

# 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

# Problem Statement

Team **39 SIR** of the **Disease And Treatment Agency**, division of **Societal Cures In Epidemiology and New Creative Engineering (DATA-SCIENCE)** is to effectively plan the deployment of pesticides in order to mitigate the spread of the West Nile Virus in Chicago City.

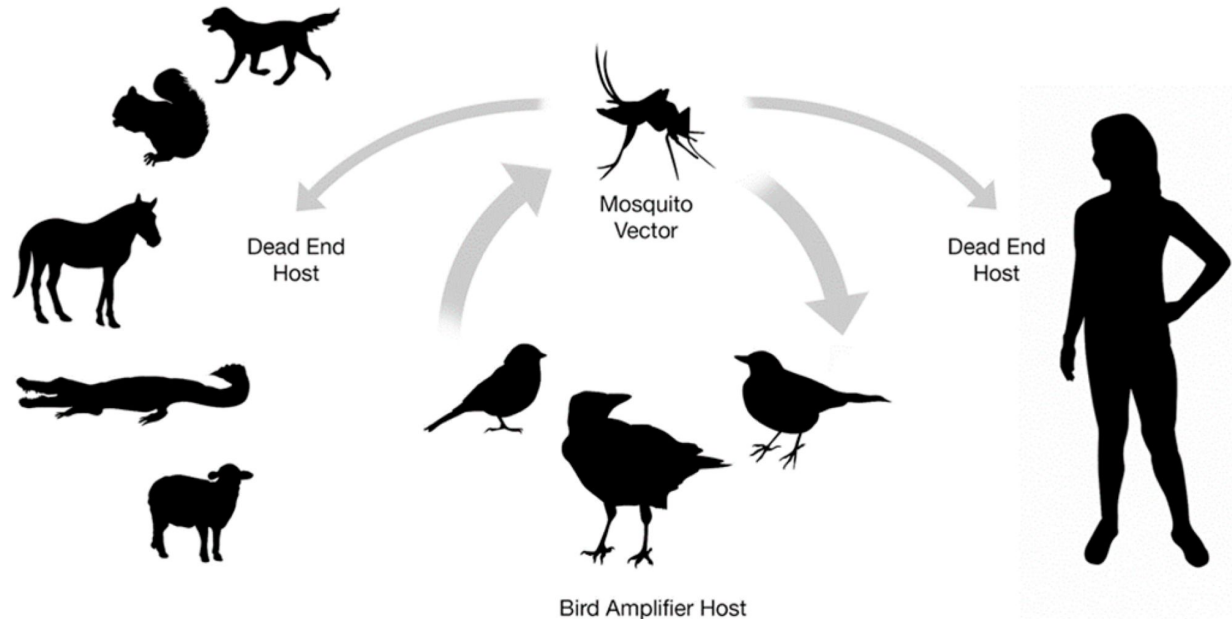
This will be done by analysing the data given by the Department of Public Health to produce the following deliverables:

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).

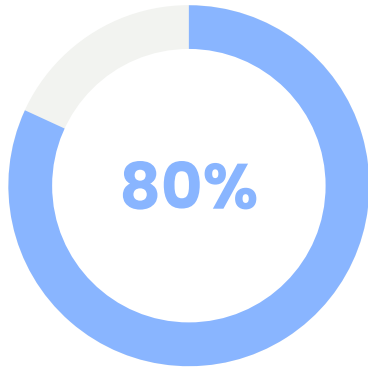


# What Is the West Nile Virus (WNV)?

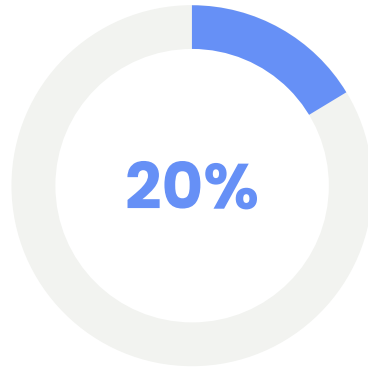
**Most common** mosquito-borne disease in the US



