

An anime-style illustration of a street scene at dusk or dawn. A bright lightning bolt strikes the sky, illuminating the clouds. The street is lined with buildings, utility poles, and cars. The overall mood is dramatic and atmospheric.

# What's Next? : Anime Recommender System

Tham Zhi Cong

# Content

1. Introduction
2. Data Cleaning and EDA
3. Content Based recommender
4. Collaborative Based recommender
5. Conclusion and future steps

# Introduction

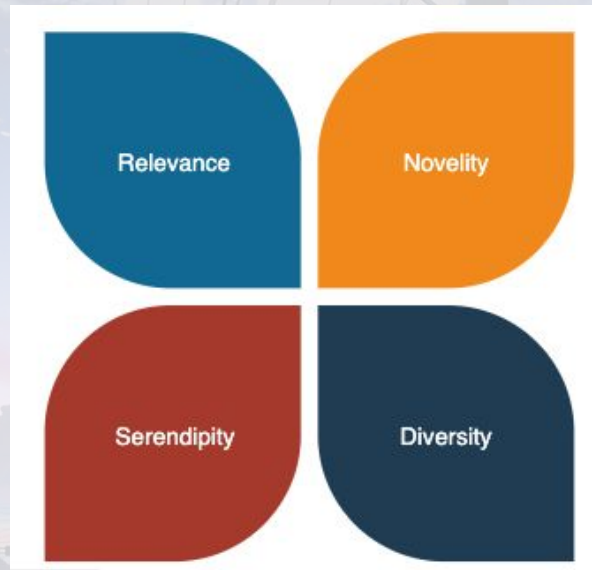
## Problem Statement

To create a recommendation system using an anime dataset and understand how 'good' the recommendations are

Evaluation: How many recommendations are relevant to users?

# Recommender Systems

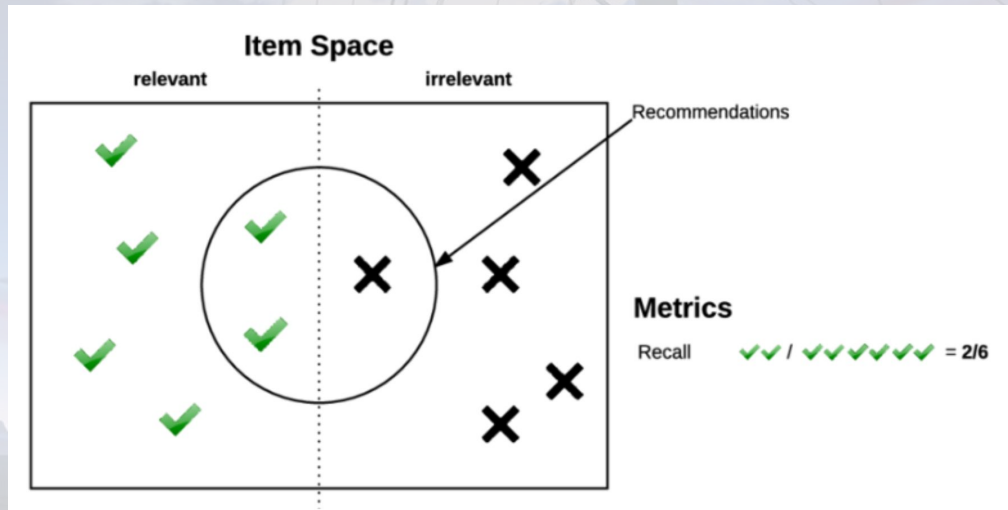
1. Relevance
2. Novelty
3. Serendipity
4. Diversity





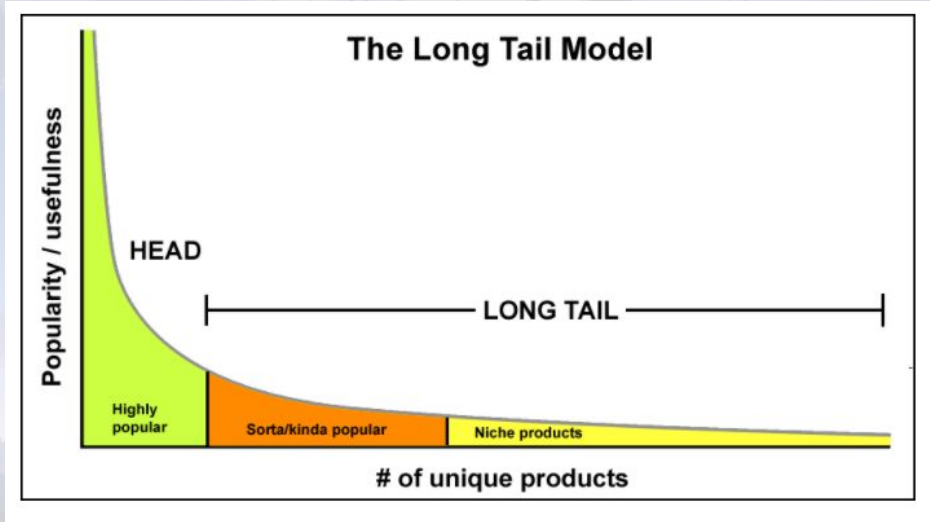
# Relevance

- How many items out of total recommendations does user like
- Similar to 'accuracy'



