



1^ο Πρότζεκτ στο μάθημα ”Διαχείριση Μεγάλων Δεδομένων”

– Ανάλυση δεδομένων για anime παραγωγές με
χρήση MongoDB –

παραδοτέο από

Κουνάδη Βασιλική (Α.Μ.: 2022202000102)
Μαζαράκης Ιωάννης (Α.Μ.: 2022202000130)
Τετράδης Αναστάσιος (Α.Μ.: 2022202000206)

18 Φεβρουαρίου 2024

Επιβλέπων: Ραυτοπούλου Παρασκευή

Τμήμα Πληροφορικής και Τηλεπικοινωνιών
Σχολή Οικονομίας και Τεχνολογίας
Πανεπιστήμιο Πελοποννήσου

Περιεχόμενα

1	Επεξεργασία των δεδομένων	3
2	Ανάλυση των δεδομένων	5
2.1	Τι ξέρουμε για τις παραγωγές του πολύ δημοφιλούς Studio Ghibli . .	5
2.1.1	Query	5
2.1.2	Αποτελέσματα	6
2.1.3	Γραφική απεικόνιση και Συμπεράσματα	7
2.2	Πόσες ετικέτες χρησιμοποιούνται συνήθως στις anime παραγωγές . .	8
2.2.1	Query	8
2.2.2	Αποτελέσματα	9
2.2.3	Γραφική απεικόνιση και Συμπεράσματα	11
2.3	Ποιες είναι οι πιο δημοφιλείς ετικέτες στις anime παραγωγές	12
2.3.1	Query	12
2.3.2	Αποτελέσματα	14
2.3.3	Γραφική απεικόνιση και Συμπεράσματα	15
2.4	Ποια στρατηγική ακολουθούν οι εταιρείες παραγωγής anime	16
2.4.1	Query	16
2.4.2	Αποτελέσματα	17
2.4.3	Γραφική απεικόνιση και Συμπεράσματα	18
2.5	Ποιος σχεδιαστής έχει δημιουργήσει τις περισσότερες anime παραγωγές	19
2.5.1	Query	19
2.5.2	Αποτελέσματα	22
2.5.3	Γραφική απεικόνιση και Συμπεράσματα	23
2.6	Σχετίζεται η δημοτικότητα μιας παραγωγής με το πλήθος των κύκλων που θα παραχθούν	24
2.6.1	Query	24
2.6.2	Αποτελέσματα	28
2.6.3	Γραφική απεικόνιση και Συμπεράσματα	32
	References	33

Οδηγίες εγκατάστασης

Παρακάτω υπάρχουν αναλυτικές οδηγίες για την εκτέλεση των query της εργασίας σε MongoDB Compass [1], καθώς και για την εγκατάσταση του ανωτέρω. Επιπλέον παραθέτουμε οδηγίες εγκατάστασης της Python με τις απαραίτητες βιβλιοθήκες, προκειμένου ο χρήστης να μπορεί να τρέξει τον κώδικα που παρέχουμε για τις γραφικές απεικονίσεις (καθώς δυστυχώς λόγω της πληθώρας των εγγραφών οι ετικέτες δεν είναι καθόλου ξεκάθαρες στις φωτογραφίες και χρειάζεται να γίνει μεγέθυνση στο διάγραμμα από το UI της Python όταν παράγεται η απεικόνιση). Για τις φωτογραφίες έχουμε χρησιμοποιήσει ένα τμήμα της κάθε γραφικής απεικόνισης.

Χειρισμός MongoDB Compass

Βήμα 1: Λήψη του Πρότζεκτ

Κατεβάστε το έργο στον τοπικό σας υπολογιστή.

Βήμα 2: Εγκατάσταση MongoDB Compass στο σύστημα σας

Επισκευθείτε τον σύνδεσμο <https://www.mongodb.com/try/download/community> και επιλέξτε το κουμπί "Select package", και από εκεί κάνετε τις κατάλληλες επιλογές για το λειτουργικό σας σύστημα. Αφού κατεβάσετε το πακέτο και το ανοίξετε ακολουθείτε κατά γράμμα τις οδηγίες της εφαρμογής μέχρι την ολοκλήρωση της εγκατάστασης.

Στη πρώτη ενεργοποίηση του προγράμματος επιλέγετε το "Connect" στον localhost σας, για να συνδεθείτε στο κύριο παράθυρο της εφαρμογής.

Βήμα 3: Εκτέλεση query στην MongoDB Compass

Από το κύριο παράθυρο της εφαρμογής σας δημιουργείτε ένα νέο database με ένα νέο collection από το JSON το οποίο προκύπτει μέσω του csv, παρεχόμενο για το πρότζεκτ, όπως περιγράφεται στην ενότητα 1.

Αφού βεβαιωθείτε ότι βρίσκεστε εντός του περιβάλλοντος εργασίας για το collection του πρότζεκτ, επιλέγετε την ενότητα "Aggregations" από το οριζόντιο μενού.

Εκεί στη συνέχεια επιλέγετε από τη δεξιά πλευρά την επιλογή `"/Text"` (αντί για το προεπιλεγμένο `"{}Stages"`).

Στο κενό πεδίο editor κάνετε επικόλληση για το query που θέλετε να τρέξετε, **μονο το κείμενο εντός των [...]** (διαφορετικά το query δεν θα λειτουργήσει, καθώς πρόκειται για άλλη βάση δεδομένων σε κάθε σύστημα). Για την εξαγωγή των αποτελεσμάτων πατήστε στο `"Export"` και επιλέξτε CSV.

Εγκατάσταση Python στο Visual Studio Code για γραφήματα

Βήμα 1: Λήψη του Πρότζεκτ

Κατεβάστε το έργο στον τοπικό σας υπολογιστή. Βεβαιωθείτε ότι το csv των αποτελεσμάτων να είναι στον ίδιο φάκελο με το .py αρχείο πριν επιχειρίσετε να τρέξετε κάποιο από αυτά.

Βήμα 2: Εγκατάσταση Python στο Visual Studio Code

Για την εφαρμογή Visual Studio code πηγαίνετε στην επιλογή `"Extensions"` όπου αναζητάτε την επέκταση `"Python"`, την οποία και κάνετε install. Έτσι εγκαθιστάτε τον compiler της Python στον υπολογιστή σας. Έπειτα ανοίγετε τον φάκελο του πρότζεκτ με `"Open Folder"`.

Βήμα 3: Εγκατάσταση βιβλιοθηκών Python για γραφήματα

Ανοίξτε ένα terminal εντός Visual Studio Code και εκτελέστε τις εντολές:

```
1 pip install matplotlib
2 pip install pandas
```

Ανοίξτε οποιοδήποτε script/.py αρχείο του πρότζεκτ στον editor και τρέξτε το προκειμένου να εμφανιστεί το ανάλογο διαγράμμα.

Σημείωση: Το ερώτημα 2.6 έχει καθυστέρηση στην εκτέλεση (3 λεπτά) λόγω του μεγέθους των δεδομένων εισόδου, και καθυστέρηση όταν γίνονται adjustments (30 δευτερόλεπτα).

1 Επεξεργασία των δεδομένων

Στα πλαίσια της εργασίας, έπρεπε να μετατρέψουμε τα δεδομένα σας σε JSON format. Προκειμένου να το επιτύχουμε αυτό μέσω του MongoDB Compass, κατευθυνθήκαμε στο αρχικό μενού της εφαρμογής και επιλέξαμε στο αριστερά μενού το συμβολο "+", δίπλα στο "Databases". Με αυτόν τον τρόπο δημιουργούμε τη βάση δεδομένων μας και της δίνουμε όνομα. Αντίστοιχα δημιουργούμε στο ίδιο παράθυρο και το πρώτο Collection της βάσης και το ονοματίζουμε (εμείς το ονομάσαμε "Anime Statistics and Metadata" το οποίο αναφέρεται στην αρχή κάθε aggregation query). Φροντίζουμε πρώτου πατήσουμε "Create Database" να μην είναι επιλεγμένο το κουτί "Time-Series".

Έπειτα αφού δημιουργηθεί η βάση με το collection θα μας δωθεί η δυνατότητα να κάνουμε "Import data". Αφού πατήσουμε το αντίστοιχο κουμπί, επιλέγουμε το "Anime.csv" για να κάνουμε Import. Στο μενού που θα μας εμφανιστεί προσέχουμε το delimiter να είναι comma, να αγνοούνται τα άδεια strings. Επίσης γίνεται λεπτομερής αναπαράσταση των data στο παράθυρο επιβεβαίωσης, όπου πρέπει να φροντίσουμε κάθε πεδίο που διαχωρίζει το Compass, να έχει τον σωστό τύπο δεδομένων, πριν κάνουμε το τελικό "Import".