# Severity of West Nile Fever (WNF)



**Asymptomatic**



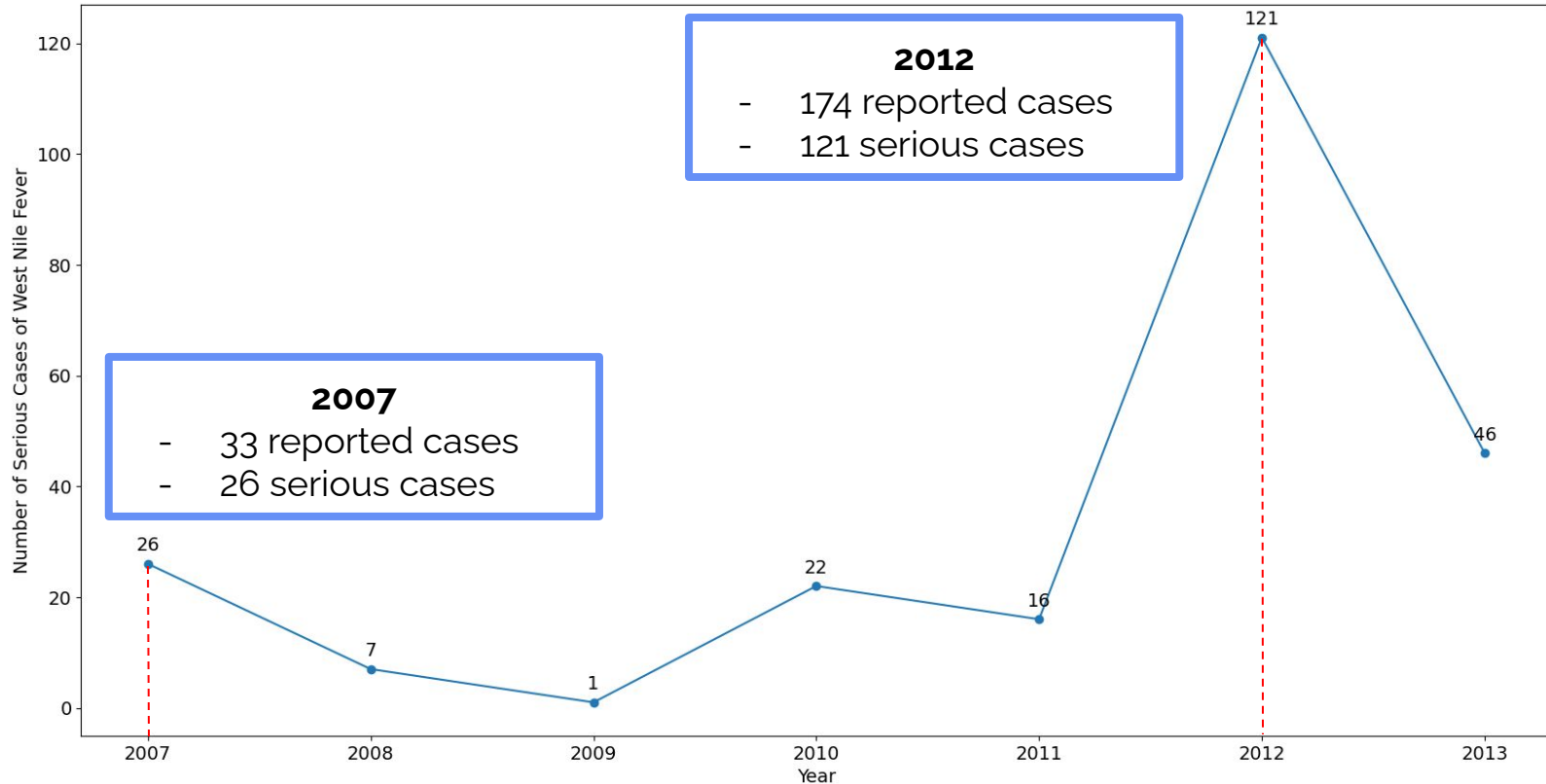
**Developed Symptoms -  
e.g. Fever**



**Developed Severe  
Illness**



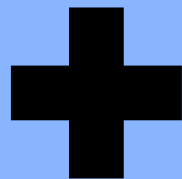
# Outbreak of WNF in Chicago in 2007 and 2012





