

Serving Recipe Recommendations

Mike Flanagan



About the Dataset

- * **Sourced from Kaggle:**
Shuyang Li, *Food.com Recipes and Interactions (2019)*
- * **Generating Personalized Recipes from Historical User Preferences**
Bodhisattwa Prasad Majumder, Shuyang Li, Jianmo Ni, Julian McAuley
EMNLP, 2019
<https://www.aclweb.org/anthology/D19-1613/>

About the Dataset

* Recipes Data:

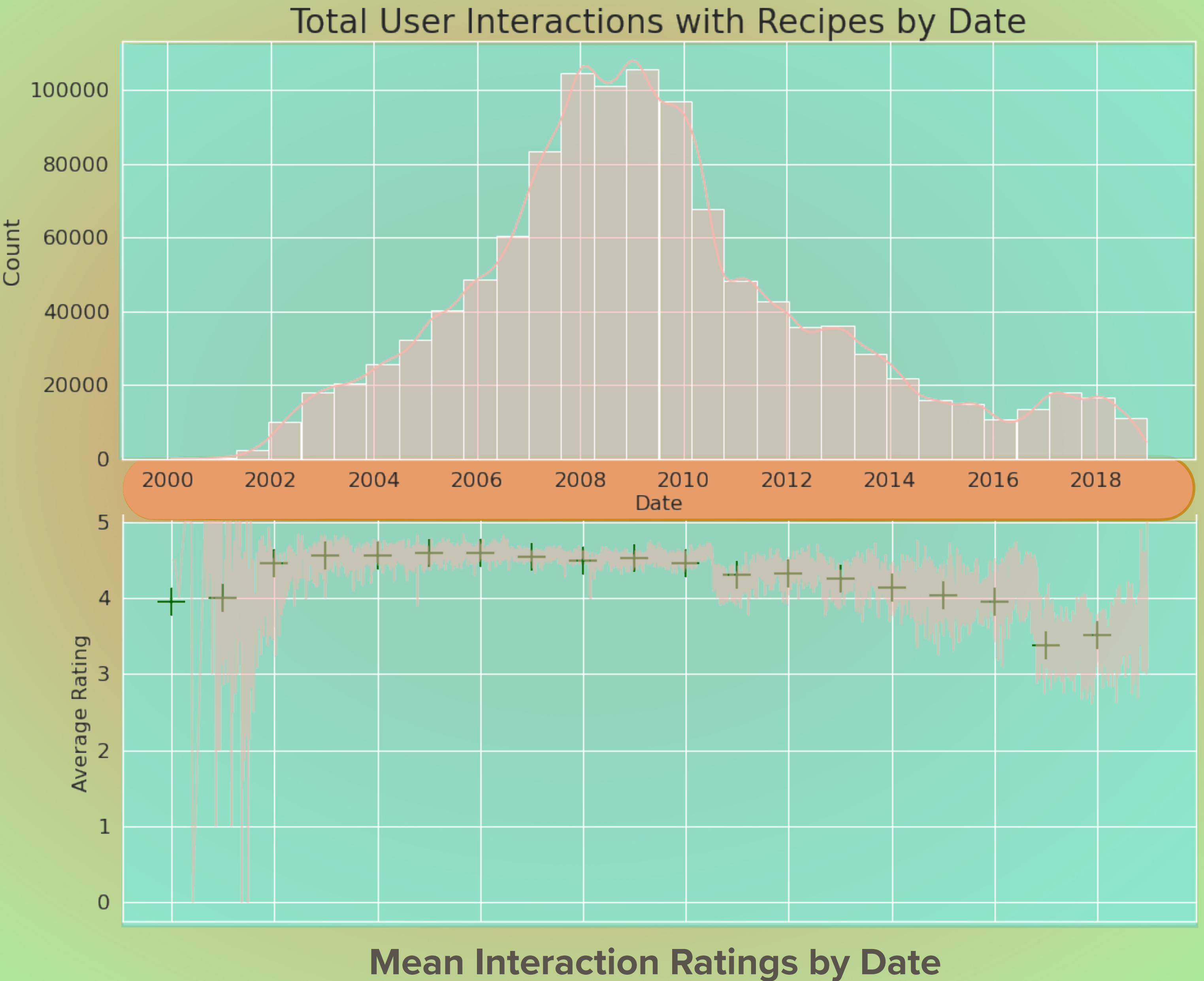
- * **231,636 rows** of recipes
- * Features include: ingredients, time to prepare, steps, nutrition, 3 user reviews, mean rating, many more.

