Beer Recommendation Model

Tomer Danon
Data Science Capstone
University of Denver

Introduction

- 5 years in beer industry.
- 3 years experience in brewing industry
  - Novel Strand Brewing

- Interesting, fun, and exciting project!
- Applicable to many platforms/apps.
  - Great for portfolio!

UNIVERSITY *of* DENVER

DANIEL FELIX RITCHIE SCHOOL
OF ENGINEERING & COMPUTER SCIENCE

Introduction

"How was you visit to ___?"

"Are you enjoying your recent purchase of___?

- Many user-based review platforms exist
  - Google, Yelp!, Trip Advisor, etc.
  - BeerAdvocate, Untapped, ratebeer.com, etc.

- Large amounts of data about user preferences.

- Opportunity! Use this data to recommend new products based on *shared preferences* → **collaborative filtering**

UNIVERSITY *of* DENVER

DANIEL FELIX RITCHIE SCHOOL
OF ENGINEERING & COMPUTER SCIENCE

The Dataset

Source

- Beer reviews from BeerAdvocate.com
  - Online beer community since 1996
  - Data ~10 years old
- Attributes:
  - Ratings for aroma, appearance, palate, taste, and overall (average) rating.
  - Brewery name, beer style, username
  - Beer ID, Brewery ID

The Dataset

Overview

| 1,586,614 Reviews | 33,388 Users | 66,055 Beers | 104 Styles | 5,743 Breweries |

	brewery_id	brewery_name	review_time	review_overall	review_aroma	review_appearance	review_profilename	beer_style	review_palate	review_taste	beer_name	beer_abv	beer_beerid
0	10325	Vecchio Birraio	1234817823	1.875	2.0	2.5	stcules	Hefeweizen	1.5	1.5	Sausa Weizen	5.0	47986
1	10325	Vecchio Birraio	1235915097	2.875	2.5	3.0	stcules	English Strong Ale	3.0	3.0	Red Moon	6.2	48213
2	10325	Vecchio Birraio	1235916604	2.875	2.5	3.0	stcules	Foreign / Export Stout	3.0	3.0	Black Horse Black Beer	6.5	48215
3	10325	Vecchio Birraio	1234725145	3.000	3.0	3.5	stcules	German Pilsener	2.5	3.0	Sausa Pils	5.0	47969

Data Cleaning & Preparation

Users DataFrame

- How many reviews?
- Scores per styles reviewed
- Beer info & scores per beers reviewed

review_profilename	num_reviews	user_styles	user_beers
0110x011	139	{'American Amber / Red Ale': 3.88, 'American B...	{'Wachusett IPA (India Pale Ale)': (23, 'Wachu...
01Ryan10	1	{'American IPA': 4.88}	{'Duet IPA': (27604, 'Alpine Beer Company', 'A...
02maxima	4	{'American Double / Imperial IPA': 4.62, 'Belg...	{'Brooklyn Black Chocolate Stout': (680, 'Broo...
03SVTCobra	3	{'English Porter': 3.5, 'Light Lager': 1.75, '...	{'Lone Star Light': (1731, 'Pabst Brewing Comp...
04101Brewer	3	{'American IPA': 3.75, 'American Pale Ale (APA...	{'Peeper Ale': (51512, 'Maine Beer Company', '...

Recommendation Model

Style or Beer

```
{'Saison / Farmhouse Ale', 'American IPA', 'American Double / Imperial IPA'}
```

```
{'Pliny The Elder', 'Heady Topper', 'Drie Fonteinen Oude Geuze', '90 Minute IPA'}
```

```
Number of users in cluster: 94
['2ndloveofmylife', 'BeerLover48Fan', 'Beervis', '
'HumuloneRed', 'Impfan', 'JeffBob', 'JoeyHotrod',
```

STYLE: American IPA

beer_beerid	beer_name	brewery_name	beer_style	review_overall
59672	Hill Farmstead Citra IPA	Hill Farmstead Brewery	American IPA	5.00
33467	High Tide Fresh Hop IPA	Port Brewing Company / Pizza Port	American IPA	5.00
67773	Puma IPA	Elk Mountain Brewing	American IPA	5.00
34436	Stone IPA Double Dry Hopped	Stone Brewing Co.	American IPA	4.88
8348	Sgt Major's IPA	Scotch Irish Brewing Company	American IPA	4.88
55580	Tomahawk Single Hop IPA	Mikkeller ApS	American IPA	4.88
55288	Amarillo Single Hop IPA	Mikkeller ApS	American IPA	4.88
39705	Russian River IPA	Russian River Brewing Company	American IPA	4.88
37395	Raasted Grilløl	Raasted & Randers Bryghus A/S	American IPA	4.88
33243	Wipeout I.P.A.	Port Brewing Company / Pizza Port	American IPA	4.81

1. New user provides beer style or beer name preferences.

2. Find users who have an average rating of 4.25 or higher for all preferred styles or beers - **CLUSTER**.

3. Find beers of the preferred styles or beer from the users in the cluster.

4. Group by beer and average score from all of the user clusters

5. Suggest top 10.

UNIVERSITY *of* DENVER

DANIEL FELIX RITCHIE SCHOOL
OF ENGINEERING & COMPUTER SCIENCE

Conclusions

- Evaluation requires **user feedback** on the recommendations they receive.
- Potential design improvement for a **machine learning** algorithm.

Thank You! Questions?

Tomer Danon

MS Data Science @ DU

Tomer.Danon@du.edu
danont@gmail.com

Connect with me on **LinkedIn**!

UNIVERSITY *of* DENVER

DANIEL FELIX RITCHIE SCHOOL
OF ENGINEERING & COMPUTER SCIENCE