

---

---

# Recepti i recenzije

Predmet: Sistemi baza podataka  
Profesor: Vladimir Dimitrieski  
Asistent: Jelena Hrnjak

Studenti: Sonja Mihajlović IN33/2020  
Luka Petrović IN26/2020

---

---

# Opis skupova podataka

Skupovi podataka sadrže:

- Informacije o svakom receptu uključuju vreme kuvanja, broj porcija, sastojke, nutritivne vrednosti, uputstva i druge detalje.
- informacije o svakoj recenziji uključuju autora, ocenu, tekst recenzije i još mnogo toga.

Broj zapisa: 522,517 recepata iz 312 različitih kategorija  
1,401,982 recenzija od 271,907 različitih korisnika

Datoteke: recipes.csv & reviews.csv

Link: <https://www.kaggle.com/datasets/irkaal/foodcom-recipes-and-reviews?select=reviews.csv>

# Semantika kolona:

Skup podataka o receptima sadrži 28 kolona.

- **RecipeId:** ID recepta
- **Name:** Ime
- **AuthorId:** ID autora
- **AuthorName:** Ime autora
- **CookTime:** Vreme kuvanja
- **PrepTime:** Vreme pripreme
- **TotalTime:** Ukupno vreme
- **DatePublished:** Datum objavljivanja
- **Description:** Opis
- **Images:** Slike
- **RecipeCategory:** Kategorija recepta
- **Keywords:** Ključne reči
- **RecipeIngredientQuantities:** Količine sastojaka recepta
- **RecipeIngredientParts:** Delovi sastojaka recepta
- **AggregatedRating:** Prosečna ocena
- **ReviewCount:** Broj recenzija
- **Calories:** Kalorije
- **FatContent:** Sadržaj masti
- **SaturatedFatContent:** Sadržaj zasićenih masti
- **CholesterolContent:** Sadržaj holesterola
- **SodiumContent:** Sadržaj natrijuma
- **CarbohydrateContent:** Sadržaj ugljenih hidrata
- **FiberContent:** Sadržaj vlakana
- **SugarContent:** Sadržaj šećera
- **ProteinContent:** Sadržaj proteina
- **RecipeServings:** Broj porcija
- **RecipeYield:** Prinos recepta
- **RecipeInstructions:** Uputstva za recept

# Semantika kolona:

Skup podataka o recenzijama sadrži 8 kolona.

- **ReviewId:** ID recenzije
- **RecipeId:** ID recepta
- **AuthorId:** ID autora
- **AuthorName:** Ime autora
- **Rating:** Ocena
- **ReviewText:** Tekst recenzije
- **DateSubmitted:** Datum podnošenja
- **DateModified:** Datum izmene

# Predlog logičke šeme baze podataka - Kolekcija: Recipes

- Recipeld: Int
- Name: String
- Author:
  - AuthorId: Int
  - AuthorName: String
- Times:
  - CookTime: String
  - PrepTime: String
  - TotalTime: String
- DatePublished: DateTime
- Description: String
- Images: [String]
- RecipeCategory: String
- Keywords: [String]
- Ingredients:
  - Quantity: String
  - Part: String
- AggregatedRating: Float
- ReviewCount: Float
- Nutrition
  - Calories: Float
  - FatContent: Float
  - SaturatedFatContent: Float
  - CholesterolContent: Float
  - SodiumContent: Float
  - CarbohydrateContent: Float
  - FiberContent: Float
  - SugarContent: Float
  - ProteinContent: Float
- RecipeServings: Float
- RecipeYield: String
- RecipeInstructions: [String]

# Predlog logičke šeme baze podataka - Kolekcija: Reviews

- ReviewId: Int
- RecipId: Int
- AuthorId: Int
- AuthorName: String
- Rating: Float
- ReviewText: String
- DateSubmitted: DateTime
- DateModified: DateTime

