Anime Analysis

Organize, Discuss, Discover

MyAnimeList

HOW MUCH ANIME HAVE YOU SEEN?

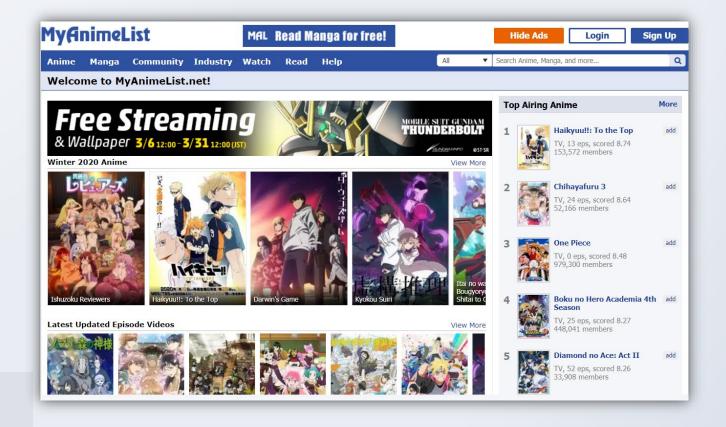


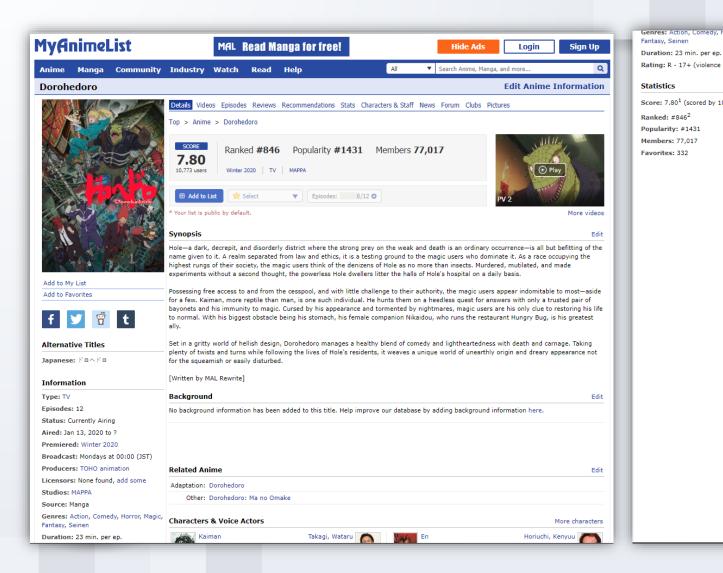


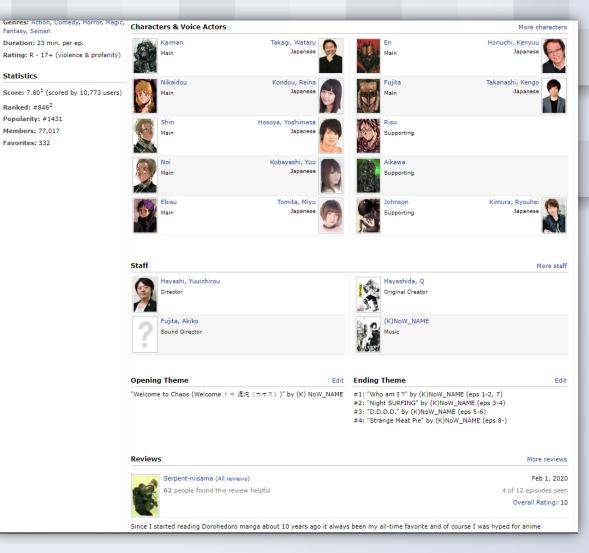
แหล่งที่มาของข้อมูล

แหล่งที่มาของข้อมูล

- Dataset: https://www.kaggle.com/azathoth42/myanimelist#anime_cleaned.csv
- ข้อมูลใน Dataset ถูกเก็บมาจากเว็บไซต์ MyAnimeList (https://myanimelist.net) ซึ่งเป็นเว็บไซต์ที่ทำการเก็บรวบรวมข้อมูลรายละเอียดของอนิเมะแต่ละเรื่อง เช่น ชื่อเรื่อง สตูดิโอผู้สร้าง ฤดูกาลฉาย หมวดหมู่ เพลงประกอบ รายชื่อตัวละคร นักพากย์ ฯลฯ