Μόλις εκτελέσουμε τα παραπάνω βήματα, έχουμε στη διάθεση μας το Anime.csv σε JSON format, με αριθμούς και strings ως τύπους των attributes (πεδία με πολλαπλές τιμές δεν είναι περασμένα ως array, αλλά ως string, και κάθε φορά που χρειαζόμαστε μεμονωμένες τις τιμές του, κάνουμε split το πεδίο με delimiter το ","). Ενδεικτικά παραθέτουμε την πρώτη εγγραφή:

```
1 {  
2   "_id": {  
3     "$oid": "65d0a2b567e0fa134945c85a"  
4   },  
5   "Rank": 1,  
6   "Name": "Demon Slayer: Kimetsu no Yaiba - Entertainment  
7     District Arc",  
8   "Japanese_name": " Kimetsu no Yaiba: Yuukaku-hen",  
9   "Type": "TV  ",  
10  "Studio": "ufotable",  
    "Release_season": "Fall  ",
```

```
11 "Tags": "Action, Adventure, Fantasy, Shounen, Demons,  
    Historical, Martial Arts, Orphans, Siblings, Swordplay,  
    Based on a Manga, Explicit Violence",  
12 "Rating": 4.6,  
13 "Release_year": 2021,  
14 "Description": "Tanjiro and his friends accompany the Hashira  
    Tengen Uzui to an entertainment district where Tengen's female  
    ninja agents were gathering information on a demon before they  
    suddenly disappeared. In order to investigate, Tanjiro and the  
    others disguise themselves as women to sneak in!",  
15 "Content_Warning": "Explicit Violence",  
16 "Related_Mange": "Demon Slayer: Kimetsu no Yaiba",  
17 "Related_anime": "Demon Slayer: Kimetsu no Yaiba, Demon Slayer:  
    Kimetsu no Yaiba Movie - Mugen Train, Demon Slayer: Kimetsu  
    no Yaiba - Mugen Train",  
18 "Voice_actors": "Inosuke Hashibira : Yoshitsugu Matsuoka,  
    Nezuko Kamado : Akari Kitou, Tanjirou Kamado : Natsuki  
    Hanae, Zenitsu Agatsuma : Hiro Shimono, Daki : Miyuki  
    Sawashiro, Tengen Uzui : Katsuyuki Konishi, Akaza : Akira  
    Ishida, Amane Ubuyashiki, Koyoharu Gotouge\nOriginal  
    Creator, Haruo Sotozaki\nDirector, Akira  
    Matsushima\nCharacter Design, Aimer\nSong Performance",  
19 "staff": "Koyoharu Gotouge : Original Creator, Haruo Sotozaki :  
    Director, Akira Matsushima : Character Design, Aimer : Song  
    Performance"  
20 }
```

2 Ανάλυση των δεδομένων

2.1 Τι ξέρουμε για τις παραγωγές του πολύ δημοφιλούς Studio Ghibli

2.1.1 Query

Το παρακάτω NoSQL query στοχεύει στην εύρεση εγγραφών που σχετίζονται με το στούντιο παραγωγής Studio Ghibli. Προβάλλει συγκεκριμένα πεδία από τις εγγραφές, φιλτράρει εγγραφές όπου το πεδίο "Studio" ισούται με "Studio Ghibli", ταξινομεί τα αποτελέσματα κατά το έτος παραγωγής σε αύξουσα σειρά. Αναλυτικά: \$filter stage: Αυτό το στάδιο φιλτράρει documents όπου το πεδίο "Studio" ισούται με "Studio Ghibli".

\$project stage: Αυτό το στάδιο προβάλλει συγκεκριμένα πεδία από τα φιλτραρισμένα documents. Τα πεδία που προβάλλονται είναι: Rank, Name, Release_year, Type, Tags, Rating και Content_Warning.

\$sort stage: Αυτό το στάδιο ταξινομεί τα παραπάνω documents κατά το πεδίο "Release_year" με αύξουσα σειρά.

```
1 db.getCollection(  
2   'Anime Statistics and Metadata'  
3 ).aggregate(  
4   [  
5     {  
6       $filter: {  
7         Studio: 'Studio Ghibli'  
8       }  
9     },  
10    {  
11      $project: {  
12        Rank: 1,  
13        Name: 1,  
14        Release_year: 1,  
15        Type: 1,  
16        Tags: 1,  
17        Rating: 1,  
18        Content_Warning: 1
```

```

19     }
20   },
21   { $sort: { Release_year: 1 } }
22 ],
23 { maxTimeMS: 60000, allowDiskUse: true }
24 );

```

Αλγόριθμος 1: Query για ερώτηση 2.1

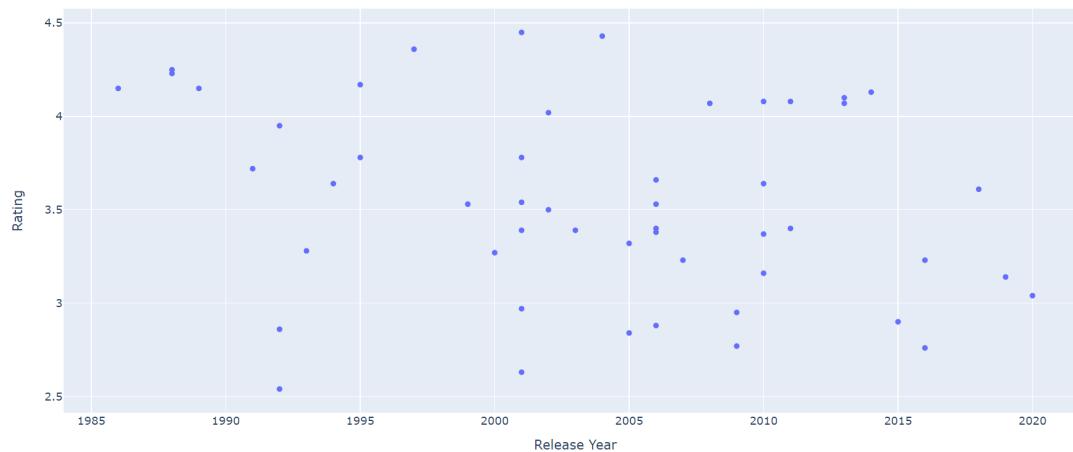
2.1.2 Αποτελέσματα

Παρακάτω παραθέτουμε ένα πίνακα με τα 20 πρώτα αποτελέσματα που λάβαμε με αυτό το query (τα ολοκληρωμένα αποτελέσματα βρίσκονται στο αρχείο "Question2_1"):

Table 1: Ταινίες του Studio Ghibli

Rank	Name	Type	Tags	Rating	Release Year	Content Warning
414	Laputa: Castle in the Sky	Movie	Adventure, Fantasy, Family Friendly	4.15	1986	
210	My Neighbor Totoro	Movie	Fantasy, Slice of Life, Contemporary Fantasy	4.25	1988	
243	Grave of the Fireflies	Movie	Drama, Autobiographies, Historical	4.23	1988	
415	Kiki's Delivery Service	Movie	Fantasy, Slice of Life, Apprenticeship	4.15	1989	
2573	Only Yesterday	Movie	Drama, Josei, Slice of Life	3.72	1991	Mature Themes
1112	Porco Rosso	Movie	Action, Adventure, Aviation	3.95	1992	
13902	Sora Iro no Tane	Other	Family Friendly, Shorts	2.86	1992	
14950	Nandarou	Other	Shorts	2.54	1992	
9170	Ocean Waves	TV Sp	Drama, Romance, Slice of Life	3.28	1993	
3231	Pom Poko	Movie	Adventure, Fantasy, Animal Protagonists	3.64	1994	
370	Whisper of the Heart	Movie	Comedy, Drama, Romance	4.17	1995	
2110	On Your Mark	Music	Angels, Family Friendly	3.78	1995	
83	Princess Mononoke	Movie	Action, Adventure, Fantasy	4.36	1997	
4678	My Neighbors The Yamadas	Movie	Comedy, Slice of Life	3.53	1999	
9257	Ghiblies	Other	Comedy	3.27	2000	
29	Spirited Away	Movie	Adventure, Fantasy, Curse	4.45	2001	
2061	Mei and the Baby Cat Bus	Movie	Adventure, Fantasy, Animal Protagonists	3.78	2001	
4575	Koro's Big Day Out	Movie	Adventure, Animal Protagonists	3.54	2001	
7093	The Whale Hunt	Movie	Fantasy, Family Friendly	3.39	2001	
13132	Lawson: Spirited Away CM	Other	Commercials, Promotional	2.97	2001	

