Recipe Site Traffic

Michael Sandilands

Business Problem

Currently the Product Manager chooses their favorite from a selection and displays that on the home page.

A popular recipe can increase traffic to the website by as much as 40% if a popular recipe is chosen.

More traffic means more subscriptions, so this is really important to the company.

Can We Do Better?

We expect random chance to correctly pick High Traffic Recipes 63.73% of the time.

The product manager gave us a target of correctly predicting High Traffic Recipes 80% of the time.

Our models can correctly predict High Traffic Recipes **80.42%** of the time on Never-Before-Seen Data, achieving our target!

The Analysis

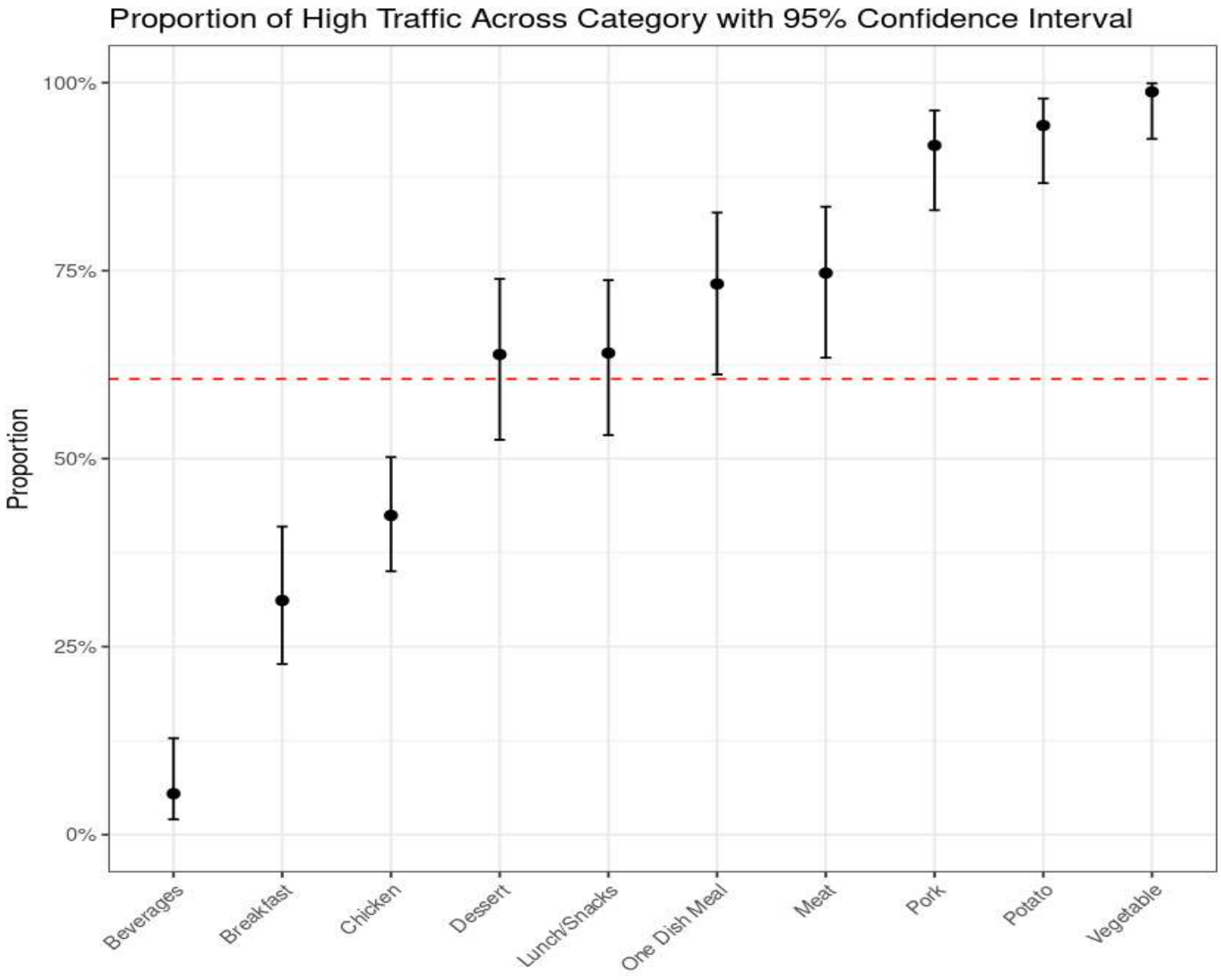
Recipe Category is the most important predictor.