# Predlog logičke šeme baze podataka - Kolekcija: Recipes

```
{
  "RecipeId": { "0": 38 },
  "Name": { "0": "Low-Fat Berry Blue Frozen Dessert" },
  "AuthorId": { "0": 1533 },
  "AuthorName": { "0": "Dancer" },
  "CookTime": { "0": "PT24M" },
  "PrepTime": { "0": "PT45M" },
  "TotalTime": { "0": "PT24H45M" },
  "DatePublished": { "0": "1999-08-09T21:46:00Z" },
  "Description": { "0": "Make and share this Low-Fat Berry Blue Frozen Dessert recipe from Food.com." },
  "Images": {
    "0": [
      "https://img.sndimg.com/food/image/upload/w_555,h_416,c_fit,fl_progressive,q_95/v1/img/recipes/38/YVclrxYl0neF1h3v3gnM_229%20berry%20blue%20fzn%20dess.jpg",
      "https://img.sndimg.com/food/image/upload/w_555,h_416,c_fit,fl_progressive,q_95/v1/img/recipes/38/AFPO0H4Tuc0bb1COp0aL_255%20berry%20blue%20fzn%20dess.jpg",
      "https://img.sndimg.com/food/image/upload/w_555,h_416,c_fit,fl_progressive,q_95/v1/img/recipes/38/IdGfBmWHT2SG6JCuzTLO_228%20berry%20blue%20fzn%20dess.jpg",
      "https://img.sndimg.com/food/image/upload/w_555,h_416,c_fit,fl_progressive,q_95/v1/img/recipes/38/Per0UN2TGSaVhs2759BA_20140722_202142.jpg",
      "https://img.sndimg.com/food/image/upload/w_555,h_416,c_fit,fl_progressive,q_95/v1/img/recipes/38/plcuaFTeN.jpg",
      "https://img.sndimg.com/food/image/upload/w_555,h_416,c_fit,fl_progressive,q_95/v1/img/recipes/38/pictzvW5.jpg"
    ]
  },
  "RecipeCategory": { "0": "Frozen Desserts" },
  "Keywords": {
    "0": ["Dessert", "Low Protein", "Low Cholesterol", "Healthy", "Free Of...", "Summer", "Weeknight", "Freezer", "Easy"]
  },
  "RecipeIngredientQuantities": {
    "0": ["4", "1/4", "1", "1"]
  },
  "RecipeIngredientParts": {
    "0": ["blueberries", "granulated sugar", "vanilla yogurt", "Lemon juice"]
  },
  "AggregatedRating": { "0": 4.5 },
  "ReviewCount": { "0": 4.0 },
  "Calories": { "0": 170.9 },
  "FatContent": { "0": 2.5 },
  "SaturatedFatContent": { "0": 1.3 },
  "CholesterolContent": { "0": 8.0 },
  "SodiumContent": { "0": 29.0 },
  "CarbohydrateContent": { "0": 37.1 },
  "FiberContent": { "0": 3.6 },
  "SugarContent": { "0": 30.2 },
  "ProteinContent": { "0": 3.2 },
  "RecipeServings": { "0": 4.0 },
  "RecipeYield": { "0": null },
  "RecipeInstructions": {
    "0": [
      "Toss 2 cups berries with sugar.",
      "Let stand for 45 minutes, stirring occasionally.",
      "Transfer berry-sugar mixture to food processor.",
      "Add yogurt and process until smooth.",
      "Strain through fine sieve. Pour into baking pan (or transfer to ice cream maker and process according to manufacturers' directions).",
      "Freeze uncovered until edges are solid but centre is soft. Transfer to processor and blend until smooth again.",
      "Return to pan and freeze until edges are solid.",
      "Transfer to processor and blend until smooth again.",
      "Fold in remaining 2 cups of blueberries.",
      "Pour into plastic mold and freeze overnight. Let soften slightly to serve."
    ]
  }
}
```

# Predlog logičke šeme baze podataka - Kolekcija: Reviews

```
{  
  "ReviewId": { "0": 2 },  
  "RecipeId": { "0": 992 },  
  "AuthorId": { "0": 2008 },  
  "AuthorName": { "0": "gayg msft" },  
  "Rating": { "0": 5 },  
  "Review": { "0": "better than any you can get at a restaurant!" },  
  "DateSubmitted": { "0": "2000-01-25T21:44:00Z" },  
  "DateModified": { "0": "2000-01-25T21:44:00Z" }  
}
```