2.1.3 Γραφική απεικόνιση και Συμπεράσματα



Εικόνα 1: Έτος παραγωγής και rating ταινιών studio Ghibli

Από την ανάλυση δεδομένων, φαίνεται ότι το Studio Ghibli, ως ένα από τα πιο διάσημα στούντιο κινουμένων σχεδίων, έχει δημιουργήσει μια μεγάλη γκάμα περιεχομένου, συμπεριλαμβανομένων ταινιών, μουσικών βίντεο, τηλεοπτικών specials, OVA και web animations. Αν και υπάρχουν διακυμάνσεις στην παραγωγή τους με την πάροδο των ετών, έχουν διατηρήσει ένα σταθερό επίπεδο παραγωγής από την ίδρυσή τους, με διάφορες συχνότητες εκδόσεων νέου υλικού. Πολλές από τις παραγωγές τους έχουν λάβει υψηλές βαθμολογίες, υποδεικνύοντας τη σταθερή ποιότητά τους και την απήχυσή τους στο κοινό, αν και δεν είναι όλα τα έργα τους top-rated. Οι παραγωγές του Studio Ghibli καλύπτουν ένα ευρύ φάσμα θεμάτων και ειδών, όπως φαντασία, δράμα, ρομαντισμό, κομμάτι της ζωής, ενηλικίωση και ιστορικά περιβάλλοντα, επιτρέποντάς τους να απευθύνονται σε ένα ευρύ κοινό. Μερικά από τα έργα τους είναι διασκευές μυθιστορημάτων, manga ή βασίζονται σε ιστορικά γεγονότα. Επίσης, ελάχιστα από τα έργα τους έχουν warning, γεγονός που δείχνει ότι είναι κατάλληλα για όλη την οικογένεια. Συνολικά, το ιστορικό του Studio Ghibli δείχνει τη δημιουργικότητα, την ευελιξία και τη δέσμευσή του να παράγουν ποιοτικά anime σε διάφορα είδη.

2.2 Πόσες ετικέτες χρησιμοποιούνται συνήθως στις anime παραγωγές

2.2.1 Query

Το παρακάτω NoSQL query στοχεύει να βρει το πλήθος των ετικετών/tags που περιγράφουν μια παραγωγή anime. Για κάθε εγγραφή στα αποτελέσματα, δείχνει την κατάταξη (Rank) του anime, την αξιολόγηση χρηστών (Rating) και τον αριθμό των ετικετών που σχετίζονται με το anime. Τα αποτελέσματα ταξινομούνται κατά τις βαθμολογίες με φθίνουσα σειρά. Αναλυτικά:

\$addField stage: Αυτό το στάδιο προσθέτει ένα νέο πεδίο που ονομάζεται "tagCount" σε κάθε document. Υπολογίζει τον αριθμό των ετικετών/tags, διαχωρίζοντας το πεδίο "Tags" με κόμμα και στη συνέχεια υπολογίζοντας το μέγεθος του πίνακα που προκύπτει.

\$project stage: Αυτό το στάδιο προβάλλει συγκεκριμένα πεδία από τα documents. Τα πεδία που προβάλλονται είναι: Name, Rank, Rating, and tagCount.

\$sort stage: Αυτό το στάδιο ταξινομεί τα documents κατά το πεδίο "Rating" με φθίνουσα σειρά.

```
1 db.getCollection(  
2   'Anime Statistics and Metadata'  
3 ).aggregate(  
4   [  
5     {  
6       $addField: {  
7         tagCount: {  
8           $size: { $split: [{ $ifNull: ["$Tags", "" ]}, "," ] }  
9         }  
10    }  
11  ],  
12  {  
13    $project: {  
14      Name: 1,  
15      Rank: 1,  
16      Rating: 1,  
17      tagCount: 1
```

```
18     }  
19   },  
20   { $sort: { Rating: -1 } }  
21 ],  
22 { maxTimeMS: 60000, allowDiskUse: true }  
23 );
```

Αλγόριθμος 2: Query για ερώτηση 2.2

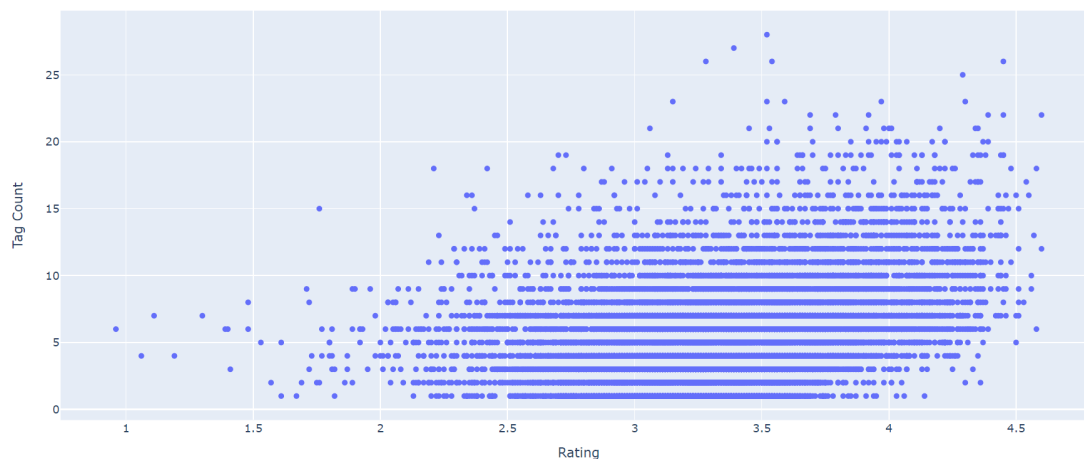
2.2.2 Αποτελέσματα

Παρακάτω παραθέτουμε ένα πίνακα με τα 20 πρώτα αποτελέσματα που λάβαμε με αυτό το query (τα ολοκληρωμένα αποτελέσματα βρίσκονται στο αρχείο "Question2_2"):

Table 2: Anime Tags

Rank	Name	Rating	Αριθμός Tag
4770	Happy Sugar Life	3.52	28
7119	Redo of Healer	3.39	27
31	Heaven Official's Blessing	4.45	26
4471	GANTZ	3.54	26
9167	Magical Girl Site	3.28	26
159	Hellsing Ultimate	4.29	25
139	Dororo (2019)	4.3	23
996	Future Diary	3.97	23
3828	A Little Princess Sara	3.59	23
4826	Gravitation	3.52	23
11158	Magical Girl Spec-Ops Asuka	3.15	23
2	Fruits Basket the Final Season	4.6	22
27	Fruits Basket 2nd Season	4.45	22
60	That Time I Got Reincarnated as a Slime	4.39	22
1247	Inuyashiki: Last Hero	3.92	22
2001	Btooom!	3.79	22
2781	Devils' Line	3.69	22
96	ERASED	4.35	21
108	Fruits Basket 1st Season	4.34	21
307	Castlevania Season 2	4.2	21

2.2.3 Γραφική απεικόνιση και Συμπεράσματα



Εικόνα 2: Αξιολογήσεις ανά πλήθος ετικετών

Με βάση την ανάλυση δεδομένων, φαίνεται ότι ο αριθμός των ετικετών για ένα anime δεν συσχετίζεται σταθερά με τη βαθμολογία χρήστη του, υποδεικνύοντας ότι η ύπαρξη των περισσότερων ετικετών δεν οδηγεί απαραίτητα στην υψηλότερη βαθμολογία. Ενώ ορισμένα anime με πολλές ετικέτες λαμβάνουν υψηλές βαθμολογίες, υπάρχουν περιπτώσεις όπου τα anime με λιγότερες ετικέτες επιτυγχάνουν εξίσου υψηλές βαθμολογίες, γεγονός που υποδηλώνει ότι η ποσότητα των ετικετών από μόνη της δεν αποτελεί αξιόπιστο δείκτη πρόβλεψης της ικανοποίησης των χρηστών. Παράγοντες πέρα από τον αριθμό των ετικετών, όπως η ποιότητα της ιστορίας, η καλλιτεχνία, η ανάπτυξη χαρακτήρων και η συνολική απόλαυση, πιθανότατα έχουν πιο σημαντική επίδραση στις αξιολογήσεις των χρηστών. Επιπλέον, τα ακραία σημεία στα δεδομένα μπορεί να αντιπροσωπεύουν μοναδικές περιπτώσεις όπου ορισμένα anime αποκλίνουν σημαντικά από τις αναμενόμενες τάσεις, πιθανώς λόγω εξειδικευμένου ή αμφιλεγόμενου περιεχομένου, εξαιρετικής ποιότητας ή λανθασμένης γενίκευσης στις ετικέτες. Σε τελική ανάλυση, ενώ η ύπαρξη πολλαπλών ετικετών μπορεί να βελτιώσει την κατανόηση του περιεχομένου ενός anime, δεν εγγυάται υψηλότερη βαθμολογία. Η ποιότητα και η ελκυστικότητα του ίδιου του anime παραμένουν οι κύριοι καθοριστικοί παράγοντες της ικανοποίησης και των αξιολογήσεων των χρηστών.

