

# **PREGLED I ANALIZA PODATAKA PROIZVODA SEPHORE**

Jovana Nedeljković, IN32-2019

Anđelija Stančić, IN19-2019

# OPIŠ SKUPA PODATAKA

- <https://www.kaggle.com/datasets/nadyinky/sephora-products-and-skincare-reviews>
- Ovaj skup podataka je prikupljen preko Python scraper-a u martu 2023.
- Sadrži:
  - informacije o svim kozmetičkim proizvodima iz Sephora online prodavnice (preko 8.000), uključujući nazive proizvoda i brendova, cene, sastojke, ocene i sve karakteristike
  - korisničke recenzije (preko 1 milion na preko 2.000 proizvoda) svih proizvoda uključujući i izgled korisnika i njihove ocene

- product\_id – jedinstveni identifikator proizvoda u Sephor
- product\_name – naziv proizvoda
- brand\_id – jedinstveni identifikator Brenda u Sephori
- brand\_name – naziv brenda
- loves\_count – broj ljudi koji su proizvod stavili u favorite
- rating – prosečna ocena proizvoda
- reviews – broj ljudi koji su ostavili komentar na proizvod
- ingredients – sastojci proizvoda
- price\_usd – cena proizvoda
- limited\_edition – da li je proizvod limitirane serije
- new – da li je proizvod nov
- online\_only – da li se prodaje samo online
- highlights – lista tagova koji opisuju proizvod
- primary\_category – primarna kategorija proizvoda
- secondary\_category – sekundarna kategorija proizvoda
- tertiary\_category – tercijalna kategorija proizvoda

## SEMANTIKA KOLONA – PODACI O PROIZVODIMA

- rating – ocena koju je korisnik ostavio za dati proizvod
- is\_recommended – da li korisnik preporučuje dati proizvod
- total\_neg\_feedback\_count – broj negativnih komentara
- total\_pos\_feedback\_count – broj pozitivnih komentara
- submission\_time – vreme ostavljanja komentara
- review\_text - komentar
- skin\_tone - boja kože korisnika koji je ostavio komentar
- eye\_color – boja očiju osobe koja je ostavila komentar
- skin\_type – tip kože korisnika koji je ostavio komentar
- hair\_color – boja kose korisnika koji je ostavio komentar
- product\_id – id proizvoda za koji je ostavljen komentar

## SEMANTIKA KOLONA – PODACI O FEEDBACK-U

# Product

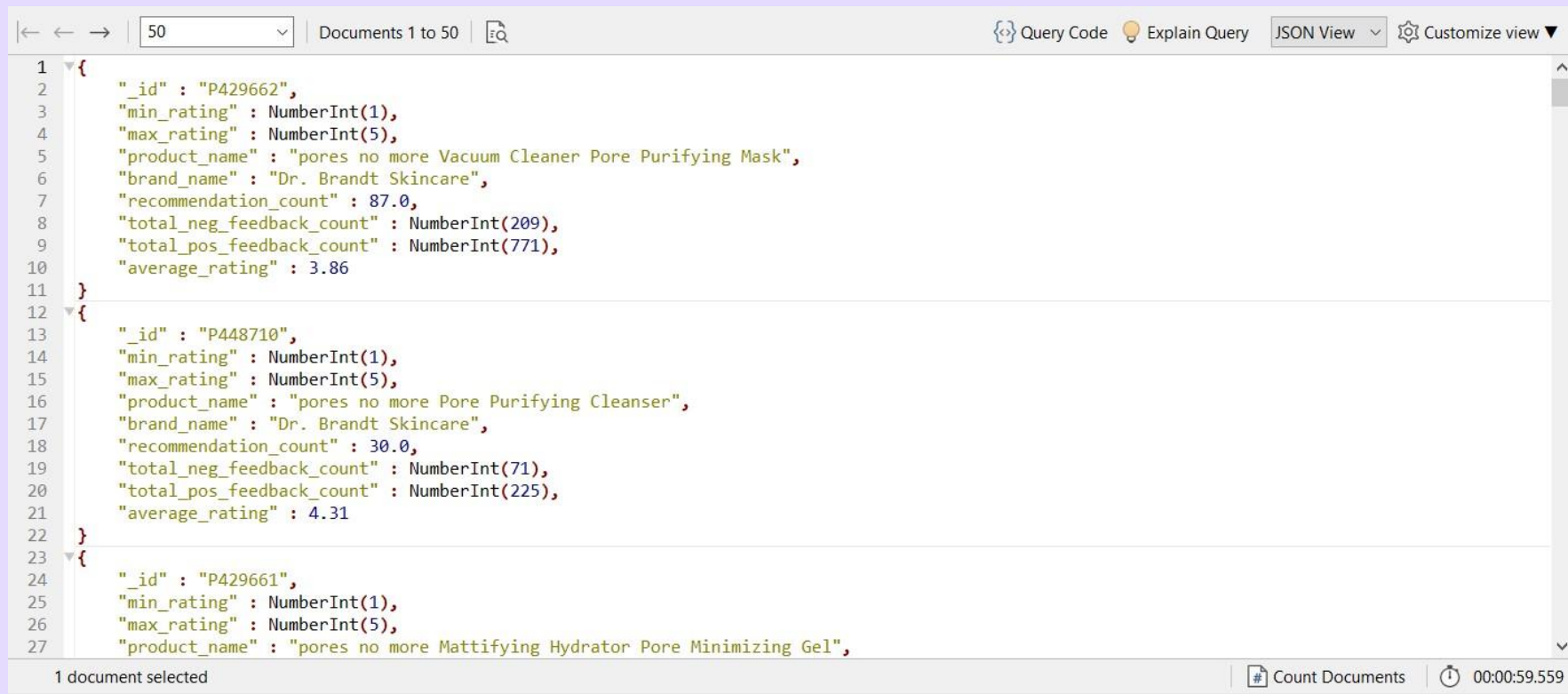
Key	Value	Type
▼ { } (5) { _id : P473658 }	{ 16 fields }	Document
{ } _id	P473658	String
{ } brand_id	6342	String
{ } brand_name	19-69	String
> { } highlights	[ 3 elements ]	Array
> { } ingredients	[ 12 elements ]	Array
{i32} limited_edition	0	Int32
{i32} loves_count	2691	Int32
{i32} new	0	Int32
{i32} online_only	1	Int32
{i23} price_usd	195.0	Double
{ } primary_category	Fragrance	String
{ } product_name	Purple Haze Eau de Parfum	String
{ } rating	3.2308	String
{ } reviews	13	String
{ } secondary_category	Women	String
{ } tertiary_category	Perfume	String

# Feedback

Key	Value	Type
▼ ⓘ (17) {_id : 33602438124}	{ 12 fields }	Document
ⓘ _id	33602438124	String
ⓘ rating	5	Int32
ⓘ is_recommended	1.0	String
ⓘ total_neg_feedback_count	0	Int32
ⓘ total_pos_feedback_count	0	Int32
ⓘ submission_time	2023-03-16T00:00:00.000Z	Date
ⓘ review_text	i love how hydrating this makes my lips feel! perfect for an overnight lip mask or ju	String
ⓘ skin_tone	lightMedium	String
ⓘ eye_color	brown	String
ⓘ skin_type	combination	String
ⓘ hair_color	brown	String
ⓘ product_id	P420652	String