* Interactions Data:

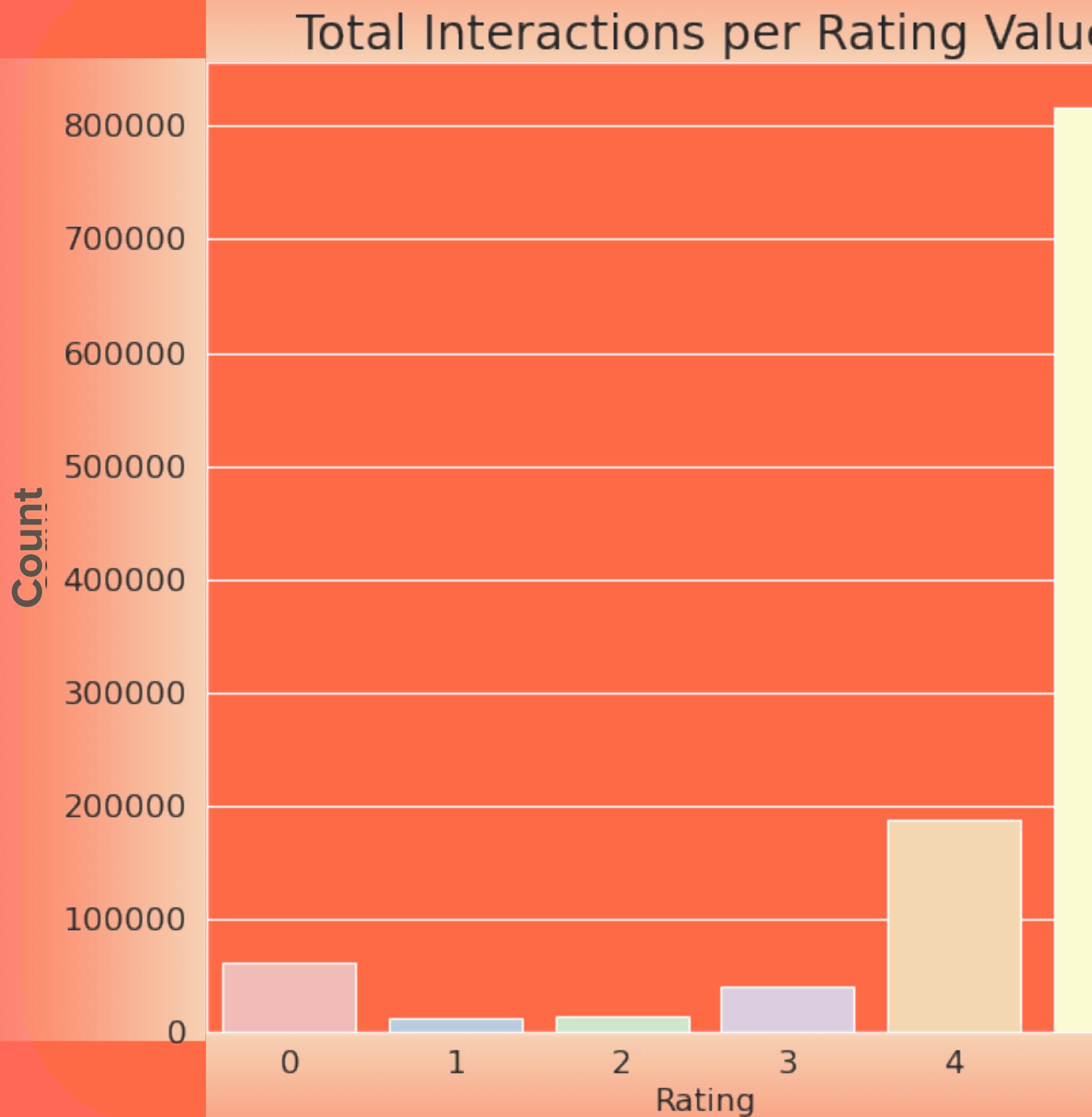
- * **1,132,367 rows** of user interactions with recipes
- * Features include user ID, recipe ID, date of interaction, rating, and review.