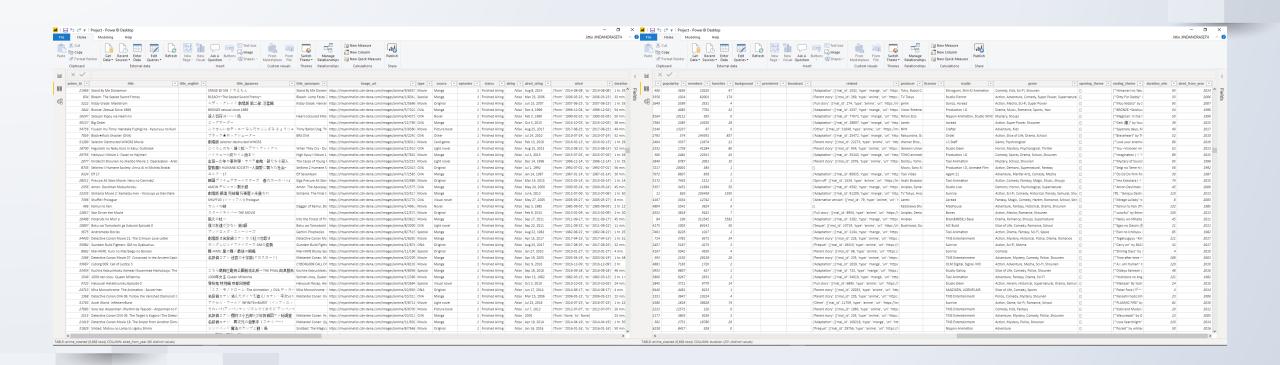




คำอธิบายของข้อมูลและคอลัมน์

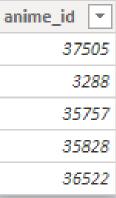
คำอธิบายของข้อมูล

- ไฟล์ anime_cleaned.csv ประกอบด้วยข้อมูลจำนวน 33 คอลัมน์ 6,668 แถว



- anime_id: key ของอนิเมะแต่ละเรื่อง เป็นข้อมูลประเภท Integer ที่สามารถอ้างอิงบน

เว็บไซต์ได้



myanimelist.net/anime/21329/

- title: ชื่อเรื่องของอนิเมะ เป็นข้อมูลประเภท String

- title_english: ชื่อเรื่องอนิเมะเป็นภาษาอังกฤษ เป็นข้อมูลประเภท String และมีข้อมูลว่าง

- title_japanese: ชื่อเรื่องอนิเมะเป็นภาษาญี่ปุ่น เป็นข้อมูลประเภท String และมีข้อมูลว่าง

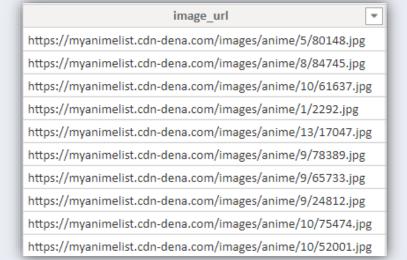
```
In [16]: df.title_japanese.head(10)

Out[16]: 0 妖狐×僕SS
1 瀬戸の花嫁
2 しゅごキャラ!!どきっ
3 プリンセスチュチュ
4 パクマン。
5 夢色パティシエール
6 ウルトラマニアック
7 灼眼のシャナII -Second-
8 のだめカンタービレ 巴里編
9 桜蘭高校ホスト部
Name: title_japanese, dtype: object
```

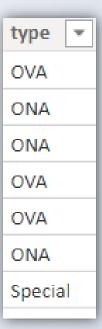
- title_synonyms: ชื่อเรื่องอนิเมะที่เป็นที่รู้จักกันโดยทั่วไป เป็นข้อมูลประเภท String

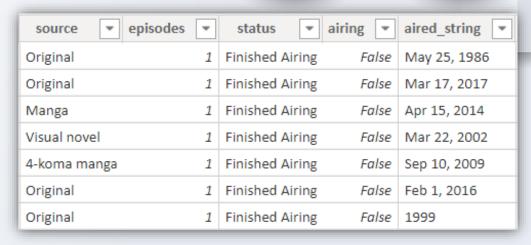
และมีข้อมูลว่าง

- image_url: ที่อยู่ของภาพบนเว็บไซต์ของอนิเมะแต่ละเรื่อง เป็นข้อมูลแบบ String และมีข้อมูลว่าง



