

PARCIALITO 4: NO SQL

Consultas de MongoDB

1. Obtener los ids, cantidad de hashtags y likes para tweets que tengan 50 o más likes ('favorite count') y hayan sido a las 3 de la tarde. Ordenar la salida de forma descendente por cantidad de likes.

```
db.tweets.find(
  {"favorite_count": {$gt: 50}, "created_at.date": /^.{11}15.*"/},
  {"favorite_count: 1, hashtag_count: {$size: "$entities.hashtags"}}
).sort({"favorite_count": -1})
```

2. Para cada hashtag obtener los usuarios que lo utilizaron además del máximo, mínimo y promedio de retweets, sólo teniendo en cuenta aquellos tweets que utilicen más de 3 hashtags (primero se deben filtrar los tweets y luego hallar los valores por cada hashtag).

```
db.tweets.aggregate([
  {
    $match: {
      $expr: { $gt: [{ $size: "$entities.hashtags"}, 3]},
    },
  },
  {
    $unwind: {
      path: "$entities.hashtags",
      preserveNullAndEmptyArrays: false,
    },
  },
  {
    $group: {
      _id: "$entities.hashtags.text",
      used_by_users: {
        $addToSet: "$user",
      },
      max_retweets: {$max: "$retweet_count"},
      avg_retweets: {$avg: "$retweet_count"},
      min_retweets: {$min: "$retweet_count"},
    },
  },
])
```

3. Dada la consulta: Anexo: Consulta ejercicio 3 o disponible en github. Explicar que sucede en cada paso del pipeline y en forma resumida que resuelve la query completa.

La consulta busca los “hilos” de conversaciones, sólo teniendo en cuenta aquellos tweets filtrados en el primer paso.

1. \$match filtra documentos que tengan lenguaje (lang) en español o portugués usando una expresión regular, y tengan como ubicación Brasil.
2. \$group agrupa los documentos por in_reply_to_status_id_str (id del tweet al que responden) y si este es nulo utiliza _id (el id del tweet original), guardando como campos: tweets (id, texto, usuario y fecha) y cantidad de retweets (error en el nombre del campo languages)
3. \$project crea campos nuevos: tweet obteniendo del listado de tweets guardados aquél que coincide su id con el de respuesta (es el original), replies un arreglo de las respuestas al tweet (el id no coincide) ordenado por la fecha de creación y la cantidad de retweets

Consultas de Neo4J

4. Investigue los crímenes cometidos en 165 Laurel Street, muestre las personas que participaron de algún crimen y si tienen relación entre ellas muéstrela.

MATCH

(n:Person)-[:PARTY_TO]->(c:Crime)-[:OCCURRED_AT]->(l:Location{address: "165 Laurel Street"})

RETURN n

5. Muestre la (o las) persona(s) que ha(n) realizado más de 7 comunicaciones telefónicas

```
MATCH (n:Person)-[:HAS_PHONE]-(f:Phone)-[]-(g:PhoneCall)
WITH n.surname, n.name, COUNT(*) AS cant
WHERE cant>7
RETURN n.name, n.surname
```