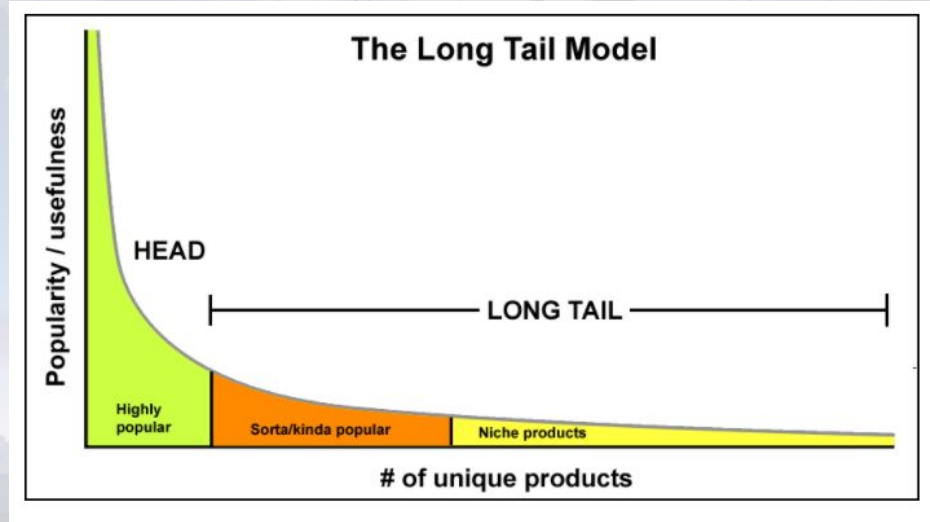
# Novelty

- Recommend what the user has not already been exposed to
- Determines how unknown recommended items are to users
- Measures how new, original or unusual the recommendation is
- Counter to accuracy



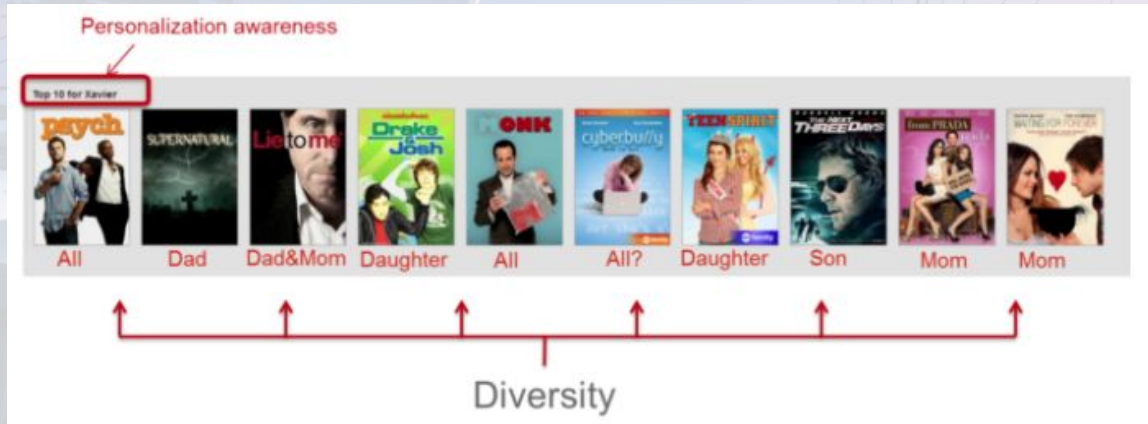
# Serendipity

- Element of 'surprise'
- Recommends beyond what is highly popular
- Solution to 'long tail' and 'cold-start' problem
- Counter to accuracy
- No agreed upon standard of measurement
- A measure of the extent to which recommendations positively surprise their users



# Diversity

- Measures how narrow or wide the spectrum of recommended products are.
- Good diversity indicates wide range of products / genres etc.







# Data Cleaning and EDA

# Data Cleaning and EDA

Datasets: anime (12,294 rows x 7 cols) – to year 2018  
synopsis (16,214 rows x 5 cols)  
rating (7,813,737 rows x 3 cols)