02

**Data  
Cleaning**





# 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



03

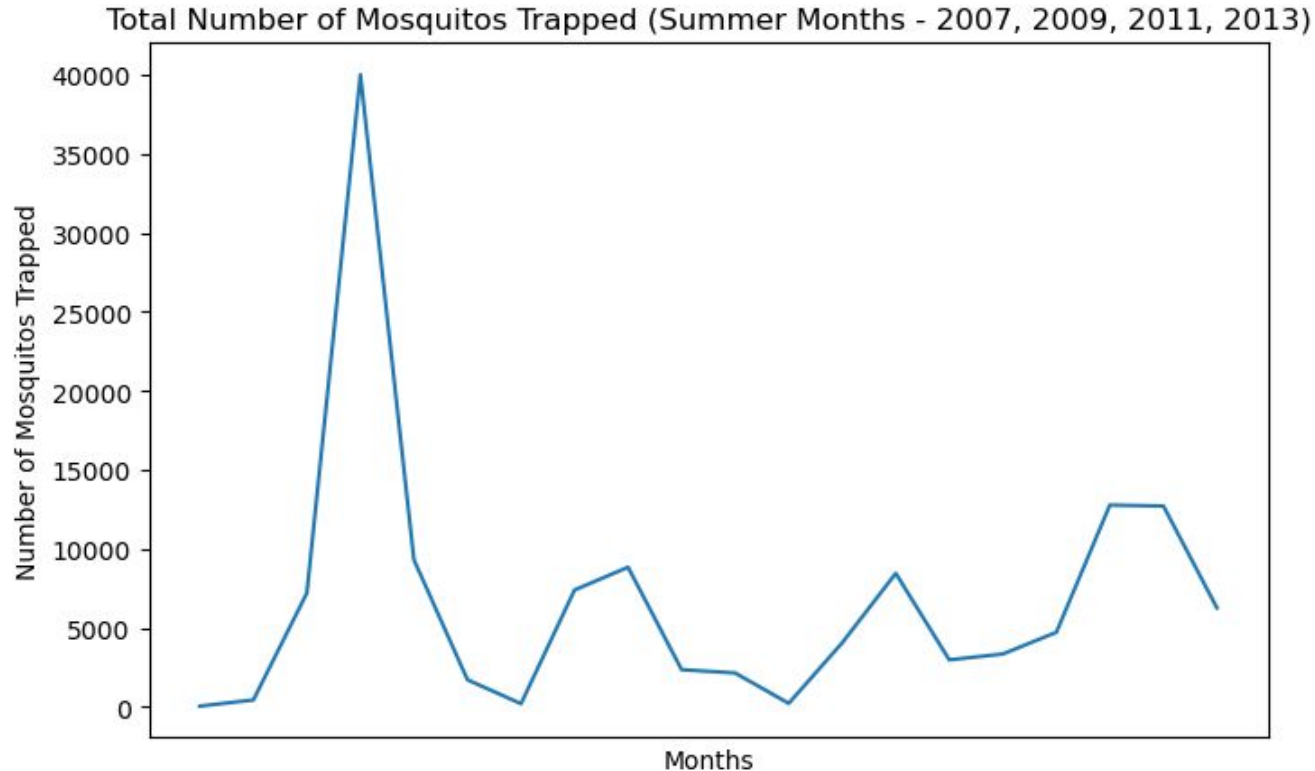
**EDA**



Exploratory Data  
Analysis

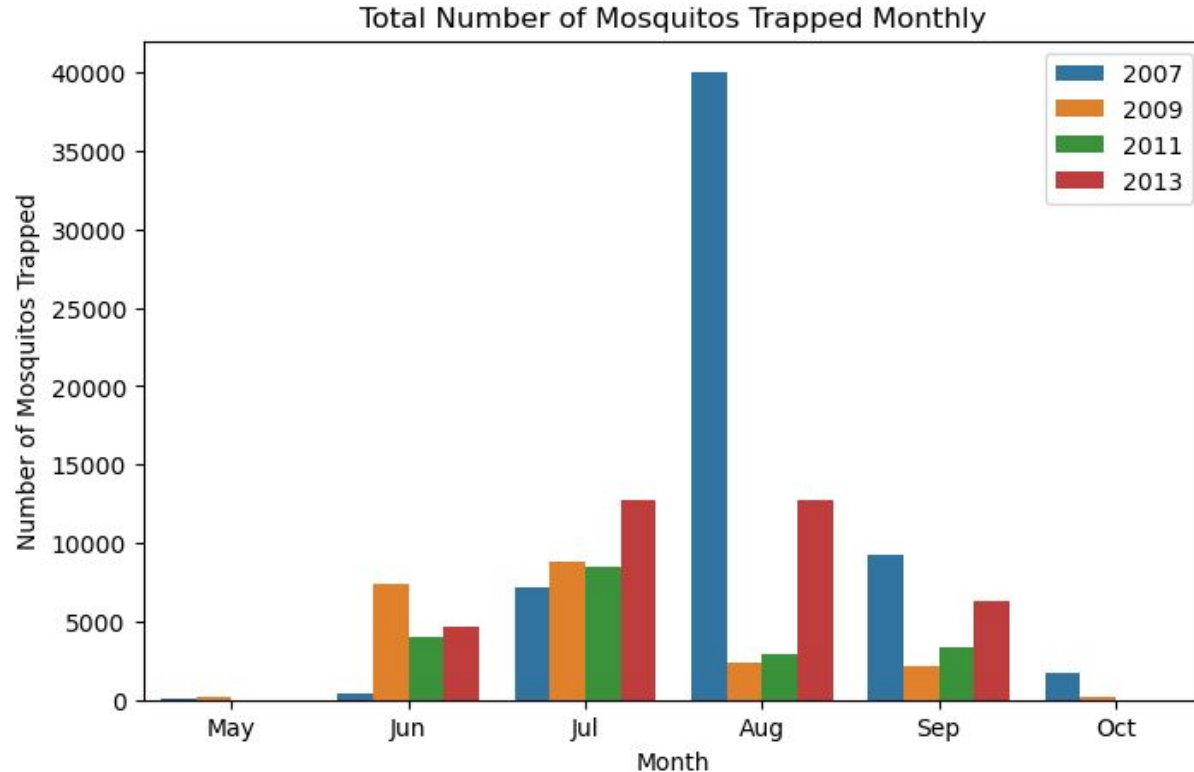
# Peak in Mosquito Activity in 2007

May 2007 – Sep 2013

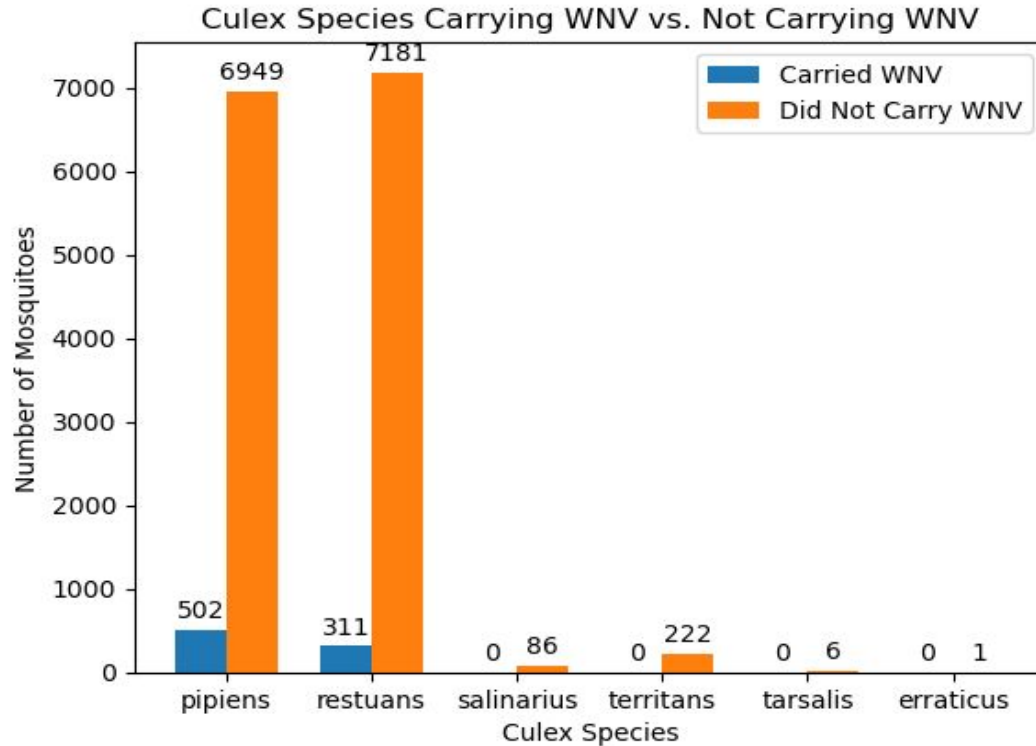


