

# Cassandra Assignment

Borja Ruiz Amantegui 100357358

- 1) Dud-land: store the number of taps for each movie and month. Then report a study on monthly taps per production country and per country and director. Take into account that many times the report is limited by the 'viewing pct' (by setting either a minimum %, a maximum, or both).
- ```
CREATE TABLE tablaCassandra1 (Title text, Month text, Country text, Director text, Viewtime double, Taps int, PRIMARY KEY(Month, Country, Director));
```
- ```
db.land.aggregate([{$unwind:"$Movies"},{$group:{_id:{Movie:"$Movies.Title",Month:{$arrayElemAt:[$$split:[$$billing, " "],0]}},ProductionCountry:{$first:"$Movies.Details.Country"},Director:{$first:"$Movies.Details.Director.Name"},Viewtime:{$avg:"$Movies.Viewing PCT"},taps:{$sum:1}}},{$out:"cassandra2"}])
```
- ```
mongoexport --db assignment --collection cassandra2 --type=csv --fields "_id.Movie,_id.Month,ProductionCountry,Director,Viewtime,taps" --out cassandra1.csv
```
- ```
COPY tablaCassandra1 (Title,Month,Country,Director, Viewtime, Taps) FROM 'cassandra1.csv' WITH HEADER = 'true' AND NULL='vacio';
```
- ```
SELECT Month, SUM(Taps) From tablaCassandra1 GROUP BY Month;
```

| month     | system.sum(taps) |
|-----------|------------------|
| June      | 10602            |
| March     | 10953            |
| October   | 11558            |
| May       | 10684            |
| September | 11188            |
| December  | 11773            |
| February  | 9829             |
| November  | 10837            |
| August    | 11344            |
| April     | 10775            |
| July      | 10770            |

2) Yearly total: store every client (DNI, full name) with monthly totals (amount, #taps, traffic in minutes displayed). Then query yearly report.

- CREATE TABLE tablaCassandra2 (TotalAmount double, DNI text, Taps int, Month text, TotalMin double, PRIMARY KEY(DNI));
- db.land.aggregate({\$project:{"TAPS":{\$concatArrays:["\$Movies","\$Series"]},  
"billing":1,"TOTAL":1,"Movies.Viewing PCT":1,"Series.Viewing PCT":1,  
MDuration:"\$Movies.Details.Duration"  
,SDuration:"\$Series.Avg duration"  
,DNI:"\$Client.DNI"  
,fullname: {\$concatArrays:[["\$Client.Name"],["\$Client.Surname"]]}  
}},  
{\$unwind:"\$TAPS"},  
{\$group:{\_id:{\_id:"\$\_id"},  
"TotalAmount":{\$addToSet:"\$TOTAL"}, "DNI":{\$addToSet:"\$DNI"},  
"FullName":{\$addToSet:"\$fullname"},  
"TAPS":{\$sum:1},  
"Month":{\$addToSet:{\$arrayElemAt:[\$\$split["\$billing", " "],0}}},  
"Min  
Series":{\$sum:{\$multiply:["\$TAPS.Avgduration"],{\$divide:["\$TAPS.Viewing  
PCT",100]}}}},  
"Min  
Movies":{\$sum:{\$multiply:["\$TAPS.Details.Duration"],{\$divide:["\$TAPS.Viewing  
PCT",100]}} } } },  
{\$project:{"DNI":1,  
"FullName":1,"TotalAmount":1,"Month":1,"TAPS":1,"TOTALMIN":{\$add:["\$Min  
Series","\$Min Movies"]}}},

```
{out:"cassandra3"}}
```

- `mongoexport --db assignment --collection cassandra3 --type=csv --fields "TotalAmount,DNI,TAPS,Month,TOTALMIN" --out cassandra2.csv`
- `COPY tablaCassandra2 (TotalAmount, DNI, Taps, Month, TotalMin) FROM 'cassandra2.csv' WITH HEADER = 'true' AND NULL='vacio';`
- `SELECT DNI, SUM(Taps), SUM(TotalAmount), Sum(TotalMin) FROM tablaCassandra2 GROUP BY DNI;`

| dni       | system.sum(taps) | system.sum(totalamount) | system.sum(totalmin) |
|-----------|------------------|-------------------------|----------------------|
| 98492841X | 12               | 744.25                  | 710.05               |
| 13292950P | 9                | 47.4                    | 500.98               |
| 77427640A | 12               | 29.65                   | 643.91               |
| 44984986E | 14               | 25.4                    | 551.79               |
| 15486142N | 8                | 385.5                   | 358.34               |
| 03615397G | 13               | 39                      | 706.33               |
| 25770592N | 5                | 37.05                   | 232.27               |
| 54100895L | 5                | 47.16                   | 366.84               |
| 59885229Z | 12               | 22.39                   | 739.63               |
| 62451751N | 10               | 31.5                    | 1007.83              |
| 26661384Z | 13               | 52.65                   | 667.34               |
| 77374999D | 15               | 29.25                   | 865.8                |
| 74731682J | 9                | 20.11                   | 560.3                |
| 14682418T | 11               | 456.57                  | 382.76               |
| 46298658A | 5                | 43.89                   | 190.52               |
| 94149574V | 4                | 12.07                   | 140.51               |
| 37366133B | 15               | 33.55                   | 948.85               |
| 12881150T | 13               | 24.19                   | 919.14               |
| 18624770Z | 11               | 532.77                  | 502.72               |
| 58437441F | 14               | 59.3                    | 880.14               |
| 63905023Y | 13               | 39                      | 563.29               |