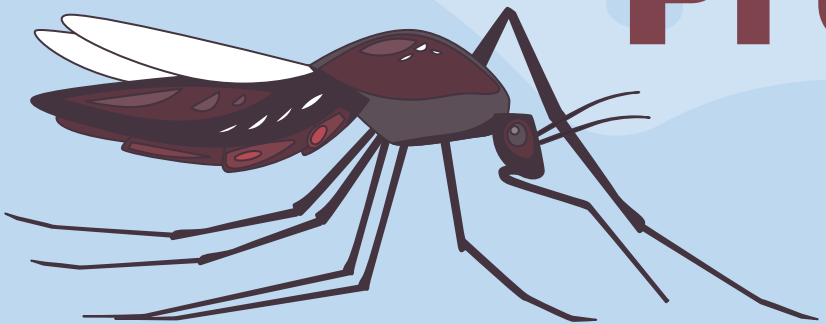
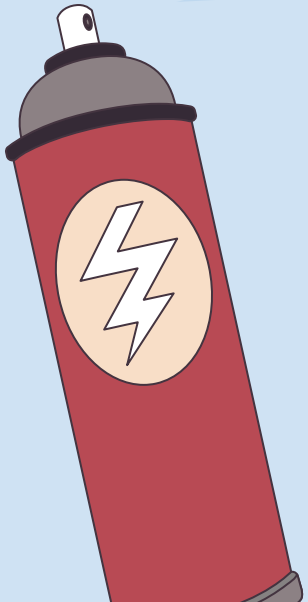
West Nile Virus Prediction



