

Assignment 3

ชุดข้อมูล movies เป็น คอลเลกชันที่มีรายละเอียดเกี่ยวกับภาพยนตร์ document แต่ละฉบับประกอบด้วยภาพยนตร์ 1 เรื่องและข้อมูลต่าง ๆ ที่เกี่ยวข้องดังแสดงในตาราง

ลำดับ	ฟิลด์ (fields)	คำอธิบาย
1.	_id	หมายเลข object
2.	plot	เนื้อเรื่องแบบย่อ
3.	genres	แนวภาพยนตร์
4.	runtime	ระยะเวลาฉาย (นาที)
5.	rated	เรทของภาพยนตร์
6.	cast	รายชื่อนักแสดง
7.	poster	โปสเตอร์
8.	title	ชื่อภาพยนตร์
9.	fullplot	เนื้อเรื่องเต็ม
10.	countries	ประเทศ
11.	released	วันที่ออกฉาย
12.	directors	ผู้กำกับ
13.	writers	ผู้เขียนบท
14.	awards	รางวัล
15.	lastupdated	อัปเดตล่าสุดเมื่อ
16.	year	ปีที่ฉาย
17.	imdb	รายละเอียดจาก IMDb ประกอบด้วย rating, votes, id
18.	type	ประเภท
19.	tomatoes	รายละเอียดจาก Rotten Tomatoes
20.	num_mflix_comments	จำนวน comment

ให้นักศึกษาแสดงคำสั่งสำหรับดำเนินการกับ document ในฐานข้อมูล MongoDB ดังนี้

1. เชื่อมต่อ database ชื่อ Mflix

```
from pymongo import MongoClient  
client = MongoClient()  
db = client["Mflix"]
```

2. คำสั่งแสดงประเภทของภาพยนตร์

```
db.movies.distinct('type')
```

โดยคำสั่งดังกล่าวให้ผลลัพธ์ดังนี้

```
['movie', 'series']
```

3. คำสั่งเพื่อแสดงตัวอย่าง document ที่สามารถรับชมภาพยนตร์ได้มากกว่า 2 ภาษา

```
db.movies.find_one({"languages.2":{"$exists":True}})
```

4. คำสั่งเพื่อนับจำนวน document ที่สามารถรับชมภาพยนตร์ได้มากกว่า 2 ภาษา

```
db.movies.count_documents({"languages.2":{"$exists":True}})
```

โดยคำสั่งดังกล่าวให้คำตอบ 2102 document

5. คำสั่งแสดงภาษา (languages) ทั้งหมดที่มี

```
db.movies.distinct("languages")
```

โดยคำสั่งดังกล่าวให้ผลลัพธ์ดังนี้

```
[' Ancient (to 1453)',  
 ' Old',  
 'Abkhazian',  
 'Aboriginal',  
 'Acholi',  
 'Afrikaans',  
 'Aidoukrou',  
 'Albanian',  
 'Algonquin',  
 'American Sign Language',  
 'Amharic',  
 'Apache languages',  
 'Arabic',  
 'Aramaic',  
 'Arapaho',  
 'Armenian',  
 'Assamese',  
 'Assyrian Neo-Aramaic',  
 'Athapascan languages',  
 'Awadhi',  
 'Aymara',  
 'Azerbaijani',  
 'Balinese',  
 'Bambara',  
 'Basque',  
 'Belarusian',  
 'Bengali',
```

'Berber languages',
'Bhojpuri',
'Bosnian',
'Brazilian Sign Language',
'Breton',
'British Sign Language',
'Bulgarian',
'Burmese',
'Cantonese',
'Catalan',
'Chechen',
'Cheyenne',
'Chinese',
'Cornish',
'Corsican',
'Cree',
'Creole',
'Creoles and pidgins',
'Croatian',
'Crow',
'Czech',
'Danish',
'Dari',
'Dinka',
'Dutch',
'Dyula',
'Dzongkha',
'Eastern Frisian',
'Egyptian (Ancient)',
'English',
'Esperanto',
'Estonian',
'Ewe',
'Faroese',
'Filipino',
'Finnish',
'Flemish',
'French',
'French Sign Language',
'Frisian',
'Fulah',
'Fur',
'Gallegan',
'Georgian',
'German',
'German Sign Language',
'Greek',
'Greenlandic',
'Guarani',
'Gujarati',
'Gumatj',
'Haitian',
'Hakka',
'Hassanya',
'Hawaiian',

