Predicting the Spread of West Nile Virus

For GA-DSI-39
By 39-SIR (Alice, Eugenia, Farhan, Ivan, Sheila)
13 Oct 2023

01

Introduction

Problem
Statement
Background

INTERNAL INFO: Feedback from Proj 3

- Write down and repeat the problem statement throughout presentation
- Remove workflow for industrial presentations
- EDA: Cross refer to third-party data points
- Creating personas for segmentation breakdown
- Any metric flashed out must be defined or provide a baseline

Contents

1. Problem Statement + Background **Eugenia**
2. Data Cleaning + EDA **Alice**
3. Modelling **Farhan**
4. Cost-Benefit Analysis **Alice**
5. Conclusion + Recommendations **Eugenia**
6. Demo **Sheila**

Problem Statement

Team **39 SIR** of the **D**isease **A**nd **T**reatment **A**gency, division of **S**ocietal **C**ures **I**n **E**pidemiology and **N**ew **C**reative **E**ngineering (DATA-SCIENCE) is to effectively plan the deployment of pesticides in order to mitigate the spread of the West Nile Virus in Chicago City.

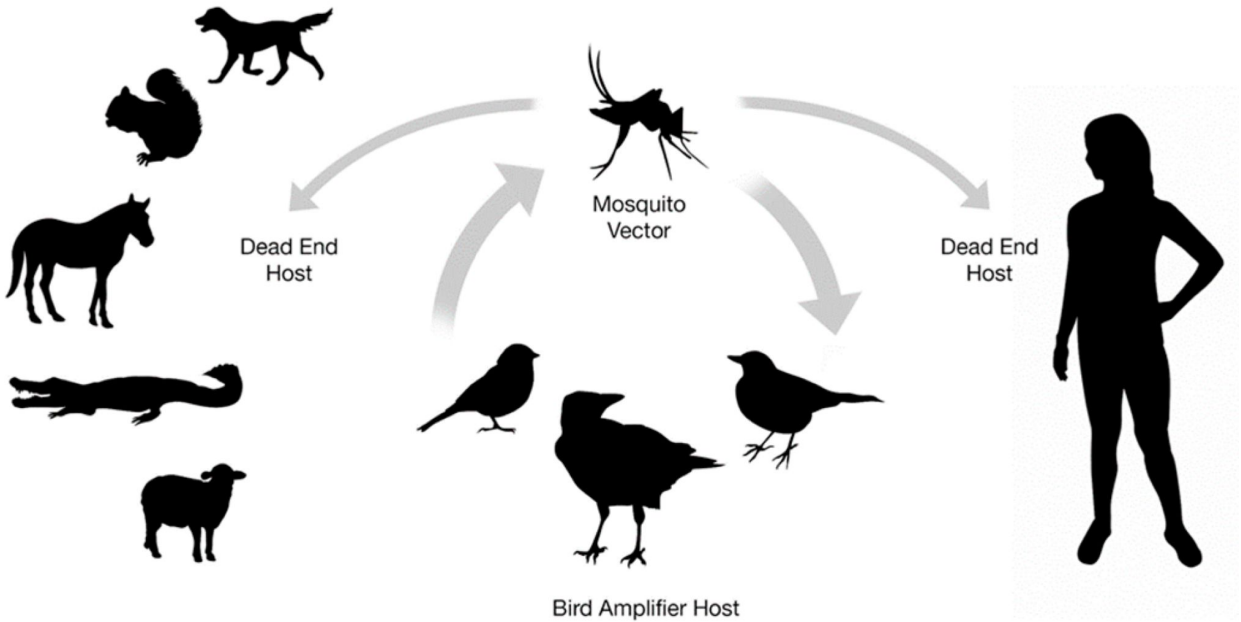
This will be done by analysing the data given by the Department of Public Health to produce the following deliverables:

1. A **predictive model** to facilitate informed decision making by the city of Chicago when it decides where to spray the pesticides.

2. **Cost-Benefit Analysis** of the annual cost projections for various levels of pesticide coverage (cost) and the effect of these various levels of pesticide coverage (benefit).

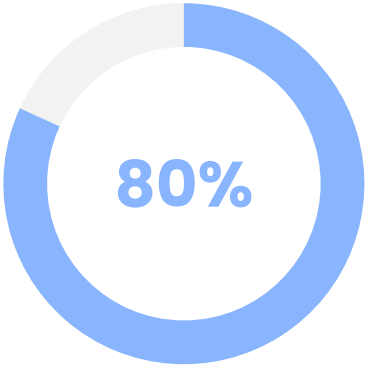
What Is the West Nile Virus (WNV)?

Most common mosquito-borne disease in the US

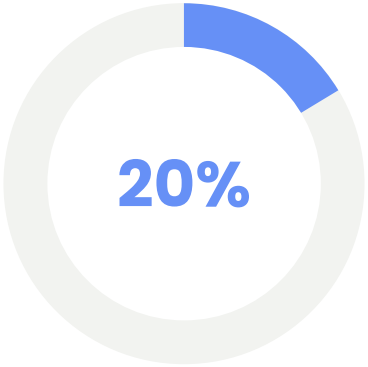


Source: https://www.mdpi.com/2076-0817/9/1/48

Severity of West Nile Fever (WNF)