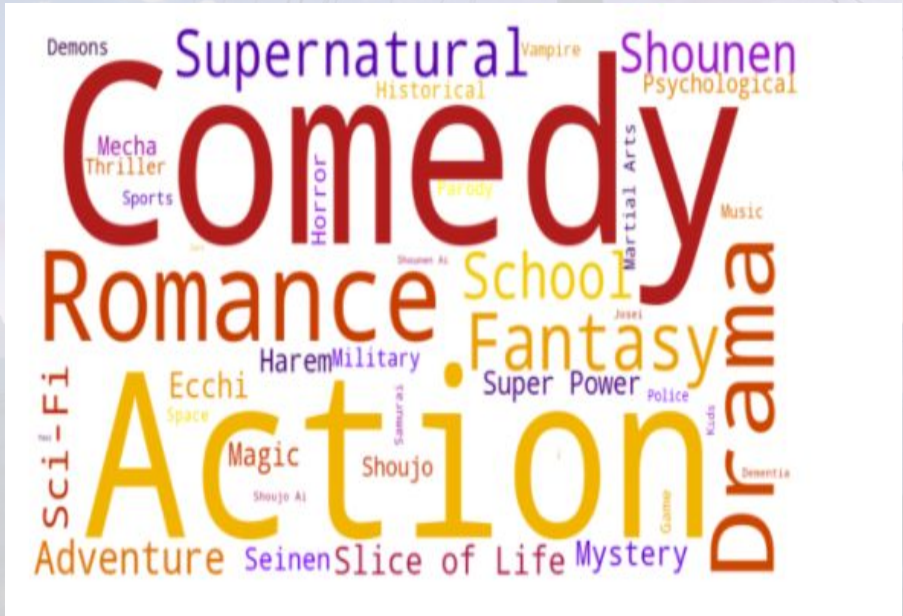
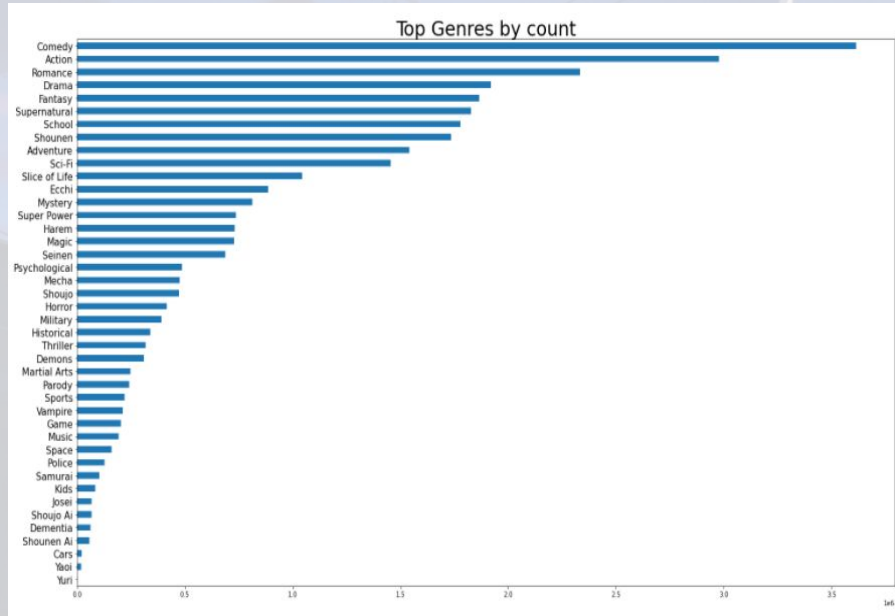
MAL_ID		Name	Score	Genres	synopsis
0	1	Cowboy Bebop	8.78	Action, Adventure, Comedy, Drama, Sci-Fi, Space	In the year 2071, humanity has colonized sever...
1	5	Cowboy Bebop: Tengoku no Tobira	8.39	Action, Drama, Mystery, Sci-Fi, Space	other day, another bounty—such is the life of ...
2	6	Trigun	8.24	Action, Sci-Fi, Adventure, Comedy, Drama, Shounen	Vash the Stampede is the man with a \$\$60,000,0...
3	7	Witch Hunter Robin	7.27	Action, Mystery, Police, Supernatural, Drama, ...	ches are individuals with special powers like ...
4	8	Bouken Ou Beet	6.98	Adventure, Fantasy, Shounen, Supernatural	It is the dark century and the people are suff...
...	...	...	...	...	...
16209	48481	Daomu Biji Zhi Qinling Shen Shu	Unknown	Adventure, Mystery, Supernatural	No synopsis information has been added to this...
16210	48483	Mieruko-chan	Unknown	Comedy, Horror, Supernatural	ko is a typical high school student whose life...
16211	48488	Higurashi no Naku Koro ni Sotsu	Unknown	Mystery, Dementia, Horror, Psychological, Supe...	Sequel to Higurashi no Naku Koro ni Gou .
16212	48491	Yama no Susume: Next Summit	Unknown	Adventure, Slice of Life, Comedy	New Yama no Susume anime.
16213	48492	Scarlet Nexus	Unknown	Action, Fantasy	Solar calendar year 2020: grotesque organisms ...

16214 rows × 5 columns

# Data Cleaning and EDA

## – Genre

Top genres are 'comedy', 'action', 'romance', 'fantasy', and 'drama'.

