DAMG 6210: Data Management and Database Design: Lab 1

Part 1 (5 points)

- Enter the normalized entities and attributes from the first part of the Normalization Exercise 8 in your design tool. You don't need to add any extra entity or attribute. (20% of grade)

- Establish the relationships between entities using the Crow's Foot notation. Please make sure each relationship line is completely visible and don't cross lines. (80% of grade)

  Note: Don't add any unnecessary relationship. An unnecessary relationship clutters the design and may degrade the database and application performance.

- Upload the ERD as an image or in the pdf format to here.

Part 2 (4 points)

Our data usage pattern is to find the average quarterly sold quantity of a color. Use the embedding design technique and the provided data to design a MongoDB database. Create the database on MongoDB Atlas that includes document(s) reflecting the data contained in the attached file. (50% of grade)

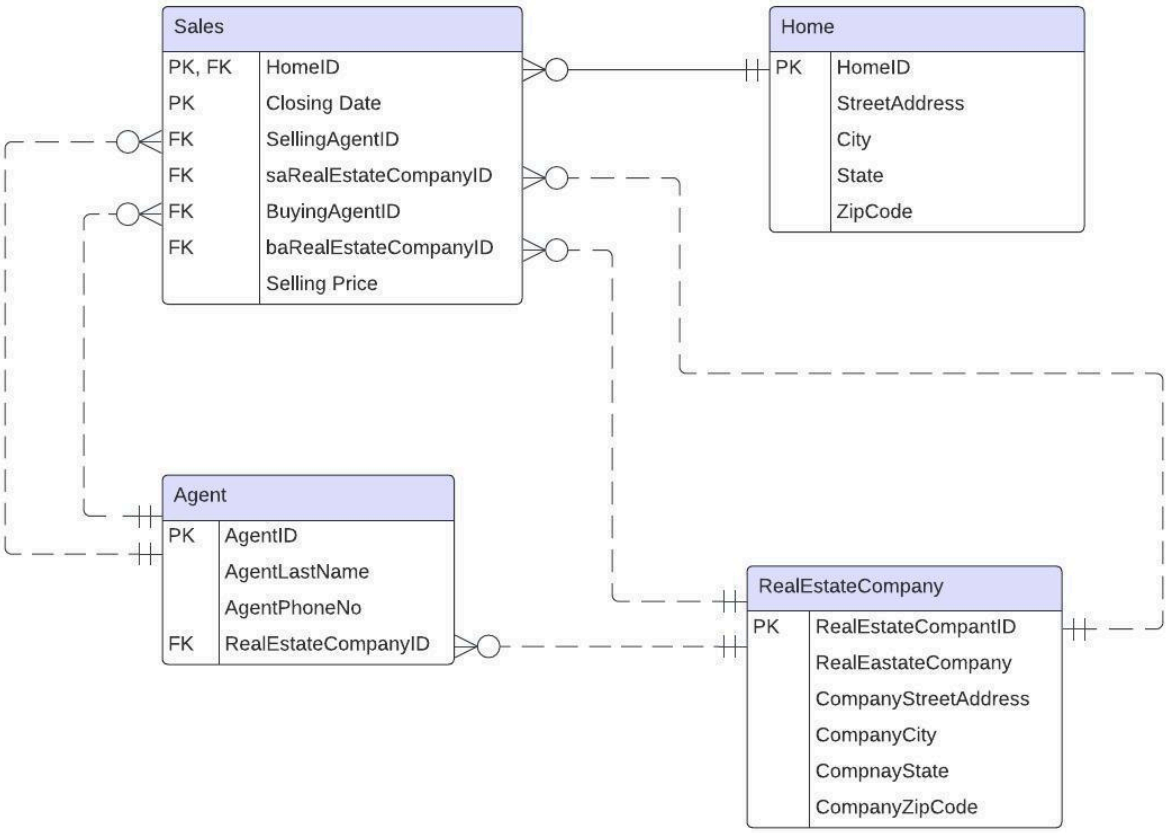
Then use the MongoDB Compass, JavaScripts and MongoDB Aggregation Pipeline to calculate the average quarterly sold quantity of the blue color. (50% of grade)

Submit your code, and a screenshot of the created documents and code execution results in the pdf format.