Emma Goh, Emi Chua,
Melissa Ng, Tiek Leong

Table of contents



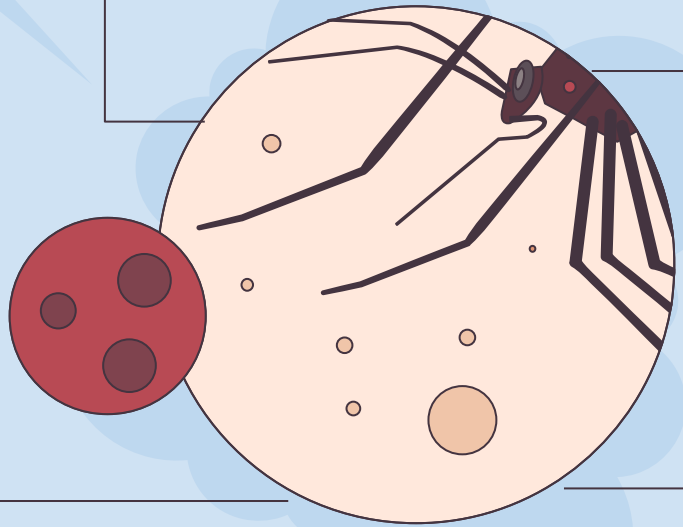
01

Problem Statement

West Nile Virus

Leading cause of mosquito-borne disease in US

1 in 5 infected develop a fever and other symptoms



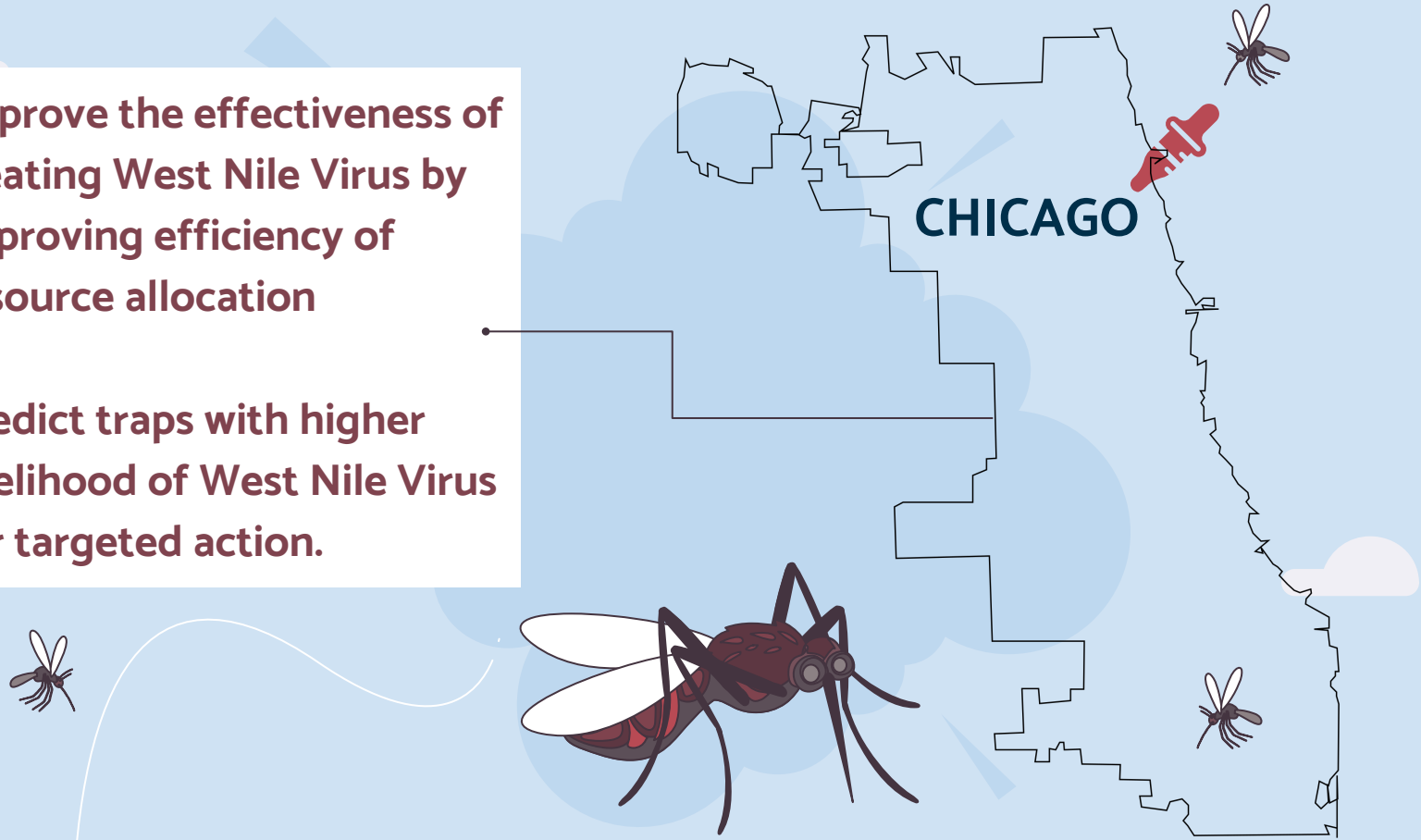
No vaccines/ medications to treat WNV

From late spring to fall, mosquitos in traps are tested for the virus

Problem Statement

Improve the effectiveness of treating West Nile Virus by improving efficiency of resource allocation

Predict traps with higher likelihood of West Nile Virus for targeted action.



