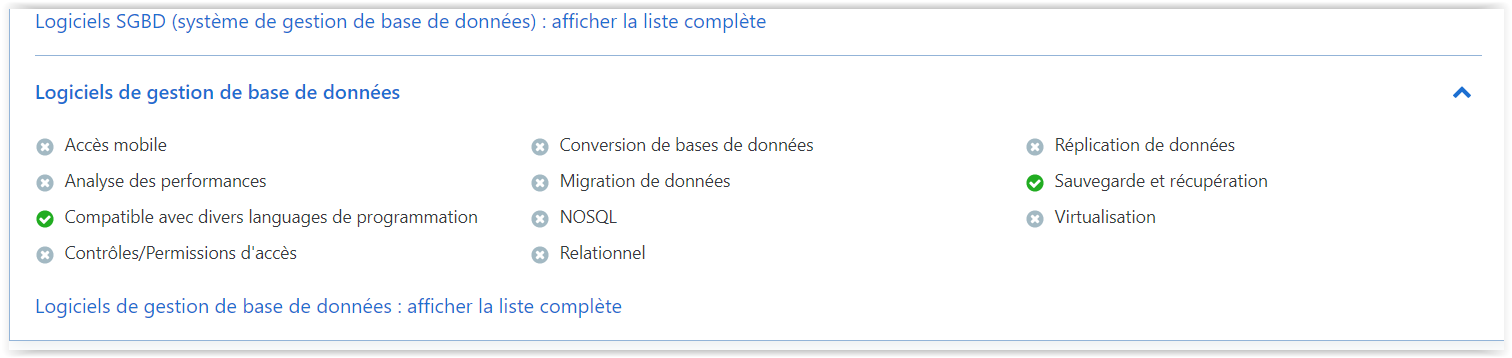
* Imaginer une architecture Big Data capable de répondre à la problématique NoSQL
  + Préciser le design architecturale, les couches fonctionnelles, les technologies permettant d’implémenter ces couches fonctionnelles et tout autre détail jugé important. Cf cours (choisir 2 axes, plus les couches)
  + Imaginer au moins deux scénarios de visualisation répondant à la problématique et concevoir la ou les structures de données adéquates. Jupiter Notebook ?
* Développer une application contenant une visualisation des scénarios. Flask API ou Django API (les scénarios :

Données: Les données viennent d’un challenge kaggle que nous avons détourné pour construire le notre, voici le lien. Libre à toi d’utiliser les données qui te semblent pertinentes.

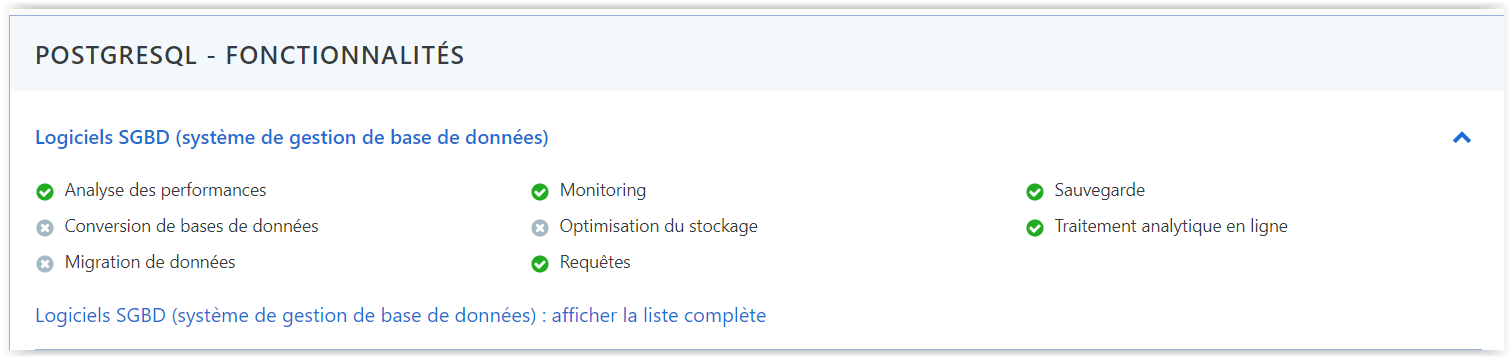
Commentaires: Notre objectif à nous est d’évaluer la démarche plus que les connaissances. Le challenge laisse volontairement beaucoup de liberté, il faudra faire des choix et les expliquer.

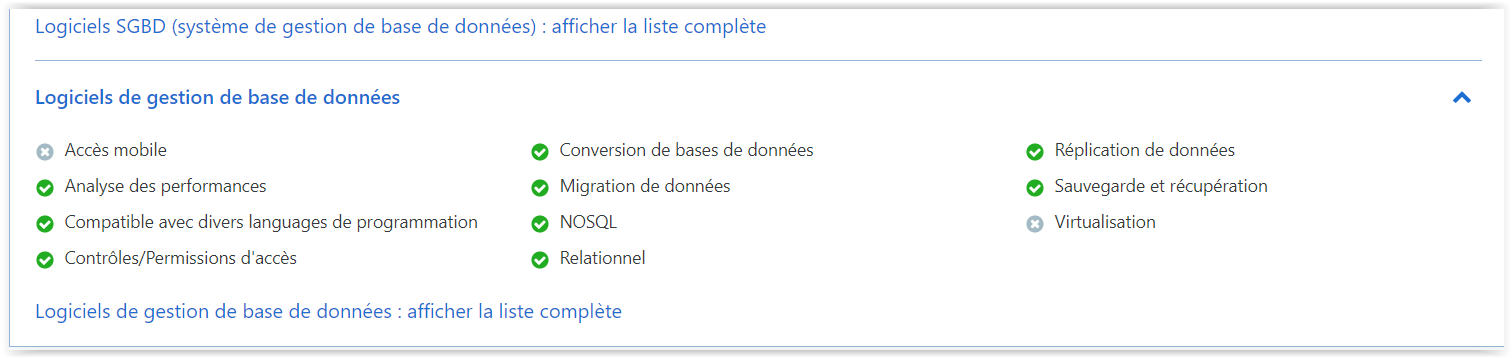
SQL LITE :



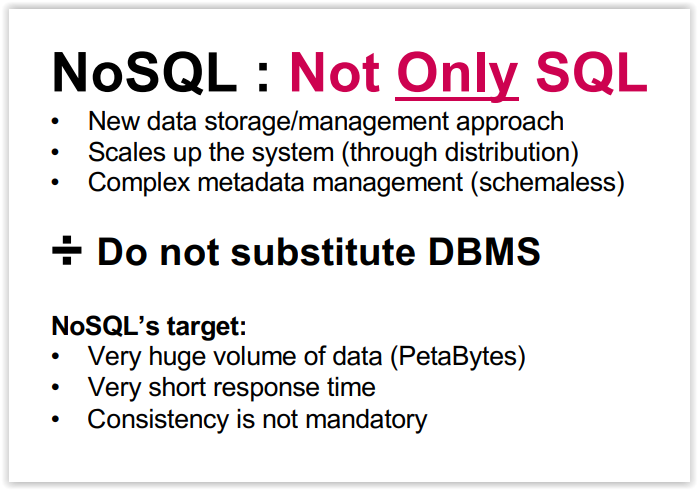


PostgreSQL : <https://www.capterra.fr/software/48791/postgresql>





Why use the NoSQL :



https://www.orsys.fr/pdf-auto/pdfcours/bd.pdf