- type: ประเภทการฉาย เป็นข้อมูลประเภท String มี 6 ประเภท ดังนี้
 - TV คือ ออกฉายทางโทรทัศน์
 - OVA คือ ไม่ได้มีการฉายออกอากาศทางทีวีหรือในโรงภาพยนตร์ แต่มีจำหน่ายเป็นแผ่น
 - ONA คือ ออกฉายทางอินเตอร์เน็ต
 - Special คือ จัดทำขึ้นมาในวาระพิเศษ
 - Movie คือ ออกฉายในโรงภาพยนตร์
 - Music คือ ออกฉายเป็นวิดีโอเพลง





- source: ต้นฉบับเนื้อหาที่มาของอนิเมะแต่ละเรื่อง เป็นข้อมูลประเภท String
- episodes: จำนวนตอนทั้งหมดของอนิเมะแต่ละเรื่องในเว็บไซต์ เป็นข้อมูลประเภท Integer
- **status:** สถานะการฉายของอนิเมะ เป็นข้อมูลประเภท String
- **airing:** แสดงว่ากำลังฉายอยู่หรือไม่ เป็นข้อมูลประเภท Boolean
- aired_string: ช่วงระยะเวลาที่ฉาย เป็นข้อมูลประเภท String

aired -	duration 🔻	rating	score 🔻	scored_by 🔻	rank 🔻
{'from': '2018-03-24', 'to': '2018-03-24'}	23 min.	R - 17+ (violence & profanity)	6.32	1486	6309
{'from': '2018-05-11', 'to': '2007-01-26'}	7 min. per ep.	PG-13 - Teens 13 or older	6.82	2803	4234
{'from': '2018-01-13', 'to': '2018-02-17'}	24 min. per ep	PG-13 - Teens 13 or older	6.89	7376	3949
{'from': '2018-01-12', 'to': '2018-03-30'}	24 min. per ep	PG-13 - Teens 13 or older	7.52	12413	1578
{'from': '2018-02-23', 'to': '2018-05-11'}	24 min. per ep	PG-13 - Teens 13 or older	7.61	1020	1330

- aired: ช่วงระยะเวลาฉาย เป็นข้อมูลประเภท String
- duration: ความยาวของอนิเมะที่ฉายต่อตอน เป็นข้อมูลประเภท String
- rating: กลุ่มอายุที่เหมาะสมของอนิเมะแต่ละเรื่อง เป็นข้อมูลประเภท String
- score: คะแนนของอนิเมะแต่ละเรื่อง เป็นข้อมูลประเภท Float
- scored_by: จำนวนคนที่ลงคะแนน เป็นข้อมูลประเภท Integer
- rank: อันดับของอนิเมะบนหน้าเว็บไซต์ เป็นข้อมูลประเภท Integer

How are MyAnimeList scores calculated?

All scores given in the database are calculated as a weighted score.

Weighted Score = (v / (v + m)) * S + (m / (v + m)) * C

- S = Average score for the anime/manga
- v = Number users giving a score for the anime/manga †
- m = Minimum number of scored users required to get a calculated score
- C = The mean score across the entire Anime/Manga database

† Note that v does not correspond to the "number of scored users" as seen on the database page. Scores from users who have not viewed 1/5 of the series upon its completion are not included. Scores given from illandimate accounts created to sway votes are also not included in the scoring algorithm.

Not Yet Aired entries have no score and will display N/A. Entries that do not meet the minimum number of scored users will also not display a calculated score.

Ton Anime/Manga Panking

The "Top Upcoming" and "Most Popular" rankings are ordered by the number of users who have added the entry to their list. All other Top Anime and Top Manga rankings are ordered by weighted score, as calcular above. Please note that while R18+ entries calculate a weiohted score, they are excluded from the rankings.

popularity 🔻	members 💌	favorites 💌	background 💌	premiered 🔻	broadcast 💌	related
5120	4442	5				$\label{eq:conditional} \ensuremath{$($'Full story': [\{'mal_id': 34279, 'type': 'anime', 'url': 'https://anime', 'url':$
4224	7233	16				{'Parent story': [{'mal_id': 80, 'type': 'anime', 'url': 'https://
2368	24865	57		Winter 2018	Saturdays at 21	$\label{eq:continuous} \begin{tabular}{ll} \b$
1790	39616	166		Winter 2018	Fridays at 01:58	$ \label{eq:continuous} \begin{tabular}{ll} \$
3861	8970	51		Winter 2018	Saturdays at 21	{'Prequel': [{'mal_id': 33605, 'type': 'anime', 'url': 'https://r