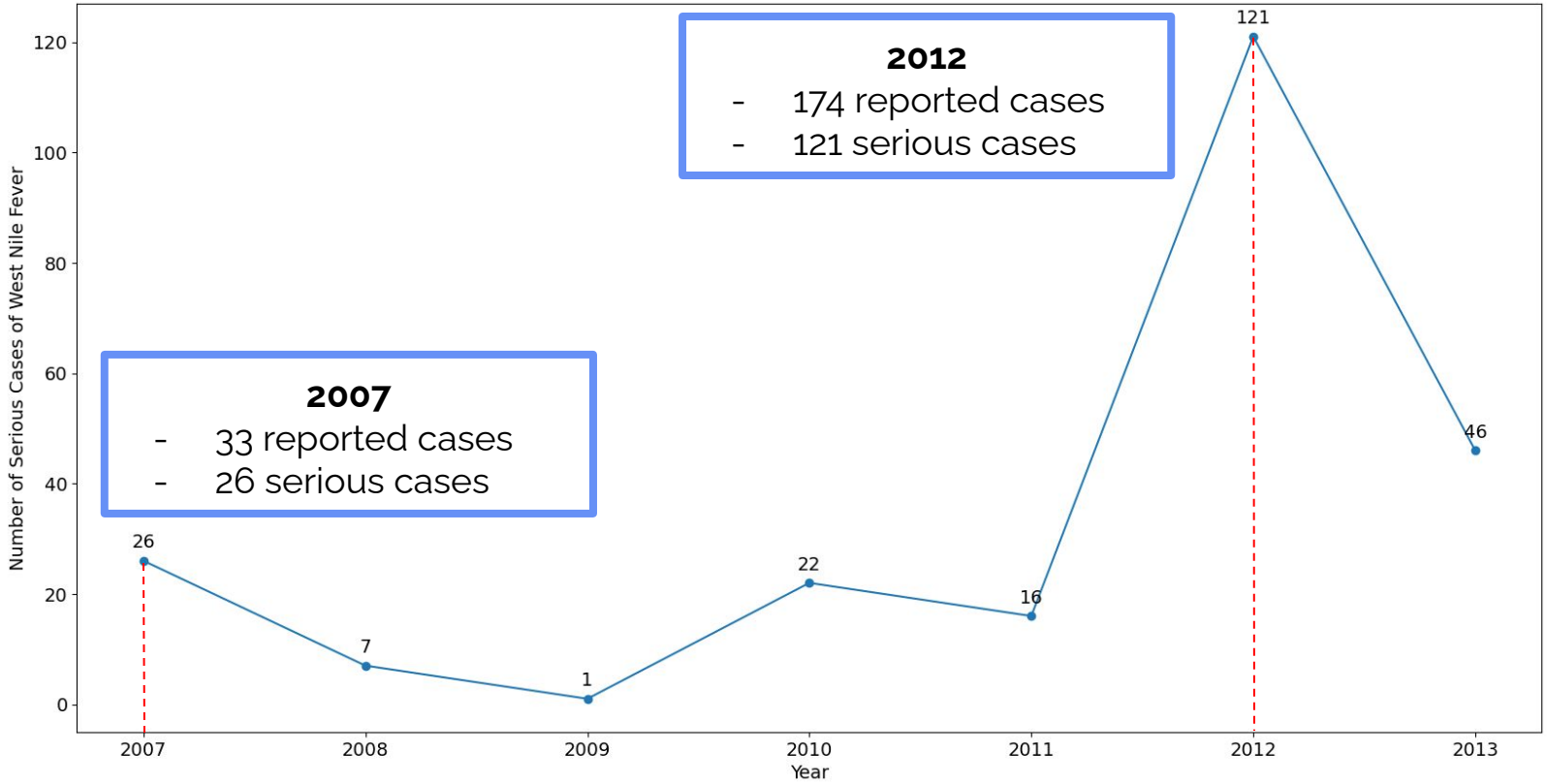
Asymptomatic



Developed Symptoms - e.g. Fever

Developed Severe Illness

Outbreak of WNF in Chicago in 2007 and 2012



02

Data Cleaning

ALICE

Overview of Data Set

- Which years were provided for train data? Years: 2007, 2009, 2011, 2013

- How many mosquito traps were set each year? 2007: 1459, 2009: 1006, 2011: 988, 2013: 1163 [Increased trapping in 2007 and 2013]

- When was spraying done? Dates: August 2011, September 2011, July 2013, August 2013, September 2013

- Traps set in Chicago City Centre (Cook County)

Data Cleaning Approach

Train data:

- Clean data, address and mosquito species are the categorical variable.
- Mosquito species are broken down into 6 subcolumns.
- Date is converted to datetime format

Weather data:

- Most data are objects and have to be converted to float
- Alphabets in ‘tavg’ is replaced by another station’s temperature

Spray data:

- Time is dropped

Test data:

- Treated similarly with train data

Data Cleaning Steps

Challenges in cleaning the following data:

Train data:
Clean data, address and mosquito species are the categorical variable. Mosquito species are broken down into 6 subcolumns.
Date is converted to datetime format

Weather data:
Most data are objects and have to be converted to float. Alphabets in ‘tavg’ is replaced by another station’s temperature