2.3 Ποιες είναι οι πιο δημοφιλείς ετικέτες στις anime παραγωγές

2.3.1 Query

Το παρακάτω NoSQL query στοχεύει στην ανάλυση των ετικετών/tags που σχετίζονται με τις παραγωγές anime. Υπολογίζει τον αριθμό των παραγωγών anime όπου εμφανίζεται κάθε ετικέτα και ταξινομεί τα αποτελέσματα με βάση τον αριθμό των anime σε φθίνουσα σειρά. Αναλυτικά:

\$addFields στάδιο: Αυτό το στάδιο προσθέτει ένα νέο πεδίο που ονομάζεται "myTags" σε κάθε document διαχωρίζοντας το πεδίο "Tags" χρησιμοποιώντας κόμματα.

\$addFields στάδιο (2ο): Αυτό το στάδιο προσθέτει ένα άλλο νέο πεδίο που ονομάζεται "myTrimTags" αντιστοιχίζοντας πάνω από τον πίνακα "myTags", περικόπτοντας κάθε tag για να αφαιρέσετε τυχόν κενά στην αρχή ή στο τέλος (για να μην υπάρξουν διπλοεγγραφές).

\$unwind stage: Αυτό το στάδιο αποδομεί/"σπάει" τον πίνακα "myTrimTags", δημιουργώντας ένα ξεχωριστό document για κάθε tag εντός του πίνακα.

\$group stage: Αυτό το στάδιο ομαδοποιεί τα έγγραφα με βάση τα trimmed tags, υπολογίζοντας τον αριθμό των παραγωγών anime για κάθε tag.

\$project stage: Αυτό το στάδιο αναδιαμορφώνει τα documents, αντικαθιστώντας το πεδίο "id" σε "myTag". Διατηρεί το πεδίο "animeCount" που αντιπροσωπεύει τον αριθμό των παραγωγών anime.

\$sort stage: Αυτό το στάδιο ταξινομεί τα documents κατά το πεδίο "animeCount" με φθίνουσα σειρά, υποδεικνύοντας πρώτα τα tags με τον υψηλότερο αριθμό παραγωγών anime.

```
1 db.getCollection(  
2   'Anime Statistics and Metadata'  
3 ).aggregate(  
4   [  
5     {  
6       $addFields: {  
7         myTags: { $split: ['$Tags', ','] }  
8       }  
9     },
```

```
10 {
11   $addFields: {
12     myTrimTags: {
13       $map: {
14         input: '$myTags',
15         as: 'item',
16         in: { $trim: { input: '$$item' } }
17       }
18     }
19   }
20 },
21 {
22   $unwind: {
23     path: '$myTrimTags',
24     preserveNullAndEmptyArrays: false
25   }
26 },
27 {
28   $group: {
29     _id: '$myTrimTags',
30     animeCount: { $sum: 1 }
31   }
32 },
33 {
34   $project: {
35     _id: 0,
36     myTag: '$_id',
37     animeCount: 1
38   }
39 },
40 { $sort: { animeCount: -1 } }
41 ],
42 { maxTimeMS: 60000, allowDiskUse: true }
43 );
```

Αλγόριθμος 3: Query για ερώτηση 2.3

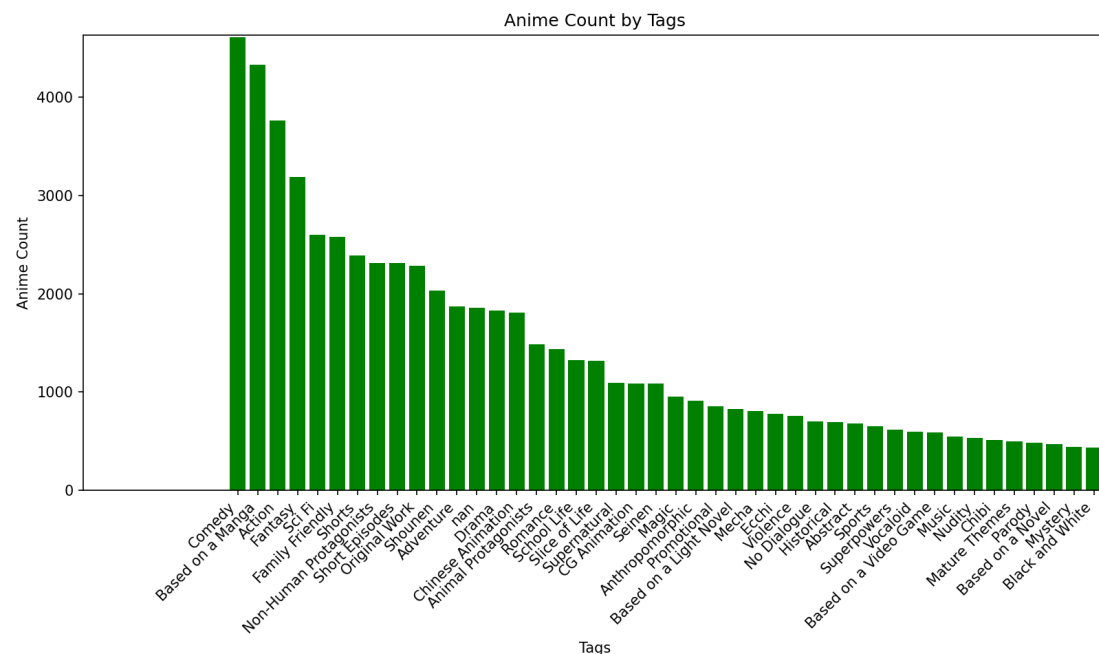
2.3.2 Αποτελέσματα

Παρακάτω παραθέτουμε ένα πίνακα με τα 20 πρώτα αποτελέσματα που λάβαμε με αυτό το query (τα ολοκληρωμένα αποτελέσματα βρίσκονται στο αρχείο "Question2_3"):

Table 3: Anime Tags Statistics

Αριθμός Anime	Tag
4614	Comedy
4328	Based on a Manga
3761	Action
3190	Fantasy
2601	Sci Fi
2580	Family Friendly
2387	Shorts
2315	Non-Human Protagonists
2314	Short Episodes
2282	Original Work
2031	Shounen
1868	Adventure
1859	
1826	Drama
1809	Chinese Animation
1484	Animal Protagonists
1438	Romance
1321	School Life
1317	Slice of Life
1095	Supernatural

2.3.3 Γραφική απεικόνιση και Συμπεράσματα



Εικόνα 3: Πλήθος άνιμε ανά ετικέτα (τμήμα)

Από την ανάλυση των ετικετών anime φαίνεται ότι δημοφιλή είδη όπως η κωμωδία, η δράση, η φαντασία και η επιστημονική φαντασία κυριαρχούν, μαζί με συχνές adaptations από manga, ελαφριά μυθιστορήματα και βιντεοπαιχνίδια. Τα anime που στοχεύουν σε συγκεκριμένο κοινό είναι εμφανή με ετικέτες όπως shounen, seinen και shoujo, ενώ τα θέματα εκτείνονται από τη μαγεία και τον ρομαντισμό έως το υπερφυσικό και το κομμάτι της ζωής, καλύπτοντας ένα ευρύ φάσμα ενδιαφερόντων των θεατών. Τα content warning tags τονίζουν τους διάφορους βαθμούς ωριμότητας και καταλληλότητας, ενώ η ποικιλομορφία των χαρακτήρων αντιπροσωπεύεται μέσω μη ανθρωπίνων πρωταγωνιστών, ανθρωπόμορφων όντων και ζώων. Η βιομηχανία παρουσιάζει διαφορετικά στυλ παραγωγής, από κινούμενα σχέδια CG έως σύντομα επεισόδια, και εξερευνά εξειδικευμένα ενδιαφέροντα με ετικέτες όπως isekai, mecha και τρόμου. Ωστόσο παρατηρούμε ότι όσο πιο συγκεκριμένο είναι το tag τόσο λιγότερες εμφανίσεις έχει, καθώς δεν μπορεί να προσαρμοστεί και να ταιριάζει με

πολλά έργα. Οι πολιτιστικές επιρροές είναι εμφανείς με αναφορές στην ιαπωνική μυθολογία, τη φεουδαρχική Ιαπωνία και τα διεθνή περιβάλλοντα.

