

# Netflix Shows Recommendation System

Guofa Shou

Self-paced

# Outline

- Netflix shows dataset
- Netflix shows recommendation system
- Conclusion

# Netflix shows dataset

- Data Source

- <https://www.kaggle.com/datasets/shivamb/netflix-shows>

- Data Exploration

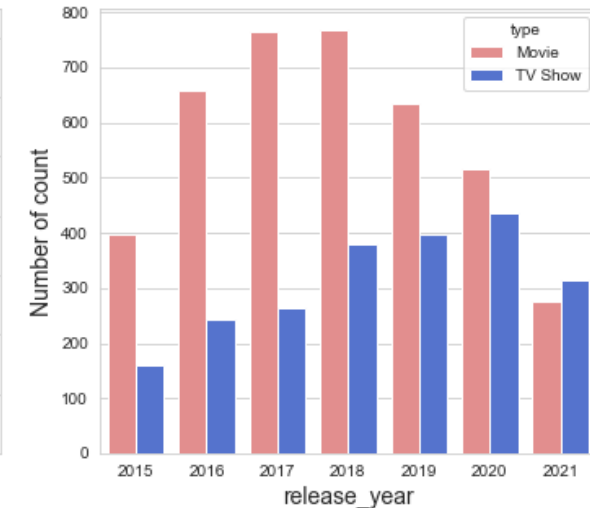
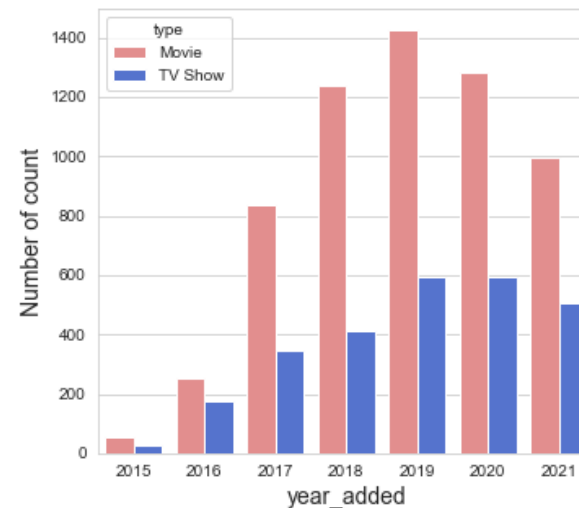
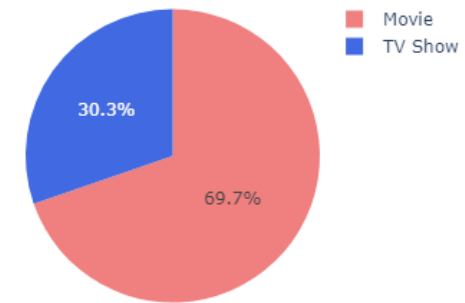
RangeIndex: 8807 entries, 0 to 8806

Data columns (total 12 columns):