Spray data:
Time is dropped

Test:
Treated similarly with train data

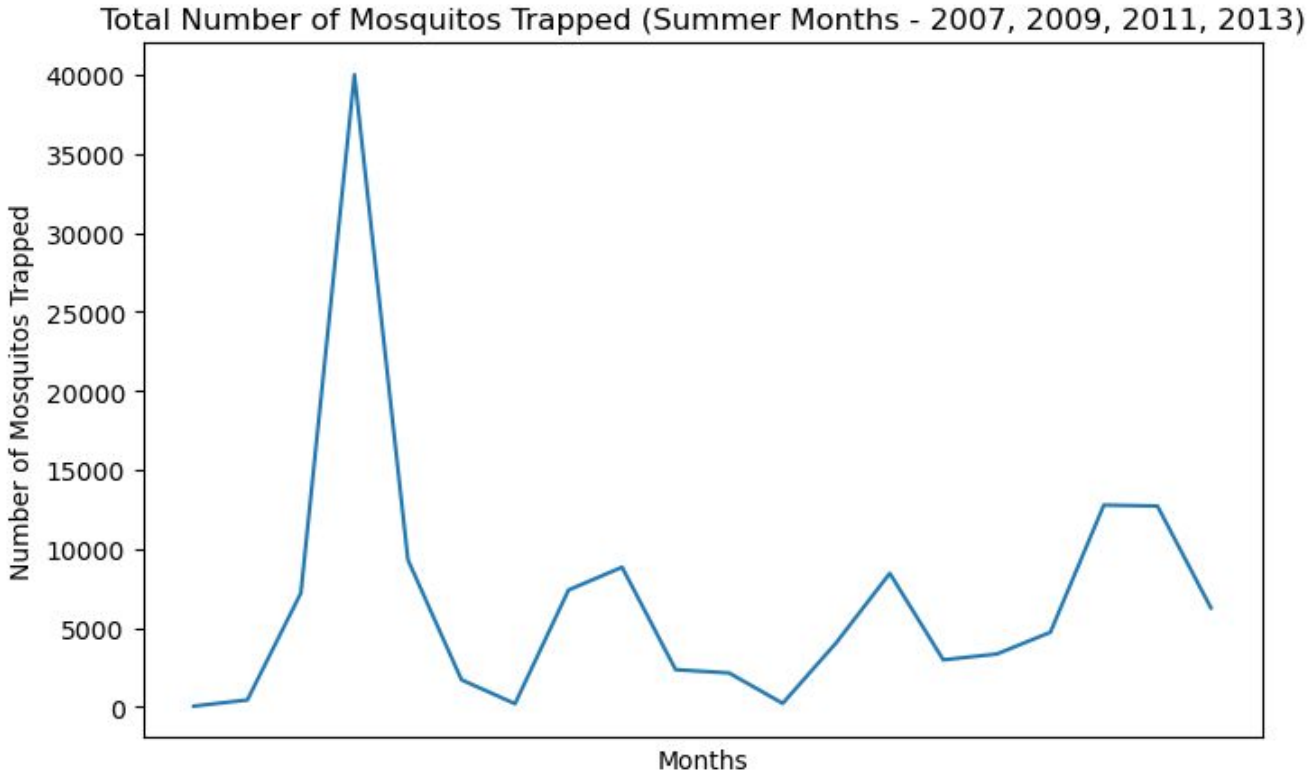
03

EDA

Exploratory Data Analysis

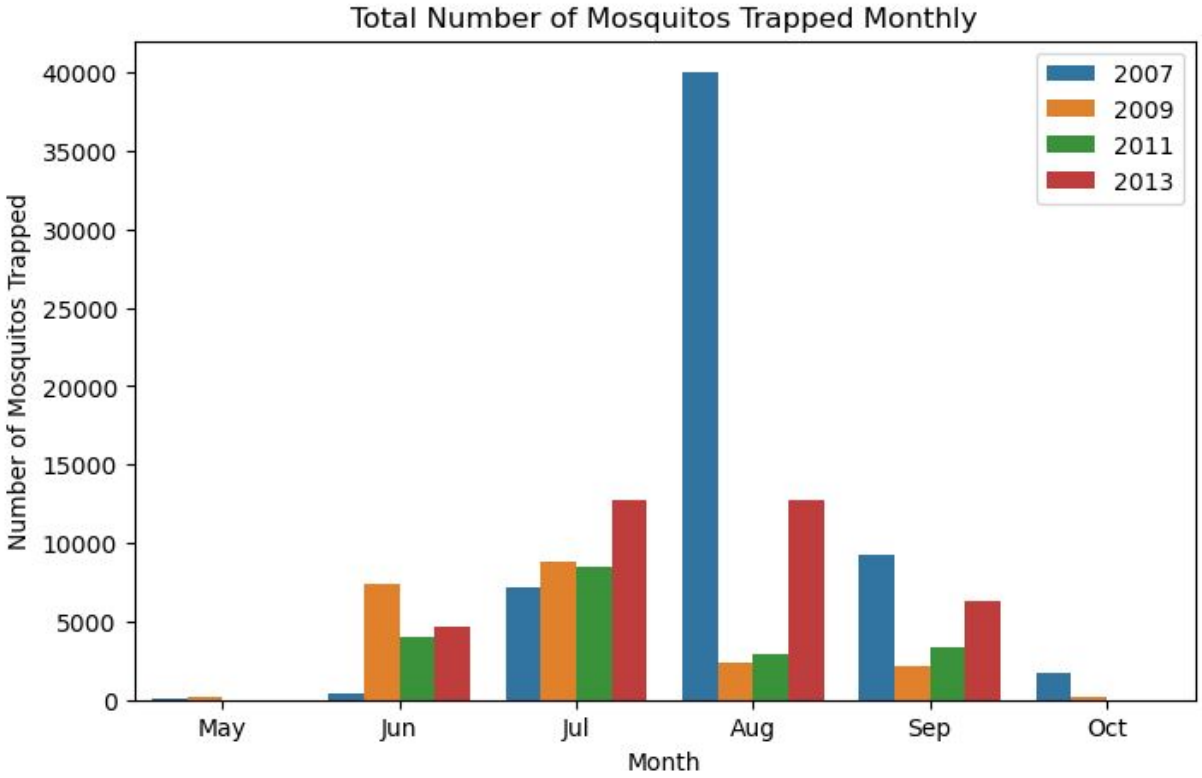
Peak in Mosquito Activity in 2007

May 2007 - Sep 2013

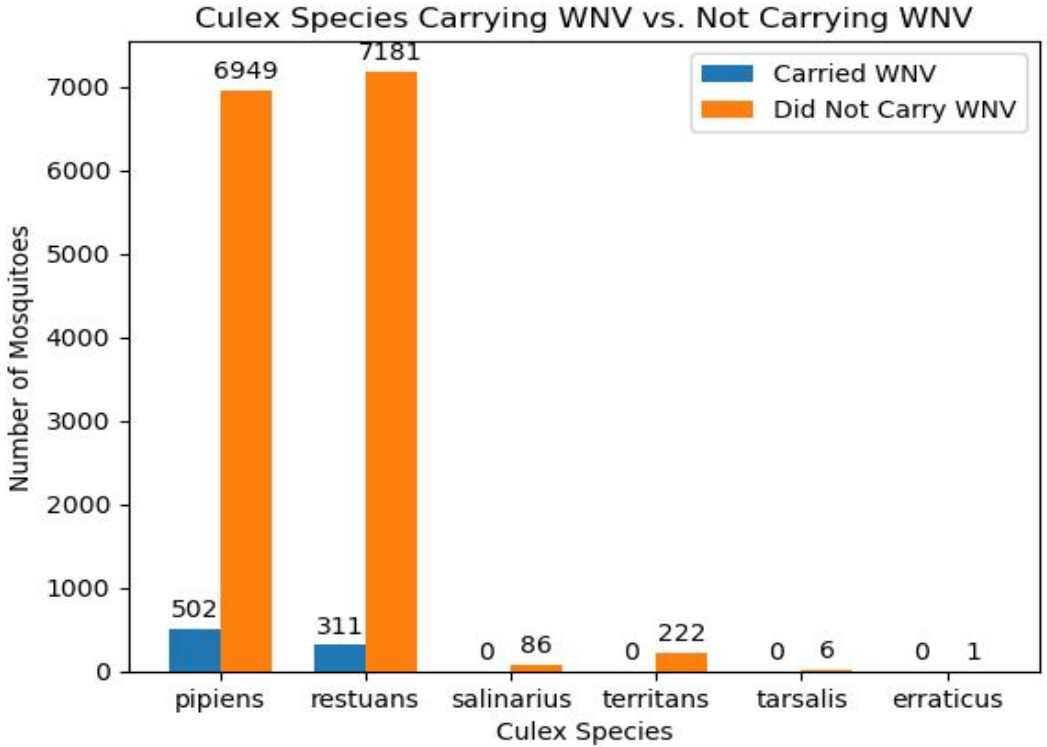


Summer Trends: July and August as Mosquito Season

May 2007 - Sep 2013



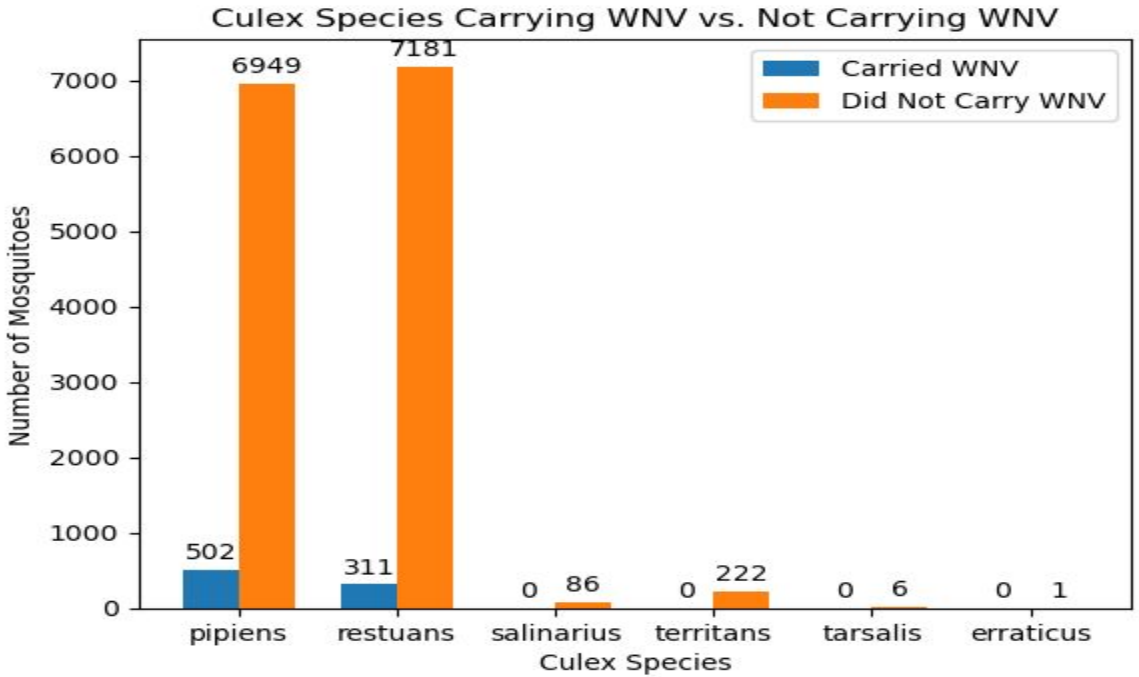
Only 2 Culex Species as WNV+ Carriers



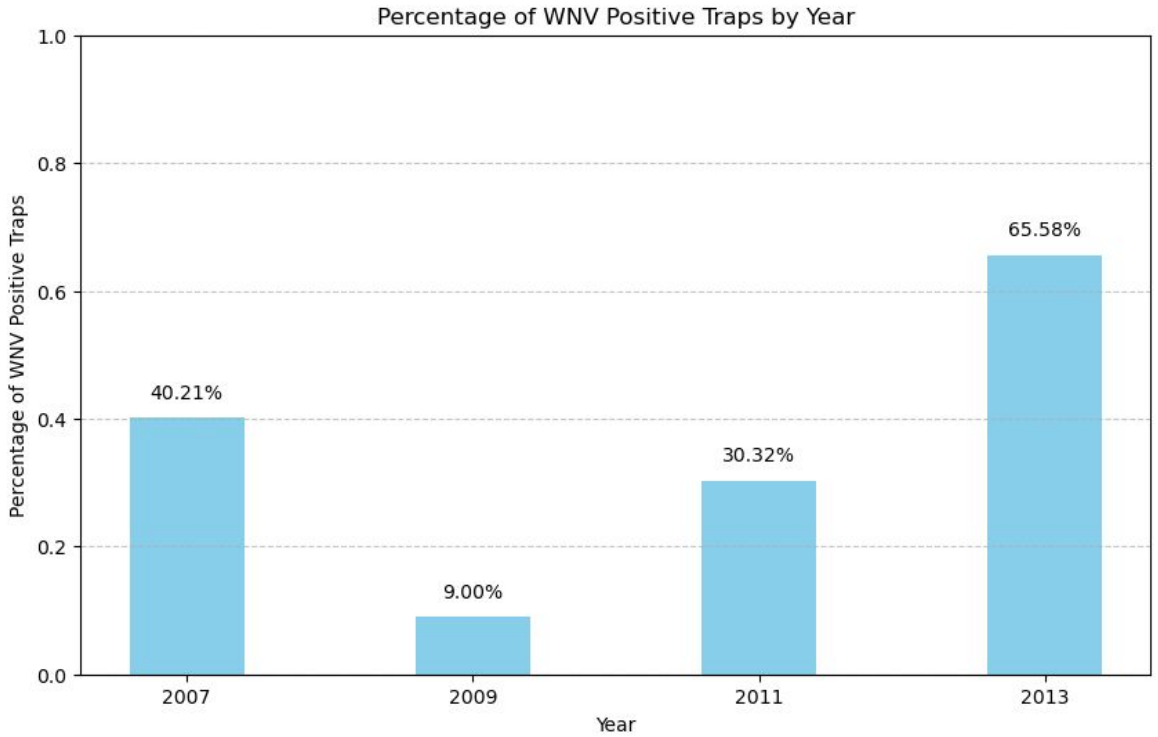
WNV cases by species

ALICE

- Which culex species are likely to be carriers?
  - May also wish to show that most culex do not carry WNV