# Predlog agregacije nad podacima - IN 33/2020

1. Izračunati prag kalorija kao polovinu prosečne kalorijske vrednosti svih recepata, zatim odrediti procenete niskokaloričnih i visokokaloričnih recepata, i identifikovati autore sa najvećim brojem recepata u obe kategorije
2. Identifikovati recepte koji imaju najmanji broj sastojaka i prikaži nazive tih recepata, sastojake, način i vreme pripreme ukoliko se mogu servirati za 4 osobe.
3. Prikazati 10 recepata koji spadaju u kategoriju sa ključnom rečju "Chicken" i imaju ocenu veću ili jednaku sa 4. Prikazati recepte ukoliko je vreme pripreme manje od 60min.
4. Grupisati recepte po ukupnom vremenu (kratki, srednje kratki, srednje dugi, dugi), potom prikazati procenete svih recepata u svakoj od kategorija. Rezultate prikazati u satima.
5. Pronaći recepte iz kategorije "Dessert", sortirati po broju recenzija. Zatim izlistati ime tri recepta , prikazati broj recenzija po receptu i tekstualne opise samih recenzija.

# Predlog agregacije nad podacima - IN 26/2020

1. Prikazati imena 10 autora sa najvećim brojem recepata koji imaju prosečnu ocenu 4.5 ili veću i nalaze se u kategoriji "Steak".
2. Prikazati ime, vreme pripreme i sastojke za 10 recepata sa najvećom prosečnom ocenom koji pripadaju kategoriji "Desert" , imaju manje od 500 kalorija i mogu biti servirani za 6 osoba i imaju bar 1 review sa 100 karaktera.
3. Prikazati naziv recepta, kolicinu kalorija i datum objavljivanja za sve recepte koji spadaju u nezdrave recepte i 'Sydney Mike' je ostavljao recenzije za njih.
4. Prikazati ime recepta, broj kalorija, količinu masnoće i proteina i odnos masnoća i proteina recepata koji imaju prosečnu ocenu veću od 4.0, sadrže barem pet sastojaka i ukupno vreme pripreme je između dva i četiri sata.
5. Prikazati ime recepta, ime autora i datum objavljivanja recepata koji su objavljeni od 2000. godine i sadrže određene sastojke. Prikazani recepti moraju imati najmanje 10 recenzija,a rezultate treba sortirati po prosečnoj oceni u opadajućem redosledu.

# REALIZACIJA PROJEKTA

# Logička šema za optimizaciju upita - Kolekcije: Recipes i Reviews

- Recipeld: Int
  - Name: String
  - Author:
    - AuthorId: Int
    - AuthorName: String
  - Times:
    - CookTime: String
    - PrepTime: String
    - TotalTime: String
  - DatePublished: DateTime
  - Description: String
  - Images: [String]
  - RecipeCategory: String
  - Keywords: [String]
  - Ingredients:
    - Quantity: String
    - Part: String
  - AggregatedRating: Float
  - ReviewCount: Float
- Nutrition
    - Calories: Float
    - FatContent: Float
    - SaturatedFatContent: Float
    - CholesterolContent: Float
    - SodiumContent: Float
    - CarbohydrateContent: Float
    - FiberContent: Float
    - SugarContent: Float
    - ProteinContent: Float
  - RecipeServings: Float
  - RecipeInstructions: [String]
  - Reviews
    - ReviewId: Int
    - Recipeld: Int
    - AuthorId: Int
    - AuthorName: String
    - Rating: Float
    - ReviewText: String
    - DateSubmitted: DateTime
    - DateModified: DateTime

# Logička šema za optimizaciju upita 2 - Kolekcije: Recipes i Reviews

- RecipeId: Int
- Name: String
- Author:
  - AuthorId: Int
  - AuthorName: String
- Times:
  - CookTime: String
  - PrepTime: String
  - TotalTime: String
- DatePublished: DateTime
- Description: String
- Images: [String]
- RecipeCategory: String
- Keywords: [String]
- Ingredients:
  - Quantity: String
  - Part: String
- AggregatedRating: Float
- ReviewCount: Float
- Nutrition
  - Calories: Float
  - FatContent: Float
  - SaturatedFatContent: Float
  - CholesterolContent: Float
  - SodiumContent: Float
  - CarbohydrateContent: Float
  - FiberContent: Float
  - SugarContent: Float
  - ProteinContent: Float
- RecipeServings: Float
- RecipeInstructions: [String]
- Reviews
  - ReviewId: Int
  - RecipeId: Int
  - AuthorId: Int
  - AuthorName: String
  - Rating: Float
  - ReviewText: String
  - DateSubmitted: DateTime
  - DateModified: DateTime
- **ProteinToFatRatio: Float**
- **IsHealthy: Boolean**