'Hebrew',
'Hindi',
'Hmong',
'Hokkien',
'Hungarian',
'Ibo',
'Icelandic',
'Indian Sign Language',
'Indonesian',
'Inuktitut',
'Inupiaq',
'Irish',
'Italian',
'Japanese',
'Japanese Sign Language',
'Jola-Fonyi',
'Kabuverdianu',
'Kabyle',
'Kannada',
'Karajè',
'Kazakh',
'Khmer',
'Kikuyu',
'Kinyarwanda',
'Kirghiz',
'Klingon',
'Konkani',
'Korean',
'Korean Sign Language',
'Kuna',
'Kurdish',
'Ladakhi',
'Ladino',
'Lao',
'Latin',
'Latvian',
'Lingala',
'Lithuanian',
'Low German',
'Luxembourgish',
'Macedonian',
'Malay',
'Malayalam',
'Maltese',
'Mandarin',
'Manipuri',
'Maori',
'Mapudungun',
'Marathi',
'Mari',
'Masai',
'Maya',
'Mende',
'Middle English',
'Min Nan',

'Mohawk',
'Mongolian',
'More',
'Nahuatl',
'Nama',
'Navajo',
'Neapolitan',
'Nenets',
'Nepali',
'Norse',
'North American Indian',
'Norwegian',
'Nyaneka',
'Nyanja',
'Occitan',
'Old English',
'Oriya',
'Panjabi',
'Pawnee',
'Persian',
'Peul',
'Polish',
'Portuguese',
'Purepecha',
'Pushto',
'Quechua',
'Quenya',
'Rajasthani',
'Romanian',
'Romany',
'Russian',
'Russian Sign Language',
'Ryukyuan',
'Saami',
'Samoan',
'Sanskrit',
'Sardinian',
'Scanian',
'Scots',
'Scottish Gaelic',
'Serbian',
'Serbo-Croatian',
'Shanghainese',
'Shanxi',
'Shona',
'Shoshoni',
'Sicilian',
'Sign Languages',
'Sindarin',
'Sinhalese',
'Sioux',
'Slovak',
'Slovenian',
'Somali',
'Songhay',

```
'Southern Sotho',
'Spanish',
'Spanish Sign Language',
'Swahili',
'Swedish',
'Swiss German',
'Syriac',
'Tagalog',
'Tajik',
'Tamil',
'Tarahumara',
'Tatar',
'Telugu',
'Thai',
'Tibetan',
'Tigrigna',
'Tok Pisin',
'Tonga',
'Tswana',
'Tulu',
'Tupi',
'Turkish',
'Turkmen',
'Tuvinian',
'Tzotzil',
'Uighur',
'Ukrainian',
'Ungwatsi',
'Urdu',
'Uzbek',
'Vietnamese',
'Visayan',
'Washoe',
'Welsh',
'Wolof',
'Xhosa',
'Yiddish',
'Yoruba',
'Zulu']
```

6. คำสั่งสำหรับนับจำนวนภาษาทั้งหมดที่มีใน ข้อ 4.

```
len(db.movies.distinct("languages"))
```

โดยคำสั่งดังกล่าวให้ผลลัพธ์ดังนี้

231

7. คำสั่งนับจำนวนภาพยนตร์ ที่มีภาษาไทย

```
db.movies.count_documents({"languages":"Thai"})
```

โดยคำสั่งดังกล่าวให้ผลลัพธ์ดังนี้

95

8. คำสั่งแสดงแนวภาพยนตร์ (genre) ของภาพยนตร์

```
db.movies.distinct("genres")
```