2.4 Ποια στρατηγική ακολουθούν οι εταιρείες παραγωγής anime

2.4.1 Query

Το παρακάτω NoSQL query στοχεύει να βρει τον μέσο αριθμό επεισοδίων και τη μέση αξιολόγηση των παραγωγών anime που παράγονται ανά studio παραγωγής. Αναλυτικά:

\$group stage: Αυτό το στάδιο ομαδοποιεί documents με βάση το πεδίο "Studio", υπολογίζει τον μέσο αριθμό επεισοδίων που παράγονται ανά studio (χρησιμοποιώντας τον τελεστή \$avg στο πεδίο "Episodes") και τον μέσο όρο των αξιολογήσεων χρηστών (χρησιμοποιώντας τον τελεστή συγκέντρωσης avg στο πεδίο "Rating"). \$project stage: Αυτό το στάδιο αναδιαμορφώνει τα documents, ώστε να προβάλλονται μόνο τα πεδία "Studio", "avgEpisodes" και "avgRating" (που αντιπροσωπεύουν αντίστοιχα τον μέσο αριθμό επεισοδίων και τον μέσο όρο αξιολογήσεων).

```
1 db.getCollection(  
2   'Anime Statistics and Metadata'  
3 ).aggregate(  
4   [  
5     {  
6       $group: {  
7         _id: '$Studio',  
8         avgEpisodes: { $avg: '$Episodes' },  
9         avgRating: { $avg: '$Rating' }  
10      }  
11    },  
12    {  
13      $project: {  
14        Studio: 1,  
15        avgEpisodes: 1,  
16        avgRating: 1  
17      }  
18    }  
19  ]  
20 )
```

```

18     }
19   ],
20   { maxTimeMS: 60000, allowDiskUse: true }
21 );

```

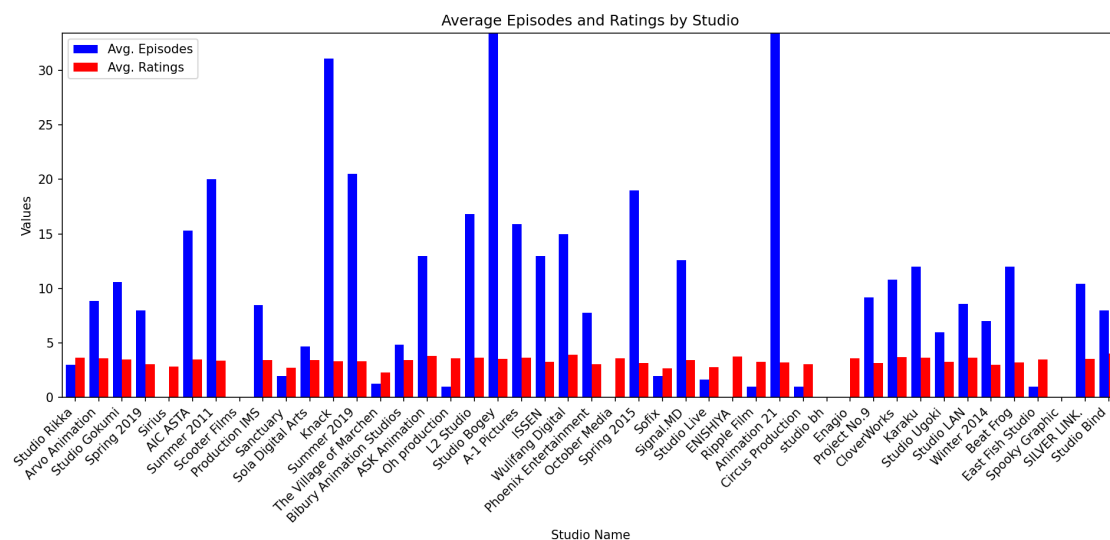
Αλγόριθμος 4: Query για ερώτηση 2.4

2.4.2 Αποτελέσματα

Παρακάτω παραθέτουμε ένα πίνακα με τα 20 πρώτα αποτελέσματα που λάβαμε με αυτό το query (τα ολοκληρωμένα αποτελέσματα βρίσκονται στο αρχείο "Question2_4"):

Studio	MO Επεισοδίων	MO Rating
Chaos Project	5	3.16
Wolf Smoke	1	3.38
Hal Film Maker	16.27	3.52
Takahashi Studio	2	3.21
Romanov Films	5	2.45
Original Dream	23.44	3.17
Studio Rikka	3	3.63
Arvo Animation	8.83	3.57
Studio Gokumi	10.6	3.46
Spring 2019	8	3.05
Sirius		2.83
AIC ASTA	15.29	3.51
Summer 2011	20	3.36
Scooter Films		
Production IMS	8.47	3.46
Sanctuary	2	2.74
Sola Digital Arts	4.66	3.43
Knack	31.08	3.3
Summer 2019	20.5	3.31

2.4.3 Γραφική απεικόνιση και Συμπεράσματα



Εικόνα 4: MO επεισοδίων και αξιολογήσεων ανά στούντιο (τμήμα)

Με βάση την ανάλυση δεδομένων, φαίνεται ότι ενώ ορισμένα στούντιο κλίνουν προς μικρότερες σειρές, άλλα επιλέγουν μεγαλύτερες, αντανακλώντας μια ποικιλία από στρατηγικές. Παραδόξως, δεν υπάρχει σαφής συσχέτιση μεταξύ του αριθμού των επεισοδίων και της μέσης βαθμολογίας, γεγονός που υποδηλώνει ότι η ικανοποίηση των θεατών μπορεί να βασίζεται περισσότερο στην ποιότητα παρά στην ποσότητα. Τα στούντιο που προτιμούν λιγότερα επεισόδια συχνά δίνουν προτεραιότητα στην ποιότητα έναντι της ποσότητας, με αποτέλεσμα υψηλότερες βαθμολογίες λόγω της πιο αυστηρής αφήγησης και του καλύτερου ρυθμού. Αυτό δείχνει μια πιθανή προτίμηση των θεατών για μικρότερες σειρές που δίνουν έμφαση στην αφοσίωση και την ποιότητα. Ωστόσο, η παρουσία μακροχρόνιων σειρών με πιστούς θαυμαστές υπογραμμίζει τη σημασία της αφοσίωσης του κοινού και της εξειδίκευσης στο στούντιο στον καθορισμό της επιτυχίας μιας σειράς. Αν και τα στούντιο τείνουν να παράγουν λίγα επεισόδια κατά μέσο όρο, υπάρχουν εξαιρέσεις όπου τα στούντιο που παράγουν περισσότερα επεισόδια λαμβάνουν επίσης υψηλές βαθμολογίες.

2.5 Ποιος σχεδιαστής έχει δημιουργήσει τις περισσότερες anime παραγωγές

2.5.1 Query

Το παρακάτω NoSQL query στοχεύει στην ανάλυση των συνεισφορών των σχεδιαστών anime. Υπολογίζει τον αριθμό των anime που έχει δουλέψει κάθε σχεδιαστής χαρακτήρων, τα στούντιο παραγωγής με τα οποία έχουν συνεργαστεί και τη μέση αξιολόγηση χρηστών ανά στούντιο παραγωγής (και ανά σχεδιαστή).

\$addField stage: Αυτό το στάδιο προσθέτει ένα νέο πεδίο που ονομάζεται "staffArray" σε κάθε document, διαχωρίζοντας το πεδίο "staff" χρησιμοποιώντας κόμματα, καθώς το "staff" περιέχει πληροφορίες σχετικά με το προσωπικό που συμμετέχει σε κάθε παραγωγή anime, άρα και τους character designers.

\$addField stage (2o): Αυτό το στάδιο προσθέτει περαιτέρω ένα νέο πεδίο που ονομάζεται "characterDesigners" φιλτράροντας το "staffArray" ώστε να περιλαμβάνει μόνο τα άτομα στο staff που εργάζονται στο "Character Design".

\$addField stage (3o): Αυτό το στάδιο προσθέτει ένα άλλο νέο πεδίο που ονομάζεται "characterDesignerName" εξάγοντας τα ονόματα των σχεδιαστών από τον πίνακα "characterDesigners". Αυτό το κάνει διαχωρίζοντας το "characterDesigners" στο σημείο ":" και κρατώντας το πρώτο μέρος (καθώς πριν από το ":" -θέση 0 στον splitted πίνακα- βρίσκεται το όνομα του σχεδιαστή, και μετά από το ":" ήταν το string "Character Design").

\$unwind stage: Αυτό το στάδιο αποδομεί/"σπάει" τον πίνακα "characterDesignerName", δημιουργώντας ένα ξεχωριστό document για κάθε σχεδιαστή.

\$group stage (1o): Αυτό το στάδιο ομαδοποιεί έγγραφα με βάση το όνομα του σχεδιαστή και το στούντιο παραγωγής, υπολογίζοντας τον μέσο όρο αξιολογήσεων και μετρώντας τον αριθμό των παραγωγών anime στις οποίες έχει δουλέψει κάθε σχεδιαστής για κάθε στούντιο.

\$group stage (2o): Αυτό το στάδιο ομαδοποιεί έγγραφα με βάση το όνομα του σχεδιαστή μόνο, ωθώντας (push) το "αντικείμενο" {στούντιο παραγωγής, μέση βαθμολογία και τον αριθμό anime} σε έναν πίνακα που ονομάζεται "mystudios" [2].

Υπολογίζει επίσης τον συνολικό αριθμό των παραγωγών anime του σχεδιαστή, και τον αριθμό των στούντιο παραγωγής με τα οποία έχει δουλέψει κάθε σχεδιαστής.

\$addField stage (4o): Αυτό το στάδιο προσθέτει νέα πεδία και αναδιαμορφώνει τα documents, συμπεριλαμβανομένων μόνο των πεδίων που είναι απαραίτητα για την τελική έξοδο, και μετονομάζει ορισμένα πεδία για λόγους σαφήνειας.