02

Data Cleaning/ Exploratory Data Analysis

Data Cleaning / Merge

Spray

- Remove duplicates & outliers

Train

- Remove duplicates
- Sum number of mosquitoes based on locations and year

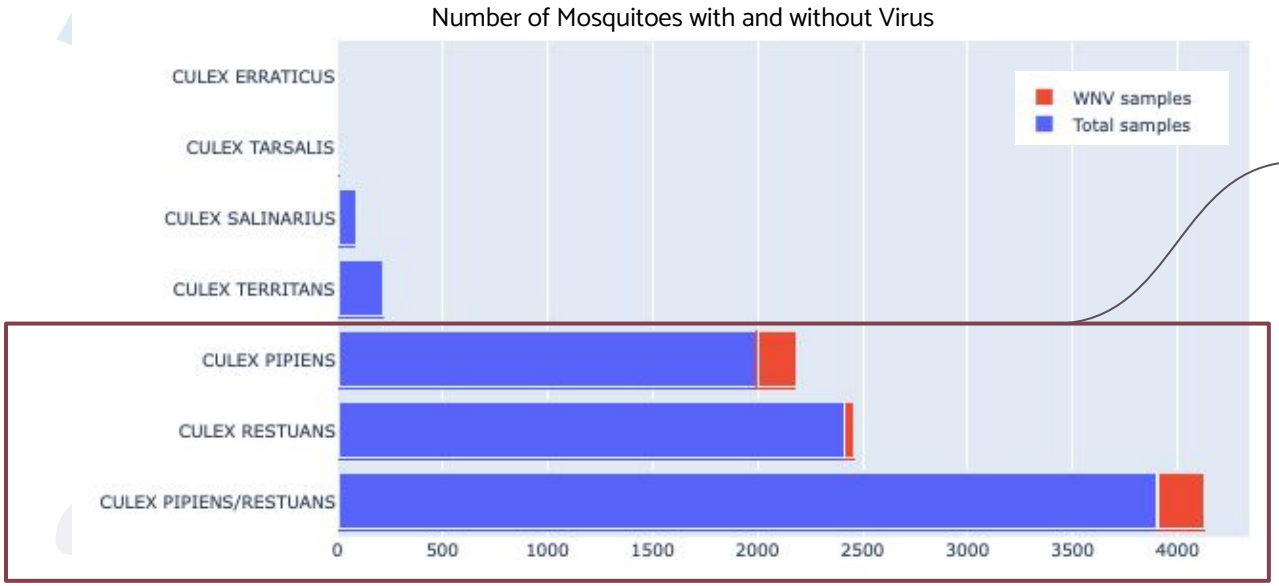
Weather

- Impute missing values

Test

- Status Quo

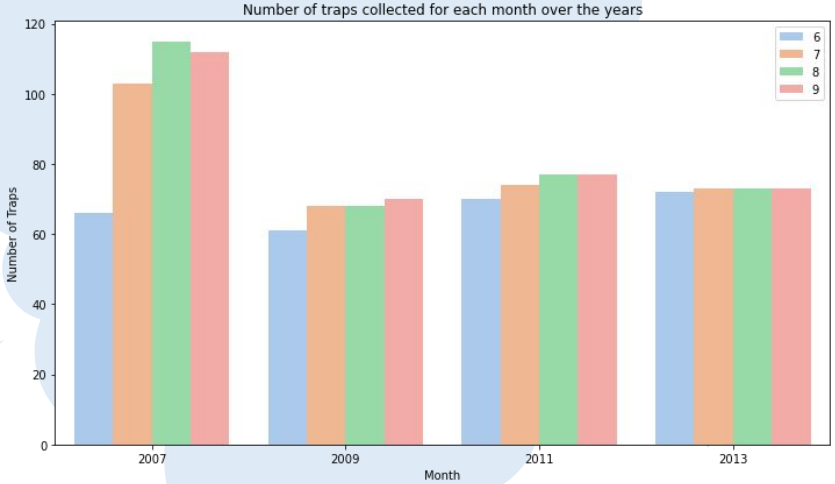
Species carrying WNV

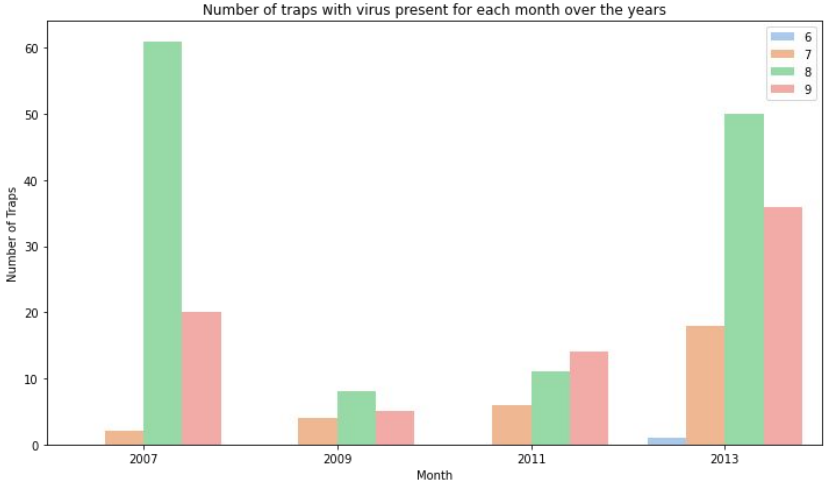


WNV only found in 2 species:

Culex Pipiens & Culex Restuans

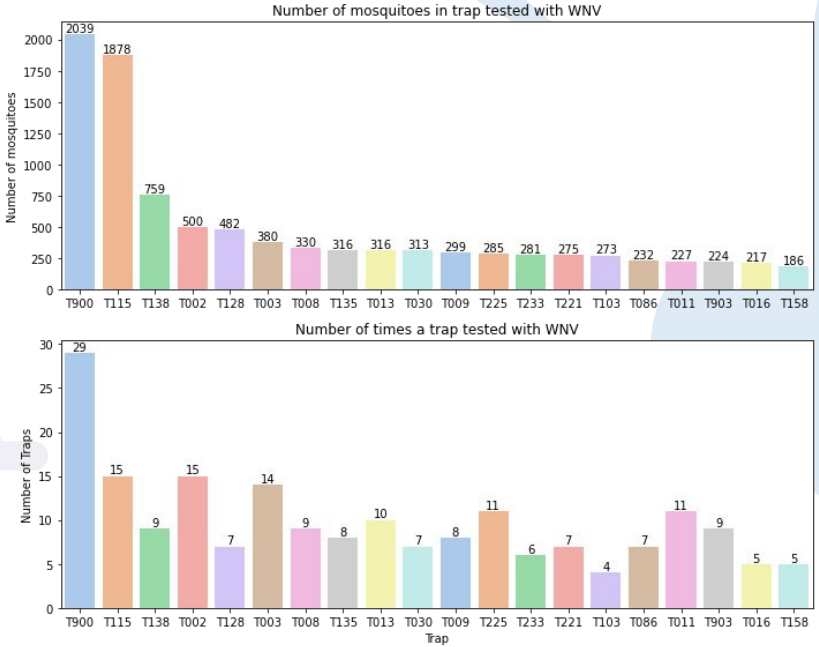
Traps Collected vs WNV Present over the years

