



Finding what covariates most accurately predict foot traffic in seafood restaurants

Isabel Beaulieu, Easton R. White, Philip Ramsey

isabel.beaulieu@unh.edu

Quantitative Marine Ecology Lab

University of New Hampshire, Department of Mathematics & Statistics,
Department of Biological Sciences

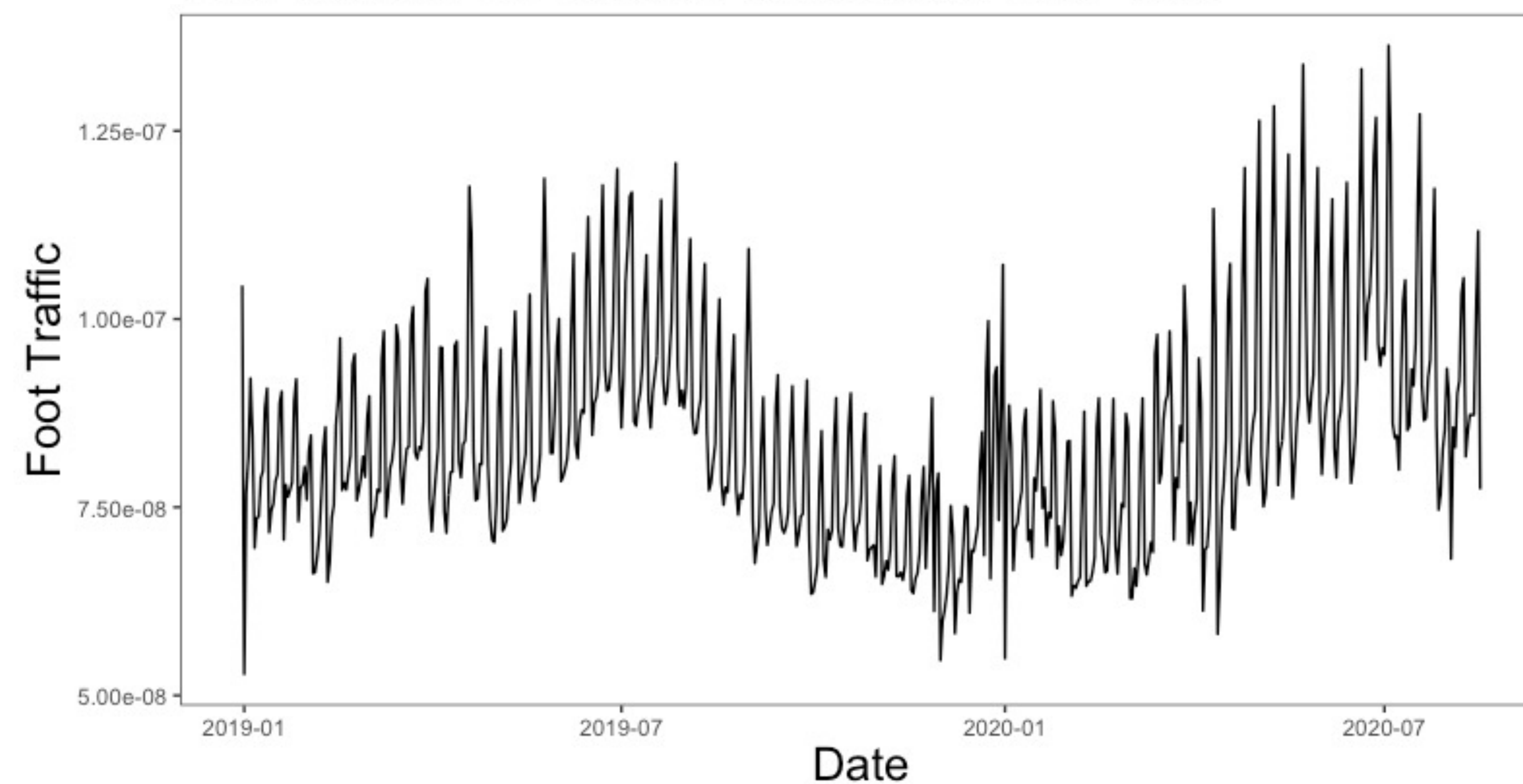


University of
New Hampshire

Introduction

- Foot traffic data is expensive, hard to get, and not available for many months after it is recorded
- Because COVID-19 caused state/nationwide lockdowns and people's lives were altered, the restaurant industry was and still is greatly affected. Foot traffic in restaurants has changed, but could other factors could have predicted this?
- The seafood industry had suffered greatly from COVID-19. There is not as much funding, seafood is perishable, somewhat seasonal in some places, and import/exports were affected from COVID-19