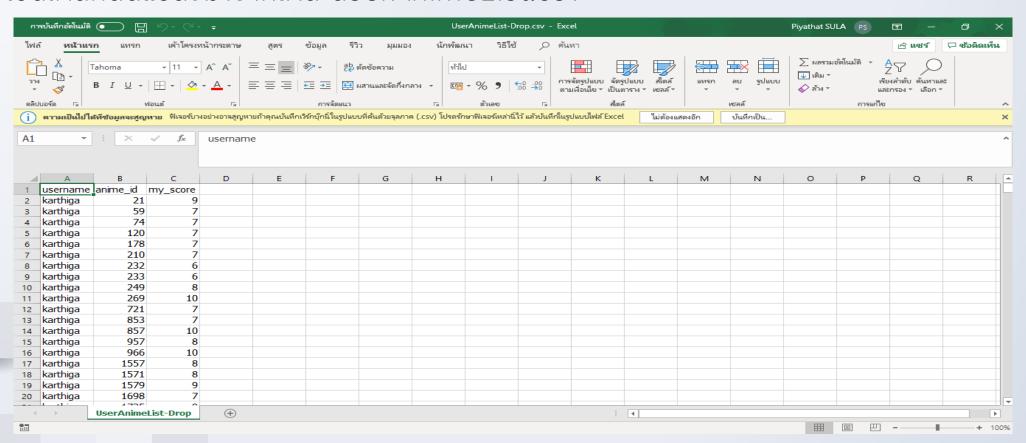
- popularity: อันดับความนิยมบนหน้าเว็บ เป็นข้อมูลประเภท Integer
- **members:** ยอดจำนวนคนดูอนิเมะแต่ละเรื่องในเว็บไซต์ เป็นข้อมูลประเภท Integer
- favorites: จำนวนคนที่ชอบอนิเมะแต่ละเรื่องบนหน้าเว็บไซต์ เป็นข้อมูลประเภท Integer
- **background:** ประวัติความเป็นมา เป็นข้อมูลประเภท String
- **premiered**: ฤดู และ ปีที่ทำการฉาย เป็นข้อมูลประเภท String
- broadcast: วันเวลาออกอากาศ เป็นข้อมูลประเภท String
- **related:** เชื่อมโยงข้อมูล เป็นข้อมูลประเภท String

producer 💌	licensor 💌	studio	genre	opening_theme 🔻	ending_theme 💌	duration_min 🔻	aired_from_year 🔱
		A-1 Pictures	Action, Fantasy	0	0	23	2018
Sotsu	Nozomi Enter	Sunrise	Action, Military, Space, Mecha		['#1 "Black Stars" by	7	2018
Haoliners Anin	Funimation	Emon, Blade	Adventure, Fantasy	["Mamori Tsunagu ([""Kimi Kimi Kimi (君	24	2018
TBS, Grooove,		8bit	Slice of Life, Comedy, Supernatural	["Fushigi na Tabi wa	["Rosetta Stone (🗆	24	2018
Tencent Anima		Haoliners Animation League	Action, Comedy, Magic, Shounen Ai, Supernatural	[""Taion (体温)" by Sh	[""I'll be there" by Ro	24	2018

- producer: ผู้อำนวยการผลิต เป็นข้อมูลประเภท String
- licensor: ผู้ถือลิขสิทธิ์ เป็นข้อมูลประเภท String
- studio: สตูดิโอที่จัดทำ เป็นข้อมูลประเภท String
- genre: ประเภทเนื้อหาของอนิเมะ เป็นข้อมูลประเภท String
- opening_theme: ซื่อเพลงเปิดที่ใช้ประกอบในอนิเมะ เป็นข้อมูลประเภท String
- ending_theme: ชื่อเพลงปิดที่ใช้ประกอบในอนิเมะ เป็นข้อมูลประเภท String
- duration_min: ความยาวของอนิเมะที่ฉายต่อตอน (นาที) เป็นข้อมูลประเภท Float
- aired_from_year: ปีที่เริ่มฉาย เป็นข้อมูลประเภท Integer