High incidence of WNV+ traps in 2013



04

Modeling

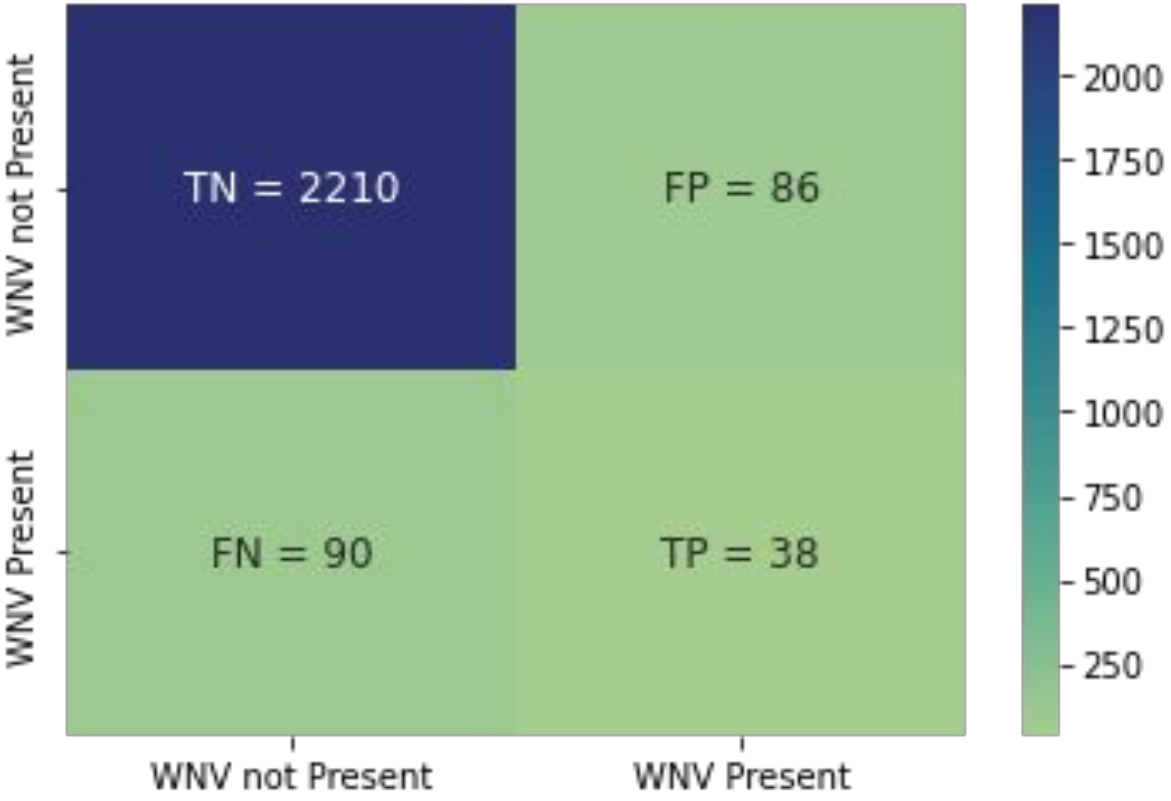
Modeling Process

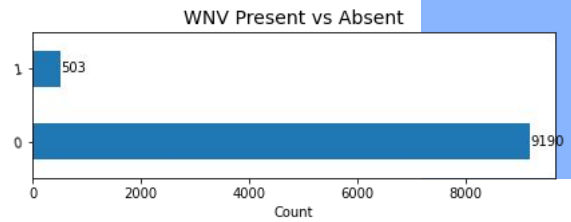
Rapid Modeling	Pre-processing	Hyperparameter Tuning	Analysis
Lazy Predict • LazyClassifier (25 models)	**Pipeline** • StandardScaler • SMOTE • XGBoost	• GridSearch	• Confusion Matrix • Top Predictors • Final score

Lazy Predict Top 5

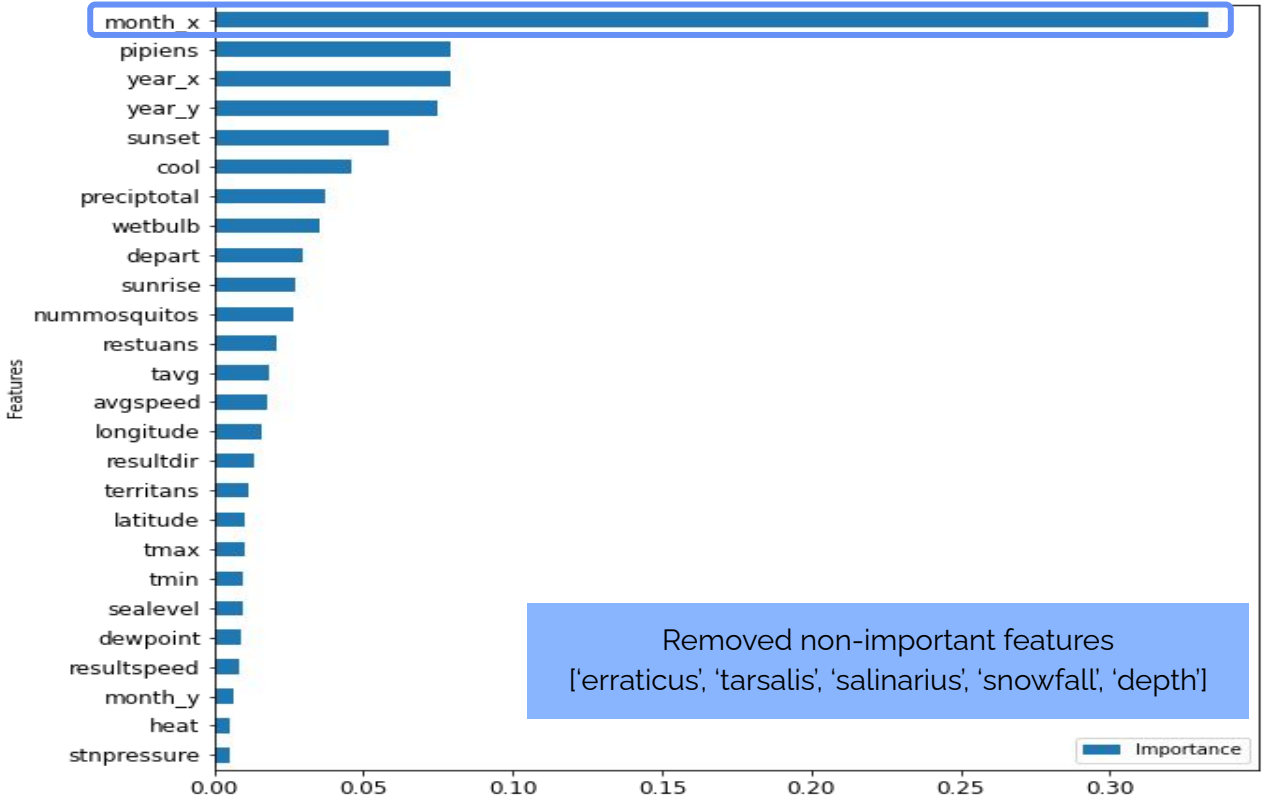
Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
LinearSVC	0.951	0.500	0.500	0.922	0.214
XGBClassifier	**0.950**	0.585	**0.585**	0.931	**0.023**
SVC	0.947	0.500	0.500	0.922	0.618
RidgeClassifierCV	0.947	0.500	0.500	0.922	0.030
DummyClassifier	0.947	0.500	0.500	0.922	0.008

Confusion Matrix





Top Predictors



Removed non-important features
['erraticus', 'tarsalis', 'salinarius', 'snowfall', 'depth']

Final Score

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
XGBClassifier	0.950	0.585	0.585	0.931	0.023
XGBClassifier (after GridSearch)	0.996	0.996	**0.999** 70.8% improvement	0.962	0.290

05

Cost Benefit Analysis

Cost-Benefit Analysis

The Pros and Cons of Mosquito Spraying: Helpful or Hazardous?

