

TASTY BYTES

RECIPES AND SITE TRAFFIC

01



>

Problem Definition

We need to Increase the Traffic on the website by choosing a popular recipe. More traffic means more subscriptions so this is really important to the company.

so our goal is to predict which recipe will lead to high traffic

Problem Definition

To solve this problem there are 2 approaches :

- we need to know which recipes tend to produce high traffic
- Build a Predictive model to capture the popular recipes

DATA DESCRIPTION

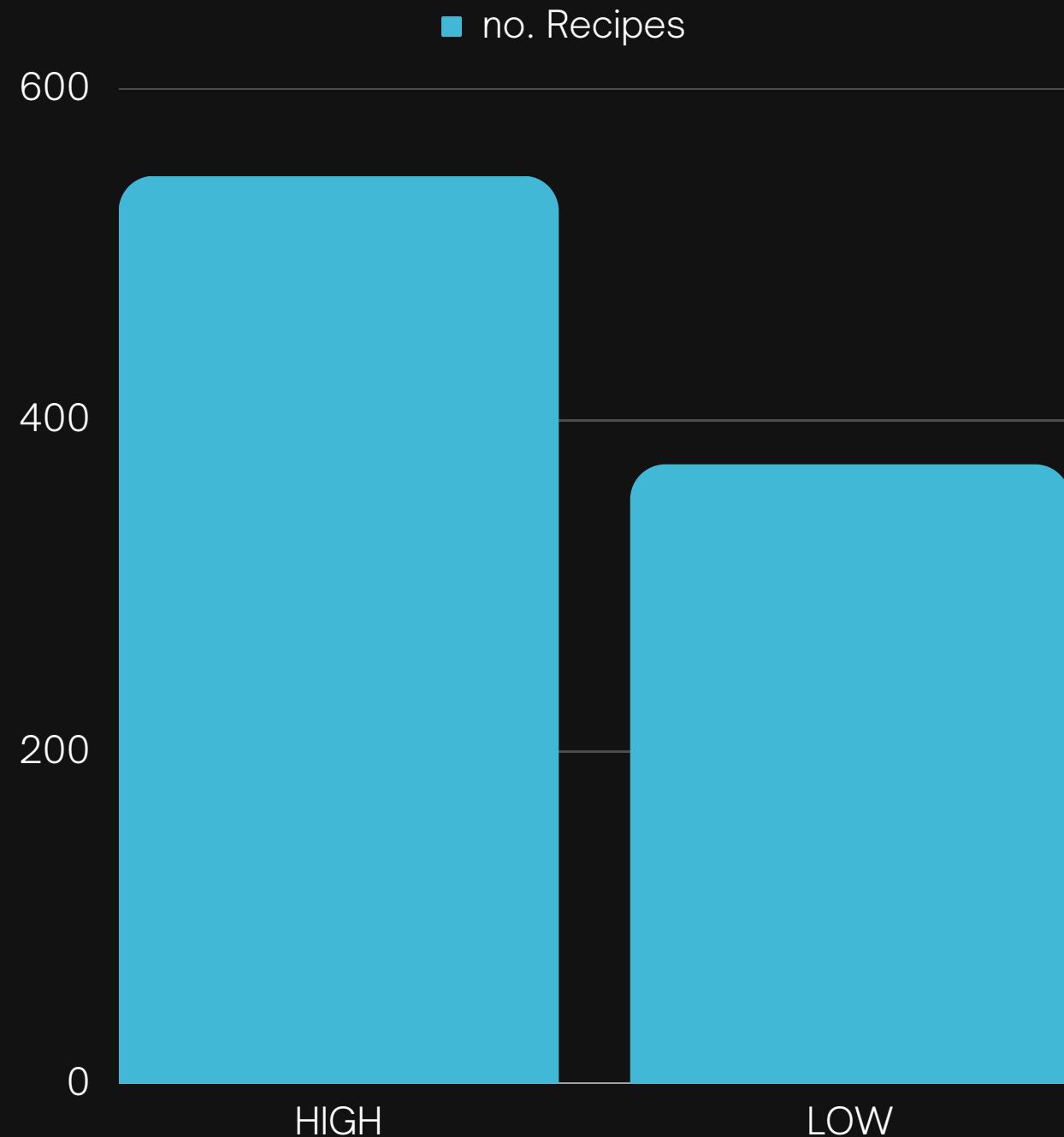


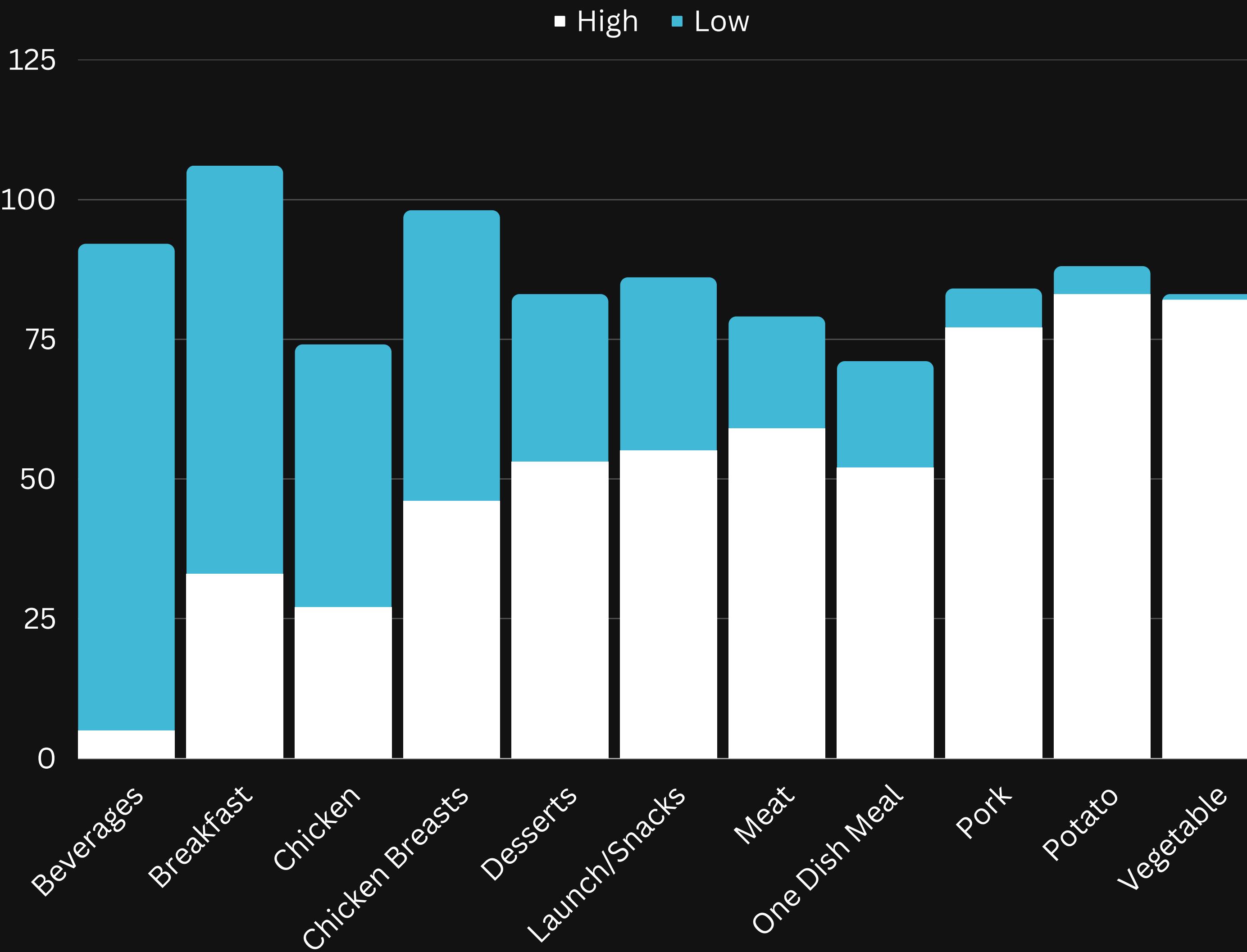
Data features

- we had nutritional information about each recipe which are calories, carbohydrates, protein and sugar
- we had the category of each recipe out of 11 categories we had in our data
- the number of servings of each recipe
- whether the recipe lead to high traffic or not 

Data Distribution

We have 547 of high traffic recipes while 373 of low traffic recipes So more recipes tend to produce high traffic data





After Analysis

category affects the traffic

after performing statistical methods we found that There is a significant association between categories and traffic

Nutritons affect the traffic

after performing statistical methods we found that There is a significant association between nutritons (calories, carbohydrates , protein and sugar) and traffic. there is big difference between the means of these nutritons at high-traffic population and low-traffic population