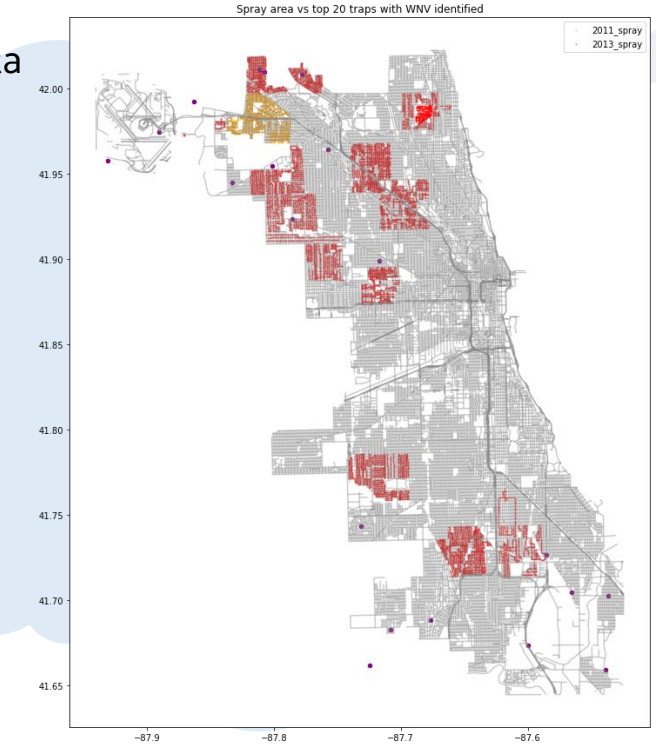




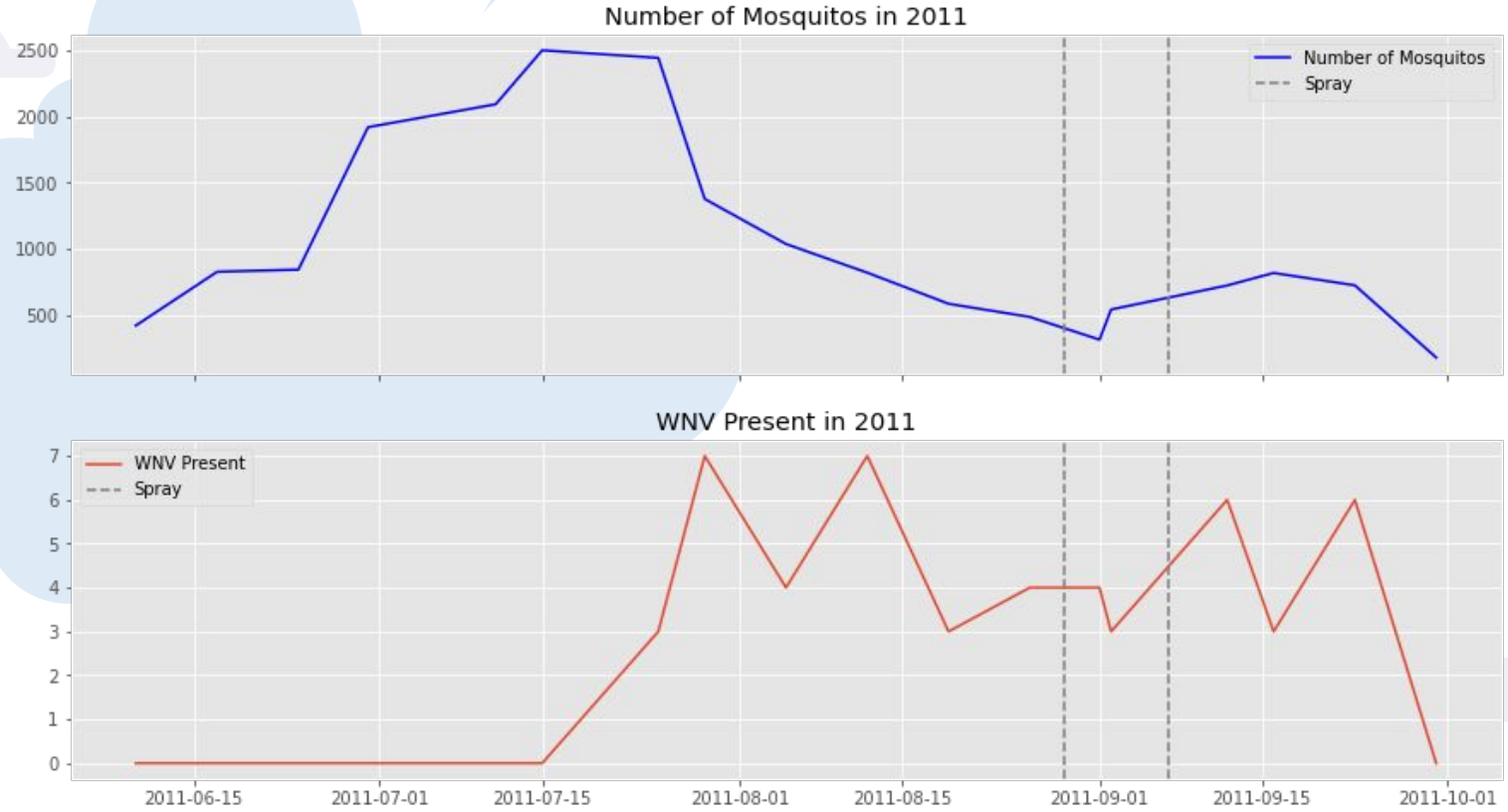
Top 20 traps found with WNV

Traps with the mosquitoes found with WNV were mainly in located in areas not shown on spray data