# Uređivanje podataka

```
def convert_to_minutes(time_str):
    hours = 0
    minutes = 0
    match = re.match(r'PT(?:\d+)H(?:\d+)M?', time_str)
    if match:
        if match.group(1):
            hours = int(match.group(1))
        if match.group(2):
            minutes = int(match.group(2))
    total_minutes = hours * 60 + minutes
    return total_minutes
```

```
# Function to process and write a bucket of recipes to the output file
def process_and_write_bucket(bucket, output_file):
    restructured_recipes = []
    for recipe in bucket:
        restructured_recipe = process_recipe(recipe)
        restructured_recipes.append(restructured_recipe)

    with open(output_file, 'a', encoding='utf-8') as file:
        for recipe in restructured_recipes:
            json.dump(recipe, file)
            file.write('\n')

# Function to read data in buckets from the input file
def read_in_buckets(file_path, bucket_size):
    with open(file_path, 'r', encoding='utf-8') as file:
        bucket = []
        for line in file:
            recipe = json.loads(line)
            bucket.append(recipe)
            if len(bucket) >= bucket_size:
                yield bucket
                bucket = []
        if bucket:
            yield bucket
```

# Uređivanje podataka

```
reviews_dict = {}  
for _, row in reviews_df.iterrows():  
    review = row.to_dict()  
    recipe_id = review.pop('RecipeId')  
    if recipe_id not in reviews_dict:  
        reviews_dict[recipe_id] = []  
    reviews_dict[recipe_id].append(review)
```

Denormalizacija

```
def process_recipes(input_file, output_file):  
    with open(input_file, 'r', encoding='utf-8') as file:  
        recipes = [json.loads(line) for line in file]  
  
    protein_to_fat_ratios = set()  
    healthy_count = 0  
    unhealthy_count = 0  
  
    for i, recipe in enumerate(recipes):  
        recipes[i], protein_to_fat_ratio, is_healthy = compute_fields(recipe)  
        protein_to_fat_ratios.add(protein_to_fat_ratio)  
        if is_healthy:  
            healthy_count += 1  
        else:  
            unhealthy_count += 1  
  
    # Save the modified recipes to a new JSON file  
    with open(output_file, 'w', encoding='utf-8') as file:  
        for recipe in recipes:  
            json.dump(recipe, file)  
            file.write('\n')
```

# Upiti - IN 33/2020

1. Izračunati prag kalorija kao polovinu prosečne kalorijske vrednosti svih recepata, zatim odrediti procenete niskokaloričnih i visokokaloričnih recepata, i identifikovati autore sa najvećim brojem recepata u obe kategorije

```
var lowCalorieThresholdResult = db.merged_recipes.aggregate([
  { $group: { _id: null, avgCalories: { $avg: "$Nutrition.Calories" } } },
  { $project: { _id: 0, lowCalorieThreshold: { $multiply: ["$avgCalories", 0.5] } } } ]).next();
var lowCalorieThreshold = lowCalorieThresholdResult.lowCalorieThreshold;
var pipeline = [
  { $facet: { lowCalorieCount: [
    { $match: { "Nutrition.Calories": { $lte: lowCalorieThreshold } } }, { $count: "count" } ],
    highCalorieCount: [
    { $match: { "Nutrition.Calories": { $gt: lowCalorieThreshold } } }, { $count: "count" } ],
    totalRecipesCount: [ { $group: { _id: null, count: { $sum: 1 } } } ],
    lowCalorieAuthor: [
    { $match: { "Nutrition.Calories": { $lte: lowCalorieThreshold } } },
    { $group: { _id: "$Author.AuthorName", count: { $sum: 1 } } }, { $sort: { count: -1 } }, { $limit: 1 } ],
    highCalorieAuthor: [
    { $match: { "Nutrition.Calories": { $gt: lowCalorieThreshold } } },
    { $group: { _id: "$Author.AuthorName", count: { $sum: 1 } } }, { $sort: { count: -1 } }, { $limit: 1 } ] ] },
  { $project: { lowCaloriePercentage: { $multiply: [
    { $divide: [ { $arrayElemAt: ["$lowCalorieCount.count", 0] }, { $arrayElemAt: ["$totalRecipesCount.count", 0] } ] }, 100 ] },
    highCaloriePercentage: { $multiply: [
    { $divide: [ { $arrayElemAt: ["$highCalorieCount.count", 0] }, { $arrayElemAt: ["$totalRecipesCount.count", 0] } ] }, 100 ] },
    lowCalorieAuthor: {
      authorName: { $arrayElemAt: ["$lowCalorieAuthor._id", 0] },
      recipeCount: { $arrayElemAt: ["$lowCalorieAuthor.count", 0] } },
    highCalorieAuthor: {
      authorName: { $arrayElemAt: ["$highCalorieAuthor._id", 0] },
      recipeCount: { $arrayElemAt: ["$highCalorieAuthor.count", 0] } } } ] ];
var results = db.merged_recipes.aggregate(pipeline).next();
printJson(results);
```