NEW FEATURES



New Features

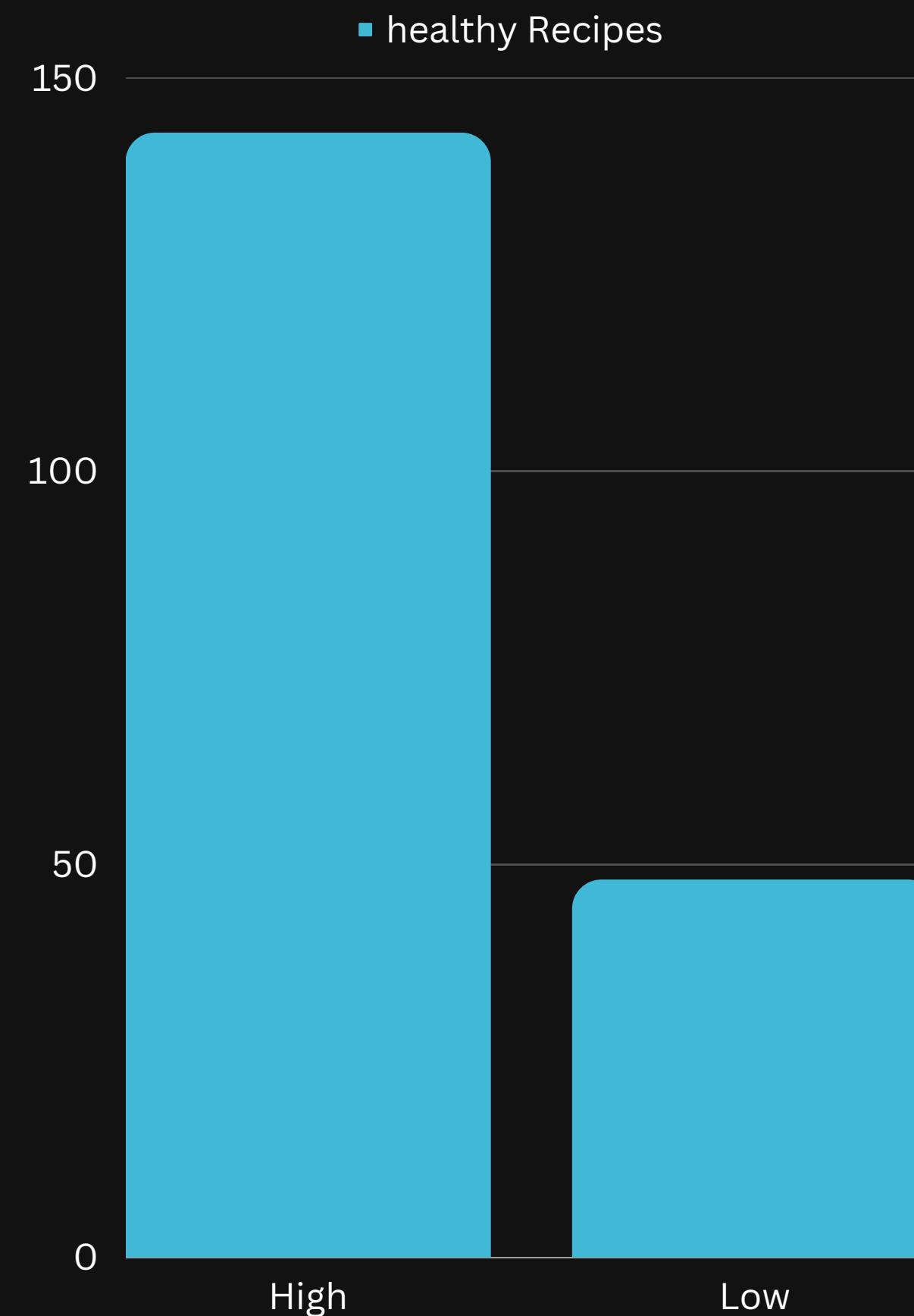