# Summer Trends: July and August as Mosquito Season

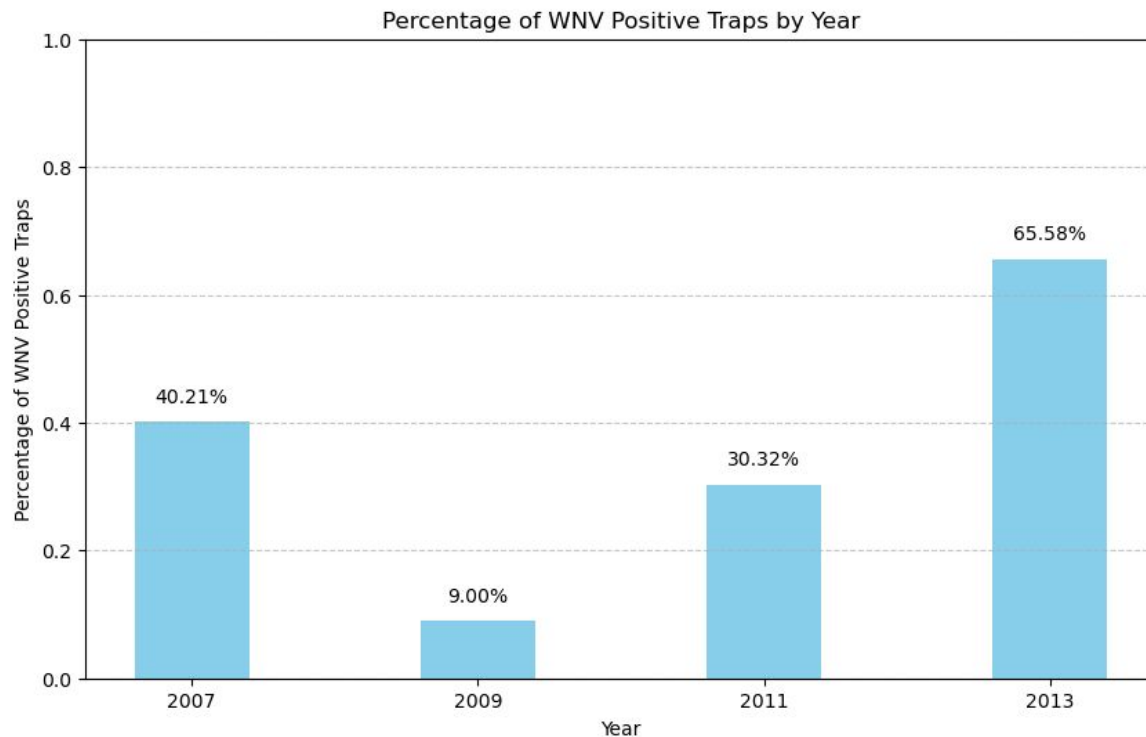
May 2007 – Sep 2013



# Only 2 Culex Species as WNV+ Carriers



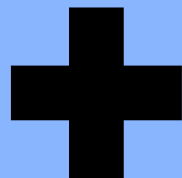
# High incidence of WNV+ traps in 2013





04

# Modeling



# Modeling Process

## Rapid Modeling

### Lazy Predict

- LazyClassifier  
(25 models)

