



# PREVENTING FOODBORNE ILLNESS

KNOW BEFORE ITS TOO LATE!

Data used was pulled from  
University of Pittsburgh's  
Tycho Database.

Data scientists utilized the following  
auto-regressive models to accurately  
predict trends for future years.

- Autoregressive Integrated Moving Average (ARIMA)
- Seasonal Autoregressive Integrated Moving Average (SARIMA)
- Seasonal Autoregressive Integrated Moving Average Exogenous (SARIMAX)

When predicting count values, models  
taking into account seasonality (SARIMA  
and SARIMAX) provided the most accurate  
results.

Below are two models created by the  
researchers which predict trends and  
count values for the conditions salmonella  
and campylobacter.

## RESEARCH GOAL

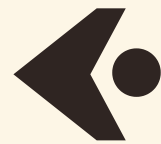
The goal of this research was to predict trends  
and/or potential outbreaks of foodborne  
illnesses in Pennsylvania. If the spread of  
foodborne illnesses can be prevented, there  
would be both public health and economic  
benefits.

## PUBLIC HEALTH IMPLICATIONS

If significant uptakes in foodborne illnesses  
can be accurately predicted, then the proper  
agencies can help to allocate the needed  
resources and reduce any healthcare burden  
before it arises.

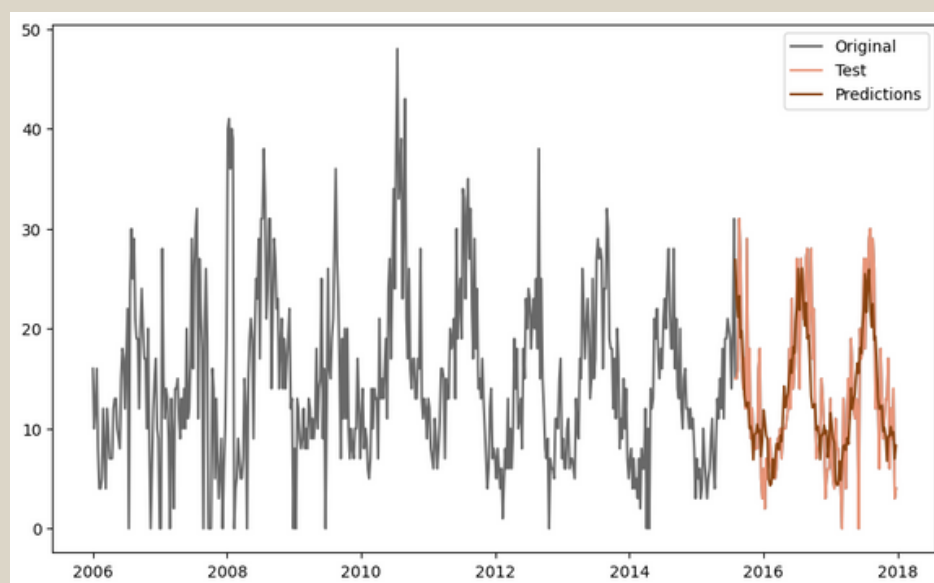
## ECONOMIC IMPLICATIONS

Outbreaks of foodborne illnesses can be very  
costly to both establishments and consumers.  
In 2021, the USDA reported the financial  
burden of the 15 most common pathogens to  
be \$17.6 billion.



## SALMONELLA

- SARIMA Model
- Mean Squared Error (MSE): 24.66
- Accuracy: 83.98%

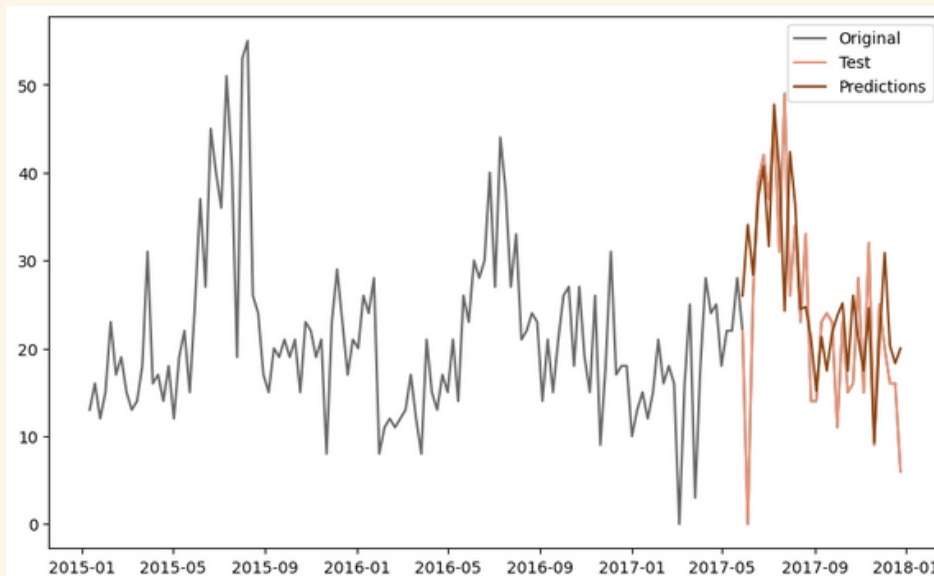


Salmonella is commonly found in  
contaminated food or water. People  
often associate salmonella with raw  
poultry and egg products.

With how close the predictions are  
to the count actual values, this  
model could help prevent up to 1,621  
cases of salmonella each year in  
Pennsylvania (PADOH).

## CAMPYLOBACTER

- SARIMA Model
- Mean Squared Error (MSE): 99.76
- Accuracy: 79.62%



Campylobacter is often associated  
with consuming raw or undercooked  
poultry or consuming anything that  
has been touched by it.

Based on how close the predictions  
are to the actual count values, this  
model could help prevent up to  
2,464 cases of campylobacter each  
year in Pennsylvania (PADOH).