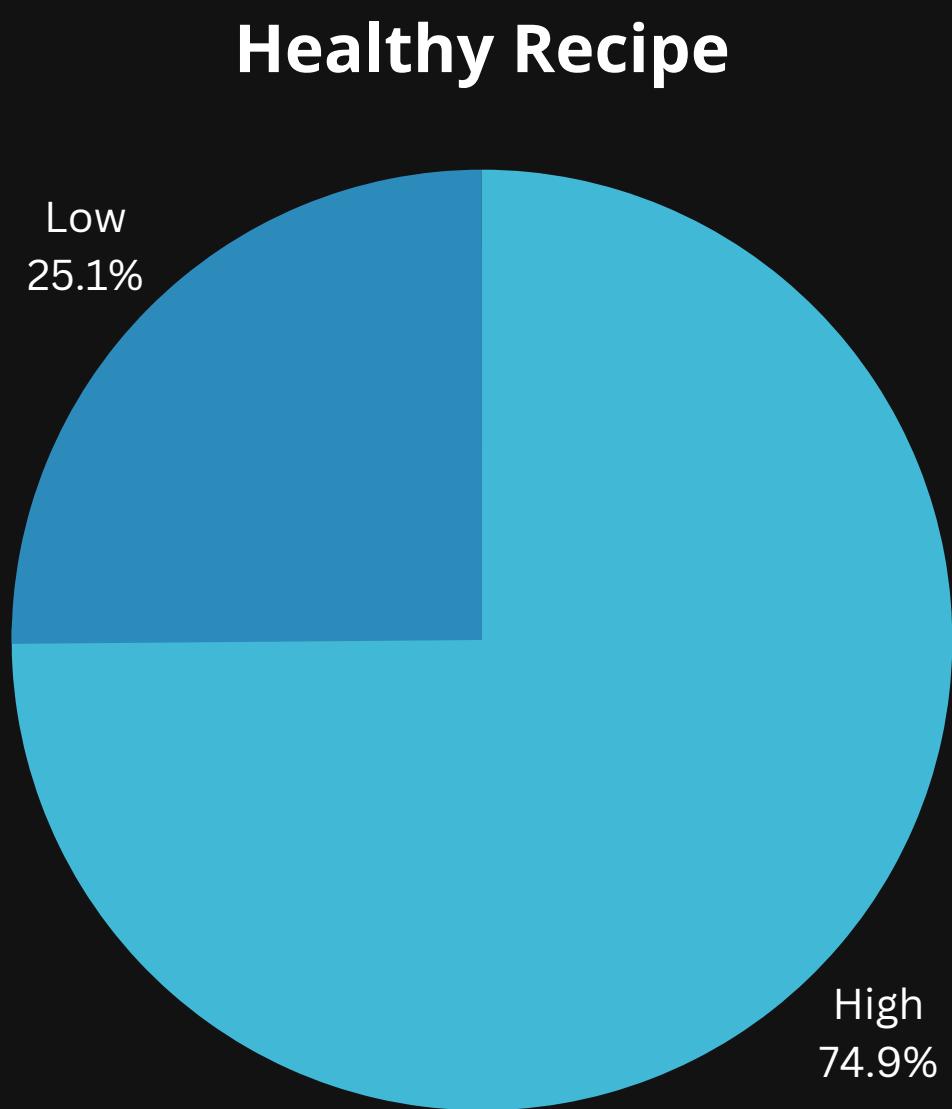
after Further analysis extracted 3 New Features :