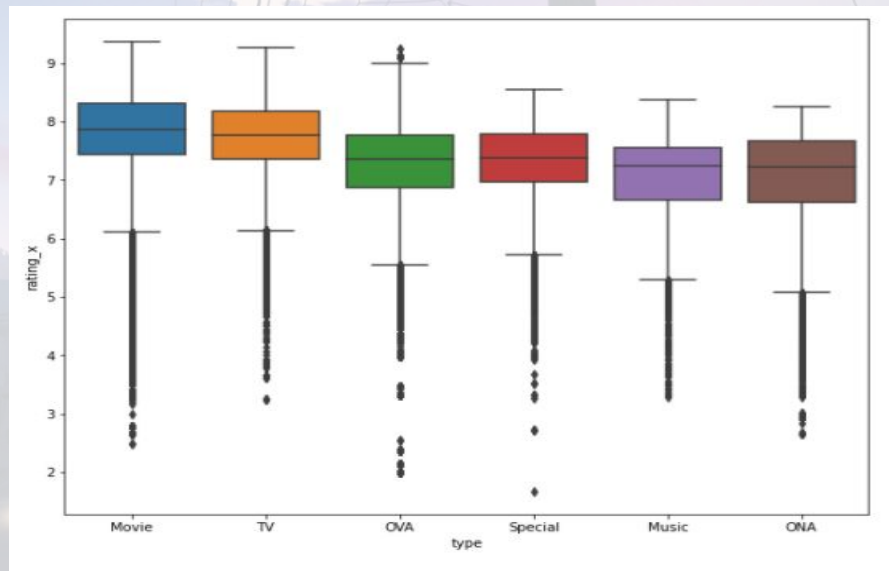
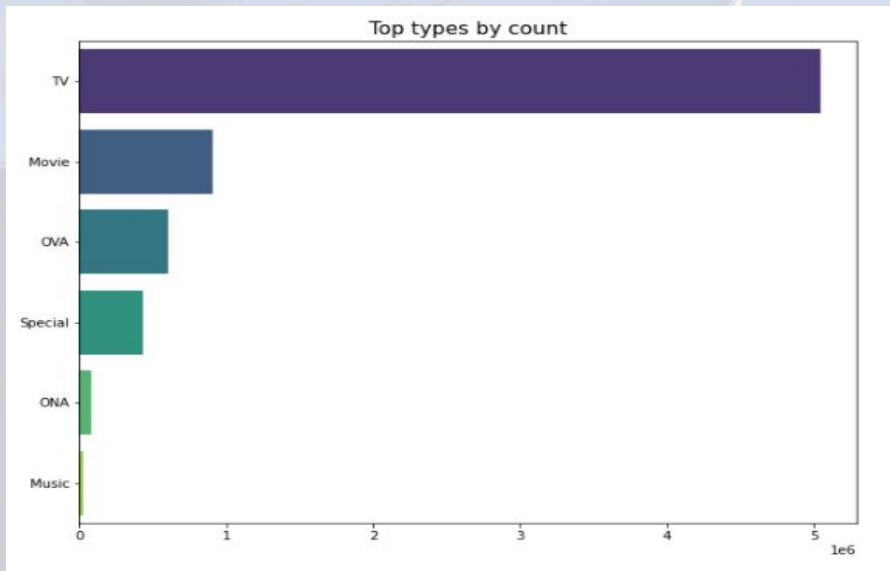


# Data Cleaning and EDA

- Type

'TV' is the most common type, with 'music' the least common.

Movies have slightly higher average score compared to other types.



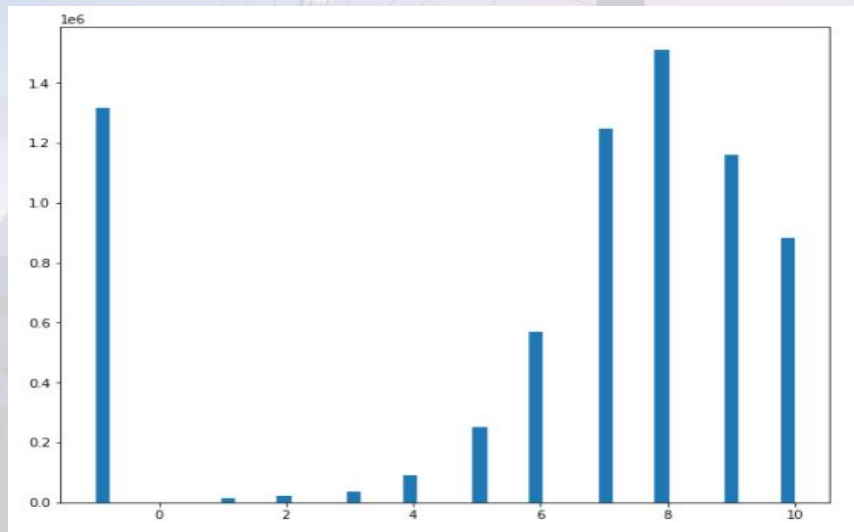
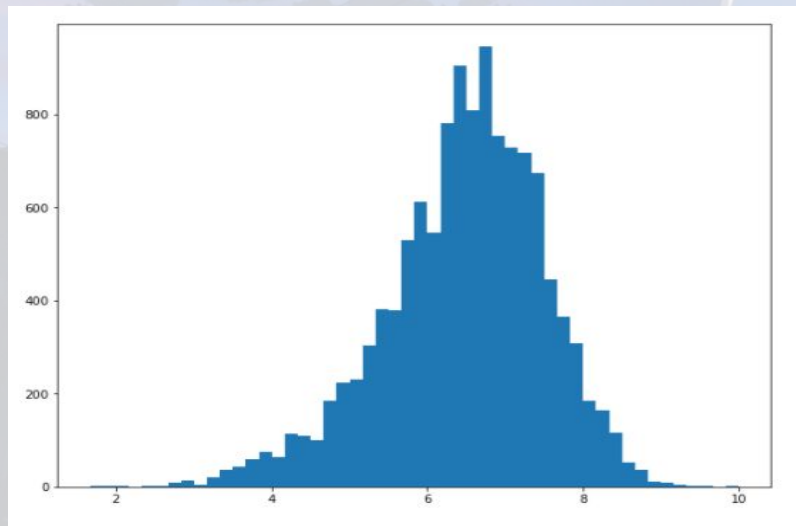