# Upiti - IN 33/2020

2. Identifikovati recepte koji imaju najmanji broj sastojaka ako spadaju u kategoriju "Vegetable" i prikaži nazive tih recepata, sastojake, način i vreme pripreme ukoliko se mogu servirati za 4 osobe.

```
db.merged_recipes.aggregate([
  { $match: { RecipeServings: 4, "RecipeIngredient.RecipeIngredientParts": { $exists: true, $not: { $size: 0 } },
    $expr: { $gte: [{ $size: "$RecipeIngredient.RecipeIngredientParts" }, 2] }, RecipeCategory: "Vegetable" } },
  { $addFields: { numIngredients: { $size: "$RecipeIngredient.RecipeIngredientParts" } } },
  { $sort: { numIngredients: 1 } },
  { $limit: 1 },
  { $project: { Name: 1, "RecipeIngredient.RecipeIngredientParts": 1, Description: 1, TotalTime: 1, RecipeInstructions: 1 } }
]);
```

# Upiti - IN 33/2020

3. Prikazati 10 recepata koji spadaju u kategoriju sa ključnom rečju "Chicken" i imaju ocenu veću ili jednaku sa 4. Prikazati recepte ukoliko je vreme pripreme manje od 60min.

```
db.merged_recipes.aggregate([
  { $match: { "Keywords": "Chicken", "AggregatedRating": { $gte: 4 }, "Times.Preptime": { $lt: 60 } } },
  { $addFields: { numIngredients: { $cond: { if: { $isArray: "$RecipeIngredient.RecipeIngredientParts" },
    then: { $size: "$RecipeIngredient.RecipeIngredientParts" }, else: { $cond: { if: { $eq: ["$RecipeIngredientParts", ""] }, then: 0, else: 1 } } } } } },
  { $match: { "numIngredients": { $gt: 0 } } },
  { $project: { "Name": 1, "Author.AuthorName": 1, "RecipeCategory": 1, "RecipeIngredient.RecipeIngredientParts": 1,
    "Nutrition.Calories": 1, "Nutrition.ProteinContent": 1, "RecipeServings": 1, "RecipeInstructions": 1, "numIngredients": 1 } },
  { $sort: { "Nutrition.Calories": -1, "Nutrition.ProteinContent": -1 } },
  { $limit: 10 }
]);
```