Color	Quarter	Sold Quantity
Blue	1	230
Blue	2	452
Blue	3	6351
Blue	4	5280
Red	1	2453
Red	2	3486
Red	3	1535
Red	4	1250
Silver	1	814
Silver	2	1039
Silver	3	5788
Silver	4	4511

Part 1: ERD - Normalization Exercise 8 (Crow's Foot Notation)

Track each Sale of Home



Assumptions:

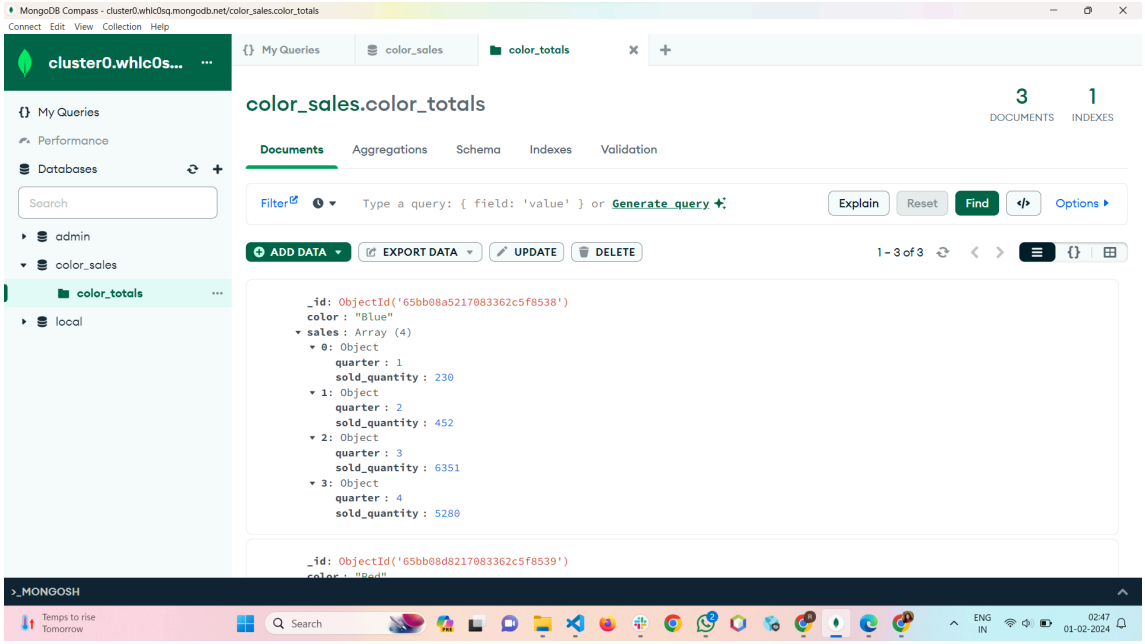
- Only 1 Selling Agent & 1 Buying Agent for each sale
- Only the Last Name of Agent is tracked
- Only 1 Phone Number of Agent is tracked
- 1 Agent is associated with only 1 Real Estate Company
- A Home can exist without any of its Sales
- A Real Estate company can exist without any Agent present in the company
- An Agent can exist without participating in a Selling Sale
- An Agent can exist without participating in a Buying Sale
- A Real Estate company can exist without participating in a Selling Sale
- A Real Estate company can exist without participating in a Buying Sale

Part 2: MongoDB - Design & Aggregation Pipeline

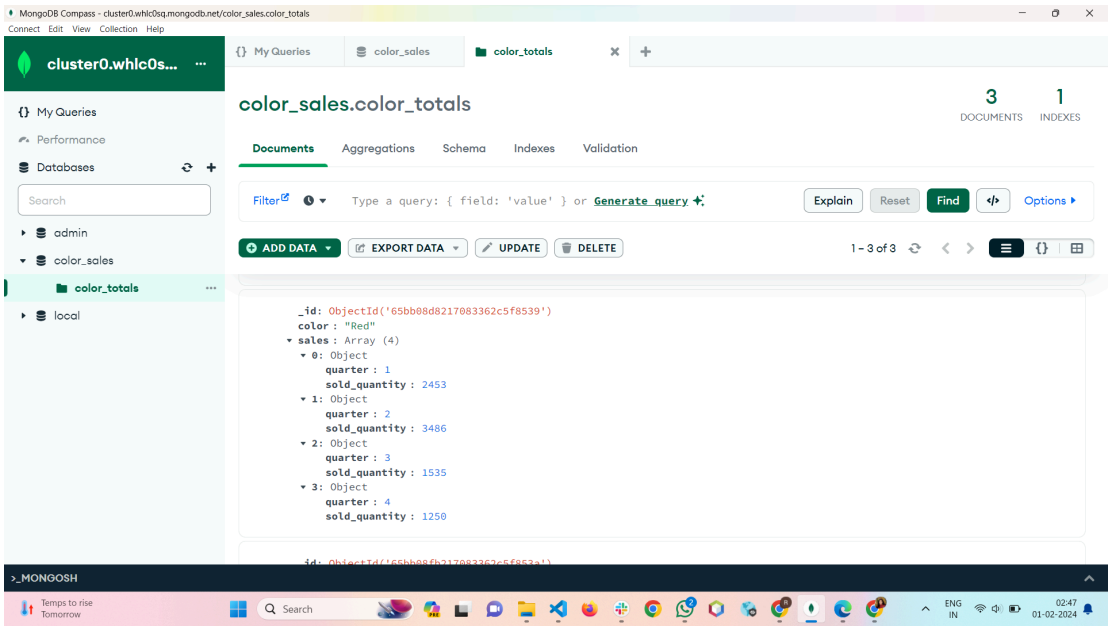
Data usage pattern: Find the average quarterly sold quantity of a color
(Output: Average Quarterly Sold Quantity of Blue Color = 3078.25)