Mosquitos are the world's deadliest animal.

1 million deaths yearly due to mosquitoes

spread life-threatening diseases like malaria, yellow fever, and dengue.

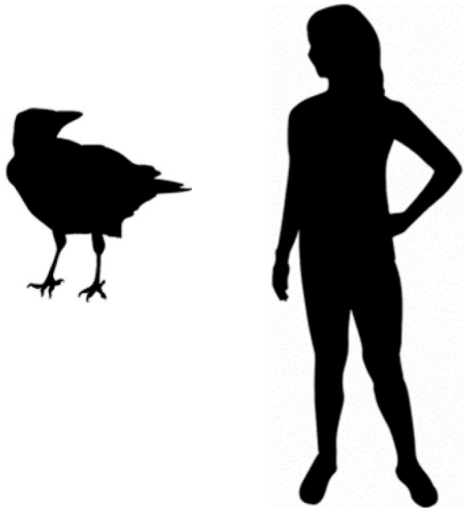
Mosquitoes outnumber humans by about **16,000 to 1**

Reference: The Pros and Cons of Mosquito Spraying: Helpful or Hazardous? - Tabor Pest Control

Monitoring and Controlling Spread of WNV:

WNV Carrier:
Monitor WNV presence by trapping mosquitoes for testing

WNV Hosts:
Monitor infections through testing and reporting



Larval Mosquito
- Larvicide in drains
- Eliminating stagnant water

Adult Mosquito
- Spraying of insecticides

Types of Mosquito Control

Biological control is used to eliminate existing larvae and adult mosquitoes.

- Introduction of predators such as birds, fish, and frogs has been shown to have some impact on controlling mosquito populations.
- Larviciding introduces low concentrations of chemicals to still water to kill off mosquito larvae and pupae without impact on other water-inhabiting organisms.
- Adulticiding uses trucks and aircraft to spray pesticides and kill mosquito populations.

Analysis of Mosquito Spraying

Pros:

1. effective, accessible, and localized
2. can significantly reduce mosquito populations in treated areas.
3. fairly quick and easy process

Cons:

1. short-term solution that fails to address the source of the issue
2. negative impact on non-targeted insects like bees
3. prolonged exposure to insecticides can lead to respiratory issues, skin irritations

Analysis of Mosquito Spraying

Pros:

- Effective, accessible, and localized
- Can reduce mosquito populations in treated areas
- Fairly quick and easy process

Cons:

- Short-term solution
- Impact on other wildlife - e.g. bees
- Prolonged exposure to insecticides can lead to respiratory issues, skin irritations

Cost-Benefit Analysis

What are the Pros and Cons of Mosquito Spraying?

Pros:

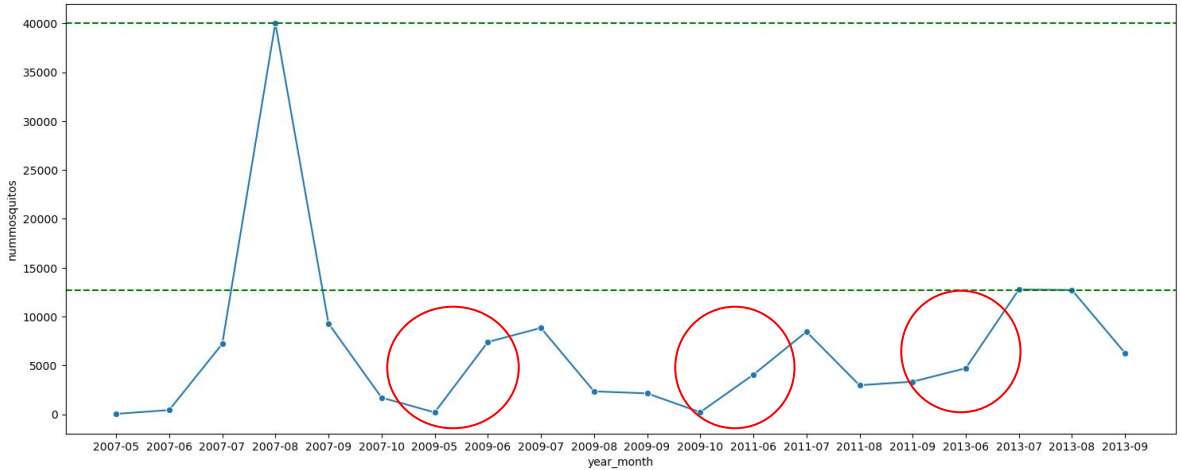
1. effective, accessible, and localized
2. can significantly reduce mosquito populations in treated areas.
3. fairly quick and easy process

Cons:

1. short-term solution that fails to address the source of the issue
2. negative impact on non-targeted insects like bees
3. prolonged exposure to insecticides can lead to respiratory issues, skin irritations

Cost-Benefit Analysis

Should Spraying of Insecticide be implemented?
If so, how much will it cost?