# Data Cleaning and EDA

- Rating

Most average scores are observed to range from 4 to 8.

A score of -1 indicates user has watched but not given a rating.

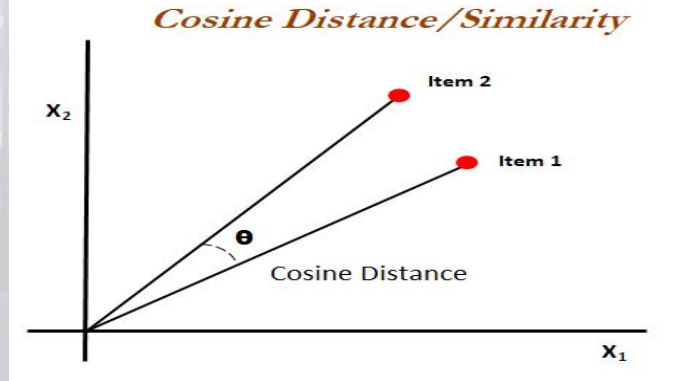
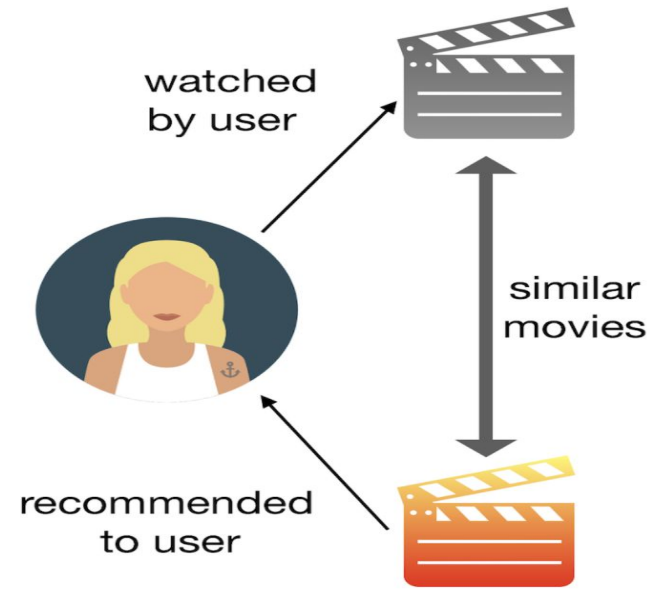


# Content Based filtering



# Content Based filtering

1. Using item features to recommend items similar to what the user likes (eg. categories, genres, types)
2. Easy to implement
3. Highly relevant to user
4. Lacks in novelty and serendipity



# Content Based filtering

## 1. Synopsis

### a. TFIDF Vectorizer (1, 4 n-grams)

### b. Cosine Similarity (similarity between word/phrase vectors)

```
1 %%time
2 synop_recommendations('Fullmetal Alchemist: Brotherhood')
```