Embedded Design MongoDB Document:

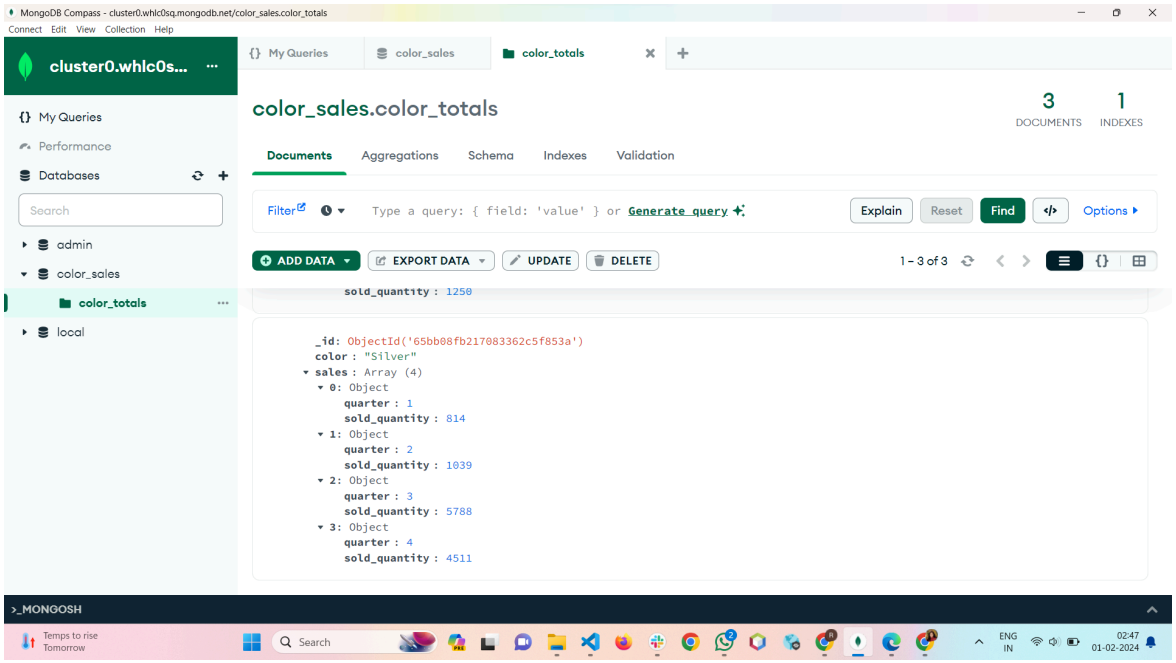
Color: Blue



Color: Red



Color: Silver

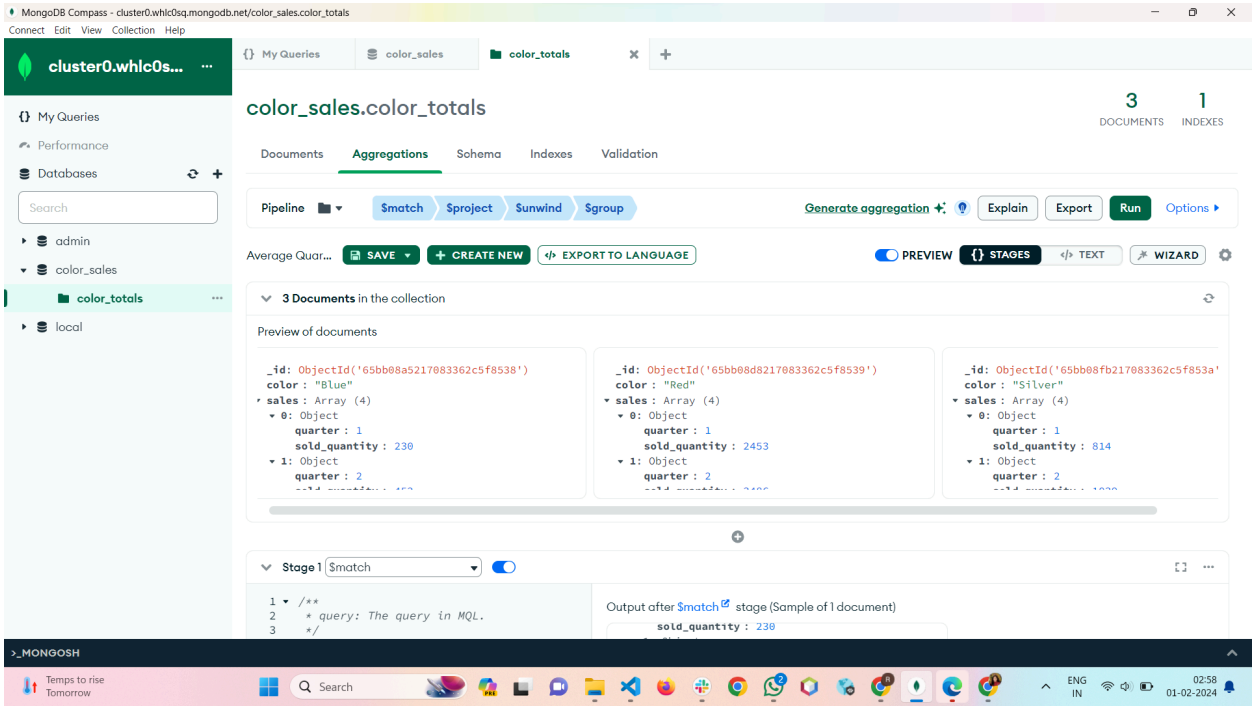


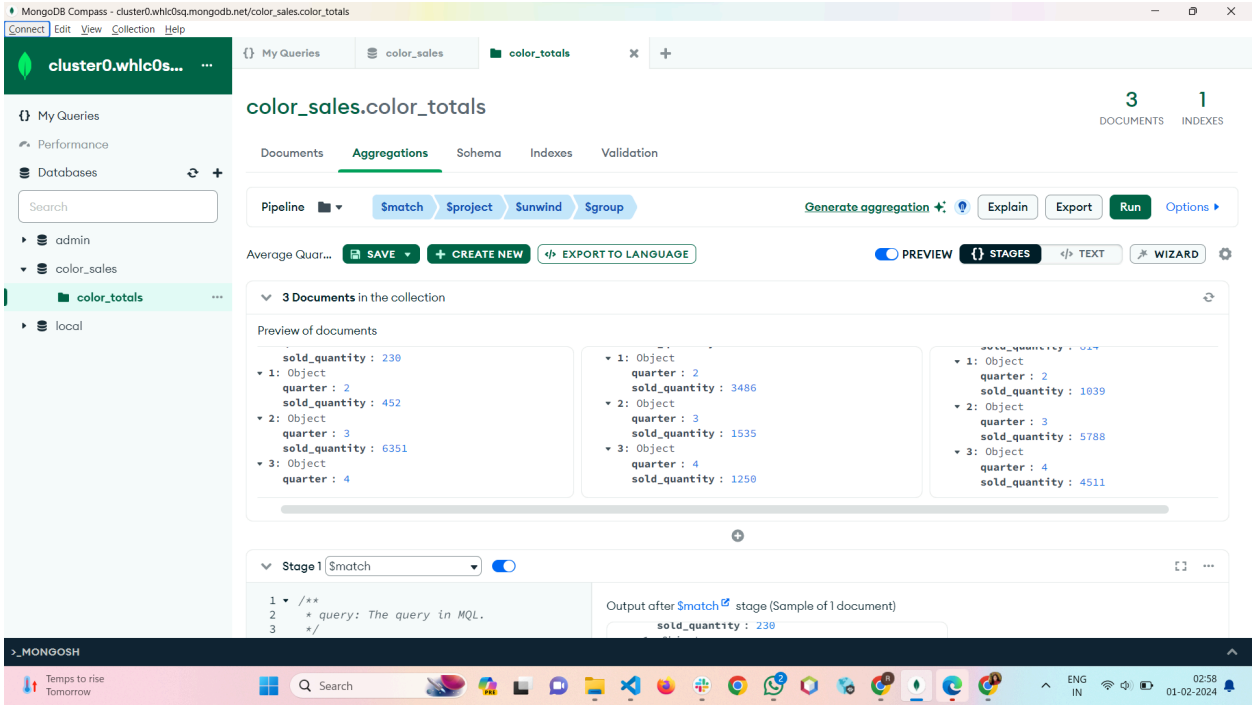
JavaScript Code: To find the average quarterly sold quantity of blue color

```
[
  {
    '$match': {
      'color': 'Blue'
    }
  }, {
    '$project': {
      'quarterly_sold_quantities': '$sales.sold_quantity'
    }
  }, {
    '$unwind': {
      'path': '$quarterly_sold_quantities'
    }
  }, {
    '$group': {
      '_id': 1,
      'avg_quarterly_sold_quantity': {
        '$avg': '$quarterly_sold_quantities'
      }
    }
  }
]
```