คำอธิบายของข้อมูล

- ไฟล์ userAmineList-Drop.csv ประกอบด้วยข้อมูลจำนวน 3 คอลัมน์ 1,048,575 แถว เป็นไฟล์ที่ดัดแปลงมาจากไฟล์ userAnimeList.csv



In [3]:	df.samp	le(10)		
Out[3]:		username	anime_id	my_score
	530083	Asta66	2923	8
	434319	PinkOuji	1241	0
	467604	CeJIDb	33988	0
	861331	jajazinha	880	10
	394429	qetou	33573	0
	353145	grklein	1023	9
	460850	tsubasaR01	1686	0
	71310	Kittyki	2106	0
	172270	Kurokihime	12711	0
	751149	Taiga-Hrk	8425	0

- **anime_id:** key ของอนิเมะแต่ละเรื่อง เป็นข้อมูลประเภท Integer ที่สามารถอ้างอิงบน เว็บไซต์ได้ เช่น <u>https://myanimelist.net/anime/37505/</u>
- **username:** ชื่อสมาชิกผู้ใช้งานในเว็บไซต์ เป็นข้อมูลประเภท string
- my_score: คะแนนที่สมาชิกแต่ละคนลงให้ในแต่ละเรื่อง โดยคะแนนจะมีตั้งแต่ 0 ถึง 10 เป็นข้อมูลประเภท Integer



สาเหตุที่มาของการเลือกข้อมูลชุดนี้

สาเหตุที่มาของการเลือกข้อมูลชุดนี้

- เนื่องจากสมาชิกภายในกลุ่มชอบดูอนิเมะ จึงอยากทำการวิเคราะห์ข้อมูลต่าง ๆ ของอนิเมะ ไม่ว่าจะเป็นความนิยมของผู้ชม ความนิยมของผู้สร้าง ความหลากหลายของเนื้อหา ฯลฯ