# Upiti - IN 33/2020

4. Grupisati recepte po ukupnom vremenu (kratki, srednje kratki, srednje dugi, dugi), potom prikazati procenete svih recepata u svakoj od kategorija. Rezultate prikazati u satima.

```
db.merged_recipes.aggregate([
  { $facet: { counts: [ { $bucket: { groupBy: "$Times.TotalTime", boundaries: [0, 370.25, 740.5, 1110.75, 1485],
    default: "Other", output: { count: { $sum: 1 } } } } ], total: [ { $group: { _id: null, totalRecipesCount: { $sum: 1 } } } ] } },
  { $unwind: "$counts" },
  { $project: { _id: 0, category: { $switch: { branches: [ { case: { $eq: ["$counts._id", 0] }, then: "do 6h" },
    { case: { $eq: ["$counts._id", 370.25] }, then: "od 6 do 12h" }, { case: { $eq: ["$counts._id", 740.5] }, then: "od 12 do 18h" },
    { case: { $eq: ["$counts._id", 1110.75] }, then: "od 18 do 24h" } ], default: "Other" } },
    percentage: { $multiply: [ { $divide: ["$counts.count", { $arrayElemAt: ["$total.totalRecipesCount", 0] } ] }, 100 ] } } }
  ]}).forEach(function(doc) { printjson({ [doc.category]: doc.percentage.toFixed(2) + "%" }); });
```

# Upiti - IN 33/2020

5. Pronaći recepte iz kategorije "Dessert", sortirati po broju recenzija. Zatim izlistati ime tri recepta , prikazati broj recenzija po receptu i tekstualne opise samih recenzija.

```
db.merged_recipes_rev.aggregate([
  { $match: { RecipeCategory: "Dessert" } },
  { $sort: { ReviewCount: -1 } },
  { $limit: 3 },
  { $project: {
    _id: 0,
    RecipeId: "$RecipeId",
    RecipeName: "$Name",
    TotalReviews: "$ReviewCount",
    Reviews: "$Reviews"
  }
}
]);
```

# Upiti - IN 26/2020

1. Prikazati imena 10 autora sa najvećim brojem recepata koji imaju prosečnu ocenu 4.5 ili veću i nalaze se u kategoriji "Steak".

```
db.final.aggregate([
  { $match: { "RecipeCategory": "Steak", "AggregatedRating": { $gte: 4.5 } } },
  { $group: { _id: "$Author.AuthorName", recipeCount: { $sum: 1 } } },
  { $sort: { recipeCount: -1 } },
  { $limit: 10 },
  { $project: { _id: 0, AuthorName: "$_id", recipeCount: 1 } }
])
```

# Upiti - IN 26/2020

2. Prikazati ime, vreme pripreme i sastojke za 10 recepata sa najvećom prosečnom ocenom koji pripadaju kategoriji "Desert", imaju manje od 500 kalorija i mogu biti servirani za 6 osoba i imaju bar 1 review sa 100 karaktera.

```
db.final.aggregate([
  {
    $match: {
      "RecipeCategory": "Dessert",
      "Nutrition.Calories": { $lt: 500 },
      "RecipeServings": 6,
      "Reviews": { $elemMatch: { "Review": { $regex: /^.{100,}$/ } } }
    }
  },
  {
    $sort: {
      "AggregatedRating": -1
    }
  },
  {
    $project: {
      _id: 0,
      Name: 1,
      "Times.TotalTime": 1,
      "RecipeIngredient.RecipeIngredientParts": 1,
      "AggregatedRating": 1
    }
  }
])
```

# Upiti - IN 26/2020

3. Prikazati naziv recepta, kolicinu kalorija i datum objavljivanja za sve recepte koji spadaju u nezdrave recepte i 'Sydney Mike' je ostavljao recenzije za njih.

```
db.final.aggregate([
  {
    $match: {
      "IsHealthy": false,
      "Reviews.AuthorName": "Sydney Mike"
    }
  },
  {
    $project: {
      _id: 0,
      Name: 1,
      "Nutrition.Calories": 1,
      DatePublished: 1
    }
  }
])
```

# Upiti - IN 26/2020

4. Prikazati ime recepta, broj kalorija, količinu masnoće i proteina i odnos masnoća i proteina recepata koji imaju prosečnu ocenu veću od 4.0, sadrže barem pet sastojaka i ukupno vreme pripreme je između dva i četiri sata.

```
db.final.aggregate([
  { $match: { "AggregatedRating": { $gt: 4.0 }, $expr: { $gte: [ { $size: "$RecipeIngredient.RecipeIngredientParts" }, 5 ] }, "Times.TotalTime": { $gte: 120, $lte: 240 } } },
  { $project: { _id: 0, Name: 1, "Nutrition.Calories": 1, "Nutrition.FatContent": 1, "Nutrition.ProteinContent": 1, "ProteinToFatRatio": 1 } },
  { $sort: { "AggregatedRating": -1 } },
  { $limit: 10 }
])
```



