

Théorème CAP

Pour chacune des situations suivantes, indiquez quelle garantie peut être relâchée (parmi la consistance, la disponibilité et le partitionnement).

1. Les données peuvent être servies sur un seul serveur.
2. Les données sont trop volumineuses pour être servies sur un seul serveur, et ne sont pas modifiées après leur création.
3. Les données sont trop volumineuses pour être servies sur un seul serveur et doivent être interrogées dans un certain délai.

Objets trouvés à la SNCF

Données (uniquement pour information) : <https://ressources.data.sncf.com>

Quelques exemples de données :

Date	Type d'objets	Nature d'objets	Gare	Code UIC
16 14:57	Bagagerie	Valise, sac sur roulettes	Cherbourg	0087444877
16 14:51	Vêtements, chaussures	Autres vêtements	Cherbourg	0087444877
16 14:50	Vêtements, chaussures	Chaussures	Montparnasse	0087391003
16 14:47	Optique	Lunettes	Montparnasse	0087391003

Considérez que l'on vous donne en temps réel les nouveaux objets trouvés.

Questions

1. Trouvez une clef de partitionnement valide (en expliquant)
2. À partir des quelques exemples utilisez votre clef de partitionnement pour les regrouper
3. Pour la suite nous allons interroger la base tel que vous l'avez organisé. Pour chaque question indiquez si la requête s'effectue "directement" sur une partition ou si plusieurs partitions doivent être interrogées. Dans ce dernier cas indiquez distinctement ce qu'il y a à faire dans la partie map et dans la partie reduce :
 - (a) Quels sont les objets trouvés aujourd'hui ?
 - (b) Quel jour de la semaine trouvons-nous le plus d'objets ?
 - (c) Quel jour de la semaine trouvons-nous le plus d'objets du type vêtement ?
 - (d) Quels sont les objets trouvés à Paris Gare de Lyon ?
 - (e) Combien d'objets trouvés le 2 janvier 2020 ont été restitués ?

Disponibilité des places Vélomagg

Données (uniquement pour information) :

<https://data.montpellier3m.fr/dataset/disponibilite-des-places-velomagg-en-temps-reel>

Quelques exemple de données :

nom	identifiant	latitude	longitude	occupées	libres	total
Jules Ferry - Saint-Roch	001	43.605366	3.881346	10	2	12
Comédie	002	43.608148	3.878778	9	15	24
Esplanade	003	43.609478	3.881293	18	14	32
Hôtel de Ville	004	43.599088	3.894866	6	10	16
Corum	005	43.613989	3.881600	11	1	12
Albert 1er - St Charles	006	43.616768	3.873375	16	15	31
Foch	007	43.610989	3.873345	7	1	8

Considérez que toutes les minutes, vous recevez la liste complète à jour.

Questions

1. Trouvez une clef de partitionnement valide (en expliquant)
2. À partir des quelques exemples utilisez votre clef de partitionnement pour les regrouper
3. Pour la suite nous allons interroger la base tel que vous l'avez organisé. Pour chaque question indiquez si la requête s'effectue "directement" sur une partition ou si plusieurs partitions doivent être interrogées. Dans ce dernier cas indiquez distinctement ce qu'il y a à faire dans la partie map et dans la partie reduce :
 - (a) Combien de places sont disponibles à l'Hotel de Ville ?
 - (b) Je suis à la latitude x et longitude y où puis-je trouver une place ?
 - (c) Quel est le graphique d'occupation de la station Place Albert 1er - St Charles ?