

# Predicting Recipe Ratings



By:  
Prasuna Mannava



# Purpose

## Problem:

Client is a vegetarian recipe creator and wants to predict his new recipes before publishing.

## Solution:

Build a model that can predict ratings using data from allrecipes.com



# Data

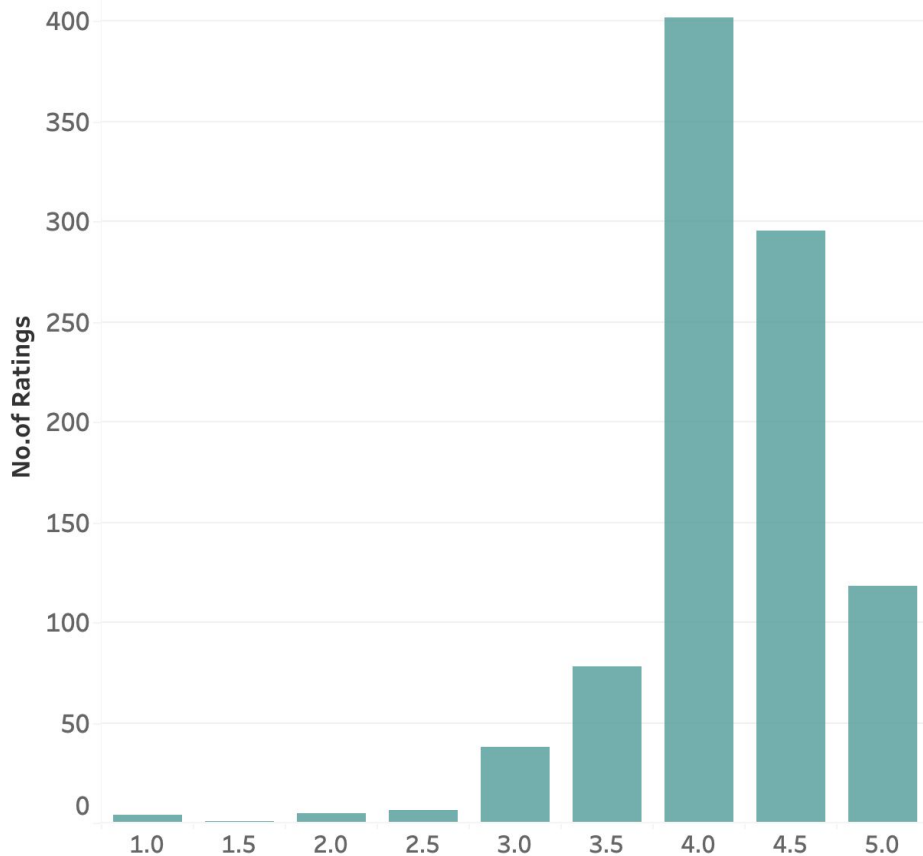
- Scrapped data from allrecipes.com using selenium and beautifulsoup
- 1000 recipes with 11 features-
  - Recipe Name
  - Ratings
  - Prep Time
  - Cook Time
  - Recipe Category
  - Yield
  - Protein
  - Carbs
  - Sodium
  - Fat
  - Cholesterol



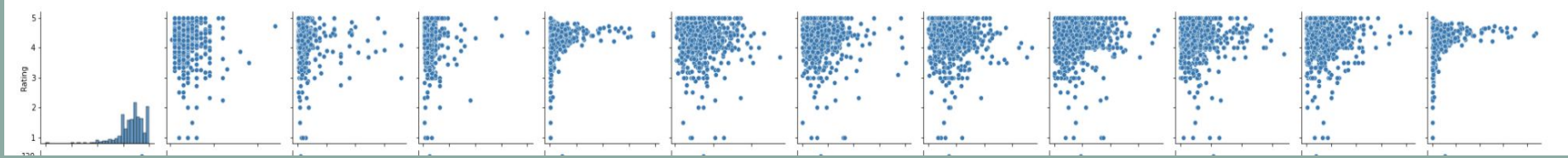
# Ratings

Number of stars	Interpretation
5	Loved it!
4	Liked it
3	It was okay
2	Didn't like it
1	Couldn't eat it

Rating Count Distribution



# Modeling



- No linear relation between target variable and features
- Boxcox transformation
- Linear Regression with transformed target variable
- Polynomial Regression with Lasso regularization
- Cross validation using KFold



# Predictions

## Linear regression with transformed target variable

```
lm.score(X_tr, y_train)
```

```
0.03633804498805526
```

```
lm.score(X_te, y_test)
```

```
0.021012311521045213
```

## Polynomial regression with Lasso

```
lasso.score(X_train_poly_scaled, y_train)
```

```
0.07252363706373843
```

```
lasso.score(X_test_poly_scaled, y_test)
```

```
-0.01473605020433122
```



# Conclusions

- Ratings may not be influenced just by the features considered in this model
- Different set of features and data required to make the model more complex
- Explore other modeling techniques where target and features are not required to be linearly related



**Thank you!**