# Upiti - IN 26/2020

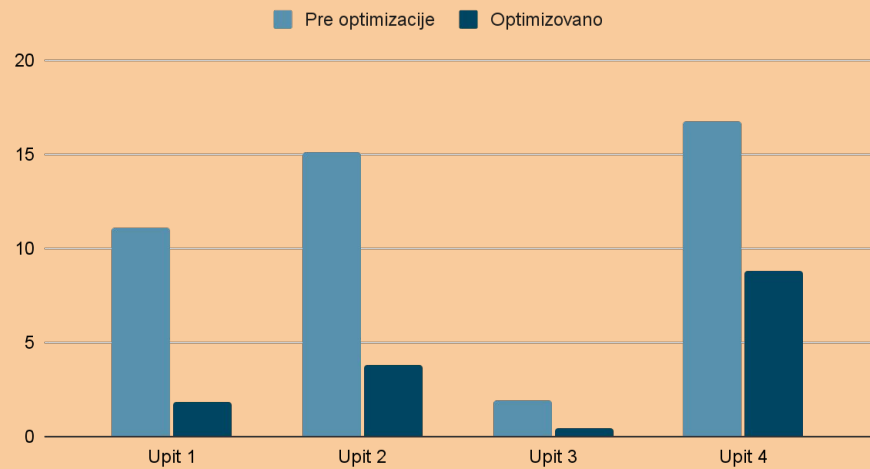
5. Prikazati ime recepta, ime autora i datum objavljivanja recepata koji su objavljeni od 2000. godine i sadrže određene sastojke. Prikazani recepti moraju imati najmanje 10 recenzija, a rezultate treba sortirati po prosečnoj oceni u opadajućem redosledu.

```
db.final.aggregate([
  {
    $match: {
      "DatePublished": { $gte: ISODate("2000-01-01"), $lt: ISODate("2021-01-01") },
      "RecipeIngredient.RecipeIngredientParts": { $in: ["eggs", "salt", "butter", "brown sugar", "granulated sugar", "vanilla extract", "flour", "pecan halves"] }
    },
  },
  { $addFields: { reviewCount: { $size: "$Reviews" } } },
  { $match: { reviewCount: { $gte: 10 } } },
  {
    $project: {
      Name: 1,
      AuthorName: "$Author.AuthorName",
      DatePublished: 1,
      "AggregatedRating" : 1,
      reviewCount: 1
    }
  }
])
```

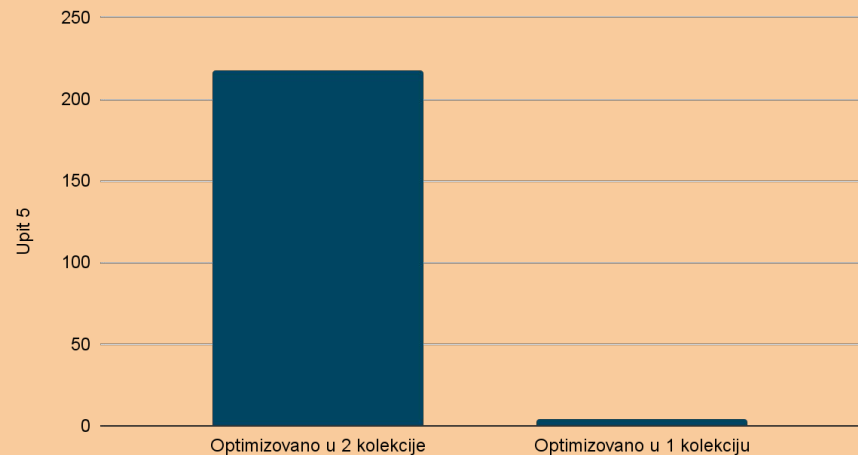
# REZULTATI OPTIMIZACIJE

# Upiti - IN 33/2020

Poređenje prva 4 upita, vreme izraženo u sekundama



Poređenje 5. upita



# Upiti - IN 33/2020

Indeksi korišćeni u upitima

```
db.merged_recipes.createIndex({ "Nutrition.Calories": 1, "Author.AuthorName": 1 });  
db.merged_recipes.createIndex({ "Keywords": "text" })  
db.merged_recipes.createIndex({ "Times.TotalTime": 1 })
```