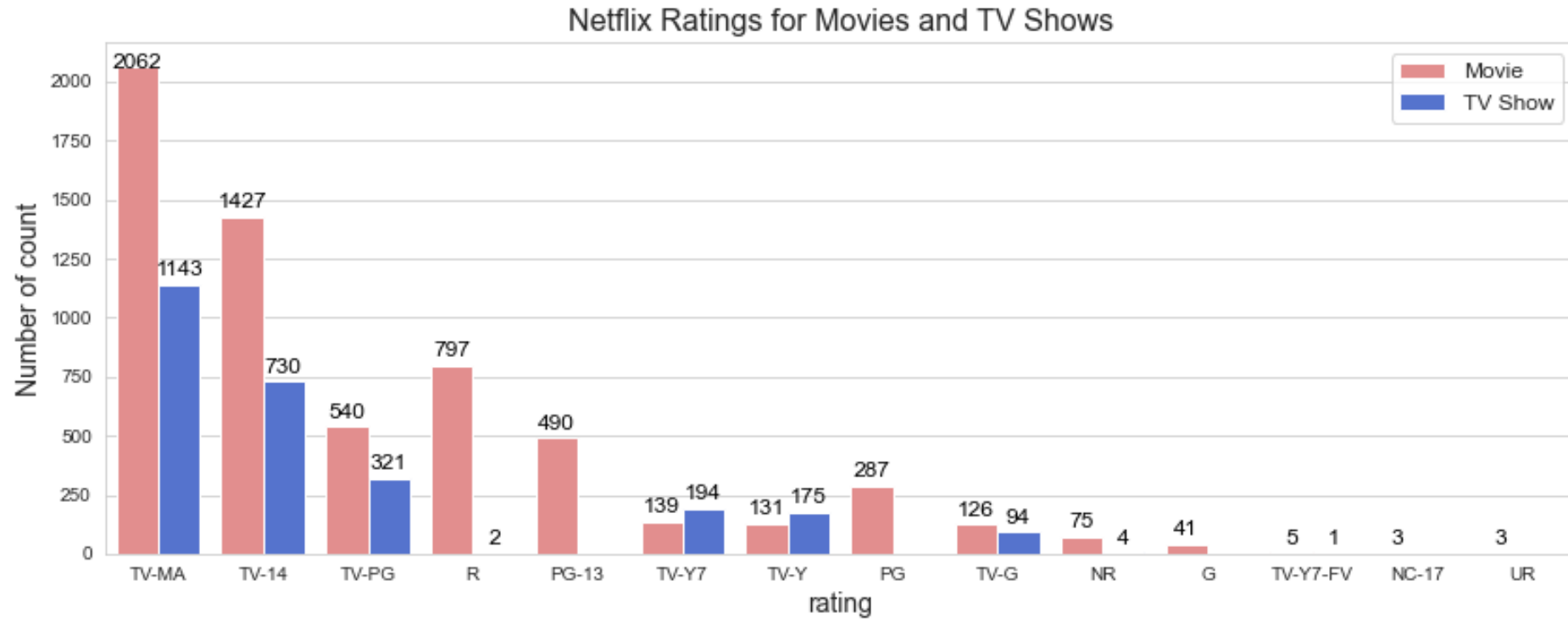
#	Column	Non-Null Count	Dtype
0	show_id	8807 non-null	object
1	type	8807 non-null	object
2	title	8807 non-null	object
3	director	6173 non-null	object
4	cast	7982 non-null	object
5	country	7976 non-null	object
6	date_added	8797 non-null	object
7	release_year	8807 non-null	int64
8	rating	8803 non-null	object
9	duration	8804 non-null	object
10	listed_in	8807 non-null	object
11	description	8807 non-null	object

dtypes: int64(1), object(11)

Movie (6126) vs. TV show (2664)



