

Zain Ul Haq

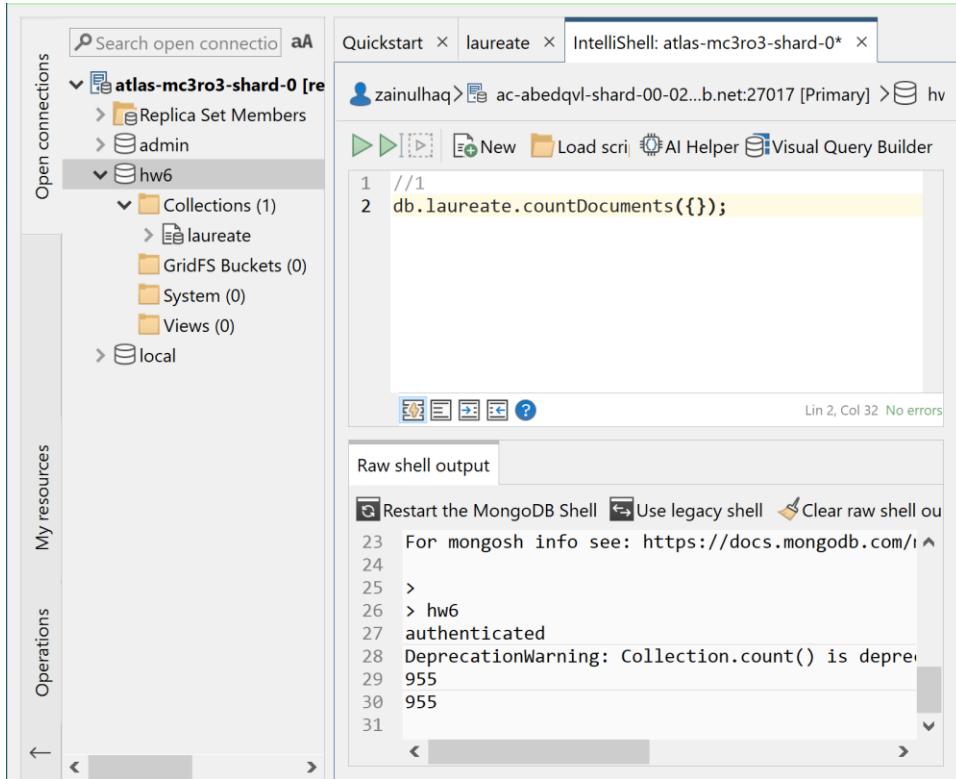
Zh05616

Data Science

Zeehasham Rasheed

Hw6

1.



The screenshot shows the MongoDB Compass interface. On the left, the 'Open connections' sidebar lists 'atlas-mc3ro3-shard-0 [re]' with its replica set members, 'admin', and 'hw6'. The 'hw6' connection is selected. The main area has tabs for 'Quickstart', 'laureate', and 'IntelliShell: atlas-mc3ro3-shard-0*'. The 'IntelliShell' tab is active, showing a command line with the following text:

```
zainulhaq> db.laureate.countDocuments({});
```

Below the command line, the 'Raw shell output' pane displays the results of the command:

```
For mongosh info see: https://docs.mongodb.com/mongosh/v1.0/introduction/using-the-shell.html
>
> hw6
authenticated
DeprecationWarning: Collection.count() is deprecated
955
955
```

This query simply returns the count of all the documents in the json file i.e. laureate, which are 955. This means we have 955 entries in the data.

2.

The screenshot shows the MongoDB Compass interface. At the top, there is a code editor window containing the following MongoDB aggregation pipeline:

```

3 //2
4 db.laureate.aggregate([
5   {
6     $group:{
7       _id: "$diedCountryCode",
8       count: {$sum:1}
9     }
10 },
11 {
12   $sort: {
13     count: -1
14   }
15 }
16 ]

```

The status bar at the bottom right indicates "Line 8, Col 23 No errors". Below the code editor is a "Raw shell output" section titled "Aggregate Query (line 4)". It shows the results of the query:

laureate > count

_id	count
null	335.0
"US"	223.0

At the bottom right of the results table, it says "47 documents" and "00:00:00.305".

This gives us the count for all the country codes where people died, ordered by descending count. Null values for died-country-code are 335 so they are at the top. The US is the second with 223, and so on. There are 47 died-country-codes used in this data.

3.

The screenshot shows the MongoDB Compass interface. At the top, there is a code editor window containing the following MongoDB aggregation pipeline:

```

18 //3
19 db.laureate.aggregate([
20 {
21   $unwind: "$prizes"
22 },
23 {
24   $group:{
25     _id: "$prizes.category",
26     countA: { $sum: 1}
27   },
28 },
29 {
30   $sort: {
31     countA: -1
32   }
33 }
34 ]

```

The status bar at the bottom right indicates "Line 19, Col 1 No errors". Below the code editor is a "Raw shell output" section titled "Aggregate Query (line 19)". It shows the results of the query:

laureate > countA

_id	countA
medicine	222.0

At the bottom right of the results table, it says "6 documents" and "00:00:00.283".

This query only returns to us the category of prize ordered by descending count. Medicine is clearly at the top with 222 entries for it and so on. There are only 6 categories.

4.

```
33 //4
36 db.laureate.aggregate([
37 {
38     $unwind: "$prizes"
39 },
40 {
41     $match: {
42         "prizes.category": "physics"
43     }
44 },
45 {
46     $group: {
47         _id: {gender: "$gender",
48             diedCountryCode:"$diedCountryCode",
49             category: "$prize.category"}, 
50             countA: { $sum: 1}
51     }
52 },
53 {
54     $sort:{ 
55         '_id.diedCountryCode': -1
56     }
57 }
58 ])
])
```

laureate > _id	
_id	countA
{} { 2 fields }	1.0
{} { 2 fields }	60.0
{} { 2 fields }	4.0
{} { 2 fields }	8.0
{} { 2 fields }	6.0
{} { 2 fields }	4.0
{} { 2 fields }	2.0
{} { 2 fields }	1.0
{} { 2 fields }	18.0
... 6 more	...

1 document selected

This data only returns to us the prize with the category physics for each gender, died-country-code and prize category, and their counts. They are ordered by died-country-code only.

6.

```

61 db.laureate.aggregate([
62   {
63     $match: {
64       "prizes.category": "physics",
65       "gender": "male"
66     }
67   },
68   {
69     $group: {
70       _id: {
71         category: "$prizes.category",
72         bornCountry: "$bornCountry"
73       },
74       totalPrizes: { $sum: 1 }
75     }
76   }
77 ])
78

```

Raw shell output | Aggregate Query (line 61) ✖

← ← → | 50 | Documents 1 to 44 | | | | | ▾

```

1  {
2    "_id" : {
3      "category" : [
4        "physics"
5      ],
6      "bornCountry" : "Italy"
7    },
8    "totalPrizes" : 5.0
9  }
10 {
11   "_id" : {
12     "category" : [
13       "physics"
14     ],
15     "bornCountry" : "Sweden"
16   },
17   "totalPrizes" : 4.0

```

| #44 documents | ⏱ 00:00:00.295

This query finds Nobel Prize laureates in the "physics" category who are male, and then groups them by the category and the country of birth. The result shows the total number of male physicists who won Nobel Prizes in each country.