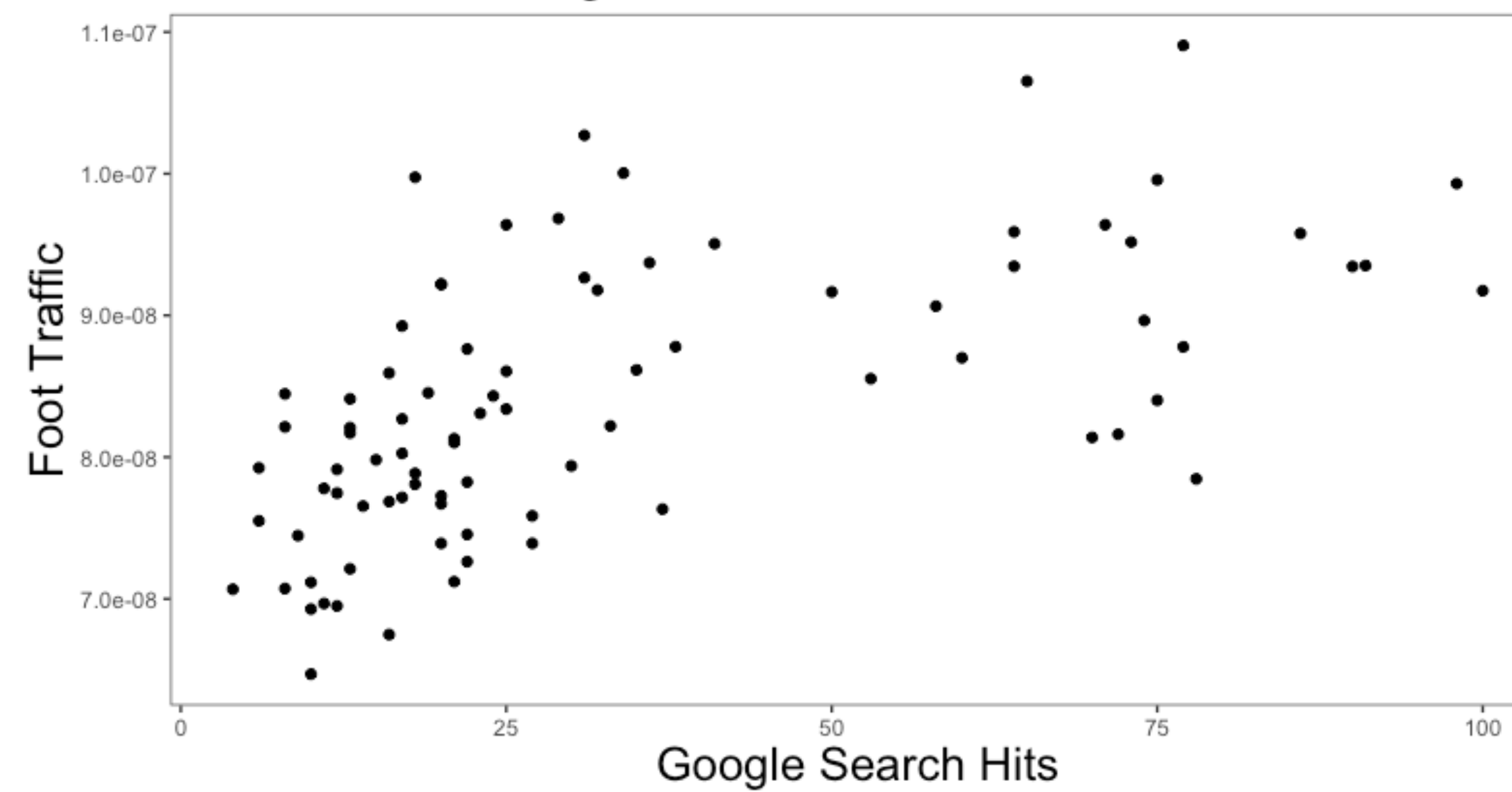
Foot Traffic in US Seafood Restaurants Over Time



Methods

- We fit different types of generalized linear models in R Studio and compared them by mean squared error and Akaike information criterion (AIC)
- Split the data up into test and training sets