- Data Exploration (cont.): ratings



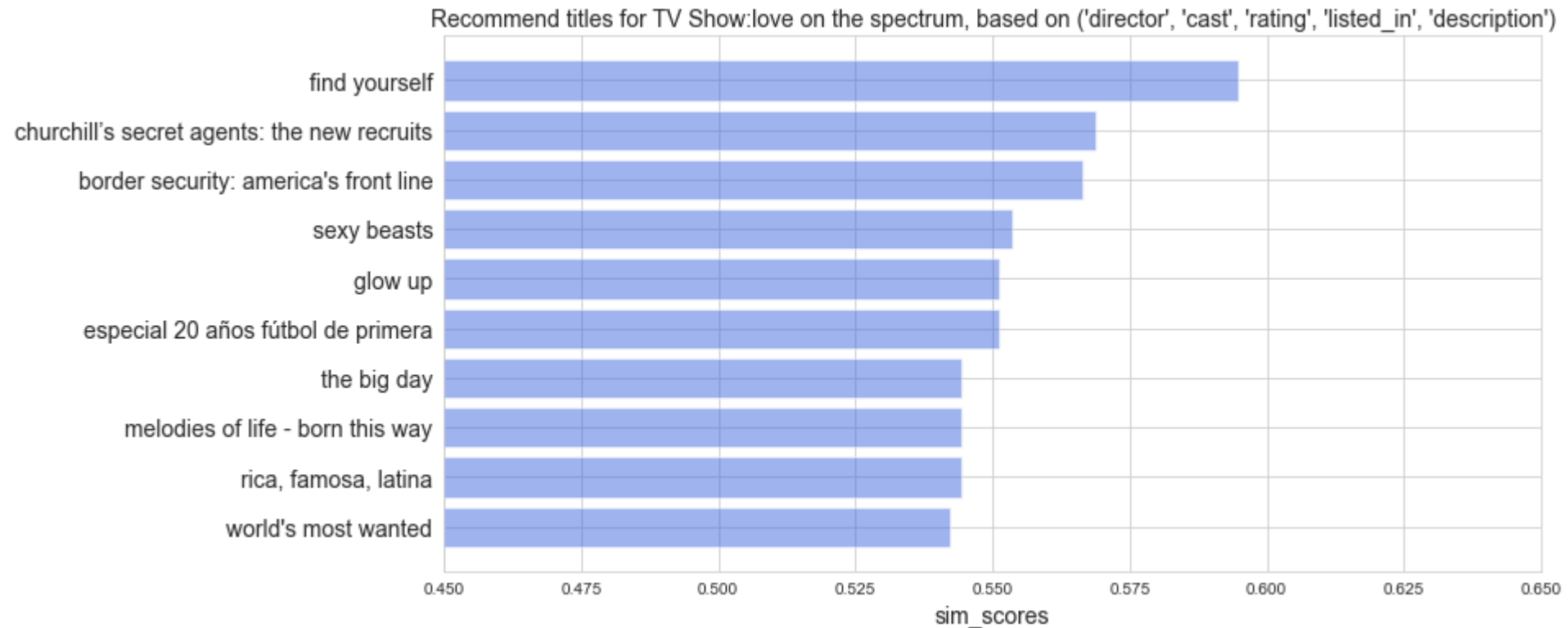


# Netflix Recommendation System

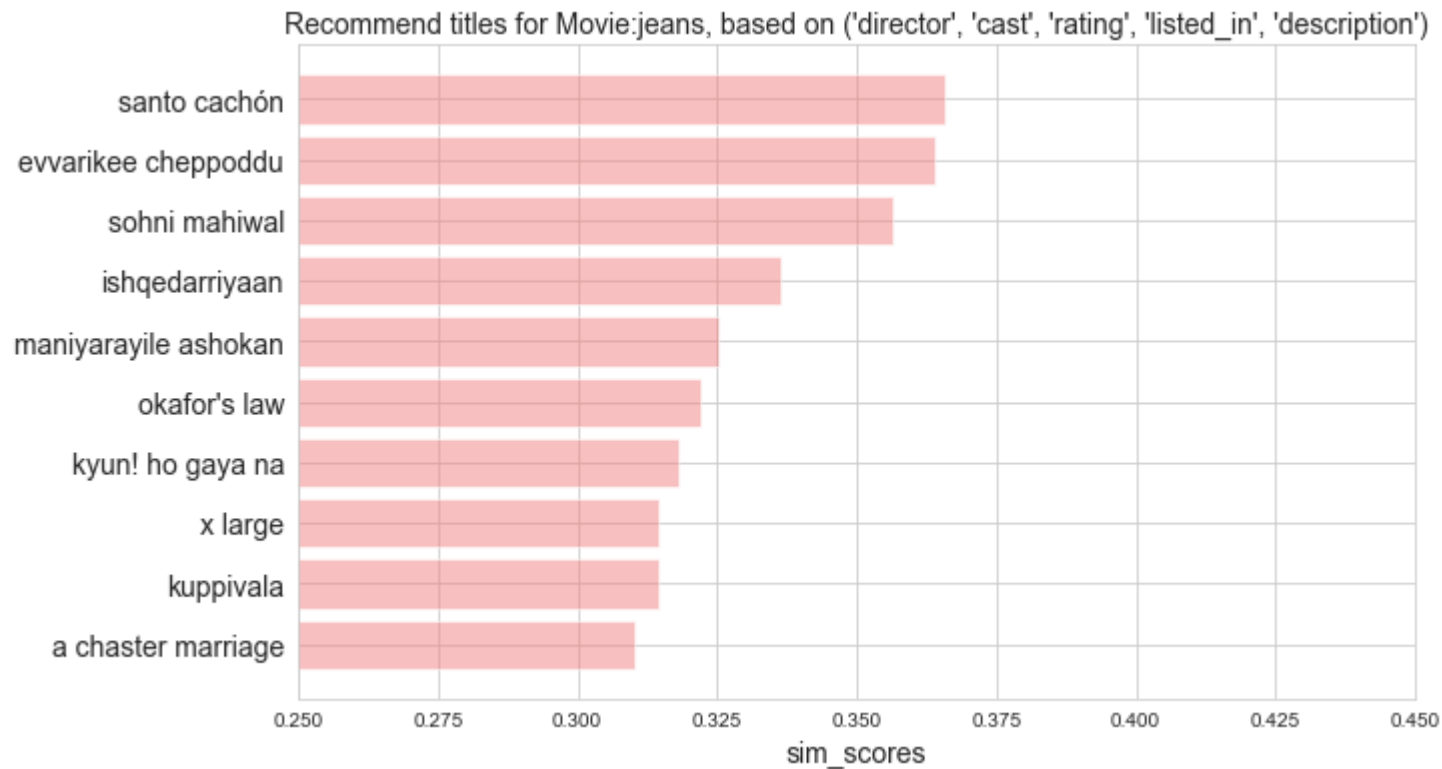
- A content-based recommendation system
  - Comparing five different combinations of features:
    - 'comb\_features0': ('director', 'cast', 'rating', 'listed\_in'),
    - 'comb\_features1': ('director', 'cast', 'rating', 'description'),
    - 'comb\_features2': ('director', 'cast', 'listed\_in', 'description'),
    - 'comb\_features3': ('director', 'rating', 'listed\_in', 'description'),
    - 'comb\_features4': ('cast', 'rating', 'listed\_in', 'description'),
    - 'comb\_features5': ('director', 'cast', 'rating', 'listed\_in', 'description')
  - An example to recommend titles for: too hot to handle: brazil

	comb_features0	comb_features1	comb_features2	comb_features3	comb_features4	comb_features5
0	back with the ex	too hot to handle	dating around: brazil	too hot to handle: latino	too hot to handle: latino	too hot to handle: latino
1	too hot to handle: latino	bling empire	too hot to handle: latino	rica, famosa, latina	dating around: brazil	dating around: brazil
2	the big day	we are the champions	rica, famosa, latina	dating around: brazil	rica, famosa, latina	rica, famosa, latina
3	baewatch: parental guidance	too hot to handle: latino	border security: america's front line	teresa	baewatch: parental guidance	baewatch: parental guidance
4	million pound menu	the movies that made us	baewatch: parental guidance	baewatch: parental guidance	back with the ex	back with the ex

- Representative example for a TV show based on com\_feature5: 'director', 'cast', 'rating', 'listed\_in', 'description'



- Representative example for a Movie based on com\_feature5: 'director', 'cast', 'rating', 'listed\_in', 'description'





# Conclusion

- A content-based recommendation system based on the input title was successfully developed
  - It could suggest existed titles if the input title is not existed
  - The recommended titles are usually in the same type of the input title

Thanks for your attention!