## Pre-processing

### Pipeline

- StandardScaler
- SMOTE
- XGBoost

## Hyperparameter Tuning

- GridSearch

## Analysis

- Confusion Matrix
- Top Predictors
- Final score



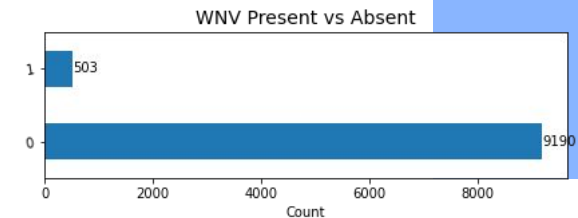
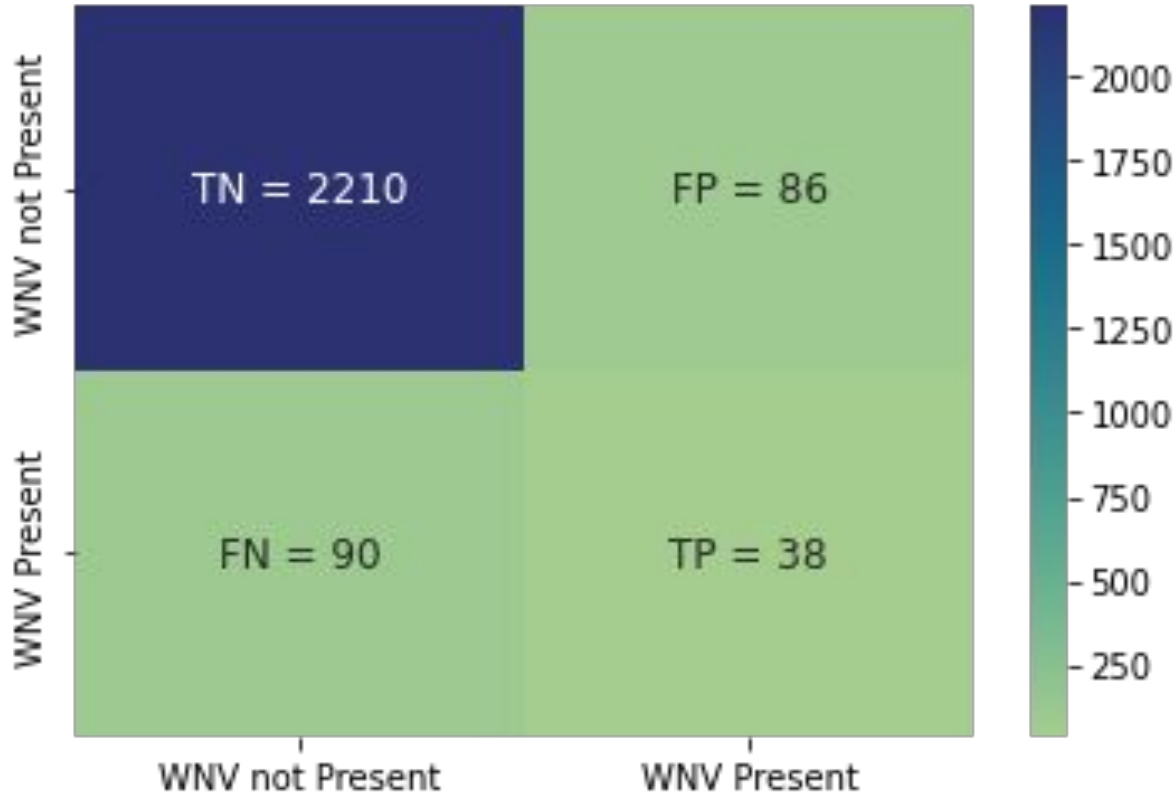
# Lazy Predict Top 5

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
LinearSVC	0.951	0.500	0.500	0.922	0.214
<b>XGBClassifier</b>	<b>0.950</b>	0.585	<b>0.585</b>	0.931	<b>0.023</b>
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

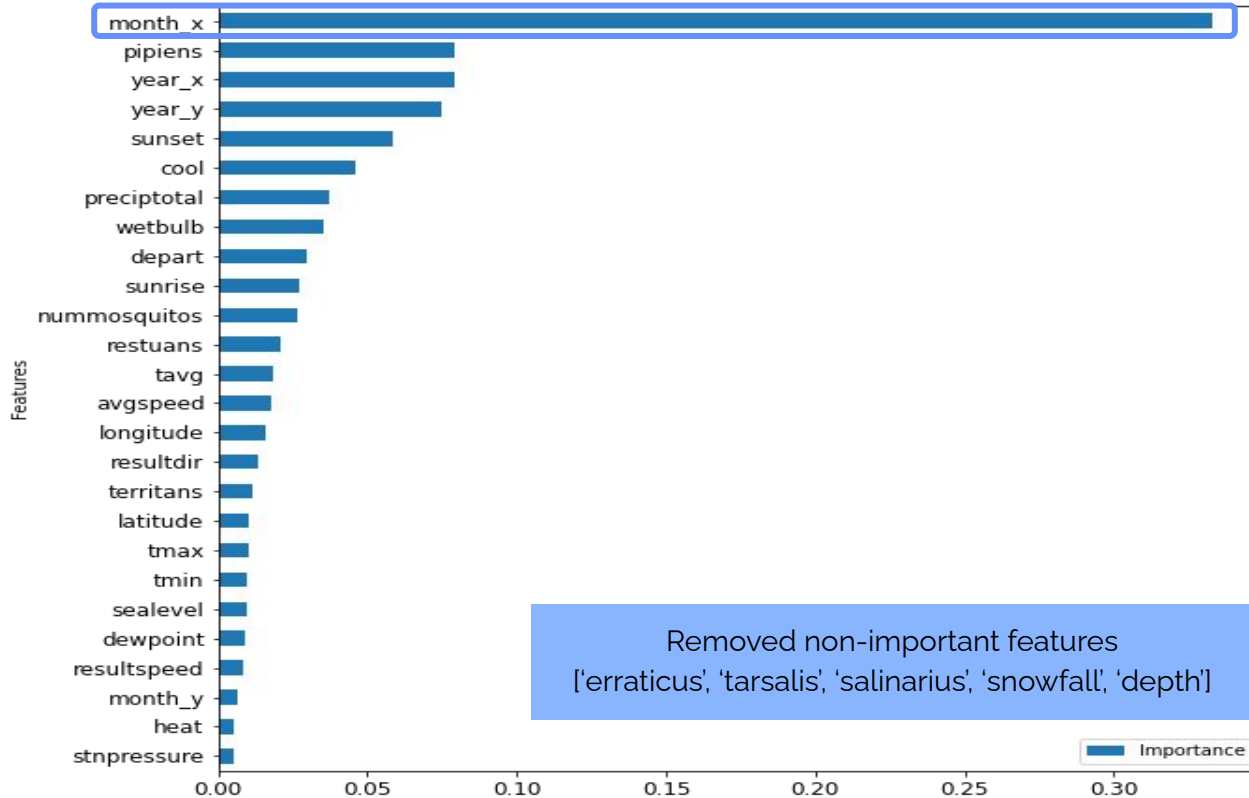




# Confusion Matrix



# Top Predictors



# 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	<b>0.999</b> 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**