Very popular recipes include:

- Vegetable
- Potato
- Pork
- Meat
- One Dish Meal

Unpopular recipes include:

- Beverages
- Breakfast
- Chicken



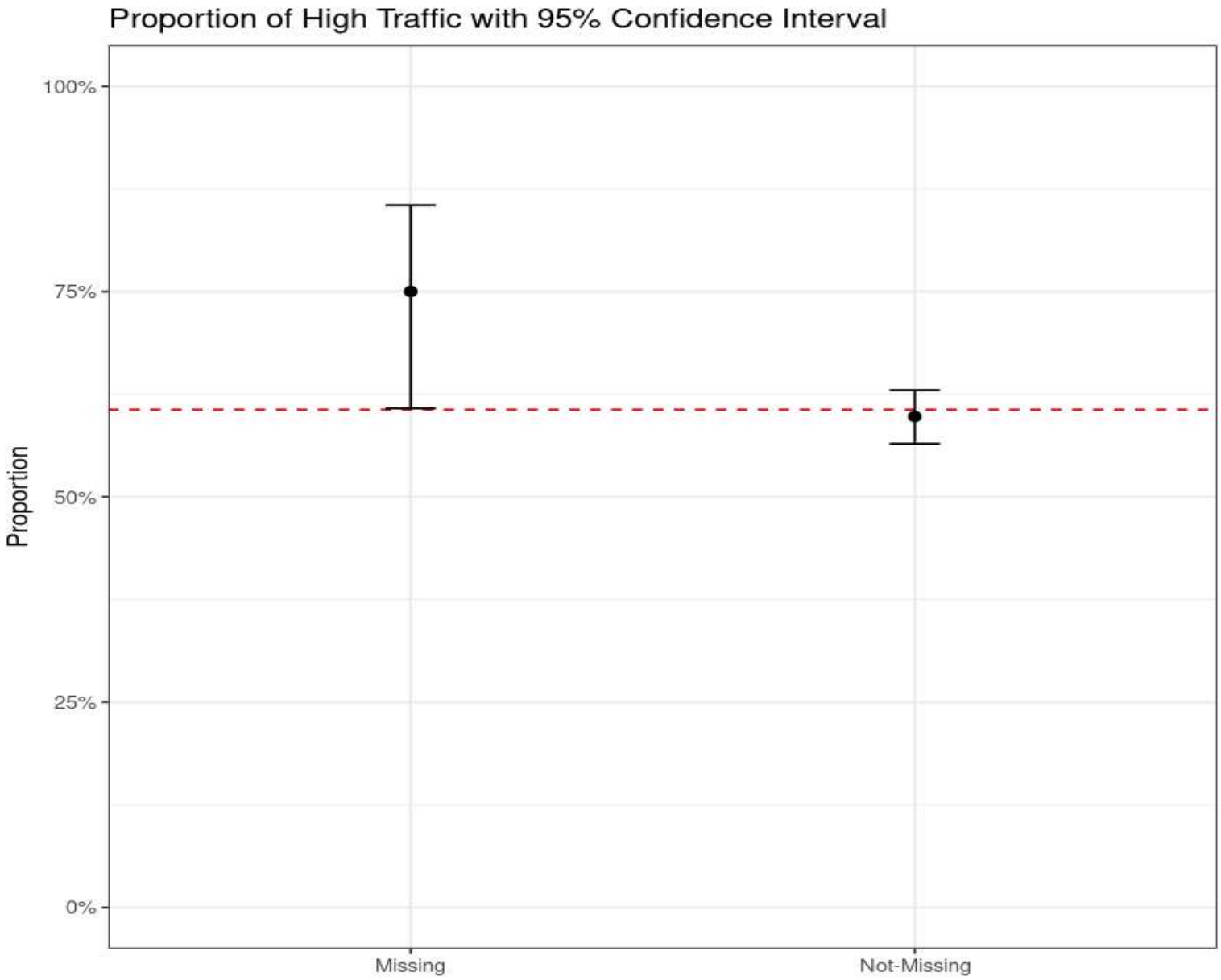
The Analysis

Missing values also have a relationship with High Traffic Recipes

Rows with missing values among the nutritional values, calories, Carbohydrate, Sugar & Protein, tend to have High Traffic Recipes

Further investigation with the product manager into these missing values will need to be pursued

Our best performing model used only these two predictors, Category and Missing



Business Value

How Can Our Models Deliver Business Value?

Explaining Gain

Without the model, we're left to random chance and we'd expect to correctly guess the global High traffic rate which is 63.73%

In other words 6 out of the 10 rows in this table would be correctly guessed.

With our model we first sort the rows by highest to lowest Predicted Probability – then we use the model to guess if a recipe will generate High Traffic.

Our model correctly guesses 10 out of 10 of the first 10 rows.