# Primeri agregacija

- Za svaki proizvod koji ima recenziju prikazati prosečnu ocenu, minimalnu ocenu i maksimalnu ocenu koju je dobio, broj pozitivnih i negativnih recenzija i koliko je puta bio preporučen od strane korisnika



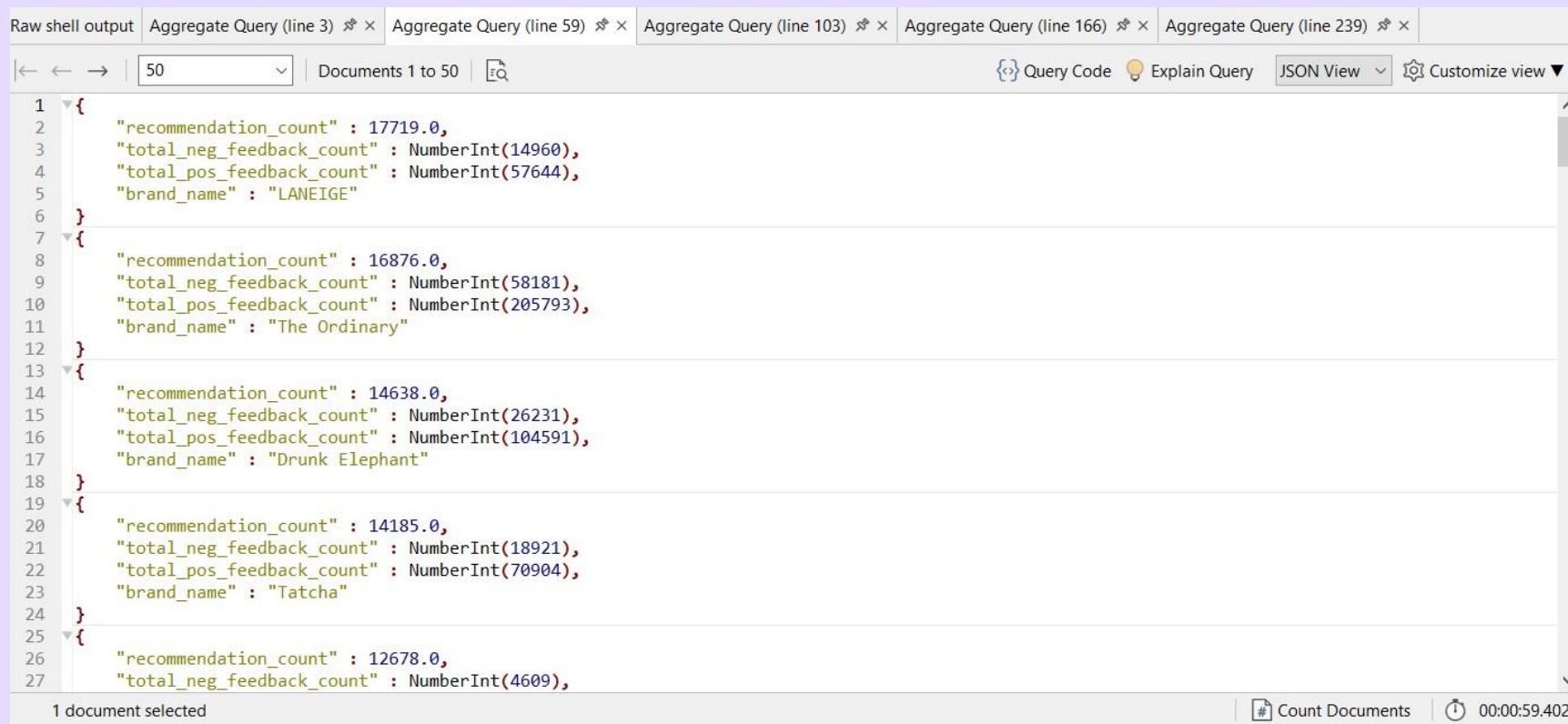
The screenshot shows a MongoDB query result in a web interface. The interface includes a top bar with navigation arrows, a document count of 50, and buttons for 'Query Code', 'Explain Query', 'JSON View', and 'Customize view'. The main area displays three JSON documents, each representing a product's aggregated data. The documents are numbered 1 through 27 on the left. The first document (lines 1-11) is for product 'P429662', the second (lines 12-22) for 'P448710', and the third (lines 23-27) for 'P429661'. Each document contains fields for '\_id', 'min\_rating', 'max\_rating', 'product\_name', 'brand\_name', 'recommendation\_count', 'total\_neg\_feedback\_count', 'total\_pos\_feedback\_count', and 'average\_rating'. The bottom status bar indicates '1 document selected' and shows a 'Count Documents' button with a timer at '00:00:59.559'.

```
1 {
2   "_id" : "P429662",
3   "min_rating" : NumberInt(1),
4   "max_rating" : NumberInt(5),
5   "product_name" : "pores no more Vacuum Cleaner Pore Purifying Mask",
6   "brand_name" : "Dr. Brandt Skincare",
7   "recommendation_count" : 87.0,
8   "total_neg_feedback_count" : NumberInt(209),
9   "total_pos_feedback_count" : NumberInt(771),
10  "average_rating" : 3.86
11 }
12 {
13   "_id" : "P448710",
14   "min_rating" : NumberInt(1),
15   "max_rating" : NumberInt(5),
16   "product_name" : "pores no more Pore Purifying Cleanser",
17   "brand_name" : "Dr. Brandt Skincare",
18   "recommendation_count" : 30.0,
19   "total_neg_feedback_count" : NumberInt(71),
20   "total_pos_feedback_count" : NumberInt(225),
21   "average_rating" : 4.31
22 }
23 {
24   "_id" : "P429661",
25   "min_rating" : NumberInt(1),
26   "max_rating" : NumberInt(5),
27   "product_name" : "pores no more Mattifying Hydrator Pore Minimizing Gel",
```

1 document selected | Count Documents | 00:00:59.559

# Primeri agregacija

- Za svaki brend čiji proizvodi imaju recenzije, izračunati koliko preporuka ukupno imaju njegovi proizvodi, i ukupan broj pozitivnih i negativnih recenzija za sve njegove proizvode



The screenshot shows a MongoDB query results interface with five tabs for different aggregate queries. The first tab, 'Aggregate Query (line 3)', is selected and displays five JSON documents. Each document represents a brand and contains aggregated feedback counts for its products. The interface includes a pagination bar showing 'Documents 1 to 50', a search bar, and various tool icons like 'Query Code', 'Explain Query', and 'JSON View'. The status bar at the bottom indicates '1 document selected' and a timer of '00:00:59.402'.

