

MONGODB ASSIGNMENT

BACK-END OF BIG DATA ANALYSIS

TONO

ANTONIO MARTÍNEZ PAYÁ

The name of the collection is “backend”.

1) Kinsmen: Full name and address of UK clients with 2nd surname.

I think it could make more sense to have separated attributes in “Name” and “Surname”, but it is asked to have the complete name, so I did the query using the function concatenate:

```
db.backend.aggregate([{$match:{"contract.country":"United Kingdom",
"Client.Surname":{"$size:2}}},
{$project:{"_id":0,"contract.address":1,
fullname:{$concat:["$Client.Name", " ", {$arrayElemAt: ["$Client.Surname", 0]}, " ",
{$arrayElemAt: ["$Client.Surname", 1]}]}]}])
```

Result:

```
{ "contract" : { "address" : "2 Our Lady of the Garden Street, Sixth floor, door A" }, "fullname" : "Denisa Rosello Bustos" }
{ "contract" : { "address" : "21 Chesterton Road, Fourth floor" }, "fullname" : "Maria de las Aguas Vivas Haya Loyola" }
{ "contract" : { "address" : "54 Milkmen Circus" }, "fullname" : "Praxedes Camones Camargo" }
```

2) Bargains: Contract ID and month(s) of invoices with at most one tap.

```
db.backend.find({$or:[{$and:[{"Movies":{"$exists:false}},{"Series":{"$exists:false}}]},
{$and:[{"Movies":{"$exists:false}},{"Series":{"$size:1}}]},
{$and:[{"Movies":{"$size:1}},{"Series":{"$exists:false}}]}]},
{"contract.contract ID":1,"billing":1,"_id":0});
```

Result:

```
{ "billing" : "January 2016", "contract" : { "contract ID" : "IS26948/01" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "BG42629/14" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "DM12156/01" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "IQ63753/01" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "US59030/06" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "KM38661/10" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "BB20236/13" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "GW43168/09" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "AO29200/02" } }
{ "billing" : "January 2016", "contract" : { "contract ID" : "NIJ00804/09" } }
```

3) Non valid emails (a string, with one 'at' char (@) and one period '.').

```
db.backend.find({"Client.Email":{"$regex": '.*@[^\.]+'.'+$', $options: 'imxs' }}, {})
```

Result: NONE

4) Dud-land: for each country, how many taps of films produced there have been laudable attempts (above 30% of its duration seen) but not achieved (displayed less than 50% of its running time).

```
db.backend.aggregate([{$unwind:"$Movies"},{$match:{"Movies.Viewing PCT":{"$gt: 30,
$lt:50}}} , { $group:{_id:"$Movies.Details.Country", "TAPS":{"$sum:1}}}]])
```

Result:

```
{ "_id" : "Spain", "TAPS" : 1 }
{ "_id" : "Canada", "TAPS" : 3 }
{ "_id" : "France", "TAPS" : 3 }
{ "_id" : "UK", "TAPS" : 10 }
{ "_id" : "South Korea", "TAPS" : 1 }
{ "_id" : "Germany", "TAPS" : 2 }
{ "_id" : "Australia", "TAPS" : 2 }
{ "_id" : "Brazil", "TAPS" : 1 }
{ "_id" : "Netherlands", "TAPS" : 1 }
{ "_id" : "USA", "TAPS" : 80 }
```

5) Yearly total: report every client (DNI, full name), with totals (amount, #taps, traffic in minutes displayed)

```
db.backend.aggregate([
  { $project: { "user": { "dni": "$Client.DNI", "name": "$Client.Name", "surname":
"$Client.Surname", "total_amount": "$TOTAL", "taps": { $concatArrays: [ "$Movies", "$Series" ] } } },
  { $unwind: "$taps",
  { $group:
    { _id: "$user", "total_taps": { $sum: 1 }, "total_traffic": { $sum: { $sum: { $sum: { $multiply: [ "$taps.Details.Duration", { $divide: [ "$taps.Viewing PCT", 100 ] } ] } } }, { $sum: { $multiply: [ "$taps.Avg duration", { $divide: [ "$taps.Viewing PCT", 100 ] } ] } } } } } } }
  ] )
```