- Is the recipe healthy or not
- the difficulty of the recipe
- the difficulty group of each recipe (easy, medium, hard)

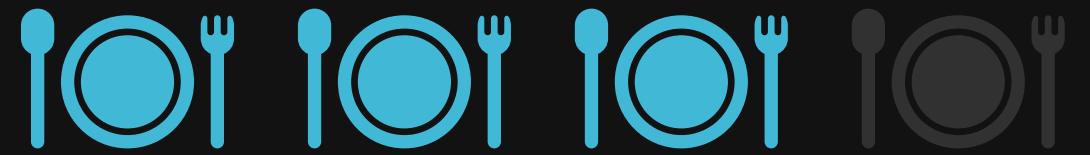
Healthy Recipe guidelines

- 1 Calories less than 400
- 2 carbohydrates between 15 & 80 gms
- 3 protein more than 10 gms
- 4 sugar less than 10 gms
- 5 Vegetable recipe

we found 191 Healthy recipes in our data
143 out of them produces high traffic
which means 75 percent of the healthy
recipes generates high traffic



Healthy food distribution



3 out of 4

healthy recipes produces high Traffic

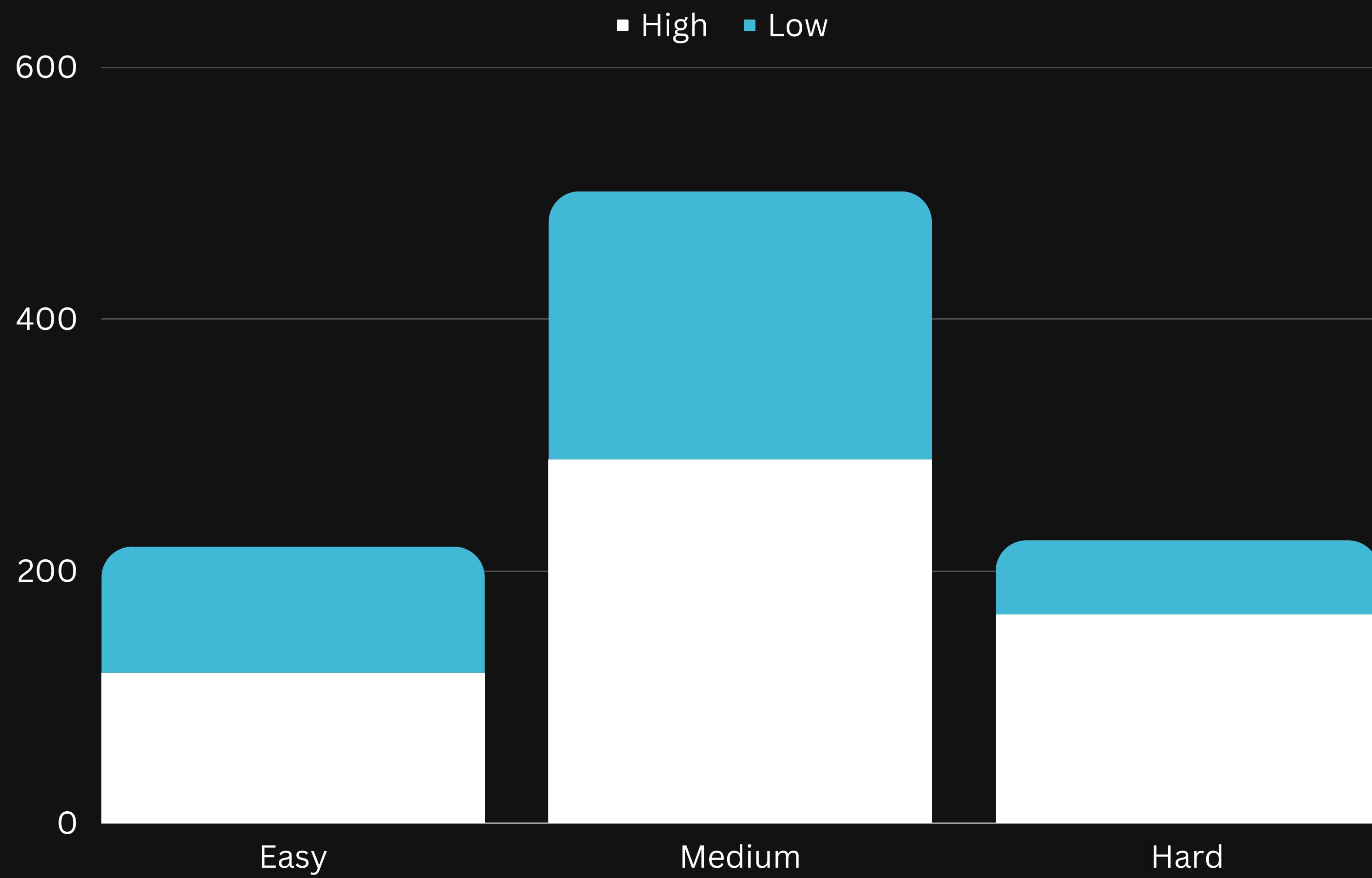


2 out of 4

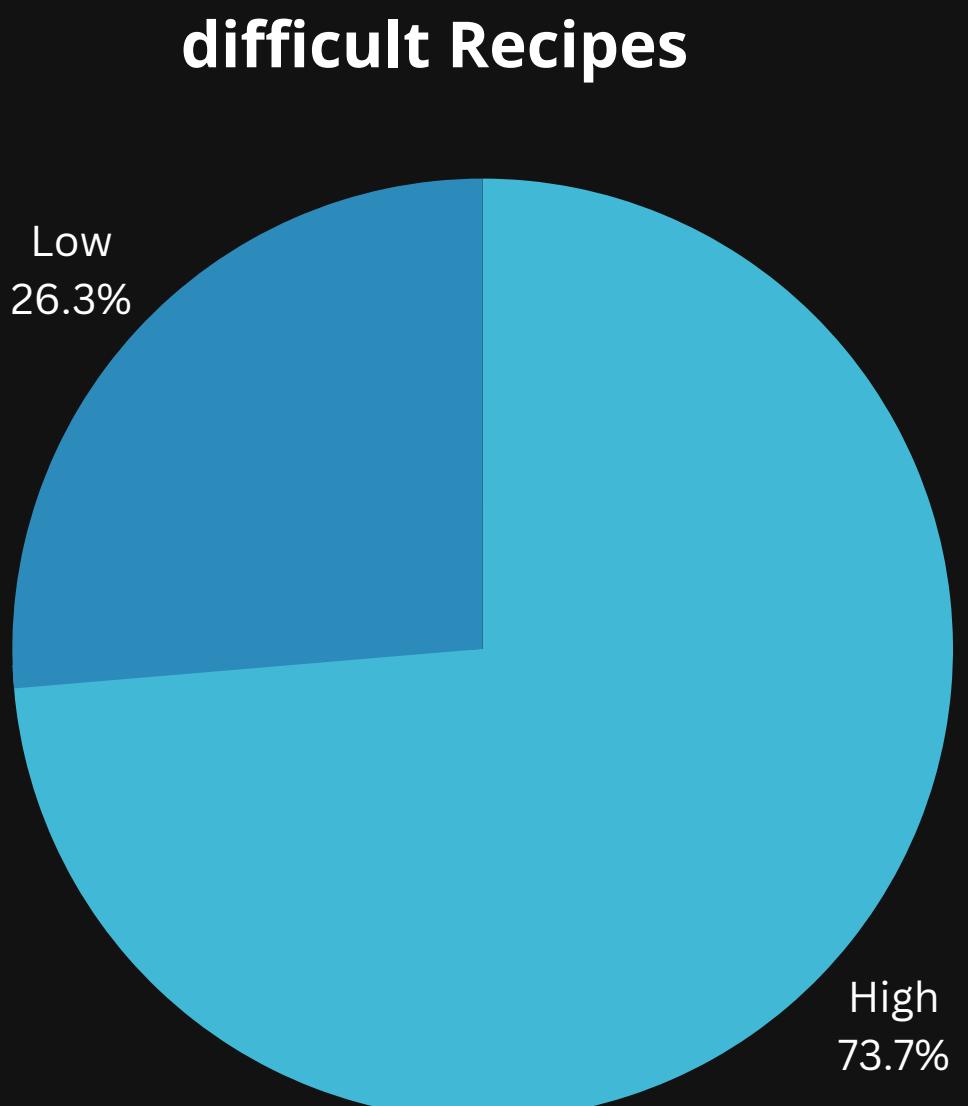
unhealthy recipes produces high traffic

Difficulty Terms

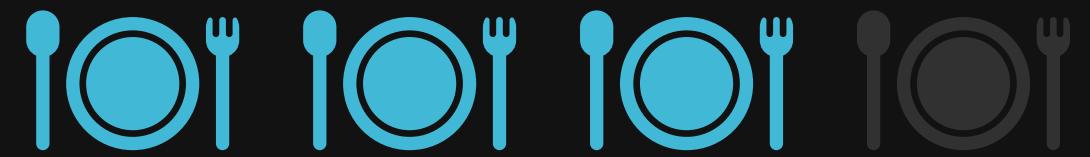
- 1 Category of the food
- 2 each category has a score
- 3 number of servings affects difficulty
- 4 add the score to number of servings
- 5 categorize them into group



- we found 224 hard-making recipes in our data 165 out of them produces high traffic which means 74% of the hard-making recipes generates high traffic
- while only 54% of easy recipes and only 57% of medium recipes generate high traffic
- this a strong indicator that the more difficult the recipe the more high-traffic it is