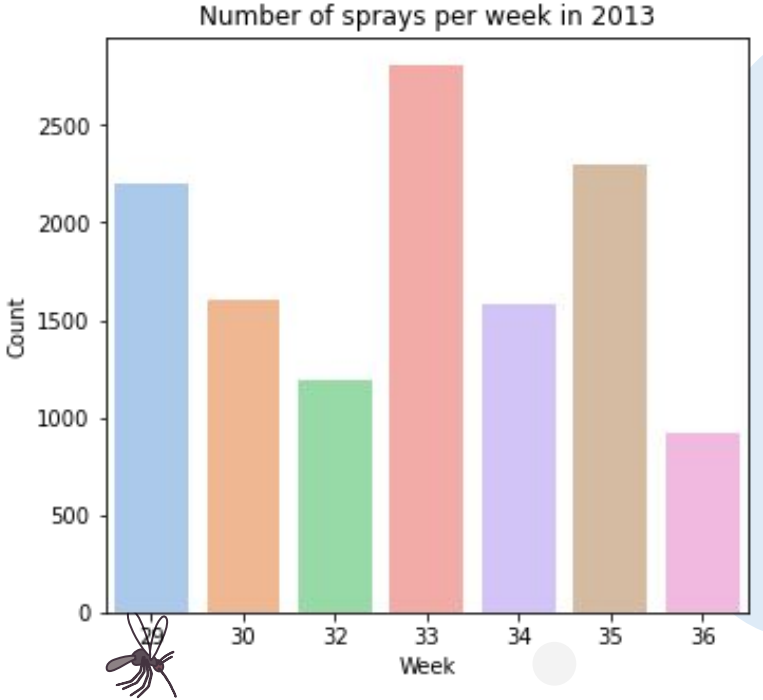


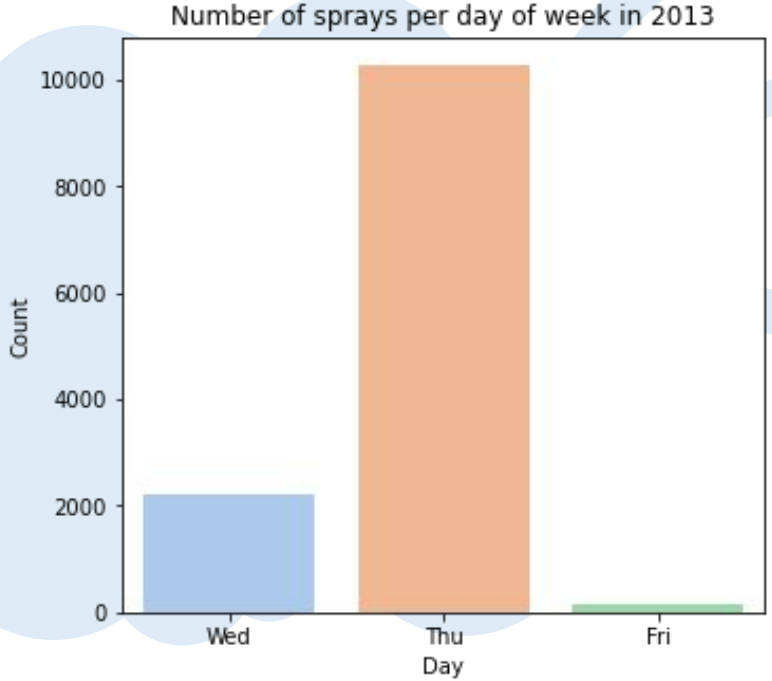
Spray 2011 vs Mozzie and WNV



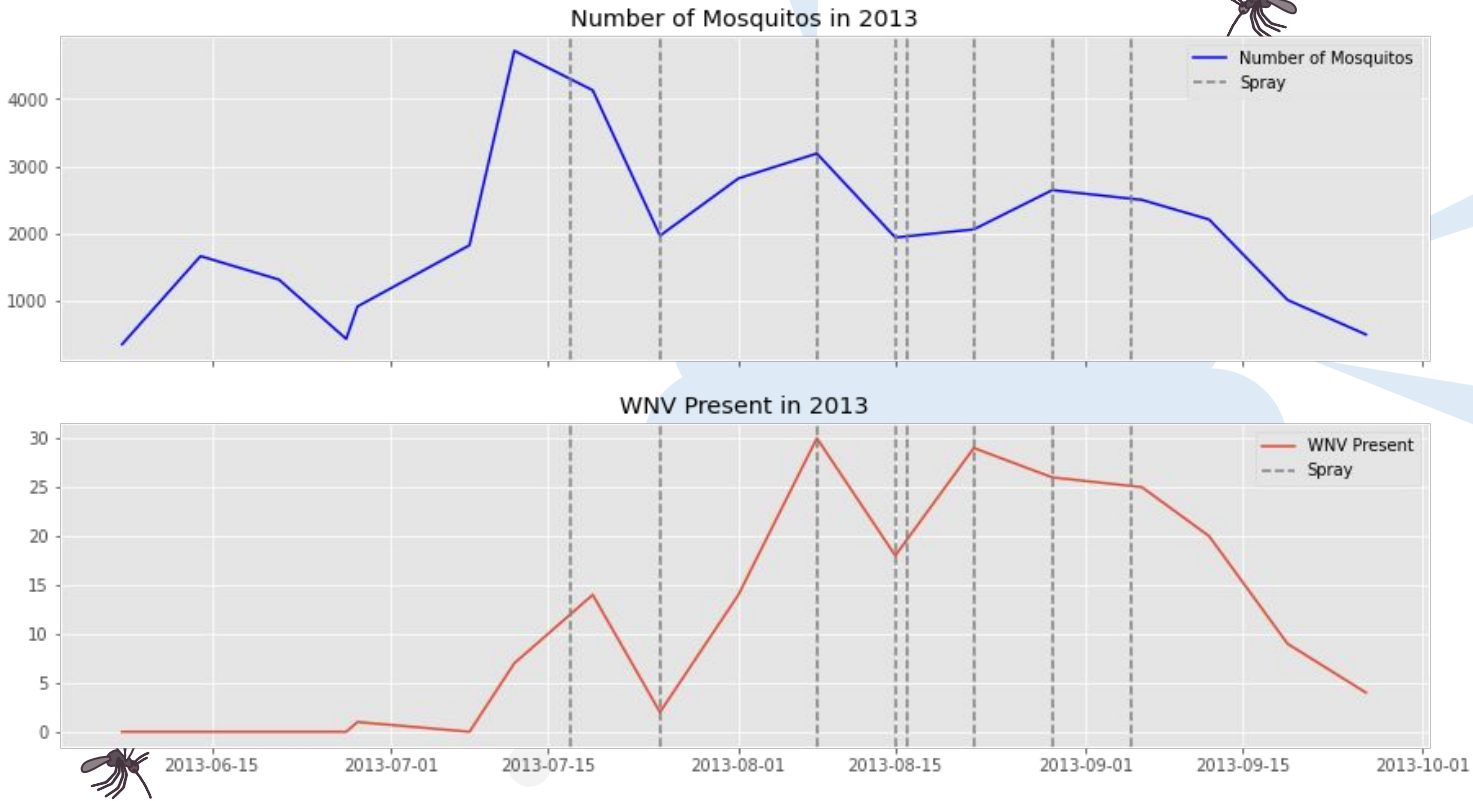
Frequency of Spray in 2013

Sprays are conducted weekly and mainly on Thursday

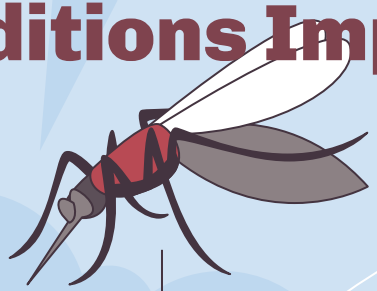




Spray 2013 vs Mozzie and WNV



Weather Conditions Impact on WNV



Temperature