# Monitoring and Controlling Spread of WNV:

## WNV Carrier:

Monitor WNV presence by trapping mosquitoes for testing



### Larval Mosquito

- Larvicide in drains
- Eliminating stagnant water

### Adult Mosquito

- Spraying of insecticides

## WNV Hosts:

Monitor infections through testing and reporting



# Analysis of Mosquito Spraying



## Pros:

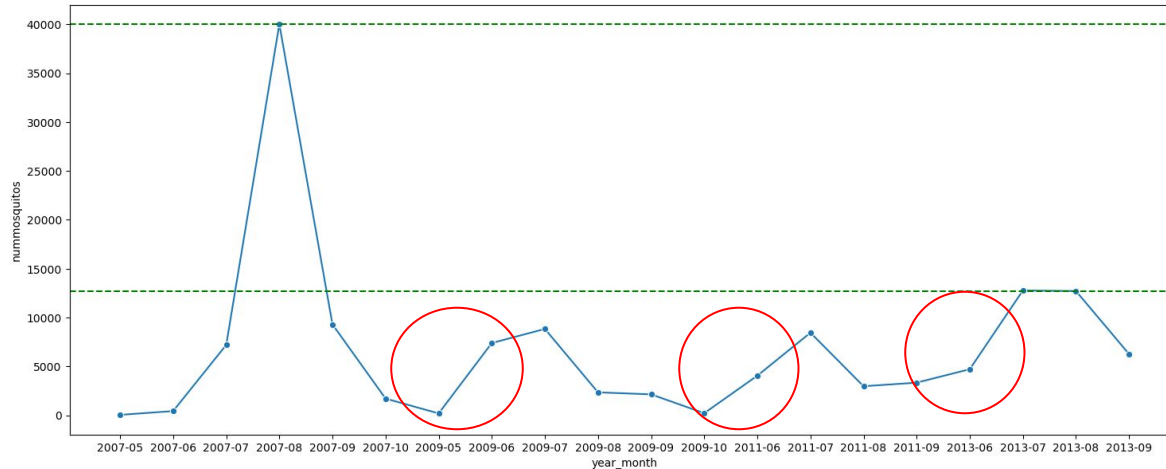
- Effective, accessible, and localized
- Can reduce mosquito populations in treated areas
- Fairly quick and easy process