Wall time: 14 ms

	Anime name	Similarity	Genre	Type	Synopsis
0	Fullmetal Alchemist	0.621646	['Action', 'Adventure', 'Comedy', 'Drama', ...]	TV	Edward Elric, a young, brilliant alchemist, ha...
1	Fullmetal Alchemist: The Sacred Star of Milos	0.434367	['Action', 'Adventure', 'Comedy', 'Drama', ...]	Movie	Chasing a runaway alchemist with strange power...
2	Fullmetal Alchemist: The Sacred Star of Milos ...	0.319905	['Comedy', 'Fantasy', 'Magic']	ONA	To mark the July 2 opening of the Fullmetal Al...
3	Fullmetal Alchemist: The Conqueror of Shamballa	0.304173	['Comedy', 'Drama', 'Fantasy', 'Historical']	Movie	In desperation, Edward Elric sacrificed his bo...
4	Fullmetal Alchemist: Brotherhood Specials	0.196841	['Adventure', 'Drama', 'Fantasy', 'Magic', ...]	Special	mazing secrets and startling facts are exposed...
5	Joker Game	0.087749	['Drama', 'Historical', 'Military']	TV	h World War II right around the corner, intell...
6	Fullmetal Alchemist: Premium Collection	0.087326	['Comedy', 'Fantasy', 'Shounen']	OVA	1. State Alchemists vs Seven Homunculi A 10 mi...
7	Diabolik Lovers More,Blood	0.080114	['Harem', 'School', 'Shoujo', 'Vampire']	TV	Yui Komori, still held captive by the Sakamaki...
8	Cowboy Bebop: Ein no Natsuyasumi	0.077855	['Adventure']	Special	Corgi can certainly dream! Taking a break from...
9	Inferius Wakusei Senshi Gaiden Condition Green	0.076809	['Action', 'Sci-Fi']	OVA	Keith, George, Edward, Yang and Shaw are Plato...

```
1 %%time
2 synop_recommendations('Death Note')
```

Wall time: 8.98 ms

	Anime name	Similarity	Genre	Type	Synopsis
0	Soul Eater	0.176649	['Action', 'Adventure', 'Comedy', 'Fantasy']	TV	Death City is home to the famous Death Weapon ...
1	Kite Liberator	0.168902	['Action', 'Drama', 'Sci-Fi']	OVA	In the previous series "Kite," a killer—Sawa—s...
2	YAT Anshin! Uchuu Ryokou 2	0.164707	['Adventure', 'Comedy', 'Sci-Fi', 'Shounen']	TV	The space travelling company, YAT, is thrown i...
3	Yami no Matsuei	0.163860	['Comedy', 'Drama', 'Fantasy', 'Horror', '...]	TV	Even after death, life is full of paperwork an...
4	Bleach: Memories in the Rain	0.145748	['Action', 'Adventure', 'Drama', 'Shounen', ...]	Special	Kurosaki Ichigo, the temporary Shinigami(Death...
5	Medarot Damashii	0.139862	['Action', 'Adventure', 'Sci-Fi']	TV	The sequel to the famous Medabots series, foll...
6	Yami no Shihosha Judge	0.133875	['Horror', 'Seinen', 'Supernatural']	OVA	Hoichiro Ohma works in an office. Everybody kn...
7	Shinigami no Ballad.	0.132780	['Drama', 'Fantasy', 'Psychological', 'Supe...]	TV	girl wrapped in white, her name is Momo...in h...
8	Ghost Messenger	0.116069	['Action', 'Fantasy']	OVA	Ghost Messengers are super-power agents from t...
9	Yume Senshi Wingman	0.115039	['Action', 'Comedy', 'Romance', 'School', '...]	TV	Hirono Kenta is your average geeky junior high...



# Content Based filtering

## 1. Genre

### a. TFIDF Vectorizer (1, 1 n-grams)

## 2. Synopsis

### a. TFIDF Vectorizer (1, 4 n-grams)

## 3. Cosine Similarity

```
1 synop_genre_recommendations('Fullmetal Alchemist: Brotherhood')
```

	Anime name	Similarity	Genre	Type	Synopsis
0	Fullmetal Alchemist	0.797745	[Action', 'Adventure', 'Comedy', 'Drama', ...	TV	Edward Elric, a young, brilliant alchemist, ha...
1	Fullmetal Alchemist: The Sacred Star of Milos	0.704106	[Action', 'Adventure', 'Comedy', 'Drama', ...	Movie	Chasing a runaway alchemist with strange power...
2	Fullmetal Alchemist: Brotherhood Specials	0.576796	[Adventure', 'Drama', 'Fantasy', 'Magic', ...	Special	mazing secrets and startling facts are exposed...
3	Fullmetal Alchemist: The Conqueror of Shamballa	0.479923	[Comedy', 'Drama', 'Fantasy', 'Historical'...	Movie	In desperation, Edward Elric sacrificed his bo...
4	Tales of Vesperia: The First Strike	0.443175	[Action', 'Adventure', 'Fantasy', 'Magic', ...	Movie	Ten years after the Great War against the demo...
5	Tide-Line Blue	0.426710	[Action', 'Adventure', 'Drama', 'Military'...	TV	The catastrophe, "Hammer of Eden," resulted in...
6	Fullmetal Alchemist: The Sacred Star of Milos ...	0.416856	[Comedy', 'Fantasy', 'Magic']	ONA	To mark the July 2 opening of the Fullmetal Al...
7	Fullmetal Alchemist: Reflections	0.406139	[Adventure', 'Comedy', 'Drama', 'Fantasy', ...	Special	flection on what happened during the FMA TV se...
8	Magi: Sinbad no Bouken (TV)	0.399622	[Action', 'Adventure', 'Fantasy', 'Magic', ...	TV	In the small, impoverished Tison Village of th...
9	Fairy Tail Movie 1: Houou no Miko	0.399445	[Action', 'Adventure', 'Comedy', 'Fantasy'...	Movie	In the mountains of north Fiore lies the Fire ...

```
1 synop_genre_recommendations('Death Note')
```