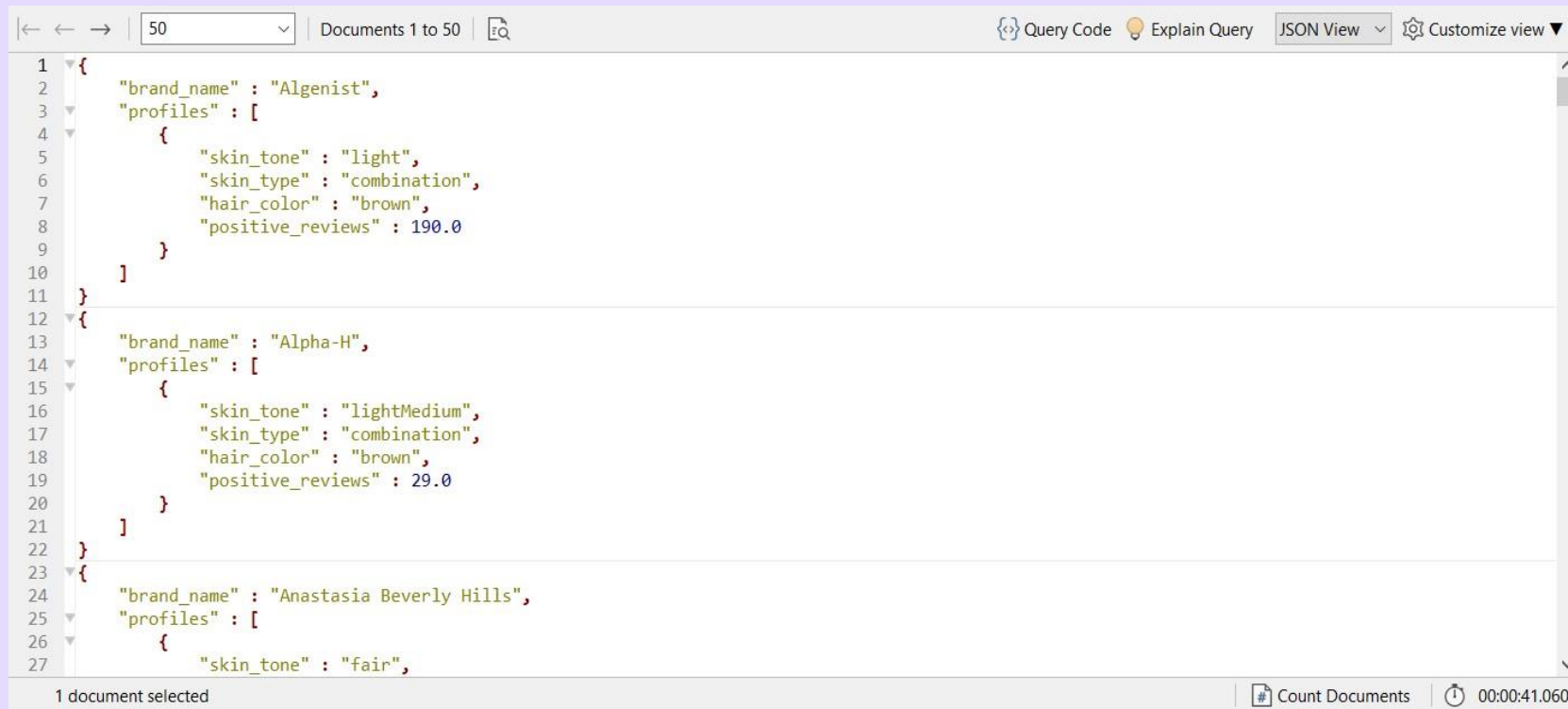
```
1 {
2   "recommendation_count" : 17719.0,
3   "total_neg_feedback_count" : NumberInt(14960),
4   "total_pos_feedback_count" : NumberInt(57644),
5   "brand_name" : "LANEIGE"
6 }
7 {
8   "recommendation_count" : 16876.0,
9   "total_neg_feedback_count" : NumberInt(58181),
10  "total_pos_feedback_count" : NumberInt(205793),
11  "brand_name" : "The Ordinary"
12 }
13 {
14  "recommendation_count" : 14638.0,
15  "total_neg_feedback_count" : NumberInt(26231),
16  "total_pos_feedback_count" : NumberInt(104591),
17  "brand_name" : "Drunk Elephant"
18 }
19 {
20  "recommendation_count" : 14185.0,
21  "total_neg_feedback_count" : NumberInt(18921),
22  "total_pos_feedback_count" : NumberInt(70904),
23  "brand_name" : "Tatcha"
24 }
25 {
26  "recommendation_count" : 12678.0,
27  "total_neg_feedback_count" : NumberInt(4609),
```

1 document selected | Count Documents | 00:00:59.402



# Primeri agregacija

- Za svaki brend koji ima recenzije svojih proizvoda pronaći profil osobe koja najčešće predlaže njegove proizvode



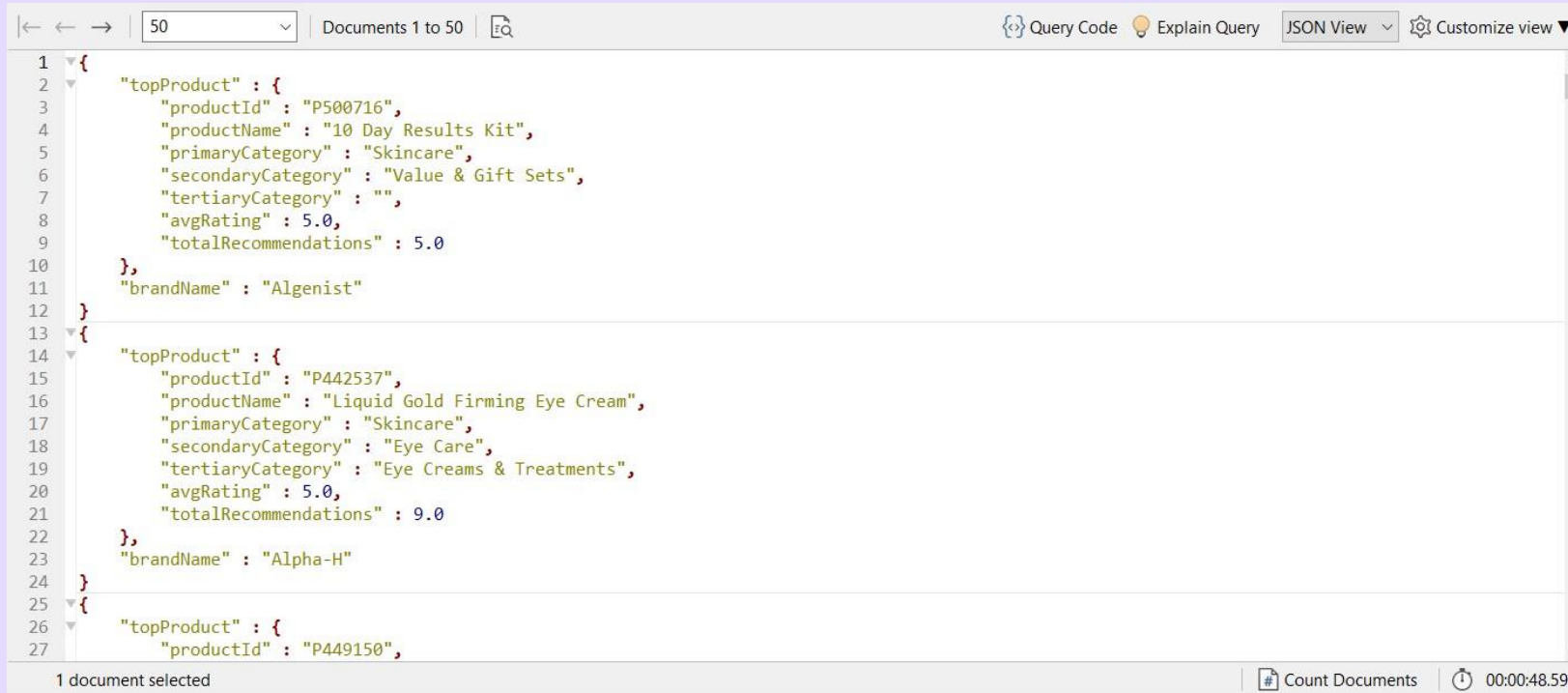
The screenshot shows a MongoDB query interface with a JSON view of documents. The interface includes a top bar with navigation controls, a document count of 50, and tabs for Query Code, Explain Query, JSON View (selected), and Customize view. The JSON documents are displayed with line numbers on the left. The first document is for 'Algenist' with a profile having 190.0 positive reviews. The second document is for 'Alpha-H' with a profile having 29.0 positive reviews. The third document is for 'Anastasia Beverly Hills' with a profile having 'fair' skin tone. The status bar at the bottom indicates '1 document selected' and provides a 'Count Documents' button and a timer showing '00:00:41.060'.

```
1 {
2   "brand_name" : "Algenist",
3   "profiles" : [
4     {
5       "skin_tone" : "light",
6       "skin_type" : "combination",
7       "hair_color" : "brown",
8       "positive_reviews" : 190.0
9     }
10  ]
11 }
12 {
13   "brand_name" : "Alpha-H",
14   "profiles" : [
15     {
16       "skin_tone" : "lightMedium",
17       "skin_type" : "combination",
18       "hair_color" : "brown",
19       "positive_reviews" : 29.0
20     }
21  ]
22 }
23 {
24   "brand_name" : "Anastasia Beverly Hills",
25   "profiles" : [
26     {
27       "skin_tone" : "fair",
```