Date

- * High Variance in Rating Mean by Date in early years
- * Only 3,012 interactions first two years
- * Year with most Interactions is 2008



Ratings

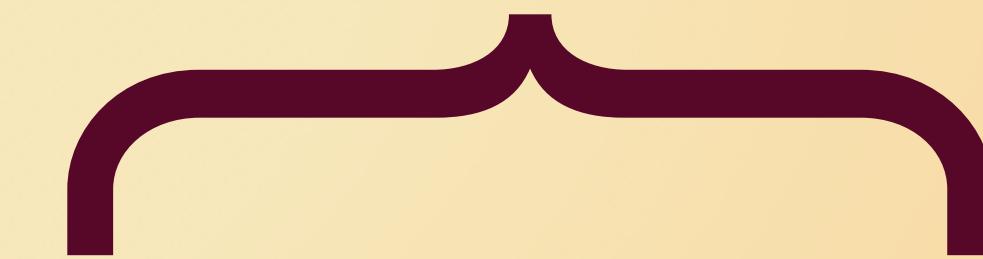


- * **816,229 Five-star Ratings**
- * **60,847 Ratings of 0**
- * **No Null Values**
- * 0 ratings managed by using *Pearson's correlation coefficient* as hyper parameter for model similarity metric.

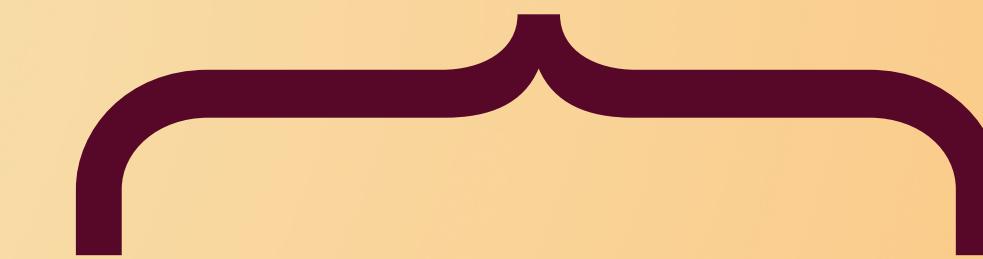
Reviews

- 169 Null Values
 - All user interactions with null review have non-null and non-zero rating

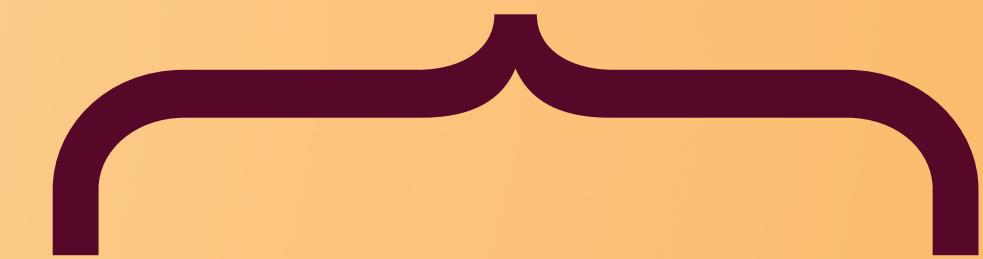
Ratings of 1 & 2



All Reviews



Ratings of 5



Modeling

Evaluating on RMSE with Pearson Correlation used for Similarity Metric

Baseline ALS: * * 1.2099 RMSE

+Best Train Time

Baseline SGD:
1.2127 RMSE

Tuned Funk's SVD: * * 1.2142

Co-clustering:
(it was worse)

Further Actions

- * A final web app that is operational, polished & accessible.
- * AB Testing
- * Return to modeling

Thank you



Mike Flanagan
flanalysis.com

<https://www.linkedin.com/in/mike-flanagan-data/>

<https://github.com/mike-flanagan>

mike-flanagan.medium.com