If we have 600 popular recipes, random chance would have gained us 6 out of those 600 in the first 10 rows (or 1%)

Our model gained us 10 out of those 600 (or 1.7%)

Actual Value	Predicted Value	Predicted Probability
High	High	92.64%
High	High	92.64%
High	High	92.64%
High	High	92.64%
High	High	92.64%
High	High	92.64%
High	High	92.64%
High	High	92.64%
High	High	92.64%
High	High	92.64%

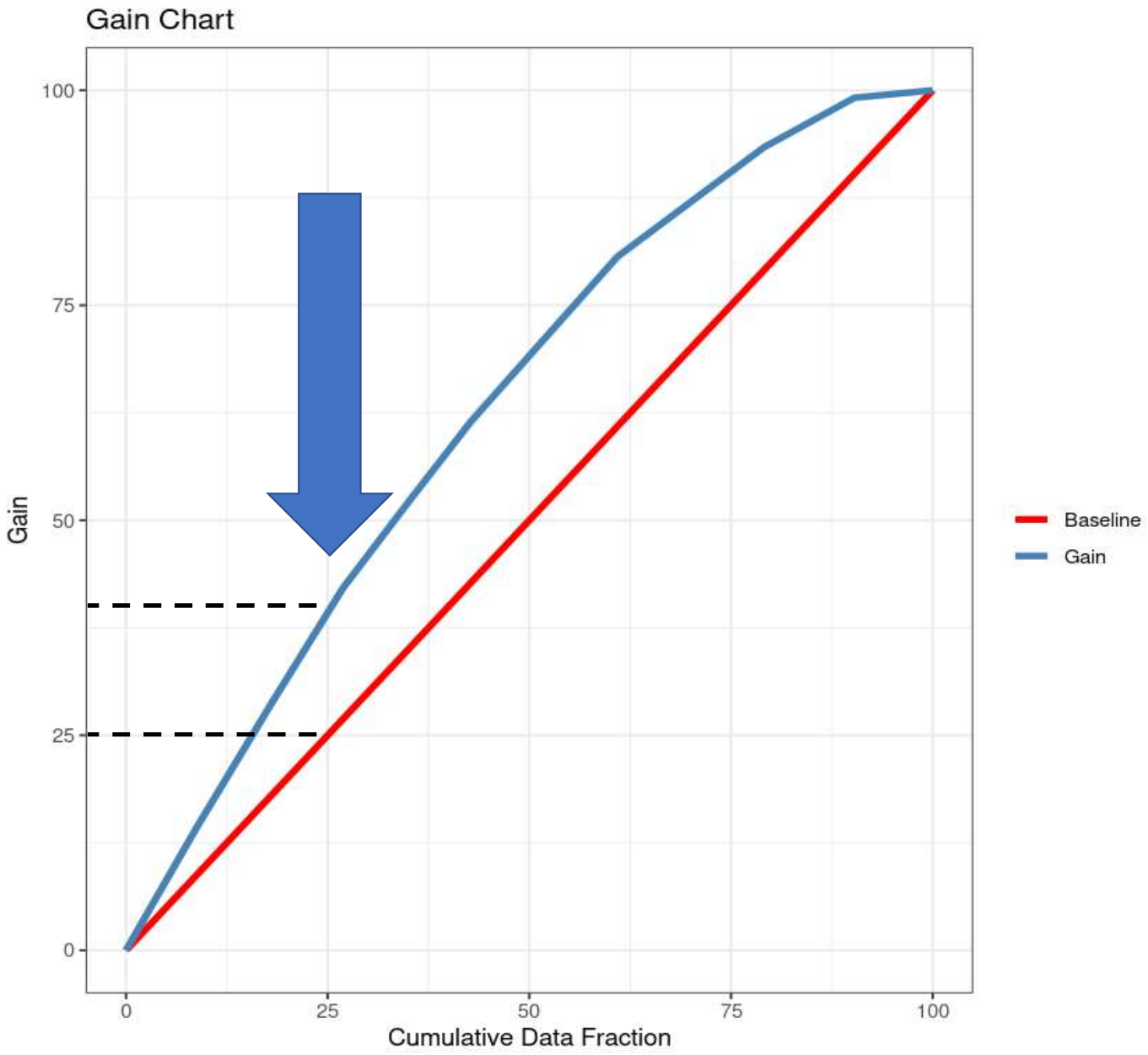
Gain Chart

The gain chart show us how much we gain as we progressively add more recipes.

With our model, the first 25% of our recipes gives us a gain of close to 40%

Meaning, out of 600 High Traffic Recipes, we just gained close to 240!

Whereas if we were using random chance, we would have gained 25% or 150 out of 600.



Recommendations

What Should The Business Do Now?

Improve The Model

Recipes may have more valuable predictors that have yet to be explored.

Predictors here that have yet to be explored are:

- The Title
- Time to Make
- Cost Per Serving
- Ingredients
- Method

Tomato Soup

Servings: 4
Time to make: 2 hours
Category: Lunch/Snack
Cost per serving: $

Nutritional Information (per serving)	
Calories	123
Carbohydrate	13g
Sugar	1g
Protein	4g

Ingredients:

- Tomatoes
- Onion
- Carrot
- Vegetable Stock

Method:

1. Cut the tomatoes into quarters....

Deploy Current Model

We've achieved the Product Managers target.

Our model can correctly predict High Traffic Recipes over 80% of the time on never-before-seen data.

While improvements can be made to our current model

It's ready to start delivering Business Value now!

Time to place our model into production