## Cons:

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

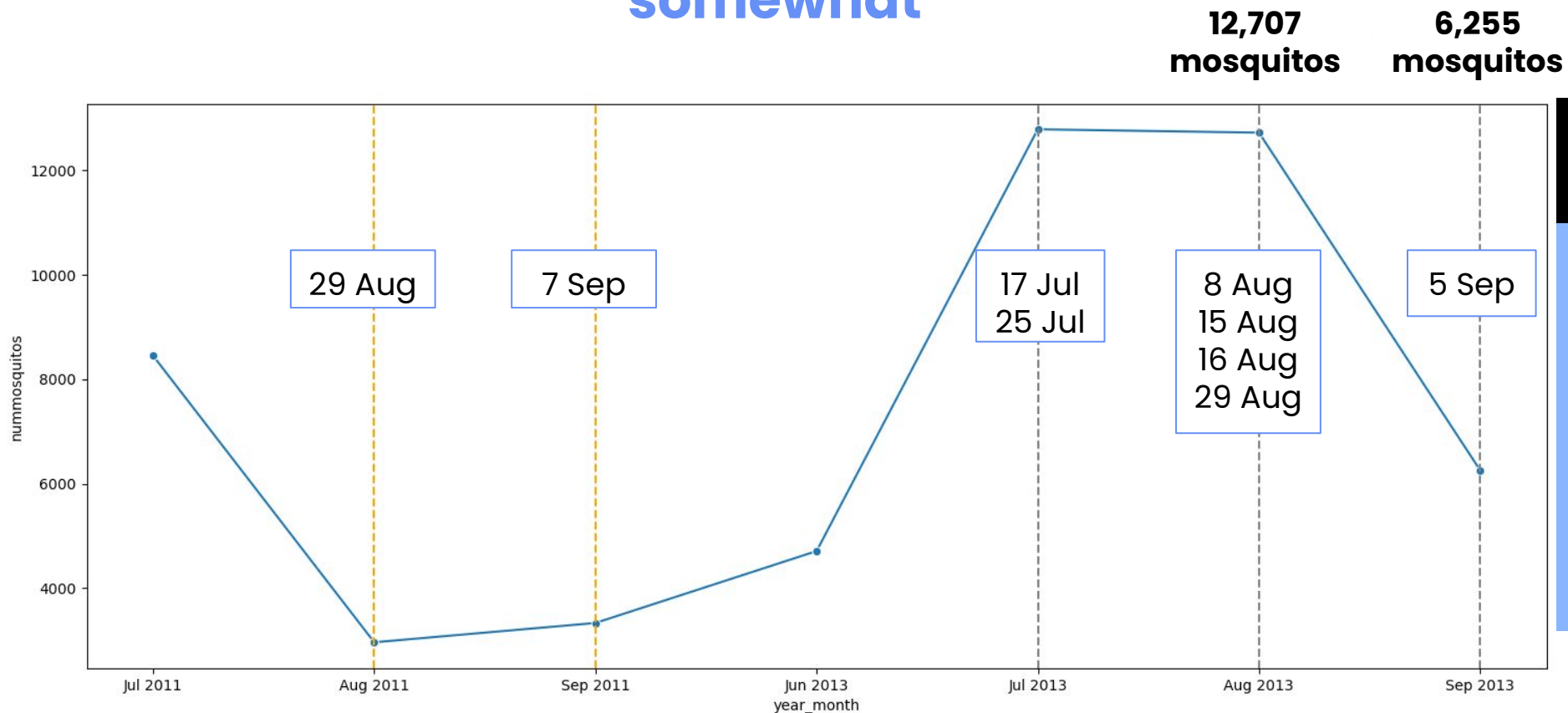
# Cost-Benefit Analysis

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

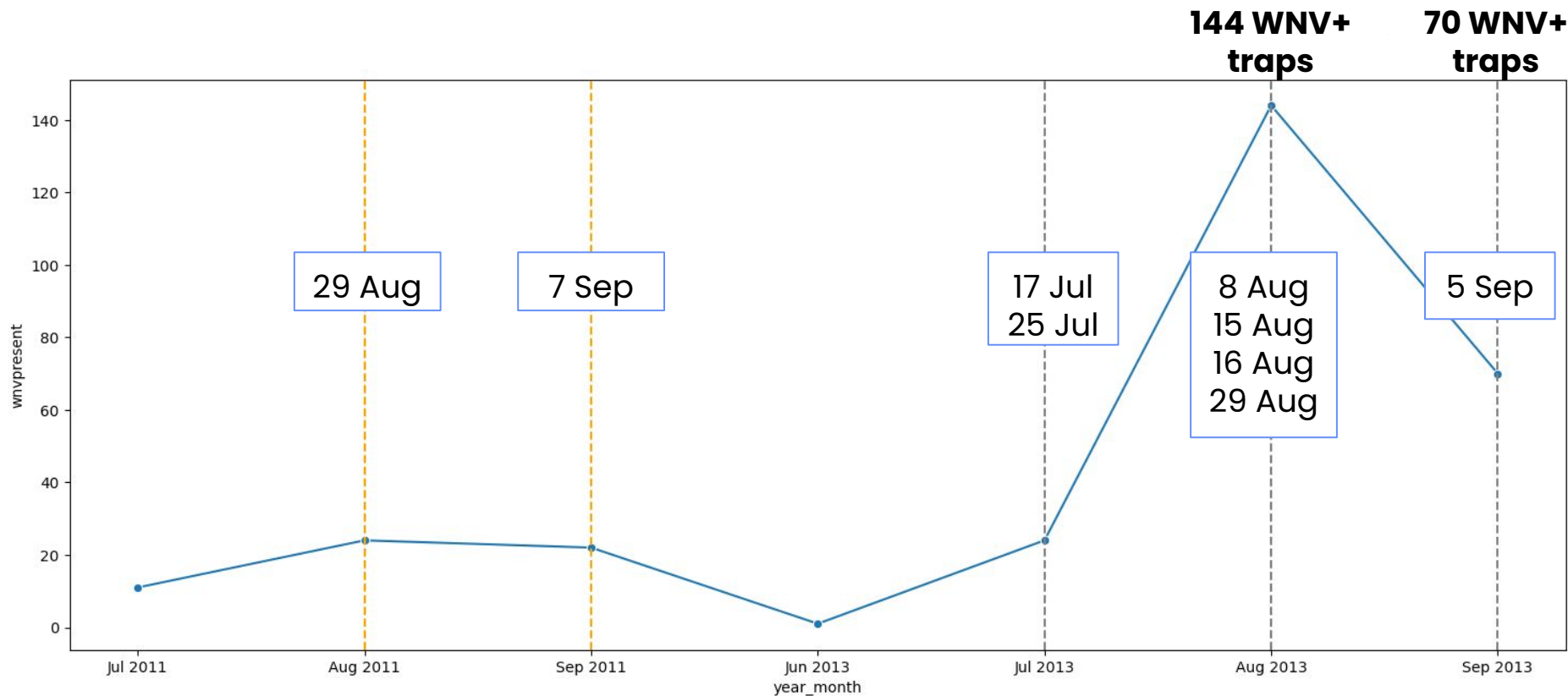




# Aggressive Spraying Reduces Mosquito Population ... somewhat



# Aggressive Spraying Reduces WNV+ ... somewhat



# 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>

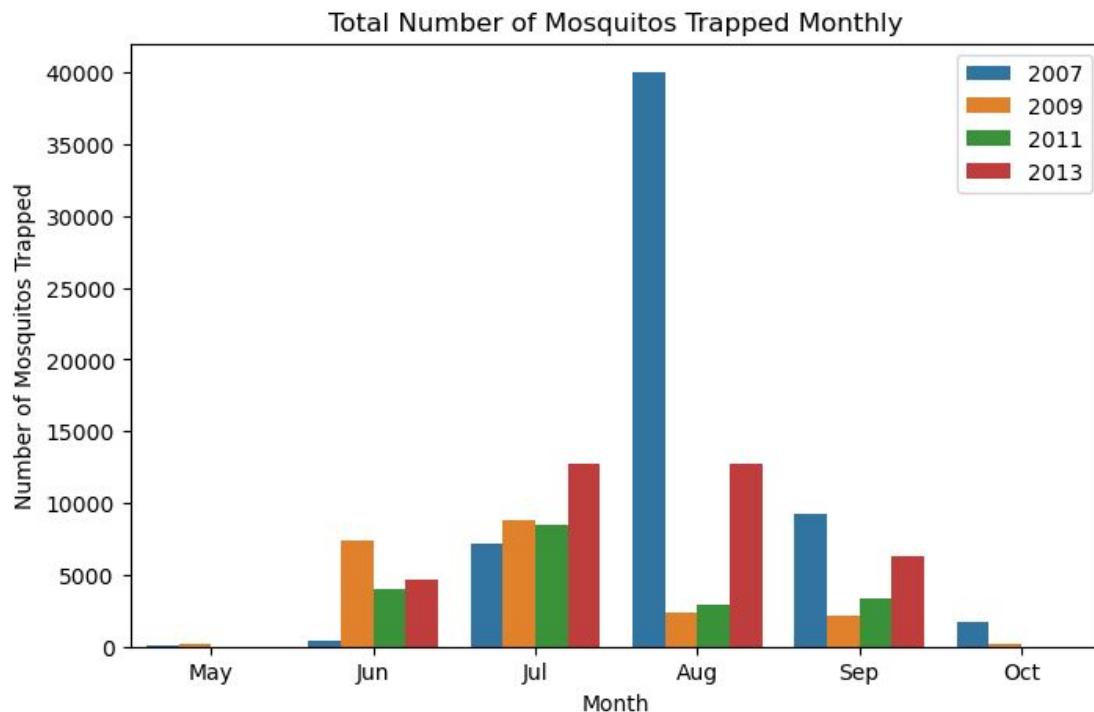
06

# Conclusion





# When



## Traps

Monitor

Action



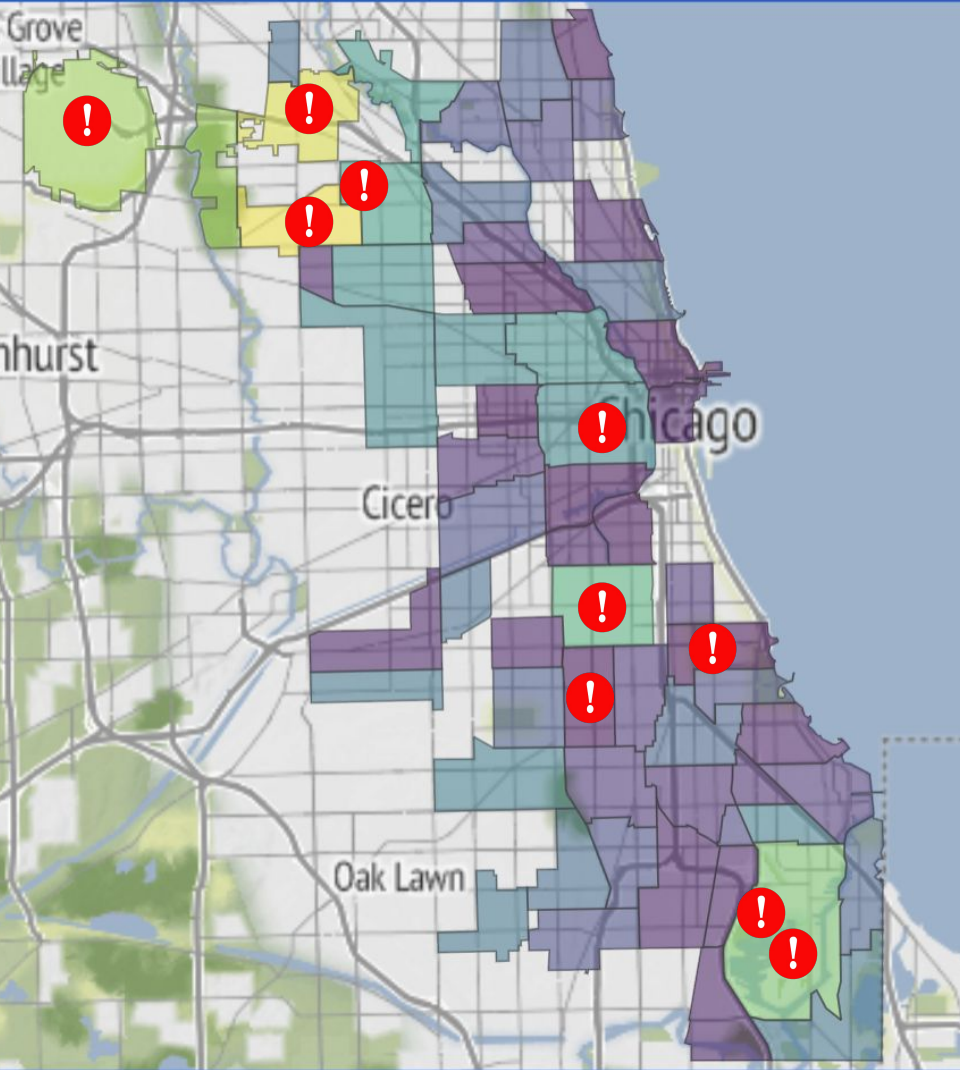
## Spray

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)



wvnpresent

20

15

10

5

0

# 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**

# 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://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://en.wikipedia.org/wiki/Community_areas_in_Chicago)
- [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjv2pnP5PCBI...\]-fact-sheet.pdf&usg=AOvVawoJET5Q\\_8Fzt5ef0CoMfzSQ&opi=89978449](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjv2pnP5PCBI...]-fact-sheet.pdf&usg=AOvVawoJET5Q_8Fzt5ef0CoMfzSQ&opi=89978449)





# Thank you!

**Do you have any  
questions?**

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



in



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](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](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**