\$project stage: Αυτό το στάδιο αναδιαμορφώνει τα documents, συμπεριλαμβανομένων μόνο των πεδίων σε σωστή σειρά.

\$sort stage: Αυτό το στάδιο ταξινομεί τα documents με βάση τον συνολικό αριθμό των παραγωγών anime με φθίνουσα σειρά.

```
1 db.getCollection(  
2   'Anime Statistics and Metadata'  
3 ).aggregate(  
4   [  
5     {  
6       $addField: {  
7         staffArray: { $split: ['$staff', ','] }  
8       }  
9     },  
10    {  
11      $addField: {  
12        characterDesigners: {  
13          $filter: {  
14            input: '$staffArray',  
15            as: 'item',  
16            cond: {  
17              $regexMatch: {  
18                input: '$$item',  
19                regex: 'Character Design'  
20              }  
21            }  
22          }  
23        }  
24      }  
25    },  
26    {  
27      $addField: {  
28        characterDesignerName: {  
29          $map: {  
30            input: '$characterDesigners',
```

```

31         as: 'designer',
32         in: {
33             $trim: {
34                 input: {
35                     $arrayElemAt: [
36                         {
37                             $split: ['$${designer}', ':']
38                         },
39                         0
40                     ]
41                 }
42             }
43         }
44     }
45 }
46 },
47 {
48     {
49         $unwind: { path: '$characterDesignerName' }
50     },
51     {
52         $group: {
53             _id: {
54                 designer: '$characterDesignerName',
55                 mystudio: '$Studio'
56             },
57             avgRating: { $avg: '$Rating' },
58             animeCount: { $sum: 1 }
59         }
60     },
61     {
62         $group: {
63             _id: '$_id.designer',
64             mystudios: {
65                 $push: {
66                     mystudio: '$_id.mystudio',
67                     avgRating: '$avgRating',
68                     animeCount: '$animeCount'
69                 }

```

```
70     },
71     totalAnimeCount: { $sum: '$animeCount' },
72     studioCount: { $sum: 1 }
73   }
74 },
75 {
76   $addFields: {
77     countTotalAnime: '$totalAnimeCount',
78     countTotalStudio: '$studioCount',
79     myGroup: {
80       addToSet: '$mystudios.mystudio'
81     },
82     myGroupRatings: {
83       addToSet: '$mystudios.avgRating'
84     }
85   }
86 },
87 {
88   $project: {
89     _id: 1,
90     countTotalAnime: 1,
91     myGroup: 1,
92     myGroupRatings: 1,
93     countTotalStudio: 1
94   }
95 },
96 { $sort: { countTotalAnime: -1 } }
97 ],
98 { maxTimeMS: 60000, allowDiskUse: true }
99 );
```

Αλγόριθμος 5: Query για ερώτηση 2.5

2.5.2 Αποτελέσματα

Παρακάτω παραθέτουμε ένα πίνακα με τα 20 πρώτα αποτελέσματα που λάβαμε με αυτό το query (τα ολοκληρωμένα αποτελέσματα βρίσκονται στο αρχείο "Question2_5"):

Η ανάλυση δεδομένων αποκαλύπτει τον Akio Watanabe ως τον τοπ σχεδιαστή χαρακτήρων, μη έχοντας δημιουργήσει τα περισσότερα anime, συνολικά 59+, και έχοντας δουλέψει με τον μεγαλύτερο αριθμό στούντιο, συνολικά 18. Τα έργα του Takahiro Kishida (όντας σε 2η θέση ως προς το πλήθος των στούντιο που έχει δουλέψει) ωστόσο λαμβάνουν την υψηλότερη μέση βαθμολογία στα διάφορα στούντιο, κυμαινόμενος πάντα μεταξύ 3.5 και 4.5, υποδεικνύοντας και τα δύο παραγωγικότητα και ποιότητα στα σχέδιά τους. Ωστόσο, η ευελιξία είναι εμφανής μεταξύ άλλων σχεδιαστών όπως ο Yoshihiko Umakoshi, ο οποίος, παρά το γεγονός ότι δημιούργησε λιγότερα anime, δούλεψε επίσης με πολλά στούντιο. Είναι ενδιαφέρον ότι οι βαθμολογίες για τα έργα κάθε σχεδιαστή ποικίλλουν μεταξύ των στούντιο, υποδηλώνοντας ότι η δυναμική συνεργασίας μεταξύ σχεδιαστών και στούντιο επηρεάζει σημαντικά την ποιότητα του τελικού προϊόντος. Ενώ οι σταθερές υψηλές βαθμολογίες του Kishida υποδηλώνουν ισχυρή φήμη και σύνολο δεξιοτήτων, τα δεδομένα υπογραμμίζουν την περίπλοκη αλληλεπίδραση παραγόντων που συμβάλλουν στην επιτυχία των σχεδιαστών χαρακτήρων anime.

2.6 Σχετίζεται η δημοτικότητα μιας παραγωγής με το πλήθος των κύκλων που θα παραχθούν

2.6.1 Query

Το παρακάτω NoSQL query στοχεύει να ομαδοποιήσει σχετικές εγγραφές anime σε ένα franchise και να αναλύσει διάφορες πτυχές τους, συμπεριλαμβανομένου του αριθμού των anime σε κάθε ομάδα, του μέσου αριθμού επεισοδίων, της χρονικής περιόδου κατά την οποία εμφανίστηκαν αυτά τα anime, και τη μέση αξιολόγηση που έλαβαν από χρήστες. Οι ομάδες/franchises και τα δεδομένα τους προκύπτουν από το συνδυασμό των πεδίων Name και Related_anime για κάθε εγγραφή, οπότε κάποια franchises περιέχουν παρόμοια anime. Αναλυτικά:

\$addFields stage: Αυτό το στάδιο προσθέτει ένα νέο πεδίο που ονομάζεται "relatedAnimeArray" σε κάθε έγγραφο διαχωρίζοντας το πεδίο "Related_anime" χρησιμοποιώντας κόμματα.

\$addFields stage (2o): Αυτό το στάδιο προσθέτει περαιτέρω ένα νέο πεδίο που ονομάζεται "relatedAnime" ,περικόποντας κάθε στοιχείο του "relatedAnimeAr-

ray” για να αφαιρέσετε τυχόν κενά στην αρχή ή στο τέλος (για να μην υπάρξουν διπλοεγγραφές).

\$unwind stage: Αυτό το στάδιο αποδομεί/”σπάει” τον πίνακα ”relatedAnime”, δημιουργώντας ένα ξεχωριστό document για κάθε σχετική καταχώρηση anime.

\$group stage: Αυτό το στάδιο ομαδοποιεί τα έγγραφα με βάση τον σχετικό τίτλο του anime ”relatedAnime” (προσπαθώντας να συνδυάσει όσους πιο πολλούς σχετικούς τίτλους anime μπορεί για franchise), προωθώντας σχετικές πληροφορίες (τίτλος franchise, αριθμός επεισοδίων, αριθμός βαθμολογιών, έτος έναρξης, έτος λήξης) σε έναν πίνακα μδε objects που ονομάζεται ”animeData”. Στο τέλος της ομαδοποίησης προκύπτουν παραπάνω από ένα document για το ίδιο franchise, ωστόσο κάποιο από αυτά περιέχει το πλήρες franchise (και λογικά στις ψηλότερες θέσεις, καθώς θα έχει τον πλήρη/maximum αριθμό anime παραγωγών για το franchise).

\$addFields stage (3ο): Αυτό το στάδιο προσθέτει επιπλέον πεδία:

- ”countAnime”: υπολογίζει το μέγεθος του πίνακα ”animeData”, αντιπροσωπεύοντας τον αριθμό των anime σε κάθε ομάδα/franchise.
- ”myGroup”: δημιουργεί ένα σύνολο μοναδικών τίτλων anime σε κάθε ομάδα (για να μαζέψει μόνο όλους τους τίτλους μαζί, χωρίς τα επιπλέον στοιχεία/πληροφορίες του αντικειμένου).
- ”avgEpisodes”: υπολογίζει τον μέσο αριθμό επεισοδίων σε όλα τα anime σε κάθε ομάδα, περικλύποντας το αποτέλεσμα σε 3 δεκαδικά ψηφία.
- ”AvgRating”: υπολογίζει τον μέσο όρο αξιολογήσεων/Rating σε όλα τα anime σε κάθε ομάδα, περικλύποντας το αποτέλεσμα σε 3 δεκαδικά ψηφία.
- ”duration”: υπολογίζει τη διάρκεια (μάξιμουμ έτος τέλους - μίνιμουμ έτος έναρξης) της σειράς anime σε κάθε ομάδα.

\$project stage: Αυτό το στάδιο αναδιαμορφώνει τα documents, συμπεριλαμβανομένων μόνο των πεδίων που είναι απαραίτητα για την τελική έξοδο, και μετονομάζει ορισμένα πεδία για λόγους σαφήνειας.