โดยคำสั่งดังกล่าวให้ผลลัพธ์ดังนี้

```
['Action',
 'Adventure',
 'Animation',
 'Biography',
 'Comedy',
 'Crime',
 'Documentary',
 'Drama',
 'Family',
 'Fantasy',
 'Film-Noir',
 'History',
 'Horror',
 'Music',
 'Musical',
 'Mystery',
 'News',
 'Romance',
 'Sci-Fi',
 'Short',
 'Sport',
 'Talk-Show',
 'Thriller',
 'War',
 'Western']
```

9. คำสั่งนับจำนวนภาพยนตร์ที่สามารถรับชมได้ทั้งภาษาไทยและอังกฤษ (ใช้ \$all)

```
db.movies.count_documents({"languages":{"$all":["Thai','English']}})
```

โดยคำสั่งดังกล่าวให้ผลลัพธ์ดังนี้

```
42
```

10. คำสั่งพิมพ์แสดงเฉพาะรายชื่อภาพยนตร์ในข้อ 9.

```
result=db.movies.find({"languages":{"$all":["Thai','English']}})
```

```
for i in result:
```

```
    print(str(i['year'])+' '+i['title'])
```

โดยคำสั่งดังกล่าวให้ผลลัพธ์ดังนี้

```
1956 The King and I
1957 The Bridge on the River Kwai
1974 The Man with the Golden Gun
1992 Police Story 3: Supercop
2000 The Beach
```

1999	Anna and the King
2000	Meet the Parents
2001	Corky Romano
2004	Bridget Jones: The Edge of Reason
2002	The Eye
2002	3 Extremes II
2004	Two Brothers
2002	Infernal Affairs
2003	Last Life in the Universe
2003	Ong-bak
2003	The Tesseract
2004	The Eye 2
2004	The Eye 2
2004	Citizen Dog
2008	Rambo
2006	Invisible Waves
2005	Typhoon
2006	Khan kluay
2008	The Amazing Truth About Queen Raquela
2008	Bangkok Dangerous
2007	Protège
2006	13: Game of Death
2007	Ploy
2007	Hotel Very Welcome
2009	Mammoth
2008	Vinyan
2008	Batman: Gotham Knight
2008	Chocolate
2008	Yami no kodomo-tachi
2011	The Burma Conspiracy
2011	The Hangover Part II
2013	Only God Forgives
2012	The Impossible
2014	The Mule
2012	Countdown
2013	The White Storm
2015	How to Win at Checkers (Every Time)

11. คำสั่งดึงข้อมูลภาพยนตร์ทั้งหมด ที่มีคะแนนผู้ชม (rating) สูงกว่า 7 บน imdb

```
rate=db.movies.find({"imdb.rating":{"$gt":7}})
for i in rate:
    print(str(i['imdb']['rating'])+' '+i['title'])
```

12. คำสั่งดึงข้อมูลภาพยนตร์ทั้งหมด ที่มีคะแนนผู้ชม (rating) สูงกว่า 4 บน Tomatoes

```
rat=db.movies.find({"tomatoes.viewer.rating":{"$gt":4}})
for i in rat:
    print(str(i['tomatoes']['viewer']['rating'])+' '+i['title'])
```


13. คำสั่งดึงข้อมูลภาพยนตร์ทั้งหมด ที่ได้รับรางวัล (award)

```
award_win=db.movies.find({"awards.wins":{"$gt":0}})

for j in award_win:

    print(str(j['awards']['wins'])+' '+j['title'])
```

14. คำสั่งค้นหาภาพยนตร์ทั้งหมดพร้อมชื่อภาพยนตร์ ภาษา ออกฉาย ผู้กำกับ นักเขียน รางวัล ปี ประเภท ระยะเวลาฉาย นักแสดง และประเทศ ที่มีการเสนอชื่อเข้าชิงอย่างน้อยหนึ่งรายการ (awards.nominations)

```
cursor2 = db.movies.find({"awards.nominations" : {"$gt":0}},
{"title":1,'awards':1,'runtime':1,'languages':1,'released':1,'directors':1,'writers':1,'countries':1,'type':1,'year':1})

for x in cursor2:

    print(str(x['awards']['nominations'])+' '+x['title']+' '+str(x['year'])+' ' +str(x['type']))
```