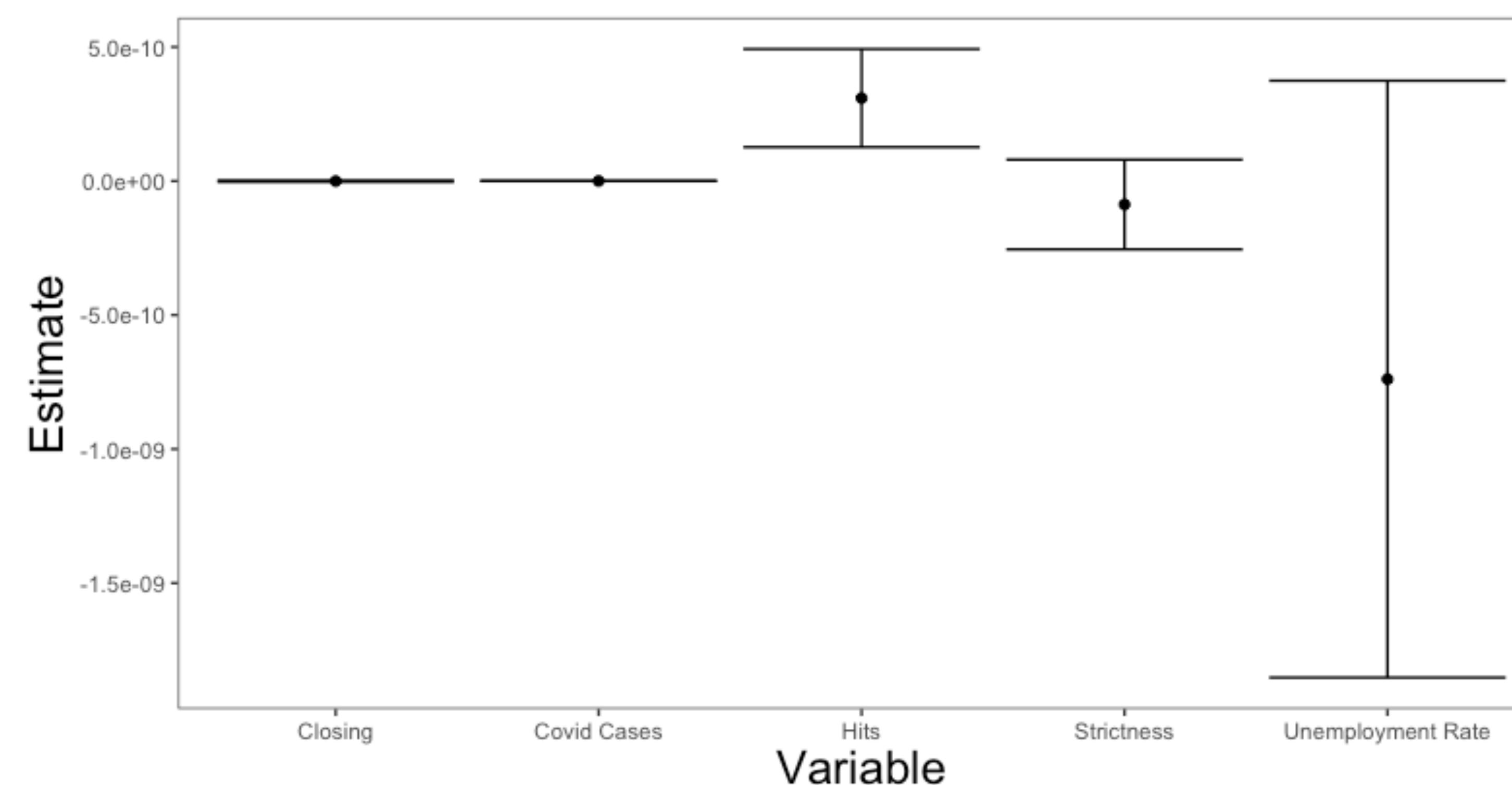
Foot Traffic vs. Google Search Hits



Results

- Google trend hits is most strongly correlated to foot traffic
- The model that had the lowest mean squared error was a simple generalized linear model with no interactions. Covariates were selected using backward selection using AIC
- The variables in the selected model: Google trend hits, unemployment rate, and covid cases

Effect Size of Covariates



Objectives

- 1) Find covariates that are easily accessible that could potentially predict foot traffic for all seafood restaurants in the United States
- 2) Create an optimal model to predict foot traffic using different methods in R
- 3) Test this model on future foot traffic data once it becomes available and make improvements if needed

Model Type	Mean Square Error (MSE)
Hits as the only covariate	8.910448e-17
Best GLM chosen by AIC (Hits, Unemployment Rate, and Covid Cases)	9.187979e-17
Ridge Regression	1.157191e-16
Best GLM chosen by AIC using the previous week's Hits (Previous Hits, Unemployment Rate, and Lockdown)	1107.19
Best GLM using interaction terms	1.429578e-16

Further Directions

- Test the model on current foot traffic data and see if the MSE is similar
- Investigate other covariates that could be useful in predicting seafood restaurant foot traffic in the United States

Acknowledgements

Thank you to Easton White for his support on this project, as well as Philip Ramsey. A big thank you to the members of our Quantitative Marine Ecology Lab for their continued support and help on this project.