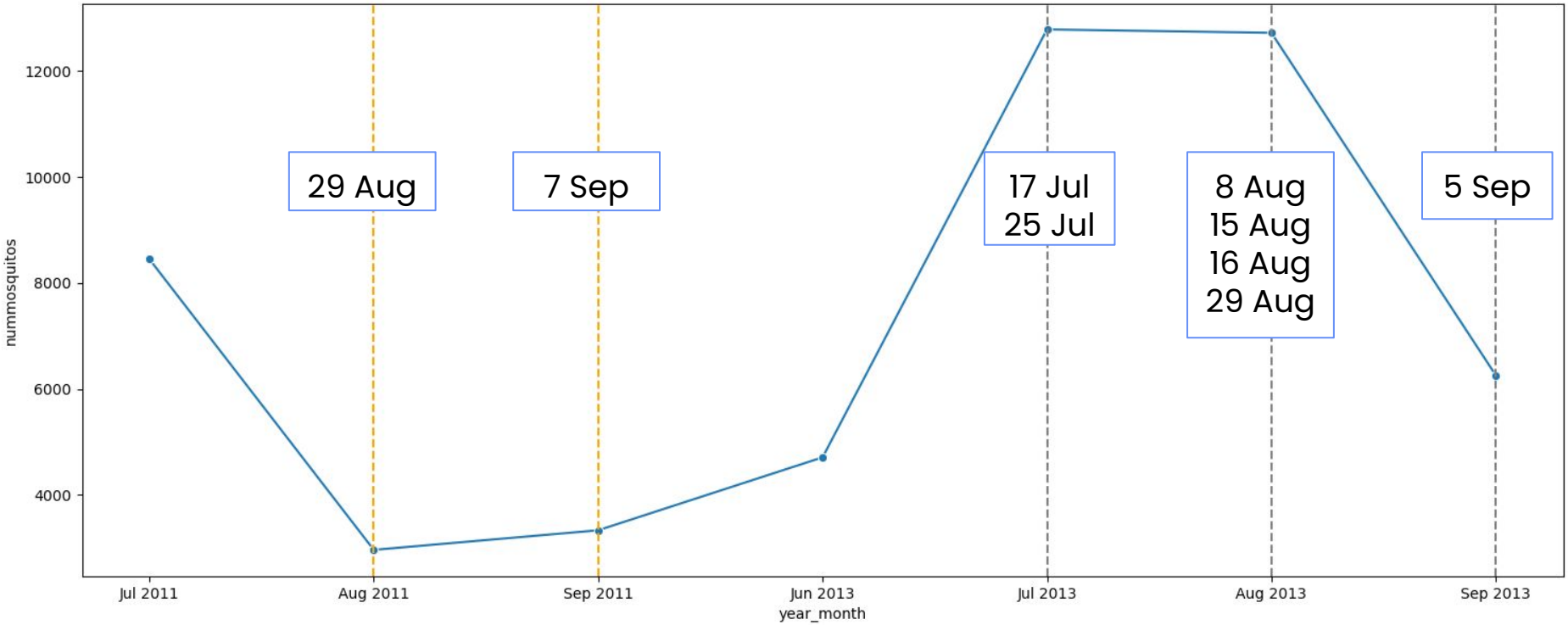


Aggressive Spraying Reduces Mosquito Population …

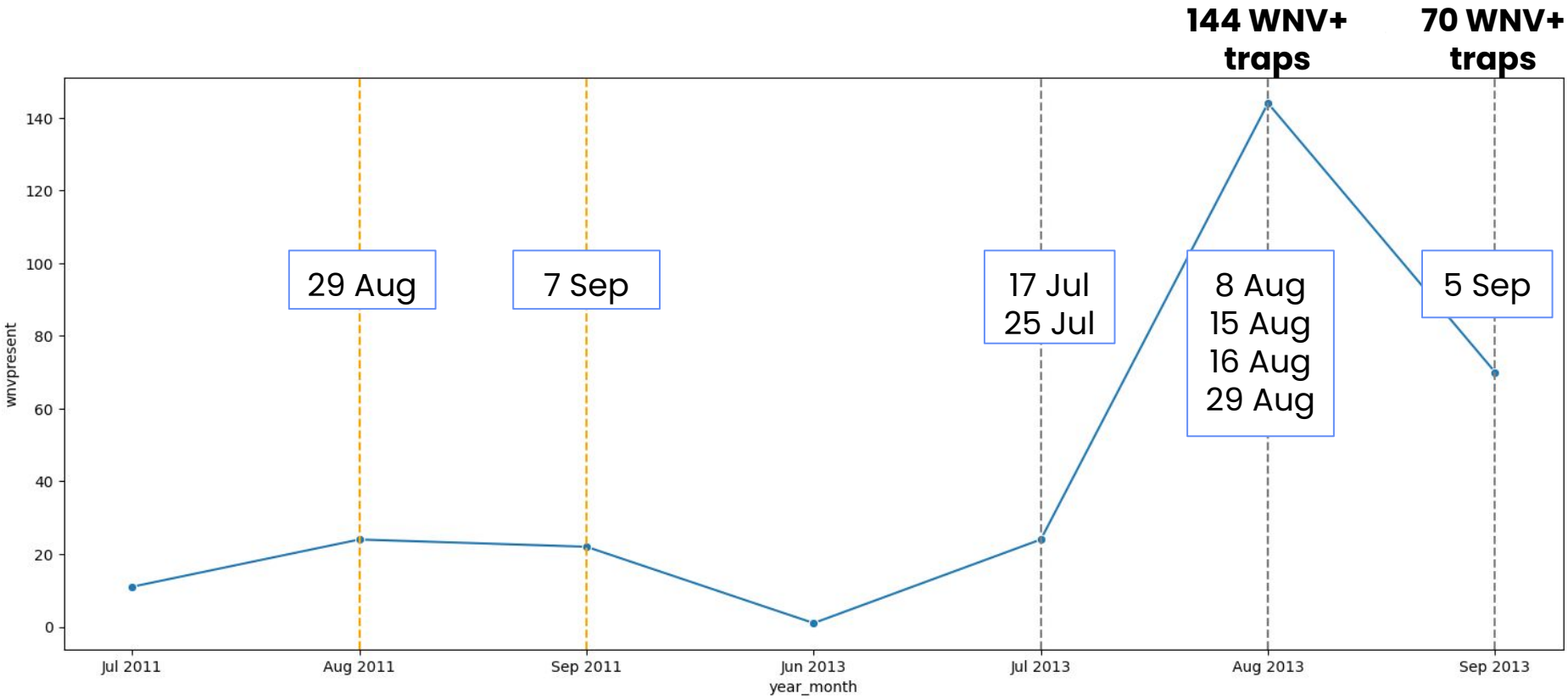
somewhat

12,707 mosquitos

6,255 mosquitos



Aggressive Spraying Reduces WNV+ … somewhat



Cost-Benefit Analysis

136 traps

Highest number of mosquitoes in the months of July and August

Factors That Affect the Cost of Spraying for Mosquitoes

1. size of the area to be treated

2. type of pesticide used

3. number of treatments required

4. Presence of standing water

5.The local mosquito population

Cost-Benefit Analysis

Assuming one treatment costs $500 per area and Zenivex used for adulticiding

It will cost $500 x 136 = $68000 for 136 locations.

If it is sprayed for months of July and August, it will cost $136 000.

Reference: How Much Does It Cost To Spray For Mosquitoes? Get the Answers Here! – bugpursuits.com

https://www.chicago.gov/content/dam/city/depts/cdph/Mosquito-Borne-Diseases/Zenivex.pdf

Proposed Measures

$1.46M saved
16% cost savings

Larvicide

Current Cost:
~$9,000,000 for 90,000 drains

Proposed Cost:
~$7,560,000 for 75,600 drains

Note: Found as 84% effective

Change: To spray when upward trend spotted

→ Cost Reduction by **$1,440,000**

Spraying

Current Cost:
~$21,730 for 11 occasions

Proposed Cost:
NA: to minimise spraying as it is **resource-intensive** to be effective

→ Cost Reduction by **$21,730**

Note:
*Larvicide (Fourstar XR®) - Placed in 90,000 catch basins but only 84% effective
*Adulticide (Zenivex ®) - Used in 11 occasions; USD$0.67 per acre; 1.5 fluid ounces per acre - no significant risk to the residents
*Benefit/Cost ratio - https://www.sciencedirect.com/science/article/pii/S0048969720313127

Evaluation – 2013

$649,658 saved
16% cost savings

Current Measures

*Larvicide (90,000 catch basins) 84% effective	$9,000,000
*Adulticide (11 occasions)	$21,730
Benefit *(Benefit/Cost ratio = 1.8)	$5,012,072
Total cost	$4,009,658

Proposed Measures

Larvicide (~75,600 catch basins, spray whenever an upward trend is spotted)	$7,560,000
Benefit (Benefit/Cost ratio = 1.8)	$4,200,000
Total cost	$3,360,000

Note:
*Larvicide (Fourstar XR®) - Placed in 90,000 catch basins but only 84% effective
*Adulticide (Zenivex ®) - Used in 11 occasions; USD$0.67 per acre; 1.5 fluid ounces per acre - no significant risk to the residents
*Benefit/Cost ratio - https://www.sciencedirect.com/science/article/pii/S0048969720313127

Proposed Measure

- Trap:
  - Use traps to monitor mosquito activities May to October
  - 70F - start to monitor regularly
- Larvicide:
  - To be applied only when mosquito numbers are increasing
  - 16% of the 90,000 basins applied did not stop mosquito breeding
- Adulticide:
  - Avoid this measure as it has been ineffective (6 more cases in 2013 than in 2012)
- More preventive measures that alleviate mosquito breeding:
  - Run campaigns more rigorously during July and August to advocate mosquito prevention at household level