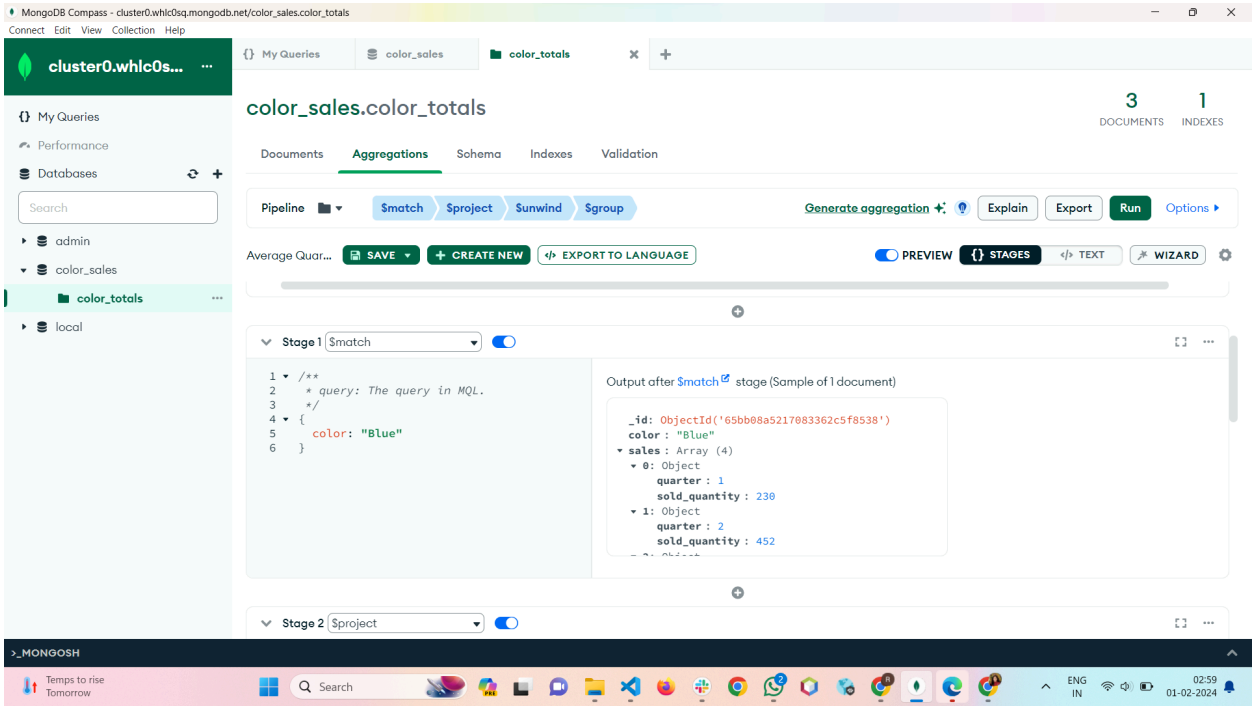
MongoDB Aggregation Pipeline

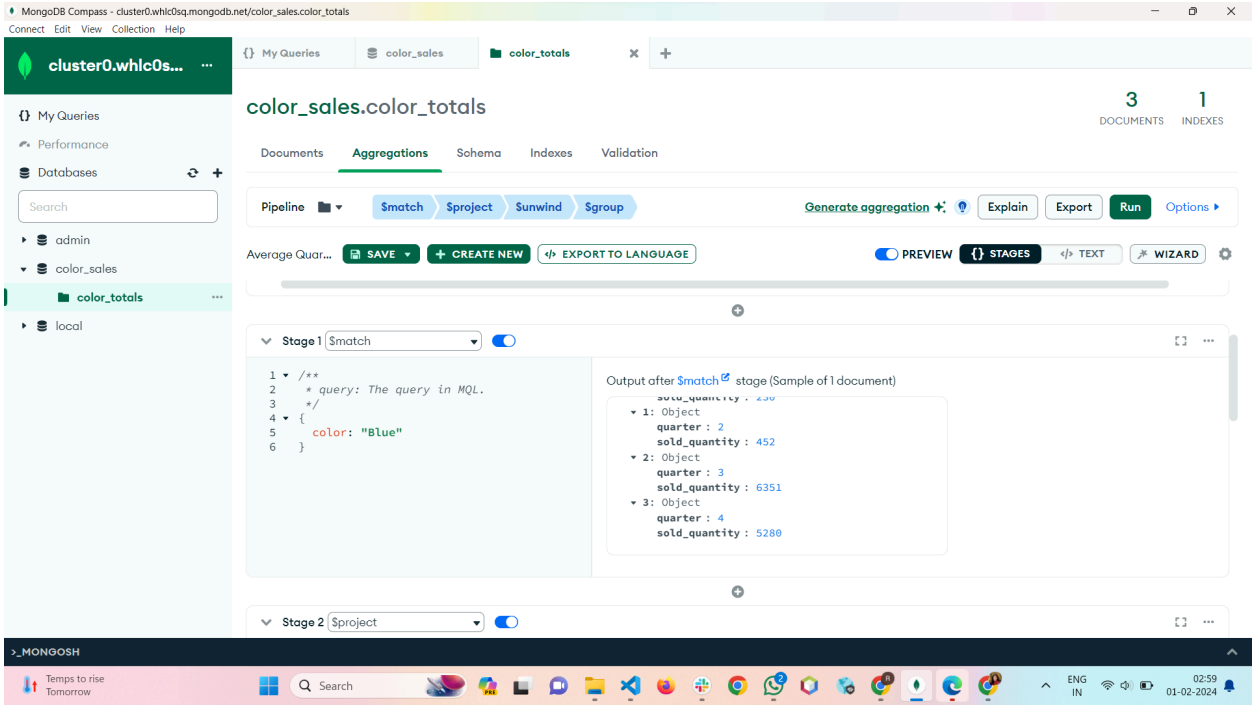
Preview of Documents