15. คำสั่งดึงข้อมูลภาพยนตร์ทั้งหมดพร้อมชื่อเรื่อง ภาษา วันเวลาที่ออกฉาย ผู้กำกับ นักเขียน รางวัล ปี ประเภท ระยะเวลาการฉาย นักแสดง ประเทศ นำแสดงโดย " Focus Jirakul"

```
cast=db.movies.find(
    {"cast":'Focus Jirakul'},
{"title":1,'languages':1,'award':1,'runtime':1,'year':1,'type':1,'released':1,'directors':1,'writers':1,'countries':1})

for i in cast:

    print(i)
```

16. คำสั่งดึงข้อมูลภาพยนตร์ทั้งหมดพร้อมชื่อเรื่อง ภาษา วันเวลาที่ออกฉาย ผู้กำกับ นักเขียน และประเทศ ที่มีคำว่า "scene" อยู่ในชื่อเรื่อง

```
db.movies.find_one(
    {"title":{"$regex":"scene",$options:'i'}},
    {"title":1,'languages':1,'released':1,'directors':1,'writers':1,'countries':1}
)
```

17. คำสั่งค้นหาภาพยนตร์ทั้งหมดพร้อมชื่อเรื่อง ภาษา วันเวลาที่ออกฉาย ผู้กำกับ ผู้ชม นักเขียน และประเทศ ที่มี rating ผู้ชมอย่างน้อย 3 และน้อยกว่า 4 บน Tomatoes

```
rating=db.movies.find(
    {"tomatoes.viewer.rating":{"$gte":3,"$lt":4}},
    {"title":1,'languages':1,'released':1,'directors':1,'writers':1,'countries':1}
)
for x in rating:
    print(x)
```

18. คำสั่งค้นหาภาพยนตร์ทั้งหมดพร้อมชื่อเรื่อง ภาษา วันเวลาที่ออกฉาย ผู้กำกับ นักเขียน และประเทศ ที่มีความยาว (runtime) ระหว่าง 60 ถึง 90 นาที

```
runtimes = db.movies.find(
    {"runtime":{"$gte":60,"$lte":90}},
    {"title":1,'languages':1,'released':1,'directors':1,'writers':1,'countries':1}
)
for x in runtimes:
    print(x)
```

19. คำสั่งค้นหาภาพยนตร์ทั้งหมดพร้อมชื่อเรื่อง ภาษา เนื้อเรื่องเต็ม ออกฉาย ผู้กำกับ นักเขียน และประเทศ ที่มีเนื้อเรื่องเต็มซึ่งมีคำว่า "fire"

```
fires = db.movies.find(
    {"fullplot":{"$regex":"fire","$options":"i"}},
    {"title":1,'languages':1,'released':1,'directors':1,'writers':1,'countries':1}
)
for j in fires:
    print(j)
```

คำแนะนำ

คำสั่งแสดงเฉพาะ field ที่ต้องการ

```
cursor = db.movies.find({ "genres": "Drama" },
                        {'_id':0,'title':1})

for movie in cursor:
    print(movie)
```

'_id':0 ไม่เลือกมา ถ้าไม่กำหนด จะนำ _id มาด้วยเสมอ

'title':1 เลือก title มา ส่วน fields อื่น ๆ ที่ไม่ระบุเป็น 1 จะไม่ถูกเรียนมากเช่นกัน

คำสั่งสำหรับแสดงเฉพาะบาง field ที่ต้องการ

```
movies2010=db.movies.find({'year':{'$gt':2010}})

for x in movies2010:
    print(str(x['year'])+'\t' + x['title'] )
```

```
2014 Tillbaka till Bromma
2014 Petals on the Wind
...
2015 Monkey King: Hero Is Back
2015 Much Loved
2015 Babai
```

References

Sample Mflix Dataset - MongoDB Atlas. (n.d.). <https://www.mongodb.com/docs/atlas/sample-data/sample-mflix/>

W3Schools.com. (n.d.). https://www.w3schools.com/python/python_mongodb_getstarted.asp