

Title: Database Assignment 5

Your Name: Jack Saunders

Date: 11/22/2024

1. Over how many years was the unemployment data collected?

The first stage groups every year.

The second stage counts the number of Year entries there are.

Answer:

Stage 1: \$group ☐

```
1 {
2   _id: "$Year"
3 }
```

Output after \$group: _id: 2015

Stage 2: \$count ☐

```
1 'Year'
```

Output after \$count: Year : 3

ALL RESULTS **OUTPUT OPTIONS**

Year : 27

2. How many states were reported on in this dataset?

The first stage groups by State.

The second stage counts the number of States.

Answer:

Stage 1: \$group ☐

```
1 {
2   _id: "$State"
3 }
```

Output after \$group stage (Sample of 10 documents): _id: "Oklahoma" _id: "Connecticut"

Stage 2: \$count ☐

```
1 "State"
```

Output after \$count: State : 47

ALL RESULTS

OUTPUT OPTIONS ▾

State : 47

3. What does this query compute?

`db.unemployment.find({Rate : {$lt: 1.0}}).count()`

The shell returns 657 when that line of code is entered

Answer:

```

>_MONGOSH
> use Unemployment
< switched to db Unemployment
> db.unemployment.find({Rate : {$lt: 1.0}}).count()
< 657
Unemployment>

```

4. Find all **counties** with unemployment rate higher than 10%

The first stage restricts the search to have only unemployment rates above 10%.

The second stage only shows the unemployment rate and county.

Answer:

Stage 1 (\$match) ☒

```

1 {
2   Rate : {$gt: 10.0}
3 }

```

Output after \$match stage (Sample of 10 documents)

_id: ObjectId('673d5ef4ab47fdb3ebf4d0a4')
Year: 2015
Month: "February"
State: "Mississippi"
County: "Kemper County"
Rate: 10.6

_id: 0
Year:
Month:
State:
County:
Rate:

Stage 2 (\$project) ☒

```

1 {
2   _id: 0,
3   Rate: 1,
4   County: 1
5 }

```

Output after \$project stage (Sample of 10 documents)

County: "Kemper County"

Rate: 10.6

County Rate:

	<div> ALL RESULTS ▼ Showing 1 – 20 of 91430 ↻ </div> <div> <div>County : "Kemper County"</div> <div>Rate : 10.6</div> </div> <div> <div>County : "Jefferson County"</div> <div>Rate : 14.3</div> </div> <div> <div>County : "Sharkey County"</div> <div>Rate : 11.1</div> </div> <div> <div>County : "Tunica County"</div> <div>Rate : 11.5</div> </div>
--	--

5. Calculate the **average unemployment rate** across all **states**.

<p>The first stage returns the average unemployment rate from all states. Assume: all the states combine to one average</p>	<div> Answer: </div> <div> <div>Stage 1 \$group ▼ ☑</div> <div> <pre>1 { 2 _id: null, 3 averageRate: {\$avg: "\$Rate"} 4 }</pre> </div> <div> Output after \$group stage (Sample of 1 document) <div> <div>_id: null</div> <div>averageRate : 5.577907</div> </div> </div> <div> ALL RESULTS ▼ Showing 1 – 1 of 1 ↻ </div> <div> <div>_id: null</div> <div>averageRate : 6.1750097115006755</div> </div> </div>
---	--

6. Find all **counties** with an unemployment rate between **5%** and **8%**.

<p>The first stage limits the query to documents with an unemployment rate of 5 - 8% inclusive.</p> <p>The second stage groups documents by county.</p>	<div> Answer: </div> <div> <div>Stage 1 \$match ▼ ☑</div> <div> <pre>1 { \$and: [{Rate: {\$gte: 5}}, 2 {Rate: {\$lte: 8}}]}</pre> </div> <div> Output after \$match stage (Sample of 10 documents) <div> <div> <div>_id: ObjectId('673d5ef4ab47fdb3ebf4d0a0')</div> <div>Year : 2015</div> <div>Month : "February"</div> <div>State : "Mississippi"</div> <div>County : "Newton County"</div> <div>Rate : 6.1</div> </div> <div> <div>_id: ObjectId('673d5ef4ab47fdb3ebf4d0a0')</div> <div>Year : 2015</div> <div>Month : "February"</div> <div>State : "Mississippi"</div> <div>County : "Monroe Co"</div> <div>Rate : 7.9</div> </div> </div> </div> <div> <div>Stage 2 \$group ▼ ☑</div> <div> <pre>1 { 2 _id: "\$County" 3 }</pre> </div> <div> Output after \$group stage (Sample of 10 documents) <div> <div>_id: "Ouray County"</div> <div>_id: "Arasstrong Co"</div> </div> </div> </div> </div>
---	---

	<div> ALL RESULTS ▼ Showing 1 – 20 of 1736 ↻ </div> <div> <div>_id: "Ouray County"</div> <div>_id: "Armstrong County"</div> <div>_id: "Izard County"</div> <div>_id: "Saint Landry Parish"</div> <div>_id: "Hawaii County"</div> </div>
--	---