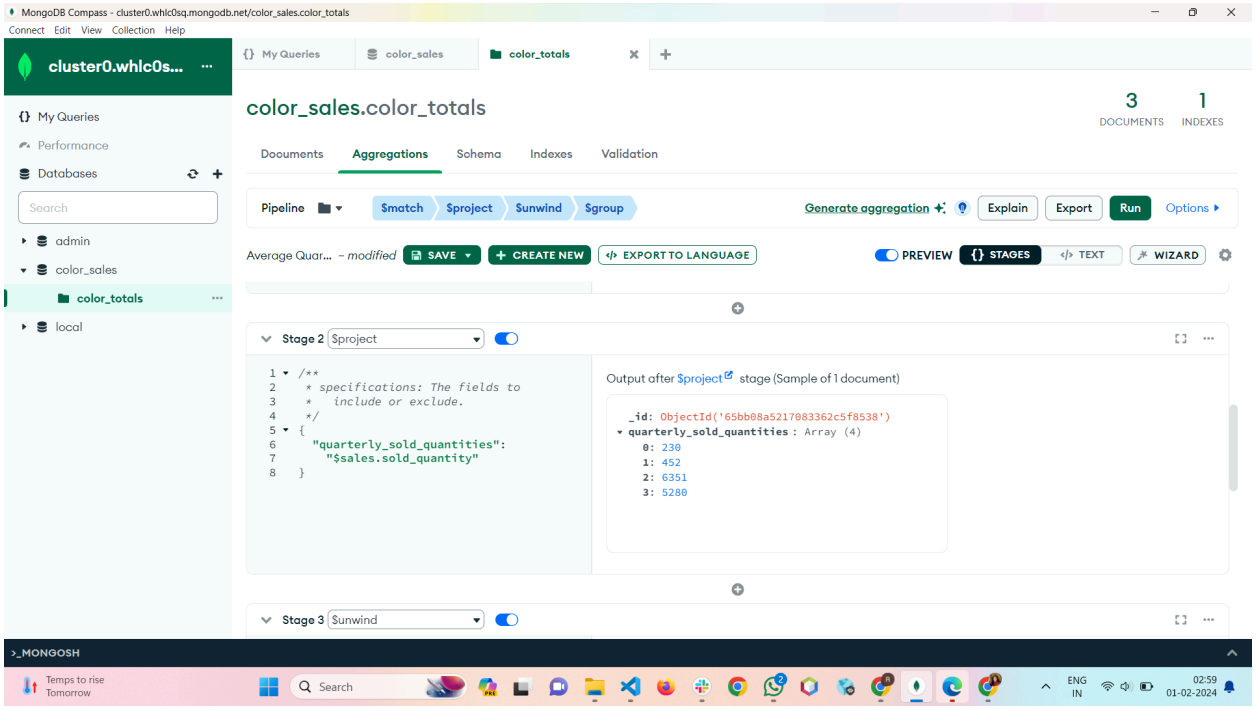


Stage 1: $match - Filters the documents to pass only the documents that match the specified condition(s) to the next pipeline stage.