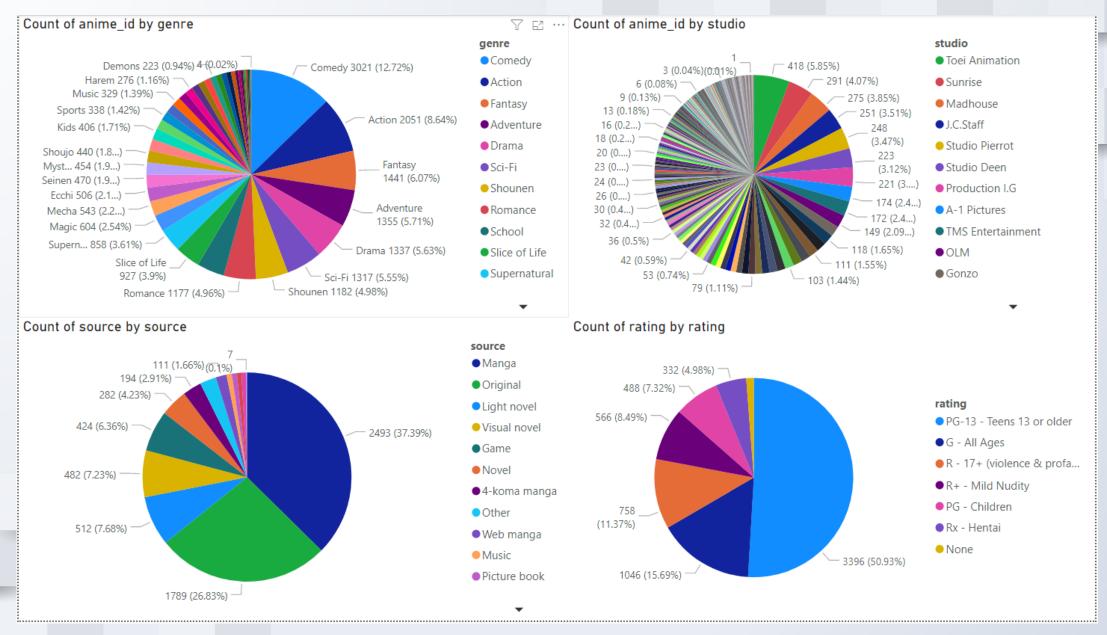




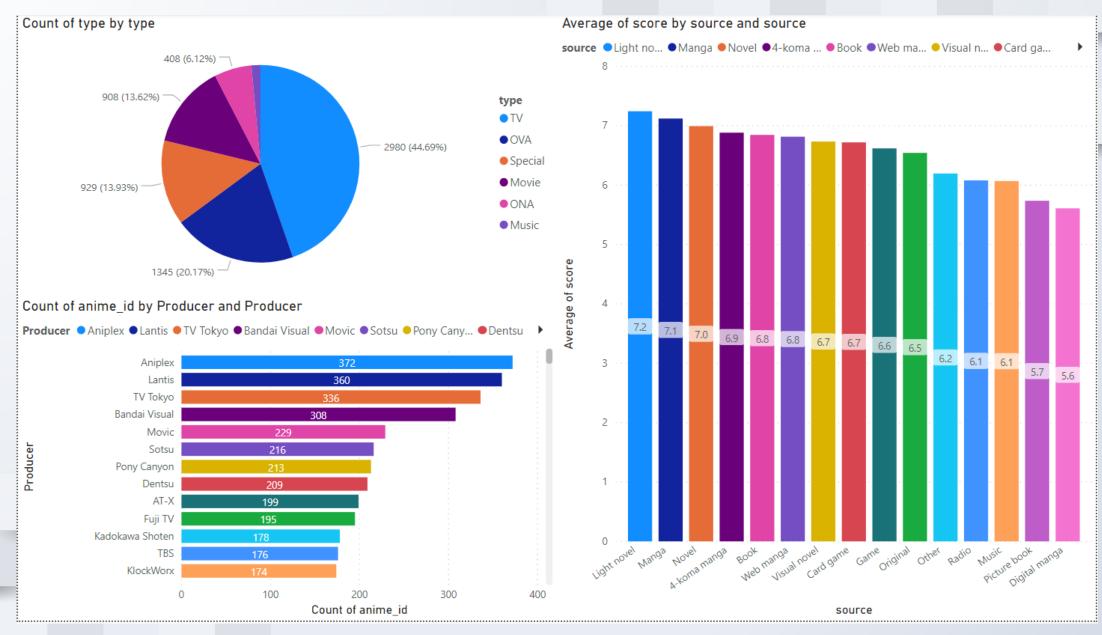


สิ่งที่น่าสนใจจาก การวิเคราะห์ข้อมูลเบื้องต้น

สิ่งที่น่าสนใจจากการวิเคราะห์ข้อมูลเบื้องต้น

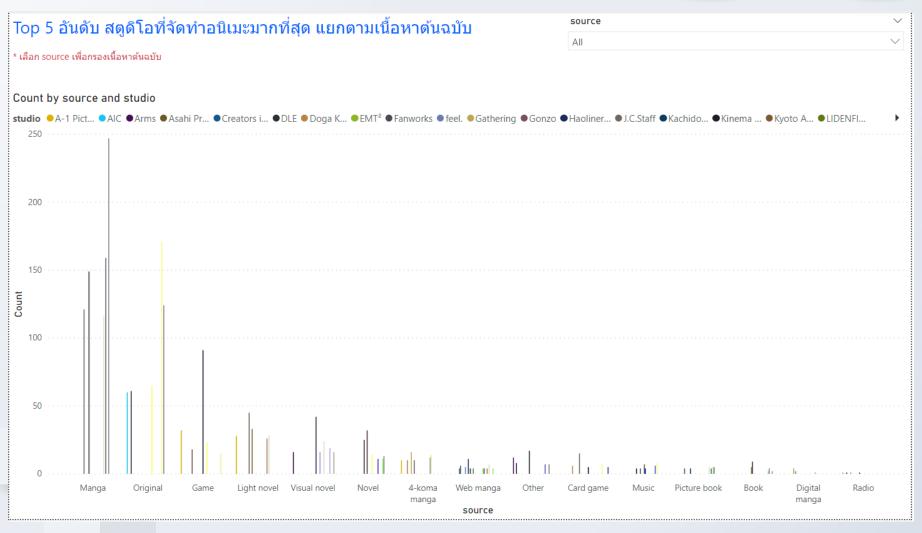


สิ่งที่น่าสนใจจากการวิเคราะห์ข้อมูลเบื้องต้น

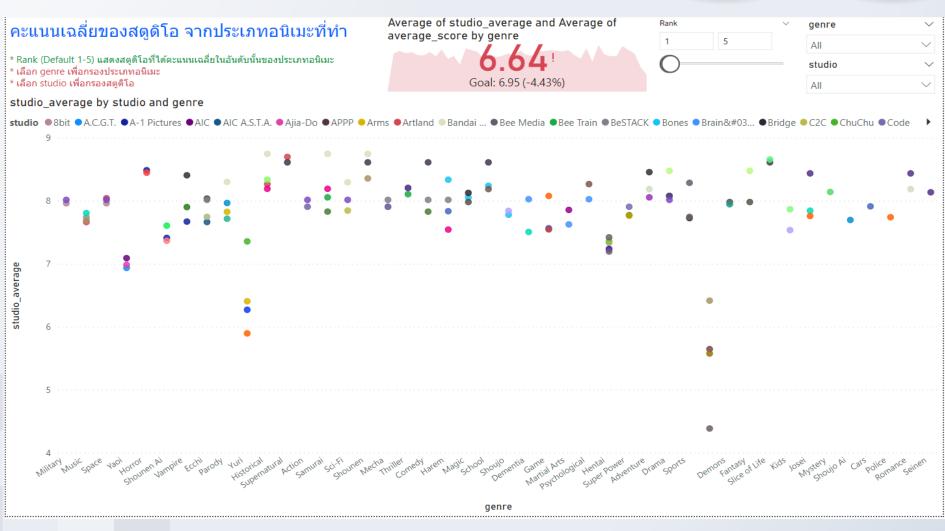




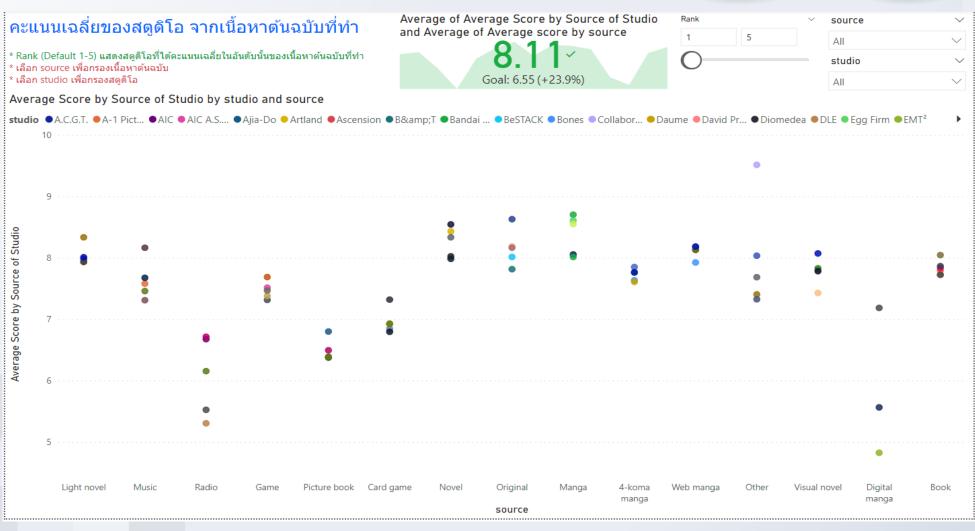
- 1. สตูดิโอจัดทำอะไรเป็นหลัก และผลงานออกมาเป็นอย่างไร
 - 1.1 เนื้อหาต้นฉบับแต่ละประเภท สตูดิโอไหนเป็นผู้จัดทำเป็นส่วนใหญ่



- 1. สตูดิโอจัดทำอะไรเป็นหลัก และผลงานออกมาเป็นอย่างไร
 - 1.2 คะแนนเฉลี่ยของสตูดิโอ จากประเภทอนิเมะที่จัดทำ



- 1. สตูดิโอจัดทำอะไรเป็นหลัก และผลงานออกมาเป็นอย่างไร
 - 1.3 คะแนนเฉลี่ยของสตูดิโอ จากเนื้อหาต้นฉบับที่จัดทำ