1 document selected | Count Documents | 00:00:41.060

# Primeri agregacija

- Pronaći najbolji proizvod za svaki brend na osnovu prosečne ocene i ukupnog broja preporuka korisnika u recenzijama



The screenshot shows a MongoDB query results interface. The top bar includes navigation controls, a document count of 50, and tabs for 'Query Code', 'Explain Query', 'JSON View', and 'Customize view'. The main area displays three JSON documents, each representing a product recommendation for a specific brand. The first document is for 'Algenist' with product ID 'P500716'. The second is for 'Alpha-H' with product ID 'P442537'. The third document is partially visible for product ID 'P449150'. Each document contains fields for 'productId', 'productName', 'primaryCategory', 'secondaryCategory', 'tertiaryCategory', 'avgRating', and 'totalRecommendations'. The interface also shows a line number indicator on the left and a status bar at the bottom indicating '1 document selected' and a timer of '00:00:48.591'.

```
1 {
2   "topProduct" : {
3     "productId" : "P500716",
4     "productName" : "10 Day Results Kit",
5     "primaryCategory" : "Skincare",
6     "secondaryCategory" : "Value & Gift Sets",
7     "tertiaryCategory" : "",
8     "avgRating" : 5.0,
9     "totalRecommendations" : 5.0
10  },
11  "brandName" : "Algenist"
12 }
13 {
14   "topProduct" : {
15     "productId" : "P442537",
16     "productName" : "Liquid Gold Firming Eye Cream",
17     "primaryCategory" : "Skincare",
18     "secondaryCategory" : "Eye Care",
19     "tertiaryCategory" : "Eye Creams & Treatments",
20     "avgRating" : 5.0,
21     "totalRecommendations" : 9.0
22  },
23  "brandName" : "Alpha-H"
24 }
25 {
26   "topProduct" : {
27     "productId" : "P449150",
```

1 document selected | Count Documents | 00:00:48.591

# Primeri agregacija

- Za svaki tip kože korisnika u bazi (dry, combination, sensitive, normal) pronaći po 2 proizvoda koja spadaju u "Moisturizers" kategoriju (tertiary\_category) a koja su ljudi sa tim tipom kože najviše preporučivali u recenzijama



The screenshot shows a MongoDB query interface with a tabbed window. The active tab is 'Aggregate Query (line 239)'. The interface includes a navigation bar with a back arrow, a search bar containing '50', and a document count 'Documents 1 to 5'. Below the navigation bar are buttons for 'Query Code', 'Explain Query', 'JSON View', and 'Customize view'. The main area displays a JSON document with the following structure:

```
1 {
2   "_id": "dry",
3   "products": [
4     {
5       "product_id": "P394639",
6       "product_name": "The True Cream Aqua Bomb",
7       "brand_name": "belif",
8       "count": 807.0
9     },
10    {
11      "product_id": "P427421",
12      "product_name": "Protini Polypeptide Firming Refillable Moisturizer",
13      "brand_name": "Drunk Elephant",
14      "count": 726.0
15    }
16  ]
17 }
18 {
19   "_id": "oily",
20   "products": [
21     {
22       "product_id": "P394639",
23       "product_name": "The True Cream Aqua Bomb",
24       "brand_name": "belif",
25       "count": 641.0
26     },
27     {
28       "product_id": "P427421",
29       "product_name": "Protini Polypeptide Firming Refillable Moisturizer",
30       "brand_name": "Drunk Elephant",
31       "count": 726.0
32     }
33   ]
34 }
```

At the bottom of the interface, it indicates '1 document selected' and shows a 'Count Documents' button with a timer at '00:00:44.363'.

# Primeri agregacija

- Za svoju boje kose, boju očiju i tip kože naći proizvode sa maksimalnom vrednošću broja korisnika koji su za te proizvode, iz tih kategorija, u toj kombinaciji, ostavili recenzije.

```
{
  "max_count" : 6319.0,
  "brand_id" : "1254",
  "hair_color" : "",
  "skin_type" : "",
  "eye_color" : ""
}
{
  "max_count" : 5735.0,
  "brand_id" : "1254",
  "hair_color" : "",
  "skin_type" : "combination",
  "eye_color" : ""
}
{
  "max_count" : 2892.0,
  "brand_id" : "6234",
  "hair_color" : "brown",
  "skin_type" : "combination",
  "eye_color" : "brown"
}
```

document selected

Activate Windows  
Go to Settings to activate Windows.

Count Documents | 00:02:51.778

# Primeri agregacija

- Nalaženje prosečne ocene iz svih feedback-ova za 2022-u godinu, proizvoda koji su preporučeni i nalaze se u trećoj kategoriji i njihovih brand-ova.

```
"brands" : [  
  {  
    "brandId" : "6329",  
    "brandName" : "FaceGym",  
    "avgRating" : 4.785714285714286  
  },  
  {  
    "brandId" : "3902",  
    "brandName" : "SEPHORA COLLECTION",  
    "avgRating" : 4.75  
  },  
  {  
    "brandId" : "6482",  
    "brandName" : "CANOPY",  
    "avgRating" : 4.891304347826087  
  },  
  {  
    "brandId" : "6001",  
    "brandName" : "NuFACE",  
    "avgRating" : 5.0  
  },  
  {  
    "brandId" : "6358",  
    "brandName" : "NuFACE",  
    "avgRating" : 4.75  
  }  
]
```

document selected

Activate Windows  
Go to Settings to activate Windows.

# Count Documents | 00:00:27.35

# Primeri agregacija

- Nalaženje profila svake osobe (boja kože, tip kože, boja očiju) koja je ostavila maksimalan broj pozitivnih feedback-ova za svaki produkt i svaki brand iz svih kategorija

```
"products" : [
  {
    "product_id" : "P385320",
    "brands" : [
      {
        "brand_name" : "NuFACE",
        "profiles" : {
          "skin_type" : "oily",
          "eye_color" : "",
          "hair_color" : "",
          "total_pos_feedback_count" : NumberInt(298)
        }
      }
    ]
  },
  {
    "product_id" : "P389097",
    "brands" : [
      {
        "brand_name" : "HUM Nutrition",
        "profiles" : {
          "skin_type" : "oily",
          "eye_color" : "",
          "hair_color" : "",
          "total_pos_feedback_count" : NumberInt(23)
        }
      }
    ]
  }
]
```

document selected

Activate Windows  
Go to Settings to activate Windows.