Warmer temperatures accelerates mosquito development, biting rates, and the incubation of the disease within a mosquito

Precipitation

Increase in precipitation increases standing water surface which is necessary for mosquito larval development. However, too much rain may wash the larval away.

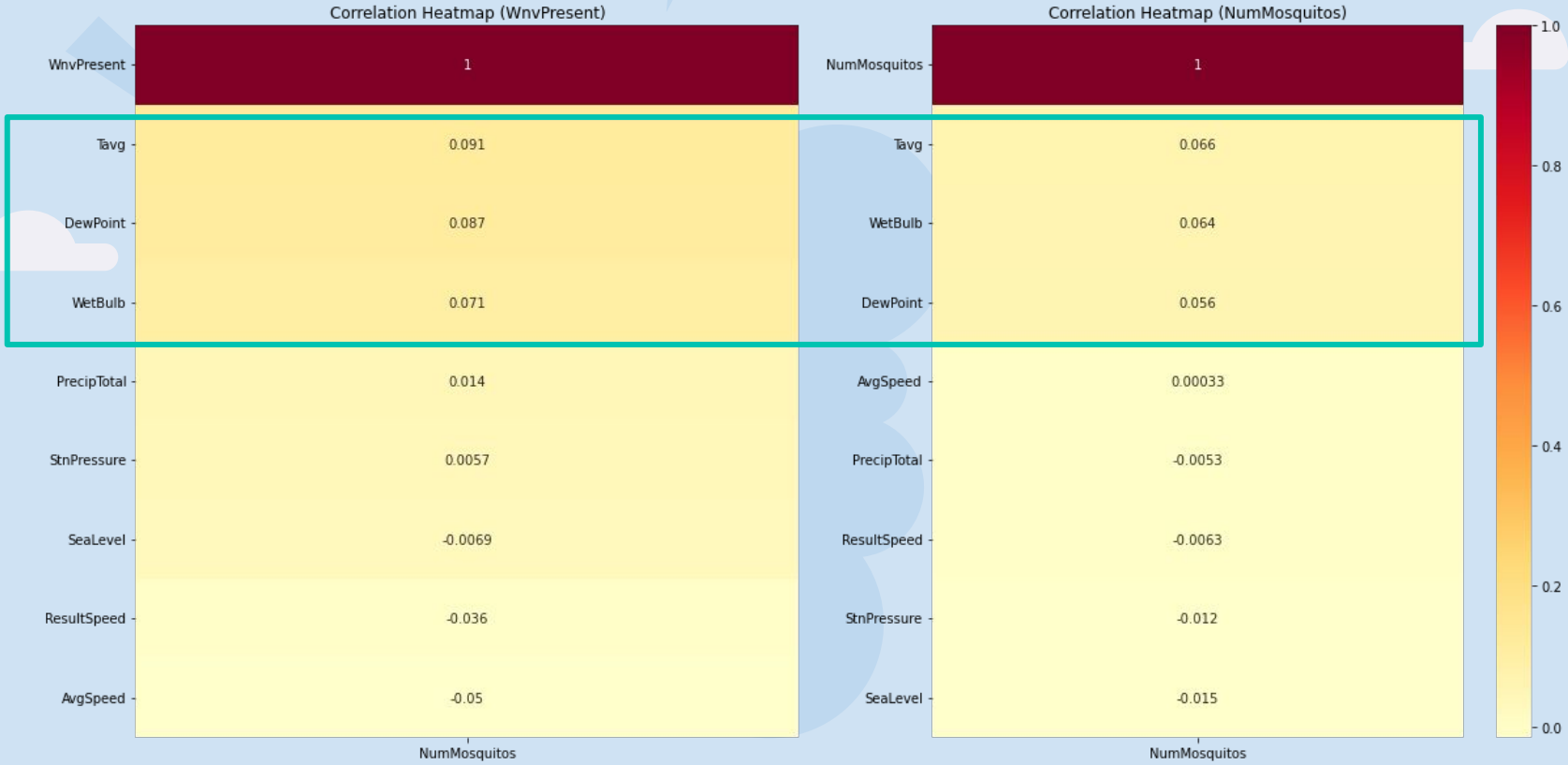
Humidity

Culex species favours dry weather. High humidity may not be conducive for the larval to develop.

