

Recipe Recommedar Assignment

Presentation by

Yukti Sharma Devesh Chandra

Aditya Verma

Steps:

• Read Ensure that all the columns are present with right data type.

• Extract Seven values are present in each row in nutrition column which represents nutrition information.

Standardize Existing values are not on same scale. We need to standardize the absolute values in nutrition to /100cal value



Convert

We will convert the tags column to array of strings.

Join

We will join recipe data and review data. Make sure that the resulting data frame will have same number of rows as in raw ratings data.

Create

we will create time based features where we CONVERT to datatype()

Exploratory Data Analysis



Steps in EDA

• Define custom functions

bucketize & clean numerical features

Add user level average features

Perform tags level EDA









Bottom 5%ile highest rated tags



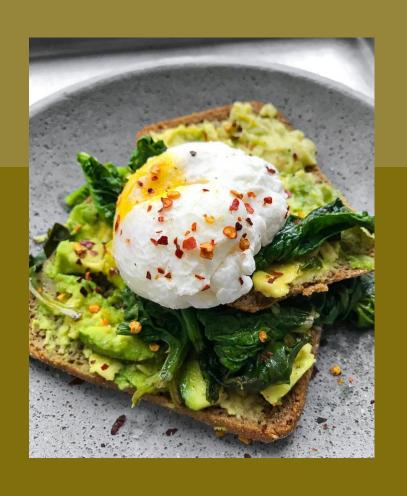
individual_tag	avg_user_rating n_us	er_ratings n_r	ecipes	in_percent_recipies i	n_percent_interactions
mothers-day	4.211538461538462	104	49	2.126680728800465	9.231241717401512E-5
simply-potatoes2	4.770642201834862	109	6	2.604098851592406	9.675051415353508E-5
cambodian	4.495412844036697	109	46	1.996475786220845E-4	9.675051415353508E-5
oatmeal	4.389380530973451	113	51	2.213484023853545	1.003009917371510
honduran 3	.7982456140350878	114	21	9.114345980573422E-5	1.011886111330550

Top 5%ile highest rated tags



individual_tag avg_user_rating n_use		recipes in percent recipies i	n_percent_interactions
ragu-recipe-contest 4.822727272727272	220	43 1.866270843641224	1.952762670988781
simply-potatoes2 4.770642201834862	109	6 2.604098851592406	9.675051415353508E-5
non-alcoholic 4.689655172413793	754	303 0.001315069920054	6.692650245116096E-4
a1-sauce 4.628491620111732	358	41 1.779467548588144	3.17767743733629E-4
labor-day 4.609756097560975	164	53 2.300287318906625	1.455695809282546E-4

thank you





Thanks