Count Documents 00:11:01.661

# Primeri agregacija

- Nalaženje brenda sa najjeftinijim proizvodima i broj pozitivnih feedback-ova iz tercijalne kategorije za tipove i boje kože

```
{
  "min_price" : 3.0,
  "brands" : [
    "SEPHORA COLLECTION",
    "BeautyBio",
    "Koh Gen Do"
  ],
  "tertiary_category" : "Face Wipes",
  "skin_tone" : "light",
  "skin_type" : "dry"
}
{
  "min_price" : 3.0,
  "brands" : [
    "Koh Gen Do",
    "BeautyBio",
    "Shiseido",
    "SEPHORA COLLECTION"
  ],
  "tertiary_category" : "Face Wipes",
  "skin_tone" : "lightMedium",
  "skin_type" : "oily"
}
{
  "min_price" : 3.0,
  "brands" : [
    "SEPHORA COLLECTION",
    "Koh Gen Do",
```

document selected

Activate Windows  
Go to Settings to activate Windows.

Count Documents 00:05:22.789

# Primeri agregacija

- Nalaženje proizvoda iz svih kategorija sa maksimalnom cenom za svaku ocenu

```
{
  "rating" : "3.0417",
  "brand" : "tarte",
  "product" : {
    "_id" : "P457468",
    "brand_id" : "5648",
    "brand_name" : "tarte",
    "highlights" : [
      ""
    ],
    "ingredients" : [
      "Water/Aqua/Eau",
      " Propylene Glycol",
      " Dihydroxyacetone",
      " PEG-6 Caprylic/Capric Glycerides",
      " Dihydroxypropyl PEG-5 Linoleammonium Chloride",
      " PEG-7 Glyceryl Cocoate",
      " Sodium Lauroyl Sarcosinate",
      " Hamamelis Virginiana (Witch Hazel) Bark/Leaf/Twig Extract",
      " Phenoxyethanol",
      " Caprylyl Glycol",
      " Potassium Sorbate",
      " Hexylene Glycol",
      " Ethoxydiglycol",
      " Erythrulose",
      " Glycerin"
    ]
  }
}
```

document selected

Activate Windows  
Go to Settings to activate Windows.

# Count Documents | 00:03:02.711



```
db.product_with_feedbacks.createIndex({ "brand_name": 1 });
db.product_with_feedbacks.createIndex({ "_id": 1 });
db.product_with_feedbacks.createIndex({ "feedback.rating": 1 });
db.product_with_feedbacks.createIndex({ "tertiary_category": 1 });
```

```
db.product_info.aggregate([
  {
    $lookup: {
      from: "feedback",
      localField: "_id",
      foreignField: "product_id",
      as: "feedbacks"
    }
  },
  {
    $out: "product_with_feedbacks"
  }
])
```

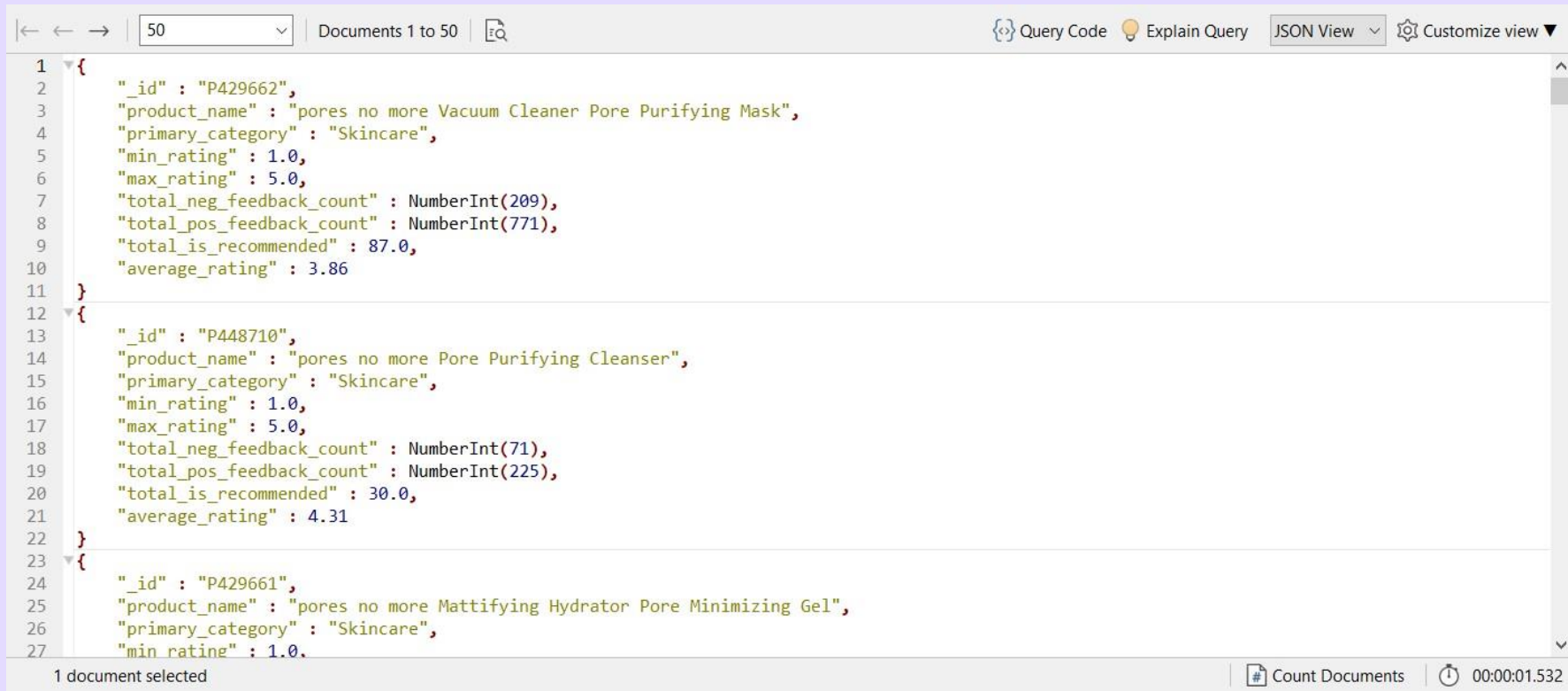
# Transformacija baze

# Product with feedback

Key	Value	Type
▼ (17) {_id : 33602438124}	{ 12 fields }	Document
_id	33602438124	String
rating	5	Int32
is_recommended	1.0	String
total_neg_feedback_count	0	Int32
total_pos_feedback_count	0	Int32
submission_time	2023-03-16T00:00:00.000Z	Date
review_text	i love how hydrating this makes my lips feel! perfect for an overnight lip mask or ju	String
skin_tone	lightMedium	String
eye_color	brown	String
skin_type	combination	String
hair_color	brown	String
product_id	P420652	String

# Primeri optimizovanih agregacija

- Za svaki proizvod koji ima recenziju prikazati prosečnu ocenu, minimalnu ocenu i maksimalnu ocenu koju je dobio, broj pozitivnih i negativnih recenzija i koliko je puta bio preporučen od strane korisnika