Wind

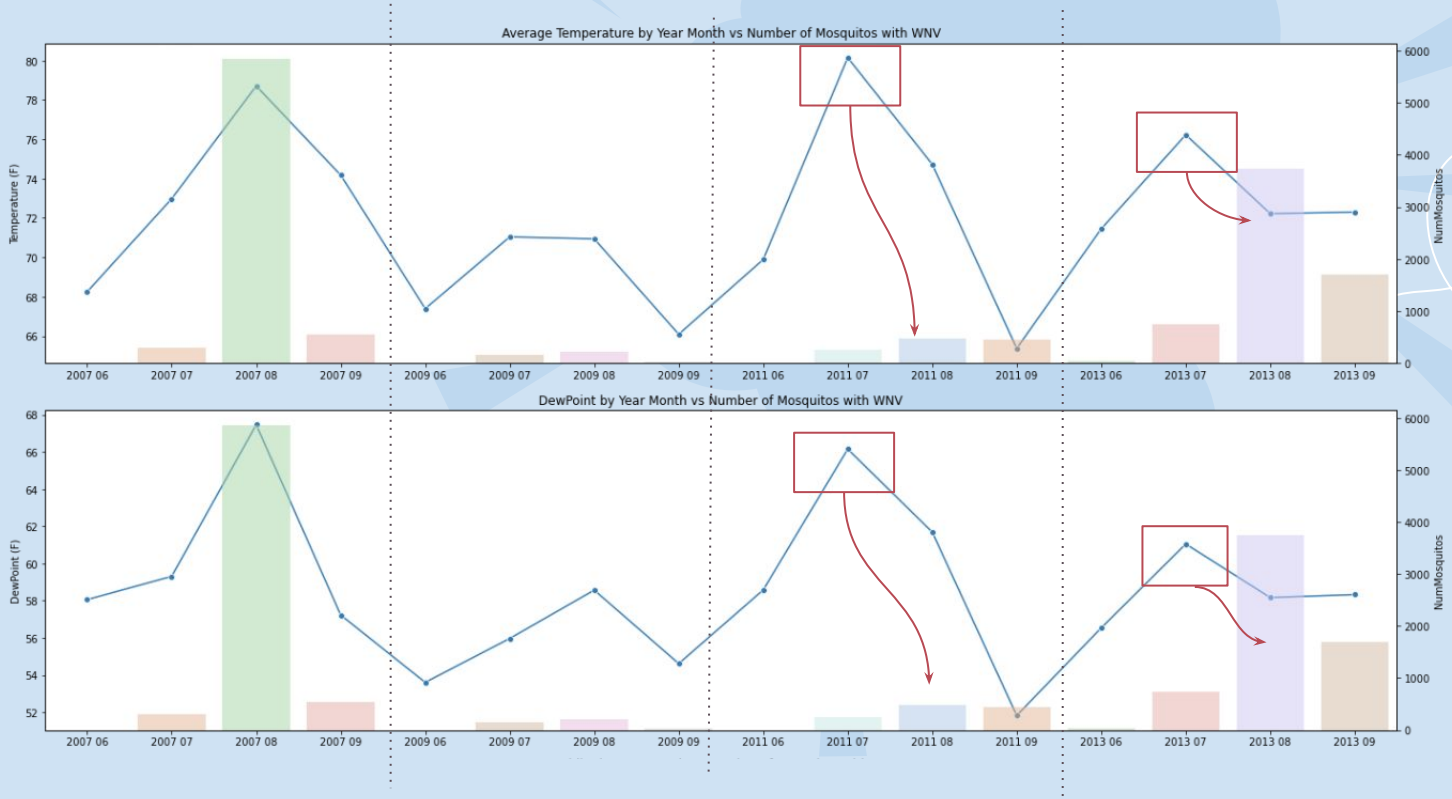
Culex species are windborne and can carry the virus further from the original point. However, strong wind works as a natural mosquito repellent as they are unable to fly.