Najbolji indeksi za optimizaciju upita nad kolekcijom merged\_recipes su:

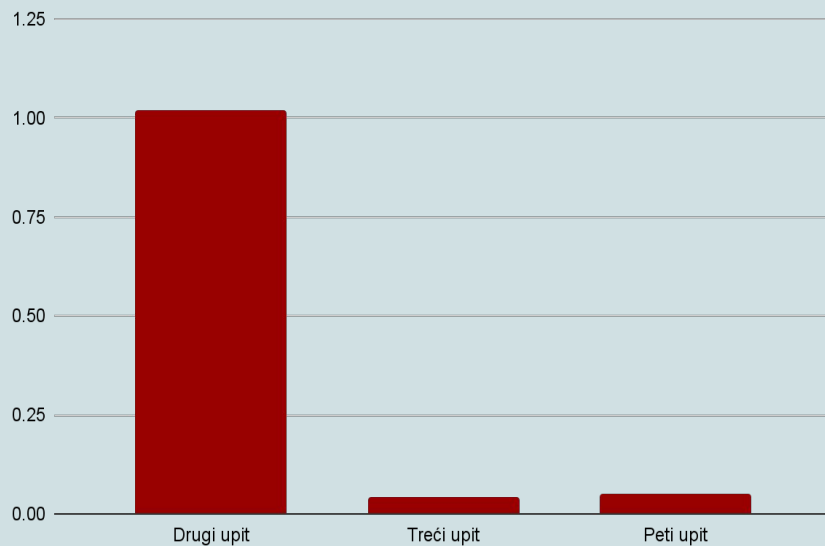
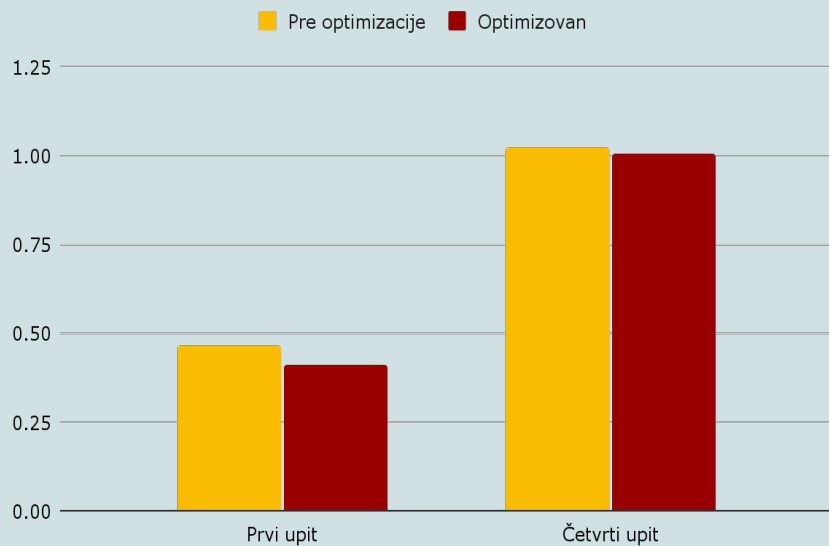
**Single Field Indexes** - Times.TotalTime

**Single Field Indexes**, Text index - Keywords

**Compound Index** koji kombinuje polja Nutrition.Calories i Author.AuthorName

# Upiti - IN 26/2020

\*Vreme je izraženo u sekundama



# Upiti - IN 26/2020

Korišćeni indexi:

*1. Compound Index*

*2. Compound Index*

*3. Single Field Index*

```
db.final.createIndex({
  "DatePublished": 1,
  "RecipeIngredient.RecipeIngredientParts": 1,
})

db.final.createIndex({
  "RecipeCategory": 1,
  "Reviews.Review": "text"
})

db.final.createIndex({
  "Reviews.AuthorName": 1
})
```

HVALA  
NA  
PAŽNJI