7. Find the state with the **highest unemployment rate**. Hint. Use **{ \$limit: 1 }**

<p>Stage 1 sorts all the rates in descending order</p> <p>Stage 2 shows only the State and unemployment Rate</p> <p>Stage 3 limits what is shown to only 1 result</p>	<div> Answer: </div> <div> <div> Stage 1 Sort ⓘ <pre>1 { 2 "Rate": -1 3 }</pre> </div> <div> Output after Sort stage (Sample of 10 documents) <div> <div> _id: ObjectId('673d5f8bab47fd34ebfa2c3f') Year: 1992 Month: "January" State: "Colorado" County: "San Juan County" Rate: 58.4 </div> <div> _id: ObjectId('673d5f8bab47fd34ebfa2c3f') Year: 1992 Month: "February" State: "Colorado" County: "San Juan County" Rate: 56.5 </div> </div> </div> </div> <div> <div> Stage 2 Project ⓘ <pre>1 { 2 _id: 0, 3 State: 1, 4 Rate: 1 5 }</pre> </div> <div> Output after Project stage (Sample of 10 documents) <div> <div> State: "Colorado" Rate: 58.4 </div> <div> State: "Colorado" Rate: 56.5 </div> </div> </div> </div> <div> <div> Stage 3 Limit ⓘ <pre>1 1</pre> </div> <div> Output after Limit stage (Sample of 1 document) <div> State: "Colorado" Rate: 58.4 </div> </div> </div>
---	--

8. Count how many **counties** have an unemployment rate above **5%**.

<p>The first stage limits the query to documents with an unemployment rate greater than 5%.</p> <p>The second stage groups documents by county.</p>	<div> Answer: </div>
---	-----------------------------

Stage 1 \$match

```
1 {
2   Rate: { $gt: 5 }
3 }
```

Output after \$match stage (Sample of 10 documents)

```
_id: ObjectId('673d5ef4ab47fdb3ebf4d0a8')
Year: 2015
Month: "February"
State: "Mississippi"
County: "Newton County"
Rate: 6.1
```

```
_id: ObjectId('673d5ef4ab47fdb3ebf4d0a9')
Year: 2015
Month: "February"
State: "Mississippi"
County: "Panola County"
Rate: 9.4
```

Stage 2 \$group

```
1 {
2   _id: "$County"
3 }
```

Output after \$group stage (Sample of 10 documents)

```
_id: "Ouray County"
```

```
_id: "Armstrong County"
```

\$match

\$group

Edit

? Explain Export

ALL RESULTS

Showing 1 - 20 of 1736

```
_id: "Ouray County"
```

```
_id: "Armstrong County"
```

```
_id: "Izard County"
```

```
_id: "Saint Landry Parish"
```

```
_id: "Hawaii County"
```

9. Calculate the average unemployment rate per state by year.

Line 2 and 3 of the first stage specifies to group the query by Year and State. This results in an object holding both values. Line 4 takes the average of the results that have been grouped.

Answer:

▼ Stage 1 \$group

1 {

2 _id: {Year: "\$Year",

3 State: "\$State"},

4 Rate: {Avg: "\$Rate"}

5 }

Output after \$group stage (Sample of 10 documents)

▼ _id: Object

Year: 2016

State: "Illinois"

Rate : 6.280880653594771

▼ _id: Object

Year: 2016

State: "West Virginia"

Rate : 7.118181818181818

ALL RESULTS ▼

Showing 1 – 20 of 1258 ↻

▼ _id: Object

Year: 2015

State: "Missouri"

Rate : 5.423260869565218

▼ _id: Object

Year: 1990

State: "Montana"

Rate : 5.305803571428571

▼ _id: Object

Year: 1997


State: "Connecticut"

Rate : 5.002083333333333

10. **(Extra Credit)** For each state, calculate the total unemployment rate across all counties (sum of all county rates).

The first stage groups the results by State. Also, it finds the sum of the unemployment. It is assumed that this sum is calculated based on all the counties within a state according to the data.


Answer:


▼ Stage1 (\$group) 

```
1 {
2   _id: "$State",
3   totalRate: {sum: "$Rate"}
4 }
```

Output after \$group stage (Sample of 10 documents)

_id: "Connecticut" totalRate : 1699.1	_id: "Louisiana" totalRate : 12575.2
--	---

ALL RESULTS 

Showing 1 – 20 of 47 

_id: "Connecticut" totalRate : 14381.8
_id: "Louisiana" totalRate : 50502.9
_id: "Oklahoma" totalRate : 130429.7
_id: "California" totalRate : 152661.6
_id: "Indiana" totalRate : 176544.6
_id: "Idaho" totalRate : 90097.8
_id: "Arkansas" totalRate : 164807.7
_id: "Mississippi" totalRate : 221059.5
_id: "Virginia" totalRate : 217485.8
_id: "New Mexico" totalRate : 75741

11. **(Extra Credit)** The same as Query 10 but for states with data from 2015 onward.

Same as query 10, but Stage 2 restricts the documents to only be from 2015 onward.

Answer:

▼ Stage 1 \$group

1 {
2 _id: {state: "\$state",
3 county: "\$county"},
4 totalRate: {\$sum: "\$rate"}
5 }

Output after \$group stage (Sample of 10 documents)

▼ _id: Object
State: "North Carolina"
County: "Pasquotank County"
totalRate : 255.9

▼ _id: Object
State: "New York"
County: "Washington County"
totalRate : 185.5

▼ Stage 2 \$match

1 {
2 Year2015: {\$gte: 2015}
3 }

Output after \$match stage (Sample of 0 documents)

No Preview Documents