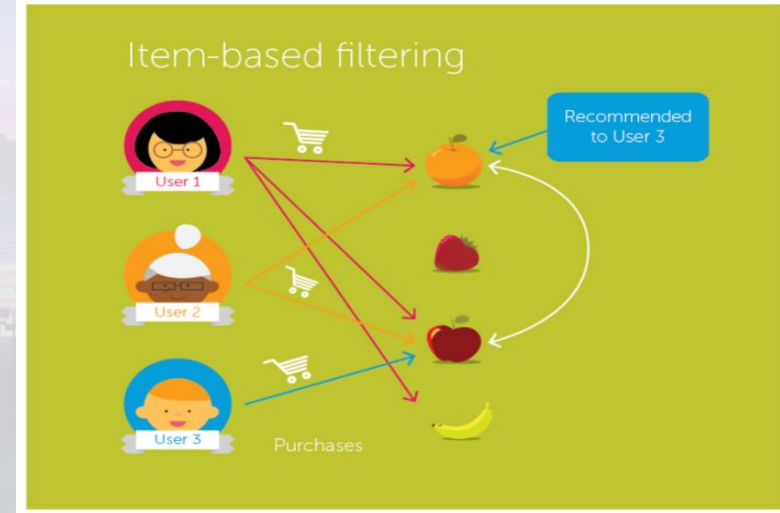
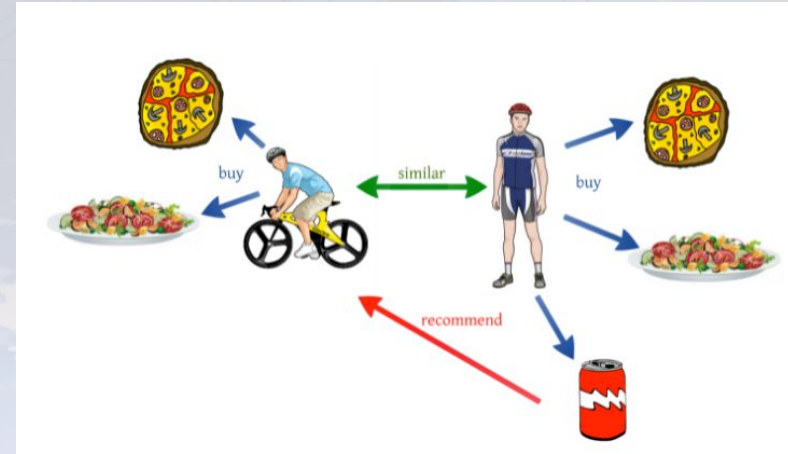
	Anime name	Similarity	Genre	Type	Synopsis
0	Mousou Dairinin	0.502951	[Drama', 'Mystery', 'Police', 'Psychologic...	TV	The infamous Shounen Bat (Lil' Slugger) is ter...
1	Higurashi no Naku Koro ni Kai	0.439441	[Mystery', 'Psychological', 'Supernatural', ...	TV	In the small village of Hinamizawa, Rika Furud...
2	Higurashi no Naku Koro ni Rei	0.435465	[Comedy', 'Mystery', 'Psychological', 'Sup...	OVA	The infamous series of unexplainable murders i...
3	Monster	0.410825	[Drama', 'Horror', 'Mystery', 'Police', '...	TV	Dr. Kenzou Tenma, an elite neurosurgeon recent...
4	Higurashi no Naku Koro ni	0.404385	[Horror', 'Mystery', 'Psychological', 'Sup...	TV	Keiichi Maebara has just moved to the quiet li...
5	Mirai Nikki (TV): Ura Mirai Nikki	0.399226	[Action', 'Comedy', 'Mystery', 'Psychologi...	Special	Unaired "Ura Mirai Nikki" anime (the short epi...
6	AD Police	0.389150	[Adventure', 'Dementia', 'Mecha', 'Mystery...	OVA	The year is 2027 in MegaTokyo, six years befor...
7	Imawa no Kuni no Alice (OVA)	0.383433	[Action', 'Psychological', 'Shounen', 'Sup...	OVA	ouhei Arisu is a high school student with no a...
8	Shigofumi	0.383216	[Drama', 'Fantasy', 'Psychological', 'Supe...	TV	There are some things that people are unable t...
9	Zankyou no Terror	0.380111	[Psychological', 'Thriller']	TV	Painted in red, the word "VON" is all that is ...