Result:

```
{ "_id" : { "dni" : "15625505H", "name" : "Manuel Ignacio", "surname" : [ "Saenz", "Nadal" ], "total_amount" : 26.35 }, "total_taps" : 12, "total_traffic" : 728.38 }
{ "_id" : { "dni" : "45016401L", "name" : "Claudio", "surname" : "Gomez", "total_amount" : 28.75 }, "total_taps" : 14, "total_traffic" : 637.89 }
{ "_id" : { "dni" : "09803499W", "name" : "Andres Daniel", "surname" : "Laureano", "total_amount" : 23.95 }, "total_taps" : 9, "total_traffic" : 599.85 }
{ "_id" : { "dni" : "12549029E", "name" : "Hermenegilda", "surname" : "Hidroso", "total_amount" : 21.1 }, "total_taps" : 9, "total_traffic" : 497.4 }
{ "_id" : { "dni" : "86681439M", "name" : "Vanesa", "surname" : "Chuan-Carlin", "total_amount" : 22.05 }, "total_taps" : 11, "total_traffic" : 400.16 }
{ "_id" : { "dni" : "38162279N", "name" : "Adelaida", "surname" : "Tello", "total_amount" : 17.25 }, "total_taps" : 6, "total_traffic" : 367.55 }
{ "_id" : { "dni" : "13473305C", "name" : "Dario", "surname" : "Rodriguez", "total_amount" : 23.95 }, "total_taps" : 11, "total_traffic" : 566.59 }
{ "_id" : { "dni" : "84581698A", "name" : "Geovanna", "surname" : "Martinez", "total_amount" : 33 }, "total_taps" : 16, "total_traffic" : 1216.03 }
{ "_id" : { "dni" : "27083261W", "name" : "Milagros", "surname" : "Orihuela", "total_amount" : 21.1 }, "total_taps" : 9, "total_traffic" : 478.66 }
{ "_id" : { "dni" : "90657001V", "name" : "Jordan", "surname" : "Peralta", "total_amount" : 24.9 }, "total_taps" : 12, "total_traffic" : 696.81 }
{ "_id" : { "dni" : "63013896Z", "name" : "Juan", "surname" : "Morillo", "total_amount" : 29.65 }, "total_taps" : 13, "total_traffic" : 1148.77 }
```

6) Top ten Spanish movies. Top made by using Facebook likes.

```
db.backend.aggregate([ { $unwind: "$Movies" }, { $match: { "Movies.Details.Country": "Spain" } },
{ $project: { "Movies.Title": 1, "Movies.Details.Facebook
likes": 1, "_id": 0 }, { $group: { "_id": { "Titulo": "$Movies.Title", "Likes": "$Movies.Details.Facebook
likes" } }, { $sort: { "_id.Likes": -1 }, { $limit: 10 } } ] )
```

Result:

```
{ "_id" : { "Titulo" : "Midnight in Paris", "Likes" : 78000 } }
{ "_id" : { "Titulo" : "The Impossible", "Likes" : 54000 } }
{ "_id" : { "Titulo" : "Agora", "Likes" : 29000 } }
{ "_id" : { "Titulo" : "Pan's Labyrinth", "Likes" : 27000 } }
{ "_id" : { "Titulo" : "Buried", "Likes" : 24000 } }
{ "_id" : { "Titulo" : "The Machinist", "Likes" : 21000 } }
{ "_id" : { "Titulo" : "[Rec]", "Likes" : 15000 } }
{ "_id" : { "Titulo" : "Vicky Cristina Barcelona", "Likes" : 14000 } }
{ "_id" : { "Titulo" : "The Orphanage", "Likes" : 11000 } }
{ "_id" : { "Titulo" : "Red Lights", "Likes" : 10000 } }
```