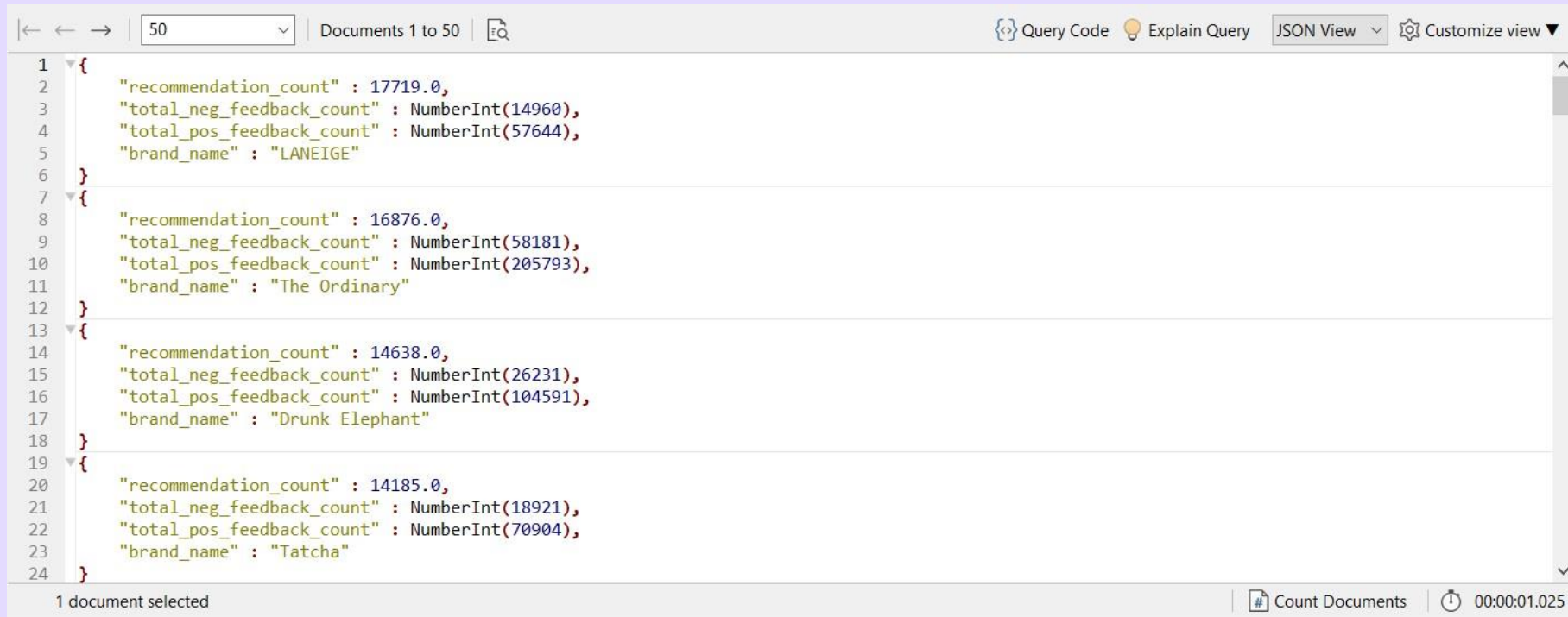
The screenshot shows a MongoDB query result interface. At the top, there are navigation controls (back, forward, search) and a dropdown menu set to 50. Below this, the query code is visible, and the results are displayed in JSON format. The results show three documents, each representing a product with aggregated feedback data. The first document is for 'pores no more Vacuum Cleaner Pore Purifying Mask', the second for 'pores no more Pore Purifying Cleanser', and the third for 'pores no more Mattifying Hydrator Pore Minimizing Gel'. Each document includes fields for product ID, name, category, min/max ratings, total negative/positive feedback counts, total recommendations, and average rating.

```
1 {
2   "_id" : "P429662",
3   "product_name" : "pores no more Vacuum Cleaner Pore Purifying Mask",
4   "primary_category" : "Skincare",
5   "min_rating" : 1.0,
6   "max_rating" : 5.0,
7   "total_neg_feedback_count" : NumberInt(209),
8   "total_pos_feedback_count" : NumberInt(771),
9   "total_is_recommended" : 87.0,
10  "average_rating" : 3.86
11 }
12 {
13   "_id" : "P448710",
14   "product_name" : "pores no more Pore Purifying Cleanser",
15   "primary_category" : "Skincare",
16   "min_rating" : 1.0,
17   "max_rating" : 5.0,
18   "total_neg_feedback_count" : NumberInt(71),
19   "total_pos_feedback_count" : NumberInt(225),
20   "total_is_recommended" : 30.0,
21   "average_rating" : 4.31
22 }
23 {
24   "_id" : "P429661",
25   "product_name" : "pores no more Mattifying Hydrator Pore Minimizing Gel",
26   "primary_category" : "Skincare",
27   "min_rating" : 1.0,
```

1 document selected | Count Documents | 00:00:01.532

# Primeri optimizovanih agregacija

- Za svaki brend čiji proizvodi imaju recenzije, izračunati koliko preporuka ukupno imaju njegovi proizvodi, i ukupan broj pozitivnih i negativnih recenzija za sve njegove proizvode