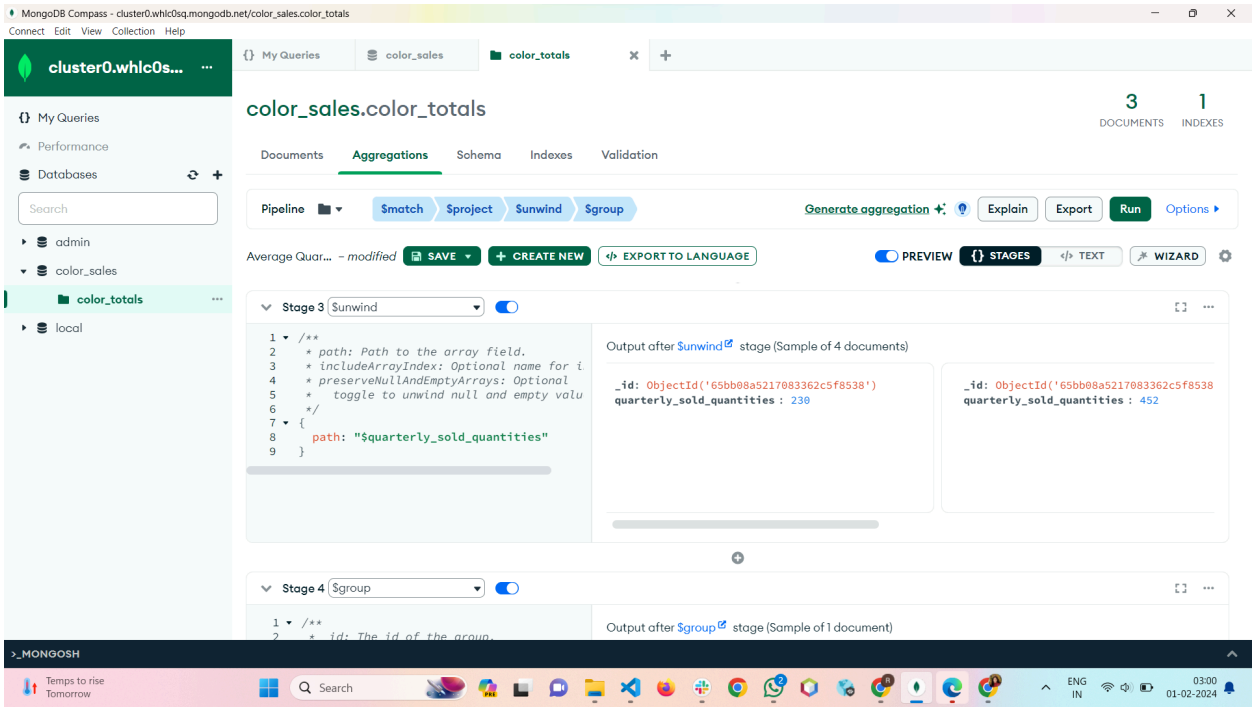


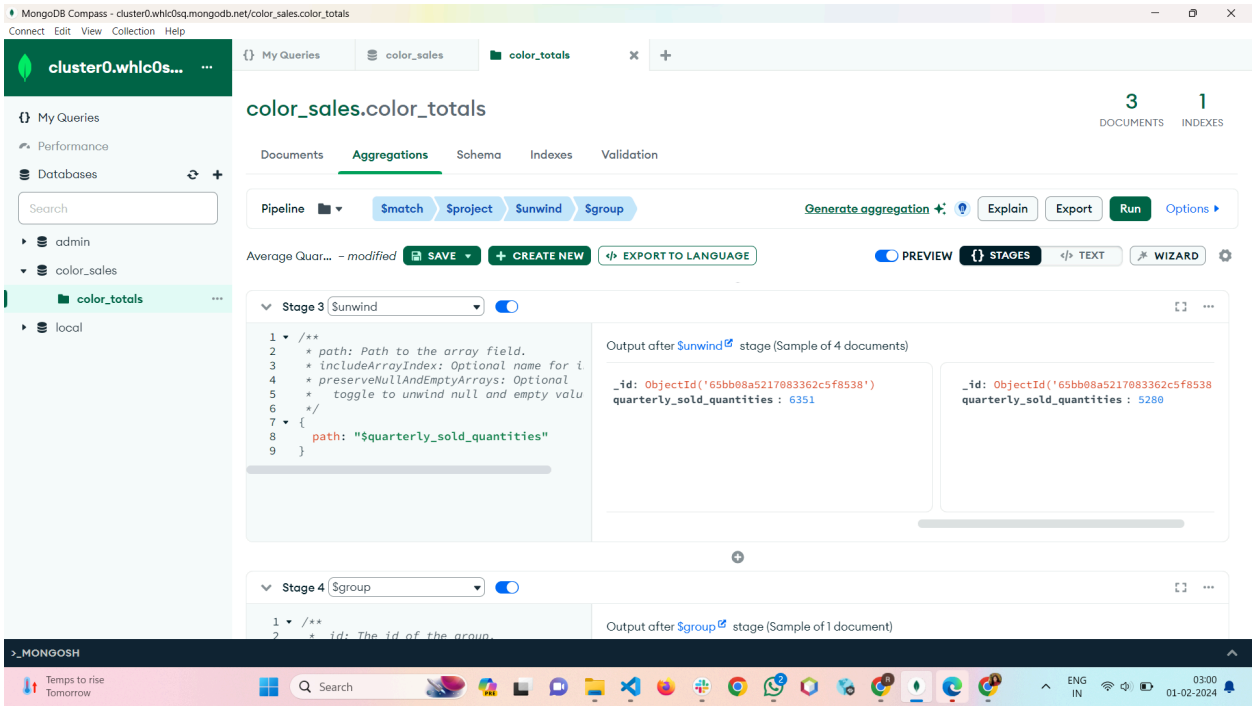


Stage 2: $project - Passes along the documents with the requested fields to the next stage in the pipeline.

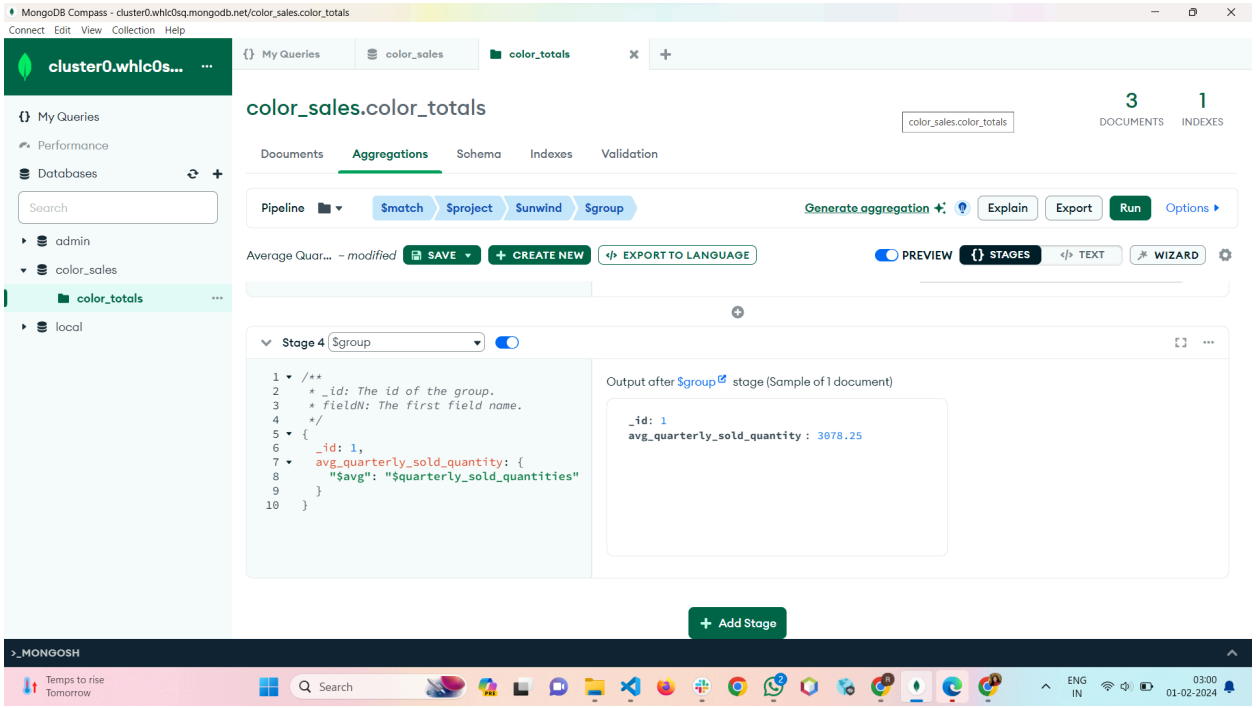


Stage 3: $unwind - Deconstructs an array field from the input documents to output a document for each element. Each output document is the input document with the value of the array field replaced by the element.





Stage 4: $group - Separates documents into groups according to a "group key". The output is one document for each unique group key.



Final Output: Average Quarterly Sold Quantity of Blue Color = 3078.25

References

- Professor's Normalization Playlist
- Professor's slides - ERD Example
- Professor's MongoDB Playlist
- MongoDB Documentation