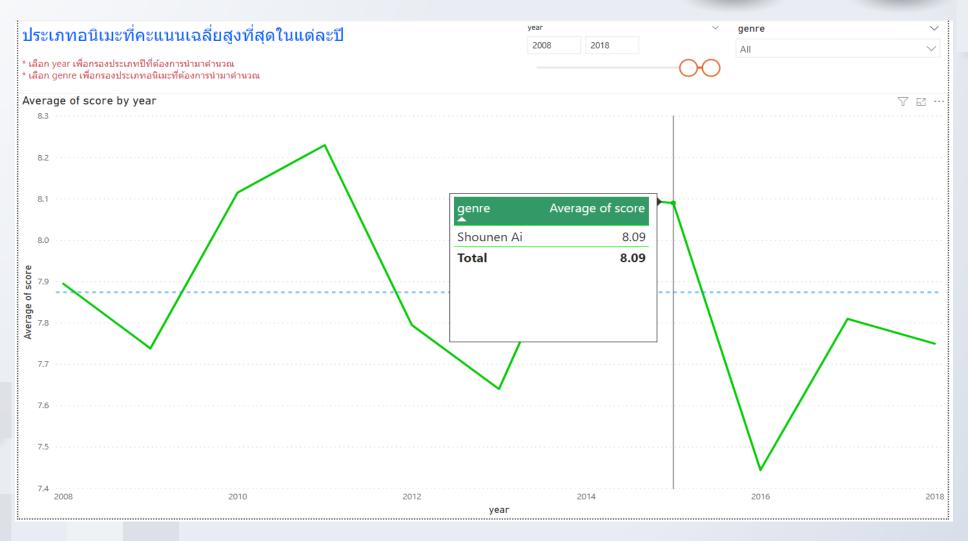
2. รายชื่ออนิเมะ 10 อันดับแรก เรียงตามคะแนนจากมากไปน้อย โดยสมาชิก

เว็บไซต์ที่ชื่นชอบอนิเมะประเภทนั้น ซึ่งมีสัดส่วนการดูอนิเมะประเภทนั้นเป็นหลัก

p 10 a	วันดับ รายชื่ออนิเมะ โดยผู้ชมที่ชื่นชอ	บประเภทอนิเมะ	นั้น			genre		
	ประเภทอนิเมะนั้น หมายถึง user ที่มีอัตราการรับชม genre นั้นเป็นห	ลัก เมื่อเทียบกับ genre อื่น				Fantasy		
เลือก genre เพื่อกรองประเภทอนิเมะที่ต้องการทราบ								
nime_id	title	score from user favorite genre	score from MAL	aired_from_year	rank	genre		
33	Kenpuu Denki Berserk	10.00	8.45	1997	132	Action, Adventure, Demons, Drama, Fantasy, Horror, Military, Romance, Seinen, Supernatural		
459	One Piece Movie 1	10.00	7.22	2000	2740	Action, Adventure, Comedy, Super Power, Fantasy, Shounen		
460	One Piece Movie 2: Nejimaki-jima no Daibouken	10.00	7.28	2001	2500	Action, Adventure, Comedy, Super Power, Fantasy, Shounen		
461	One Piece Movie 3: Chinjuu-jima no Chopper Oukoku	10.00	7.13	2002	3087	Action, Adventure, Comedy, Super Power, Fantasy, Shounen		
463	One Piece Movie 5: Norowareta Seiken	10.00	7.39	2004	2078	Action, Adventure, Comedy, Super Power, Fantasy, Shounen		
464	One Piece Movie 6: Omatsuri Danshaku to Himitsu no Shima	10.00	7.75	2005	998	Adventure, Comedy, Fantasy, Shouner		
465	One Piece Movie 7: Karakuri-jou no Mecha Kyohei	10.00	7.41	2006	1985	Adventure, Comedy, Fantasy, Shouner		
466	One Piece: Taose! Kaizoku Ganzack	10.00	6.83	1998	4207	Action, Adventure, Fantasy, Comedy, Super Power, Shounen		
969	Tsubasa Chronicle 2nd Season	10.00	7.68	2006	1161	Action, Adventure, Fantasy, Romance, Supernatural, Drama, Mystery, Shounen		

3. ประเภทอนิเมะ ปีและฤดูกาล

3.1.1 ประเภทอนิเมะที่คะแนนเฉลี่ยสูงที่สุดในแต่ละปีและฤดูกาล - จำแนกตามปี



3. ประเภทอนิเมะ ปีและฤดูกาล

3.1.2 ประเภทอนิเมะที่คะแนนเฉลี่ยสูงที่สุดในแต่ละปีและฤดูกาล - จำแนกตามปีและฤดู



- 3. ประเภทอนิเมะ ปีและฤดูกาล
 - 3.2 คะแนนเฉลี่ยอนิเมะในแต่ละปีและฤดู และประเภทอนิเมะที่นิยมจัดทำมากสุด