\$sort stage: Αυτό το στάδιο ταξινομεί τα documents με βάση τον αριθμό των anime σε κάθε ομάδα με φθίνουσα σειρά.

```
1 db.collection(  
2   'Anime Statistics and Metadata'  
3 ).aggregate(  
4   [  
5     {  
6       $addFields: {  
7         relatedAnimeArray: {  
8           $split: ['$Related_anime', ',',']  
9         }  
10      }  
11    },  
12    {  
13      $addFields: {  
14        relatedAnime: {  
15          $map: {  
16            input: '$relatedAnimeArray',  
17            as: 'item',  
18            in: { $trim: { input: '$$item' } }  
19          }  
20        }  
21      }  
22    },  
23    { $unwind: { path: '$relatedAnime' } },  
24    {  
25      $group: {  
26        _id: '$relatedAnime',  
27        animeData: {  
28          $push: {  
29            title: '$Name',  
30            numEpisodes: '$Episodes',  
31            numRating: '$Rating',  
32            startYear: '$Release_year',  
33            endYear: '$End_year'  
34          }  
35        }  
36      }  
37    },  
38    {  
39      $addFields: {
```

```
40     countAnime: { $size: '$animeData' },
41     myGroup: { addToSet: '$animeData.title' },
42     avgEpisodes: {
43       $trunc: [
44         { $avg: '$animeData.numEpisodes' },
45         3
46       ]
47     },
48     avgRating: {
49       $trunc: [
50         { $avg: '$animeData.numRating' },
51         3
52       ]
53     },
54     duration: {
55       $subtract: [
56         { $max: '$animeData.endYear' },
57         { $min: '$animeData.startYear' }
58       ]
59     }
60   }
61 },
62 {
63   $project: {
64     _id: 1,
65     myGroup: 1,
66     countAnime: 1,
67     avgEpisodes: 1,
68     avgRating: 1,
69     maxMinDuration: '$duration'
70   }
71 },
72 { $sort: { countAnime: -1 } }
73 ],
74 { maxTimeMS: 60000, allowDiskUse: true }
75 );
```

Αλγόριθμος 6: Query για ερώτηση 2.6

2.6.2 Αποτελέσματα

Παρακάτω παραθέτουμε ένα πίνακα με τα 20 πρώτα αποτελέσματα που λάβαμε με αυτό το query (τα ολοκληρωμένα αποτελέσματα βρίσκονται στο αρχείο "Question2_6"):

Table 6: Στοιχεία από Franchises

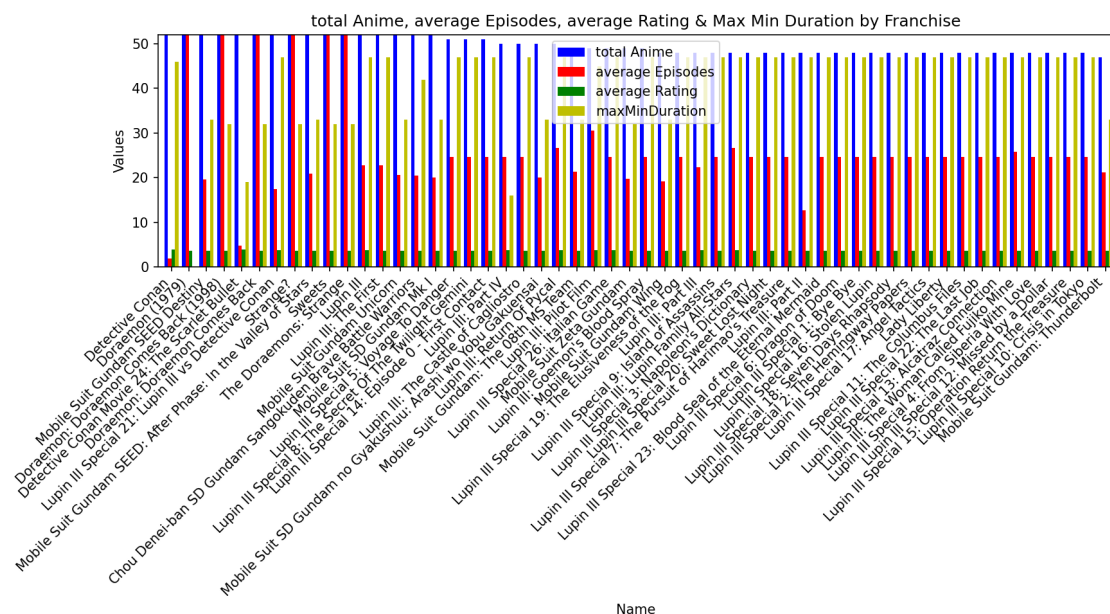
Franchise	animeCount	Group	avgEpisodes	avgRating	Duration
Detective Conan	74	Detective Conan: Episode One - The Great Detective Turned Small,Detective Conan Movie 20: The Darkest Nightmare, ..., Detective Conan: Zero's Tea Time,Detective Conan Movie 25: The Bride of Halloween,Detective Conan: Tantei-tachi no Starry Night	1.827	3.865	46
Doraemon (1979)	64	Stand by Me Doraemon,Doraemon: Nobita no Shin Kyouryuu,Doraemon Movie 31: Shin Nobita to Tetsujin Heidan - Habatake Tenshi-tachi,Doraemon Movie 38: Nobita's Treasure Island,Doraemon (2005),Doraemon Movie 37: Nobita no Nankyoku Kachi Kochi Daibouken, ...,Doraemon: Boku no Umareta Hi,"Doraemon: Boku, Momotarou no Nanna no Sa",Dorami-chan: MiniDora SOS!!!,The Doraemons: Goal! Goal! Goal!,The Doraemons: Mushimushi Pyonpyon Daisakusen!	194.666	3.567	-
Doraemon: Doraemon Comes Back (1998)	60	Stand by Me Doraemon,Doraemon: Nobita no Shin Kyouryuu,Doraemon Movie 31: Shin Nobita to Tetsujin Heidan - Habatake Tenshi-tachi, ...,Dorami-chan: Hello Kyouryuu Kids!!,Doraemon: Boku no Umareta Hi,"Doraemon: Boku, Momotarou no Nanna no Sa",Dorami-chan: MiniDora SOS!!!,The Doraemons: Goal! Goal! Goal!,The Doraemons: Mushimushi Pyonpyon Daisakusen!	253	3.561	32
Doraemon: Doraemon Comes Back	60	Stand by Me Doraemon,Doraemon: Nobita no Shin Kyouryuu,Doraemon Movie 31: Shin Nobita to Tetsujin Heidan - Habatake Tenshi-tachi, ...,Dorami-chan: Hello Kyouryuu Kids!!,Doraemon: Boku no Umareta Hi,"Doraemon: Boku, Momotarou no Nanna no Sa",Dorami-chan: MiniDora SOS!!!,The Doraemons: Goal! Goal! Goal!,The Doraemons: Mushimushi Pyonpyon Daisakusen!	253	3.562	32

Franchise	animeCount	Group	avgEpisodes	avgRating	Duration
Detective Conan Movie 24: The Scarlet Bullet	60	Detective Conan: Episode One - The Great Detective Turned Small,Detective Conan Movie 20: The Darkest Nightmare,Detective Conan Movie 6: The Phantom of Baker Street, ...,Detective Conan: A Written Challenge from the Pyramids!,Detective Conan: The Internet - The Mysterious E-mail Case,Detective Conan: Chase the Mysterious Comet Mon- ster!,Detective Conan vs. Wooo	4.814	3.899	19
Mobile Suit Gun- dam SEED Des- tiny	60	Mobile Suit Gundam: Iron-Blooded Orphans 2,Mobile Suit Gundam 00 Second Season,Mobile Suit Gun- dam: The Origin,Mobile Suit Gun- dam 00, ...,Mobile Suit Gundam ZZ: Gundam Frag.,Chou Denei-ban SD Gundam Sangokuden Brave Bat- tle Warriors,Mobile Suit Gundam 0083: Stardust Memory - The Mayfly of Space,All That Gun- dam,Mobile Suit Gundam: Twi- light Axis, Ring of Gundam,Mobile Suit Gundam SEED PROJECT Ig- nited	19.562	3.578	33
Lupin III Special 21: Lupin III vs Detective Conan	59	Lupin III: The First,Lupin III vs Detective Conan: The Movie,Lupin III: Part IV,Lupin III: The Cas- tle of Cagliostro,Detective Co- nan, ...,Lupin III: Return Of Pycal,Monkey Majik: Lupin the Third,Detective Conan vs. Wooo,Lupin III: Pilot Film,Lupin III: Pilot Film,Lupin VIII: Arsene and Cie	17.5	3.692	47
The Doraemons: Strange	57	Stand by Me Doraemon,Doraemon: Nobita no Shin Kyouryuu,Doraemon Movie 31: Shin Nobita to Tetsujin Heidan - Habatake Tenshi-tachi,Doraemon Movie 38: Nobita's Treasure Is- land,Doraemon (2005), ...,Doraemon: Boku no Umareta Hi,"Doraemon: Boku, Momotarou no Nanna no Sa",Dorami-chan: MiniDora SOS!!!,The Doraemons: Goal! Goal! Goal!,The Doraemons: Mushimushi Pyonpyon Daisakusen!	190.5	3.558	32
Strange?	57	Stand by Me Doraemon & Doraemon: Nobita no Shin Kyouryuu,Doraemon Movie 31: Shin Nobita to Tetsujin Heidan - Habatake Tenshi-tachi,Doraemon Movie 38: Nobita's Treasure Island,Doraemon (2005), ...,Doraemon Eigo Kyoushitsu,Dorami-chan: Hello Kyouryuu Kids!!,Doraemon: Boku no Umareta Hi,"Doraemon: Boku, Momotarou no Nanna no Sa",Dorami-chan: MiniDora SOS!!!,The Doraemons: Goal! Goal! Goal!,The Doraemons: Mushimushi Pyonpyon Daisakusen!	190.5	3.558	32