06

App Demo

App Demo

Chicago West Nile Virus Control
About
Trends
Surveillance
Control Measures
About
West Nile Virus
West Nile Virus (WNV) is a mosquito-borne disease that can infect humans and animals, leading to various symptoms, including fever and, in severe cases,
complications. It is transmitted through mosquito bites, and the virus has been reported in several regions worldwide. Vigilance against WNV is crucial as it can
issues, and preventing mosquito breeding and taking protective measures, such as using mosquito repellent, can reduce the risk of infection.
In Chicago
Chicago is grappling with a severe West Nile Virus (WNV) situation, standing as the second-worst-hit city in the United States. The presence of the Culex mosc
carrier of the virus, has significantly contributed to the high number of WNV cases in the area. With its dense population and ideal conditions for mosquito breed
transmission has escalated.
The city's residents face a persistent threat to their health due to the alarming spread of this mosquito-borne illness. WNV has caused numerous cases of fever,
instances, neurological complications, leading to a considerable health burden in the city.
Fork this app
BiteBalance
By 39 SIR
Insights
Analyse User
Analyse Text
Reddit User Classifier
You can use this app to classify a Reddit
Enter Reddit user's name:
Select number of posts to scrape:
100
1
200
Submit

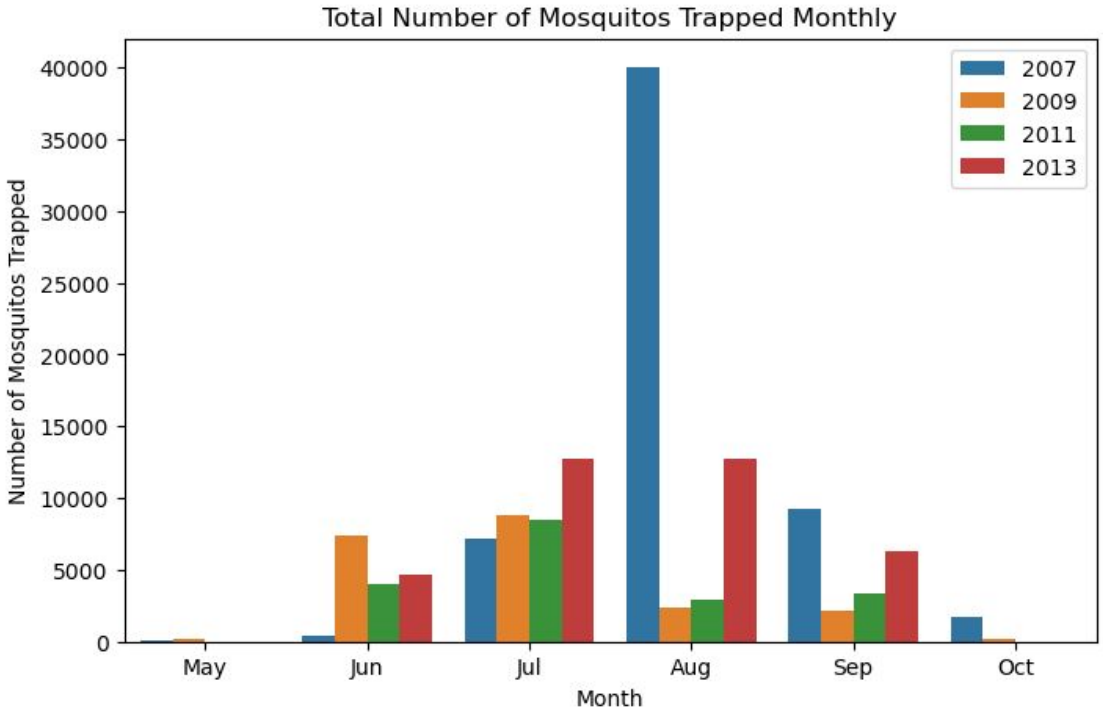
07

Conclusion

Recommendations

- Surveillance and monitoring through Geographic Information Systems (GIS) to analyse hotspots
- Establish an early warning system for detecting West Nile virus activity in mosquitoes and birds. (?)
- Introducing natural predators of mosquito larvae, such as certain species of fish and copepods, into bodies of water
- More rigorous testing during peak months (July and August)
- Public awareness

When



Traps

Monitor → Action

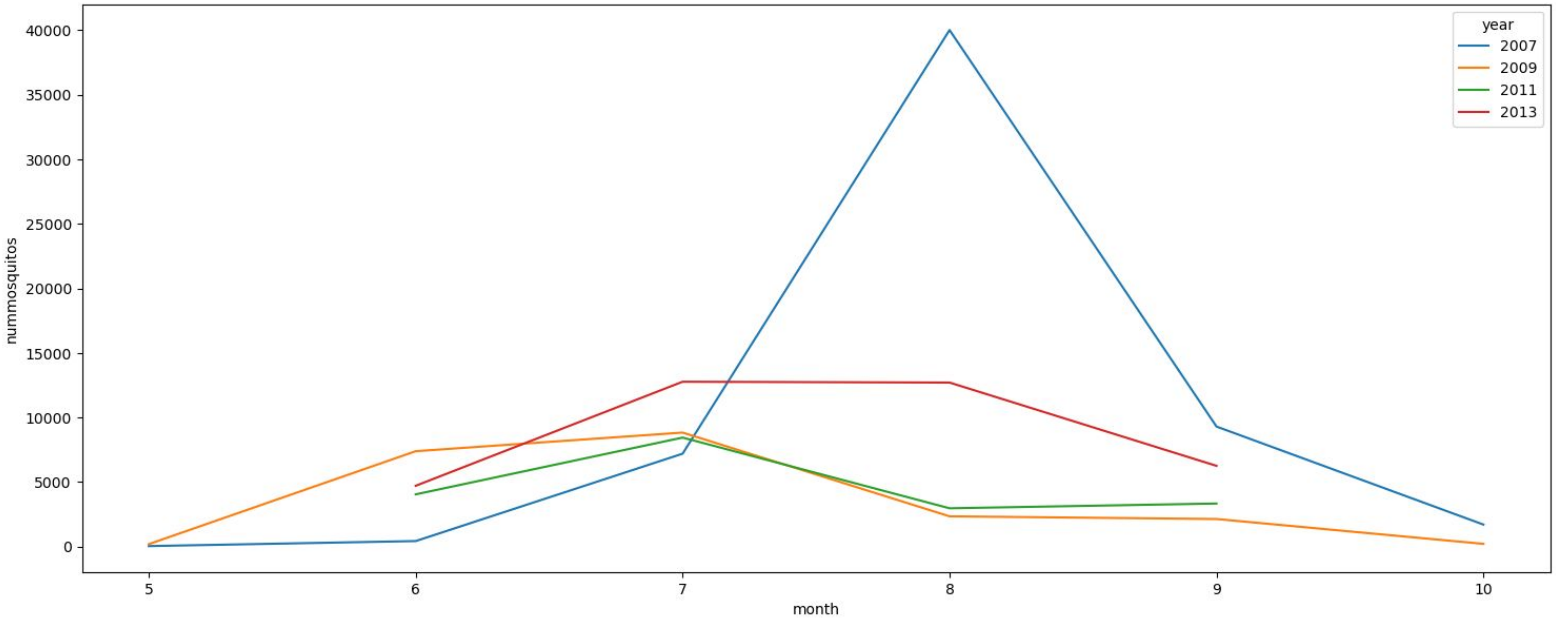
Spray

Continue from **May through Oct**

When traps hit **14% WNV-positive**

Note:
*14% threshold was determined by the average of percentage of WNV-positive traps in Aug 2007, Aug 2013 and Sep 2013 (refer to annex for figures)