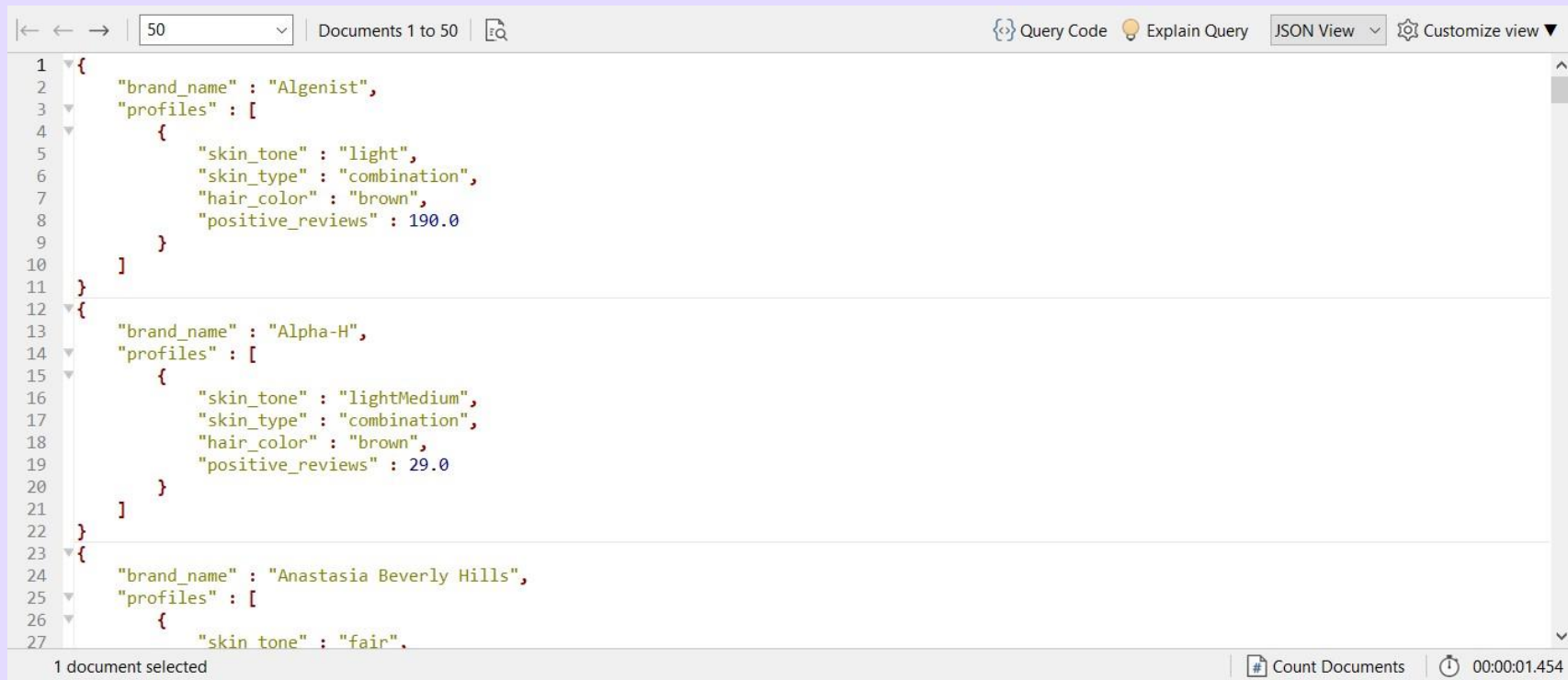
The screenshot displays a MongoDB query result in JSON format. The interface includes a top bar with navigation arrows, a document count of 50, and options for 'Query Code', 'Explain Query', 'JSON View', and 'Customize view'. The results are shown as a list of four documents, each representing a brand's aggregated feedback data. The status bar at the bottom indicates '1 document selected' and a timer of 00:00:01.025.

```
1 {
2   "recommendation_count" : 17719.0,
3   "total_neg_feedback_count" : NumberInt(14960),
4   "total_pos_feedback_count" : NumberInt(57644),
5   "brand_name" : "LANEIGE"
6 }
7 {
8   "recommendation_count" : 16876.0,
9   "total_neg_feedback_count" : NumberInt(58181),
10  "total_pos_feedback_count" : NumberInt(205793),
11  "brand_name" : "The Ordinary"
12 }
13 {
14   "recommendation_count" : 14638.0,
15   "total_neg_feedback_count" : NumberInt(26231),
16   "total_pos_feedback_count" : NumberInt(104591),
17   "brand_name" : "Drunk Elephant"
18 }
19 {
20   "recommendation_count" : 14185.0,
21   "total_neg_feedback_count" : NumberInt(18921),
22   "total_pos_feedback_count" : NumberInt(70904),
23   "brand_name" : "Tatcha"
24 }
```

1 document selected | # Count Documents | 00:00:01.025

# Primeri optimizovanih agregacija

- Za svaki brend koji ima recenzije svojih proizvoda pronaći profil osobe koja najčešće predlaže njegove proizvode

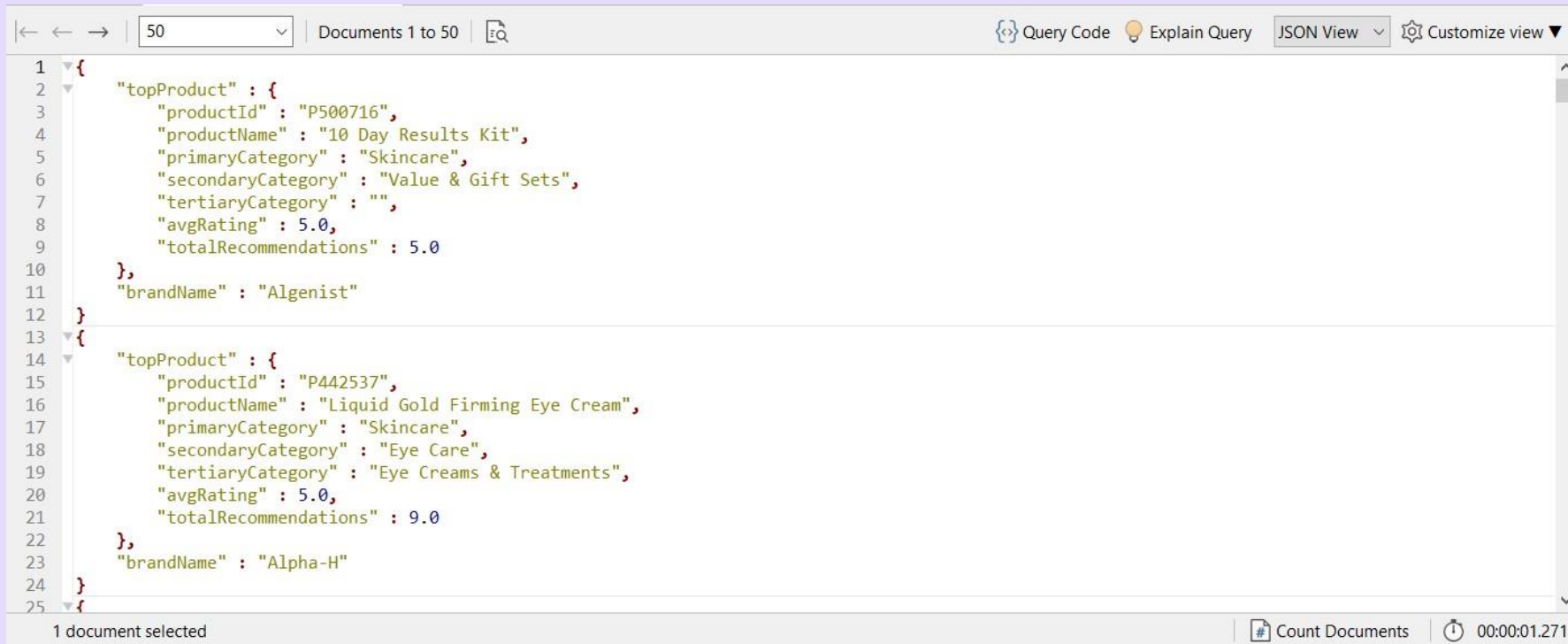


```
1 {
2   "brand_name" : "Algenist",
3   "profiles" : [
4     {
5       "skin_tone" : "light",
6       "skin_type" : "combination",
7       "hair_color" : "brown",
8       "positive_reviews" : 190.0
9     }
10  ]
11 }
12 {
13   "brand_name" : "Alpha-H",
14   "profiles" : [
15     {
16       "skin_tone" : "lightMedium",
17       "skin_type" : "combination",
18       "hair_color" : "brown",
19       "positive_reviews" : 29.0
20     }
21  ]
22 }
23 {
24   "brand_name" : "Anastasia Beverly Hills",
25   "profiles" : [
26     {
27       "skin_tone" : "fair",
```

1 document selected | Count Documents | 00:00:01.454

# Primeri optimizovanih agregacija

- Pronaći najbolji proizvod za svaki brend na osnovu prosečne ocene i ukupnog broja preporuka korisnika u recenzijama