Franchise	animeCount	Group	avgEpisodes	avgRating	Duration
Mobile Suit Gundam SEED: After Phase: In the Valley of Stars	57	Mobile Suit Gundam: Iron-Blooded Orphans 2, Mobile Suit Gundam 00 Second Season, ..., Chou Deneiban SD Gundam Sangokuden Brave Battle Warriors, Mobile Suit Gundam 0083: Stardust Memory - The Mayfly of Space, All That Gundam, Ring of Gundam, Mobile Suit Gundam SEED PROJECT Ignited	20.937	3.609	33
Sweets	57	Stand by Me Doraemon, Doraemon: Nobita no Shin Kyouryuu, Doraemon Movie 31: Shin Nobita to Tetsujin Heidan - Habatake Tenshi-tachi, ..., Dorami-chan: Hello Kyouryuu Kids!!, Doraemon: Boku no Umareta Hi, "Doraemon: Boku, Momotarou no Nanna no Sa", Dorami-chan: MiniDora SOS!!!, The Doraemons: Goal! Goal! Goal!, The Doraemons: Mushimushi Pyonpyon Daisakusen!	190.5	3.558	32
Lupin III	54	Lupin III: Part V, Lupin III: The First, Lupin III vs Detective Conan: The Movie, Lupin III: Part IV, Lupin III: The Castle of Cagliostro, ..., Lupin III: The Mystery of Mamo, Lupin III: Secret Files, Lupin III: Return Of Pycal, Monkey Majik: Lupin the Third, Lupin III 3DCG, Lupin III: Pilot Film, Lupin III: Pilot Film, Lupin VIII: Arsene and Cie	22.75	3.689	47
Lupin III: The First	54	Lupin III: Part V, Lupin III vs Detective Conan: The Movie, Lupin III: Part IV, Lupin III: The Castle of Cagliostro, Lupin III Special 21: Lupin III vs Detective Conan, ..., Lupin III: The Mystery of Mamo, Lupin III: Secret Files, Lupin III: Return Of Pycal, Monkey Majik: Lupin the Third, Lupin III 3DCG, Lupin III: Pilot Film, Lupin III: Pilot Film, Lupin VIII: Arsene and Cie	22.769	3.68	47
Mobile Suit SD Gundam Mk I	52	Mobile Suit Gundam: Iron-Blooded Orphans 2, Mobile Suit Gundam 00 Second Season, Mobile Suit Gundam: The Origin, Mobile Suit Gundam 00, ..., Mobile Suit Gundam 0083: Stardust Memory - The Mayfly of Space, All That Gundam, Ring of Gundam, Mobile Suit SD Gundam The Movie: Musha Knight Commando: SD Gundam Scramble, Mobile Suit SD Gundam Festival	19.97	3.567	33

Franchise	animeCount	Group	avgEpisodes	avgRating	Duration
Mobile Suit Gundam Unicorn	52	Mobile Suit Gundam: Iron-Blooded Orphans 2, Mobile Suit Gundam 00 Second Season, Mobile Suit Gundam: The Origin, Mobile Suit Gundam 00, ..., Chou Denei-ban SD Gundam Sangokuden Brave Battle Warriors, Mobile Suit Gundam 0083: Stardust Memory - The Mayfly of Space, All That Gundam, Mobile Suit Gundam: Twilight Axis, Ring of Gundam	20.548	3.548	33
Chou Denei-ban SD Gundam Sangokuden Brave Battle Warriors	52	Mobile Suit Gundam: Iron-Blooded Orphans 2, Mobile Suit Gundam 00 Second Season, ..., Mobile Suit SD Gundam Mk V, SD Gundam World: Sangoku Souketsuden, Mobile Suit Gundam 0083: Stardust Memory - The Mayfly of Space, All That Gundam, Ring of Gundam	20.432	3.559	42
Lupin III Special 8: The Secret Of The Twilight Gemini	51	Lupin III: Part V, Lupin III: The First, Lupin III: Part IV, Lupin III: The Castle of Cagliostro, ..., Lupin III: Secret Files, Lupin III: Return Of Pycal, Monkey Majik: Lupin the Third, Lupin III 3DCG, Lupin III: Pilot Film, Lupin III: Pilot Film, Lupin VIII: Arsene and Cie	24.583	3.68	47
Lupin III Special 14: Episode 0 - First Contact	51	Lupin III: Part V, Lupin III: The First, Lupin III: Part IV, Lupin III: The Castle of Cagliostro, ..., Lupin III: Secret Files, Lupin III: Return Of Pycal, Monkey Majik: Lupin the Third, Lupin III 3DCG, Lupin III: Pilot Film, Lupin III: Pilot Film, Lupin VIII: Arsene and Cie	24.583	3.676	47
Lupin III Special 5: Voyage To Danger	51	Lupin III: Part V, Lupin III: The First, Lupin III: Part IV, Lupin III: The Castle of Cagliostro, Lupin III Special 21: Lupin III vs Detective Conan, ..., Lupin III: Return Of Pycal, Monkey Majik: Lupin the Third, Lupin III 3DCG, Lupin III: Pilot Film, Lupin III: Pilot Film, Lupin VIII: Arsene and Cie	24.583	3.679	47
Mobile Suit SD Gundam no Gyakushuu: Arashi wo Yobu Gakuensai	50	Mobile Suit Gundam: Iron-Blooded Orphans 2, Mobile Suit Gundam 00 Second Season, ..., Mobile Suit Gundam 0083: Stardust Memory - The Mayfly of Space, All That Gundam, Ring of Gundam	19.97	3.563	33

2.6.3 Γραφική απεικόνιση και Συμπεράσματα



Εικόνα 6: Στατιστικά των franchises (τμήμα)

Το σύνολο δεδομένων περιλαμβάνει μια ποικιλία από franchise anime, που παρουσιάζουν διακυμάνσεις στον αριθμό των anime σε κάθε franchise, τον μέσο αριθμό επεισοδίων, τις μέσες βαθμολογίες και τα έτη διάρκειας. Διάφορα μοτίβα προκύπτουν κατά την ανάλυση: τα franchise με πολυάριθμες συμμετοχές anime τείνουν να έχουν μεγαλύτερο διάστημα ετών, πιθανότατα υποδηλώνοντας διαρκή δημοτικότητα ή συνεχή παραγωγή. Επιπλέον, τα franchise με υψηλότερη μέση βαθμολογία μπορεί να υποδεικνύουν σταθερή ποιότητα σε όλες τις παραγωγές τους, συμβάλλοντας ενδεχομένως στη μακροζωία τους. Αντίθετα, τα franchise με χαμηλότερες βαθμολογίες ενδέχεται να έχουν μικρότερη διάρκεια ή λιγότερες συμμετοχές, πιθανώς λόγω της μειωμένης απήχησης του κοινού. Ο μέσος αριθμός επεισοδίων μπορεί επίσης να αντικατοπτρίζει το βάθος αφήγησης, με τις σειρές μεγαλύτερης διάρκειας να εξερευνούν πιθανώς πιο περίπλοκες αφηγήσεις. Επιπλέον, η παρουσία ειδικών, ταινιών και spin-off εντός των franchise υποδηλώνει μια προσπάθεια επέκτασης του σύμπαντος ή αξιοποίησης της δημοτικότητας των κεντρικών χαρακτήρων ή των ιστοριών.

References

- [1] G. Script, “How to install mongodb 7 on windows 10/11 [2023 update] mongodb compass.” <https://www.youtube.com/watch?v=h2x6BmUi5zQ&t=308s>, 2023. Πρόσβαση στις 10-02-2024.
- [2] Stackoverflow, “Mongodb group values by multiple fields.” <https://stackoverflow.com/questions/22932364/mongodb-group-values-by-multiple-fields>, 2014. Πρόσβαση στις 10-02-2024.