When



Traps

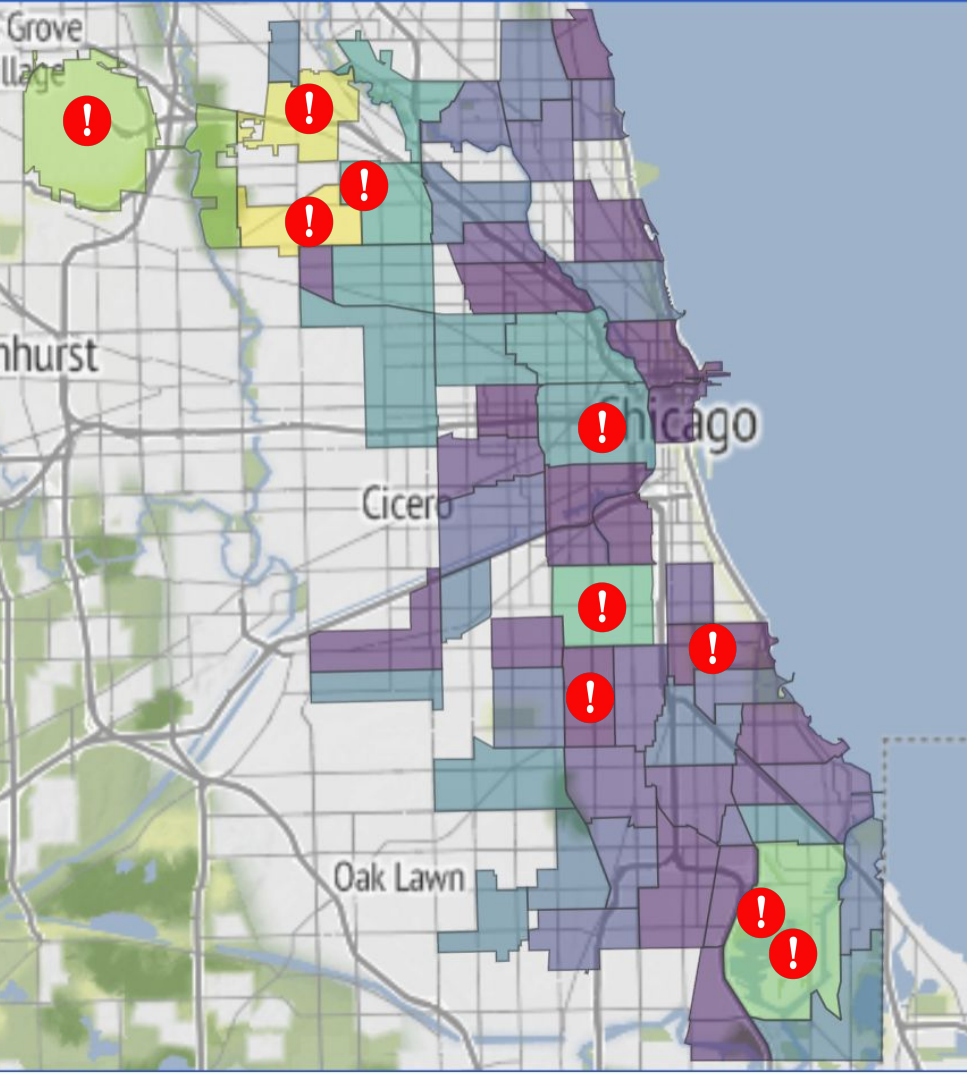
Monitor → Action

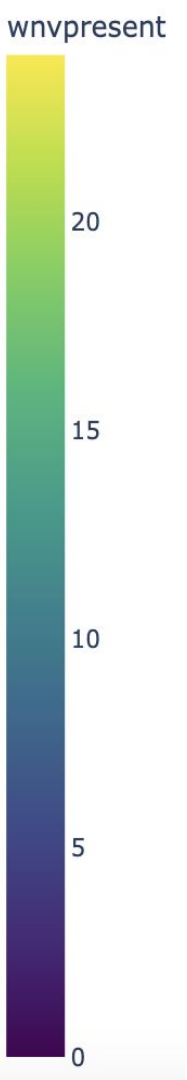
Spray

Continue from **May through Oct**

When traps hit **14% WNV-positive**

Note:
*14% threshold was determined by the average of percentage of WNV-positive traps in Aug 2007, Aug 2013 and Sep 2013 (refer to annex for figures)





Where

1. **ORD Terminal 5, O'Hare International Airport**
2. **South Doty Avenue**
3. **4100 North Oak Park Avenue**
4. **South Stony Island Avenue**
5. **4600 Milwaukee Avenue**
6. **8200 South Kostner Avenue**
7. **2400 East 105th Street**
8. **3600 North Pittsburgh Avenue**
9. **O'Hare Court, Bensenville**
10. **7000 North Moselle Avenue**

Future Research

- Analyse effect of birds on WNV infection
  - Birds are amplifying hosts (Environmental Research and Public Health, 2020)

- Analyse the severity of WNV cases
  - Look at total no. of cases instead of binary outcomes

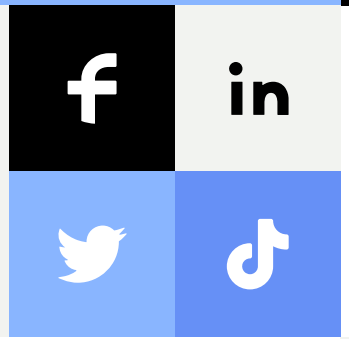
References

- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7344584/
- https://twitter.com/CDCgov/status/1694801796764451240
- https://www.cdc.gov/westnile/resources/pdfs/13_240124_west_nile_lifecycle_birds_plainlanguage_508.pdf
- https://en.wikipedia.org/wiki/Community_areas_in_Chicago
- https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjv2pnP5PCB[…]-fact-sheet.pdf&usg=AOvVaw0JET5Q_8Fzt5ef0CoMfzSQ&opi=89978449

Thank you!

Do you have any questions?

39sir@datascience.il
+65 8765 4321
datascience.il/39sir



CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**

Please keep this slide for attribution

08

Annex

Spray Cost Estimate

Address	*Community	*Area (acres)	*Cost
ORD Terminal 5, O'Hare International Airport	O'Hare	8537.49	5720.12
South Doty Avenue	Burnside	390.43	261.59
4100 North Oak Park Avenue	Dunning	2379.63	1594.35
South Stony Island Avenue	South Chicago	2137.46	1432.10
4600 Milwaukee Avenue	Portage Park	2527.89	1693.69
8200 South Kostner Avenue	Scottsdale	837.69	561.25
2400 East 105th Street	East Side	1907.654	1278.13
3600 North Pittsburgh Avenue	Dunning	2379.63	1594.35
O'Hare Court, Bensenville	O'Hare	8537.4909	5720.12
7000 North Moselle Avenue	Norwood Park	2797.23	1874.14

Note:
*'Community' and 'Area': https://en.wikipedia.org/wiki/Community_areas_in_Chicago
*'Cost': https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjv2pnP5PCB[…]-fact-sheet.pdf&usg=AOvVaw0JET5Q_8Fzt5ef0CoMfzSQ&opi=89978449

Threshold for Spraying

9% — **August 2007**
200 out of 2050 traps

17% — **August 2013**
144 out of 834 traps

14% — **September 2013**
70 out of 834 traps

14% (Mean) — **RECOMMENDED**

Speed of Results

- Within a short period, the application of insecticides can lead to a noticeable reduction in mosquito activity, providing immediate comfort to affected communities.

Disease Prevention and Public Health

- Mosquito spraying plays a vital role in preventing mosquito-borne diseases, thereby safeguarding public health.
- By reducing mosquito populations, it minimizes the transmission of diseases such as malaria, dengue fever, Zika virus, and West Nile virus.

Localization

- Mosquito spraying allows for targeted control efforts, focusing on specific areas with high mosquito activity or disease prevalence.
- This localized approach optimizes the allocation of resources, effectively addressing mosquito-related concerns where they are most needed.

Resistance Development

- Frequent and indiscriminate use of insecticides can lead to the development of resistance in mosquito populations.
- Over time, mosquitoes may become less susceptible to the chemicals used in spraying, rendering control efforts less effective in the long run.

Failure to Address the Source

- Mosquito spraying primarily targets adult mosquitoes, often overlooking the breeding sites where mosquitoes reproduce.
- Neglecting the elimination or treatment of stagnant water sources can result in a continuous cycle of mosquito populations and re-infestation.

Lack of Sustainability of Results

- Spray treatments generally last 4-8 weeks before populations return to normal levels.
- This means that the treatments have to be regular and continuous.