The screenshot shows a MongoDB query result in a web interface. The interface includes a top bar with navigation arrows, a document count of 50, and buttons for 'Query Code', 'Explain Query', 'JSON View', and 'Customize view'. The main area displays two JSON documents, each representing a product recommendation for a specific brand. The first document is for the brand 'Algenist' and the second is for 'Alpha-H'. Each document contains a 'topProduct' object with fields for 'productId', 'productName', 'primaryCategory', 'secondaryCategory', 'tertiaryCategory', 'avgRating', and 'totalRecommendations'. The status bar at the bottom indicates '1 document selected' and 'Count Documents'.

```
1 {
2   "topProduct" : {
3     "productId" : "P500716",
4     "productName" : "10 Day Results Kit",
5     "primaryCategory" : "Skincare",
6     "secondaryCategory" : "Value & Gift Sets",
7     "tertiaryCategory" : "",
8     "avgRating" : 5.0,
9     "totalRecommendations" : 5.0
10  },
11  "brandName" : "Algenist"
12 }
13 {
14   "topProduct" : {
15     "productId" : "P442537",
16     "productName" : "Liquid Gold Firming Eye Cream",
17     "primaryCategory" : "Skincare",
18     "secondaryCategory" : "Eye Care",
19     "tertiaryCategory" : "Eye Creams & Treatments",
20     "avgRating" : 5.0,
21     "totalRecommendations" : 9.0
22  },
23  "brandName" : "Alpha-H"
24 }
25 }
```

1 document selected

Count Documents 00:00:01.271

# Primeri optimizovanih agregacija

- Za svaki tip kože korisnika u bazi (dry, combination, sensitive, normal) pronaći po 2 proizvoda koja spadaju u "Moisturizers" kategoriju (tertiary\_category) a koja su ljudi sa tim tipom kože najviše preporučivali u recenzijama



```
1 {
2   "_id" : "normal",
3   "products" : [
4     {
5       "product_id" : "P394639",
6       "product_name" : "The True Cream Aqua Bomb",
7       "brand_name" : "belif",
8       "count" : 628.0
9     },
10    {
11      "product_id" : "P427421",
12      "product_name" : "Protini Polypeptide Firming Refillable Moisturizer",
13      "brand_name" : "Drunk Elephant",
14      "count" : 482.0
15    }
16  ]
17 }
18 {
19   "_id" : "oily",
20   "products" : [
21     {
22       "product_id" : "P394639",
23       "product_name" : "The True Cream Aqua Bomb",
24       "brand_name" : "belif",
25       "count" : 641.0
26     },
27     {
28       "product_id" : "P427421",
29       "product_name" : "Protini Polypeptide Firming Refillable Moisturizer",
30       "brand_name" : "Drunk Elephant",
31       "count" : 482.0
32     }
33   ]
34 }
```

1 document selected | Count Documents | 00:00:00.263

# Primeri optimizovanih agregacija

- Za svoju boje kose, boju očiju i tip kože naći proizvode sa maksimalnom vrednošću broja korisnika koji su za te proizvode, iz tih kategorija, u toj kombinaciji, ostavili recenzije.

```
{
  "max_count" : 6319.0,
  "brand_id" : "1254",
  "hair_color" : "",
  "skin_type" : "",
  "eye_color" : ""
}
{
  "max_count" : 5735.0,
  "brand_id" : "1254",
  "hair_color" : "",
  "skin_type" : "combination",
  "eye_color" : ""
}
{
  "max_count" : 2892.0,
  "brand_id" : "6234",
  "hair_color" : "brown",
  "skin_type" : "combination",
  "eye_color" : "brown"
}
{
  "max_count" : 2430.0,
  "brand_id" : "6234",
  "hair_color" : "black",
  "skin_type" : "combination",
  "eye_color" : ""
}
```

document selected

Activate Windows  
Go to Settings to activate Windows.

Count Documents | 00:00:05.656



# Primeri optimizovanih agregacija

- Nalaženje prosečne ocene iz svih feedback-ova za 2022-u godinu, proizvoda koji su preporučeni i nalaze se u trećoj kategoriji i njihovih brand-ova.

```
"brands" : [  
  {  
    "brandId" : "6309",  
    "brandName" : "goop",  
    "avgRating" : 4.714285714285714  
  },  
  {  
    "brandId" : "7017",  
    "brandName" : "iluminage",  
    "avgRating" : 4.8  
  },  
  {  
    "brandId" : "5668",  
    "brandName" : "Dr. Dennis Gross Skincare",  
    "avgRating" : 4.742857142857143  
  },  
  {  
    "brandId" : "6136",  
    "brandName" : "DERMAFLASH",  
    "avgRating" : 4.836601307189542  
  },  
  {  
    "brandId" : "6256",  
    "brandName" : "Mount Lai"  
  }  
]
```

current selected

Activate Windows  
Go to Settings to activate Windows.

Count Documents | 00:00:03.034

# Primeri optimizovanih agregacija

- Nalaženje profila svake osobe (boja kože, tip kože, boja očiju) koja je ostavila maksimalan broj pozitivnih feedback-ova za svaki produkt i svaki brand iz svih kategorija

```
"products" : [
  {
    "product_id" : "P385320",
    "brands" : [
      {
        "brand_name" : "NuFACE",
        "profiles" : {
          "skin_color" : "",
          "skin_type" : "oily",
          "eye_color" : "",
          "hair_color" : "",
          "total_pos_feedback_count" : NumberInt(298)
        }
      }
    ]
  },
  {
    "product_id" : "P389097",
    "brands" : [
      {
        "brand_name" : "HUM Nutrition",
        "profiles" : {
          "skin_color" : "",
          "skin_type" : "oily",
          "eye_color" : ""
        }
      }
    ]
  }
]
```

document selected

Activate Windows  
Go to Settings to activate Windows.

Count Documents | 00:00:00.076

# Primeri optimizovanih agregacija

- Nalaženje brenda sa najjeftinijim proizvodima i broj pozitivnih feedback-ova iz tercijalne kategorije za tipove i boje kože

```
1
{
  "min_price" : 3.0,
  "total_pos_feedback" : NumberInt(104),
  "brands" : [
    "BeautyBio",
    "Koh Gen Do",
    "SEPHORA COLLECTION"
  ],
  "tertiary_category" : "Face Wipes",
  "skin_tone" : "light",
  "skin_type" : "dry"
}
{
  "min_price" : 3.0,
  "total_pos_feedback" : NumberInt(58),
  "brands" : [
    "SEPHORA COLLECTION",
    "Koh Gen Do",
    "BeautyBio",
    "Shiseido"
  ],
  "tertiary_category" : "Face Wipes",
  "skin_tone" : "lightMedium",
  "skin_type" : "oily"
}
{
  "min_price" : 3.0,
  "total_pos_feedback" : NumberInt(113),
  ...
}
```

document selected

Activate Windows  
Go to Settings to activate Windows.

Count Documents | 00:00:02.317

# Primeri optimizovanih agregacija

- Nalaženje proizvoda iz svih kategorija sa maksimalnom cenom za svaku ocenu

```
{
  "rating" : "3.0417",
  "brand" : "tarte",
  "product" : {
    "_id" : "P457468",
    "brand_id" : "5648",
    "brand_name" : "tarte",
    "highlights" : [
      ""
    ],
    "ingredients" : [
      "Water/Aqua/Eau",
      " Propylene Glycol",
      " Dihydroxyacetone",
      " PEG-6 Caprylic/Capric Glycerides",
      " Dihydroxypropyl PEG-5 Linoleammonium Chloride",
      " PEG-7 Glyceryl Cocoate",
      " Sodium Lauroyl Sarcosinate",
      " Hamamelis Virginiana (Witch Hazel) Bark/Leaf/Twig Extract",
      " Phenoxyethanol",
      " Caprylyl Glycol",
      " Potassium Sorbate",
      " Hexylene Glycol",
      " Ethoxydiglycol",
      " Erythulose",
      " "
    ]
  }
}
```

document selected

Activate Windows  
Go to Settings to activate Windows.

Count Documents 00:00:39.794

# Optimizovano vreme izvršavanja