# Collaborative Based filtering

# Collaborative Based filtering

1. Using similarity between users / items to recommend items
2. Requires user rating for recommendation
3. Prone to cold start problem
4. Provides novelty and serendipity, lacks in relevance





# Collaborative Based filtering

1. Memory based techniques
  - a. rely heavily on simple similarity measures (Cosine similarity, Pearson correlation, Jaccard coefficient etc) to match similar people or items together
  - b. considered more memory intensive as compared to model based techniques
2. Model based techniques
  - a. “guess” how much a user will like an item that they did not encounter before
  - b. utilize machine learning algorithms to train on the vector of items for a specific user



# Collaborative Based filtering

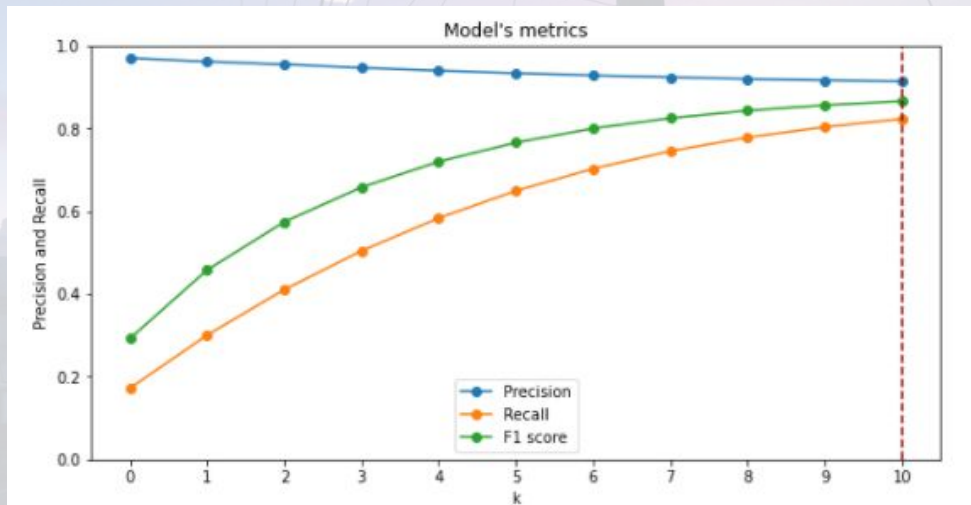
## SURPRISE recommender library

1. Memory based techniques
  - a. KNNBasic – a basic collaborative filtering algorithm
  - b. KNNWithMeans – takes into account mean ratings of each user
2. Model based techniques
  - a. SVD – similar to Probabilistic Matrix Factorization
  - b. BaselineOnly – predicts baseline estimate for user and item
  - c. NMF – very similar to SVD, user and item factors are kept positive

# Collaborative Based filtering

1. Metrics used: RMSE, MAP@K
2. Users  $\geq 200$  ratings
3. 100,000 random rows
4. Threshold at score of 7
5. GridSearchCV

	RMSE	MAE	MAP@10	MAR@10	F1@10
KNNBasic	1.519	1.176	0.901	0.719	0.800
KNNWithMeans	1.406	1.066	0.922	0.735	0.818
BaselineOnly	1.302	0.996	0.886	0.832	0.858
SVD	1.276	0.976	0.917	0.804	0.856
NMF (Non Negative MF)	1.692	1.360	0.999	0.608	0.756



# Collaborative Based filtering

Model Used: SVD

User id: 65108

Top N Recommendations (User's animes not included):

Ginga Eiyuu Densetsu

Gintama°

Gintama: Yorinuki Gintama-san on Theater 2D

One Piece: Episode of Merry - Mou Hitori no Nakama no Monogatari

Kimi no Na wa.

Gintama

Romeo no Aoi Sora

Yojouhan Shinwa Taikei

Tamala 2010: A Punk Cat in Space

Kaguya-hime no Monogatari

User's Animes:

Top wo Nerae! Gunbuster

Kamisama Hajimemashita

Kyoukai no Kanata

Sidonia no Kishi

Mobile Suit Zeta Gundam: A New Translation III - Love Is the Pulse of the Stars

Date A Live: Date to Date

Shigatsu wa Kimi no Uso

Shaman King

Hamatora The Animation

Air in Summer

Model Used: SVD

User id: 22479

Top N Recommendations (User's animes not included):

Mobile Suit Gundam

Yojouhan Shinwa Taikei

Seikai no Senki

Romeo no Aoi Sora

Kaze Tachinu

Doraemon

Project A-Ko 4: Final

Mobile Suit Zeta Gundam: A New Translation - Heir to the Stars

Digimon Adventure: Bokura no War Game!

Zetsuen no Tempest

User's Animes:

Clannad

Hajime no Ippo: Rising

Gintama: Shinyaku Benizakura-hen

Diamond no Ace

Ao no Exorcist

Space☆Dandy

Baka to Test to Shoukanjuu: Matsuri

Kekkai Sensen: Soresaemo Saitei de Saikou na Hibi

Dakara Boku wa, H ga Dekinai.

Zettai Karen Children



# Conclusion and future steps

1. Incorporate more features into content based filtering eg voice actors, animation studios etc.
2. Try out more models such as co-clustering, SVDpp.
3. Deploy hybrid based filtering
4. Do user testing for recommendations





Thank You!