Recipe difficulty distribution



3 out of 4

Difficult recipes produces high Traffic



2 out of 4

easy and medium recipes produces
high traffic

PREDICTIVE MODEL



Predict popular Recipes

Model Chosen

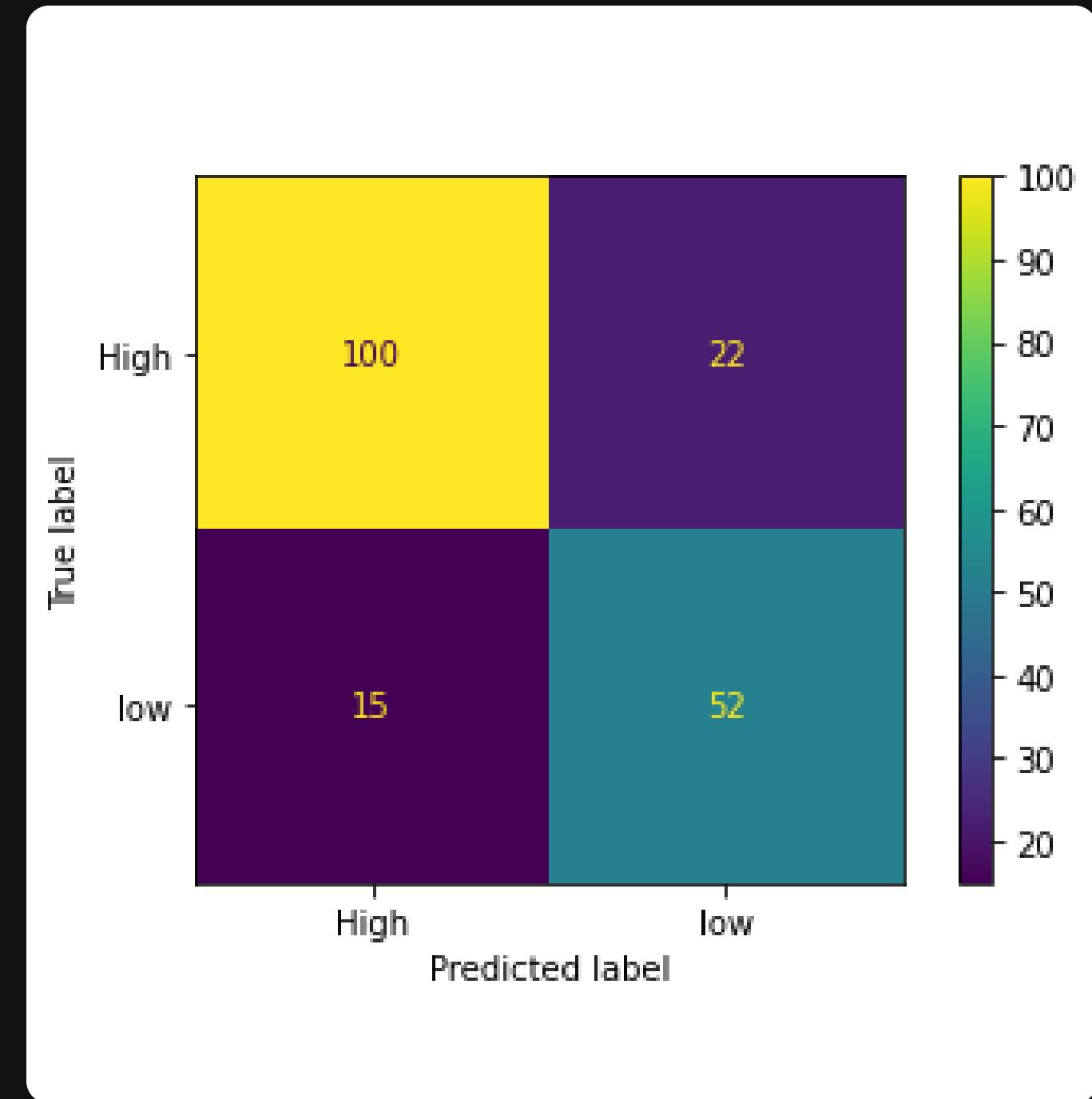
Our goal is to be able to predict the high_traffic recipe 80% correctly so after my Expirements, I chose the Random forest Model as it satisfies our goals

Model perfomance

the model has 80% accuracy on Test data and was able to precisely predict the high_traffic recipe at 78% which is very close to our goal

Model Results

- the model correctly predicted 100 recipes as high traffic and only missed 15 recipes this means the model can accurately identify high-traffic data
- the model only predicted 22 low-traffic recipes as high-traffic which is also a small percentage which means the model can distinguish between low and high-traffic recipes



Final Recommendations

Focus on Healthy Recipes it has a 75% chance to increase the site traffic

reduce certain categories like Beverages and Breakfast and focus on vegetables meat and pork

Hard-making and complicated recipes have a 73% chance to increase the site traffic

The AI model can know the popular recipe 78% of the time that minimize the chance of showing unpopular recipes

Thank you!