Correlation with Weather



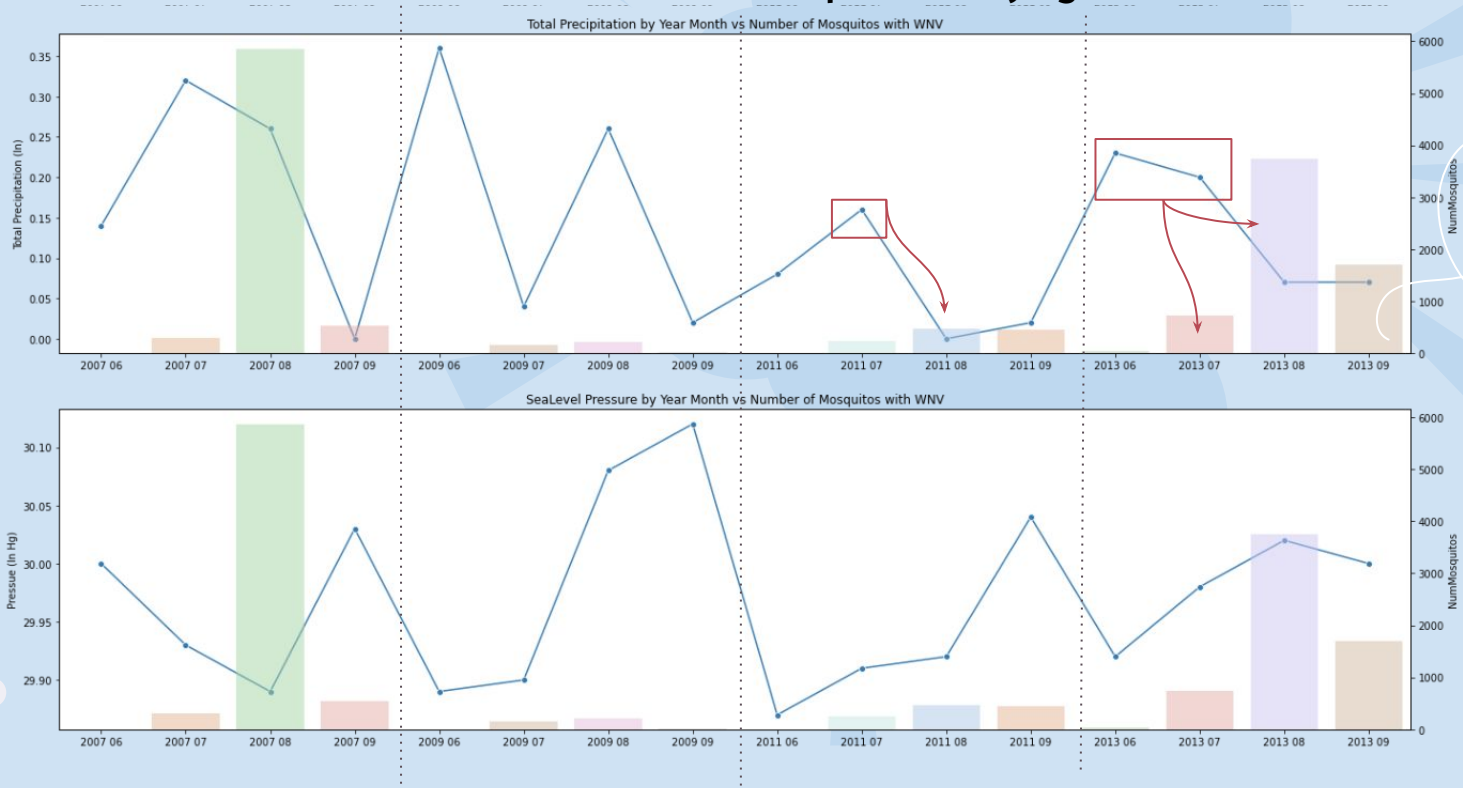
Weather VS Mozzies

The higher the temperature and the lower the humidity (dryer the air) seems to result in higher the number of mosquitoes carrying WNV the following month.



Weather VS Mozzies

Whilst the higher the precipitation seems to result in higher the number of mosquitoes carrying WNV the following month, sea level pressure does not seem to have a distinct effect on the number of mosquitoes carrying WNV.

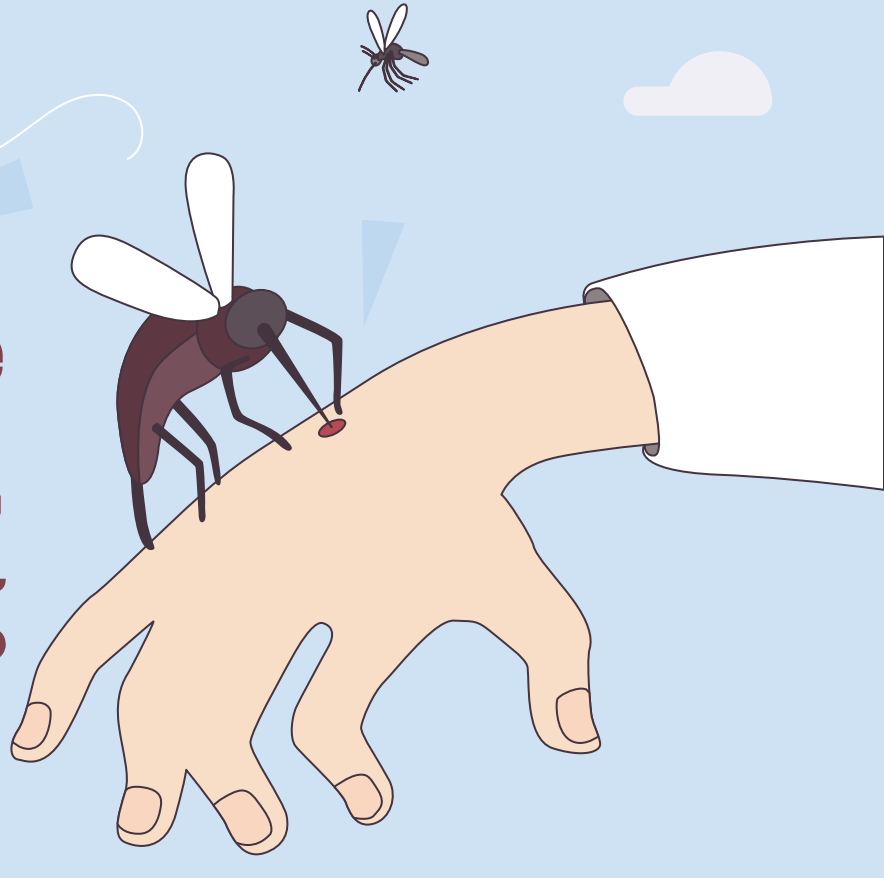


Weather VS Mozzies

The average and resultant wind speeds do not seems to have a consistent affect on the number of mosquitoes carry WNV.

03

Feature Engineering, Preprocessing



Feature Engineering

Convert mosquito species to an ordinal feature

Species	Probability	Ordinal	Value
Cullex Pipiens	9.22%	3	
Cullex Pipiens/Restuans	5.77%	2	
Cullex Restuans	1.99%	1	
All Others	0.00%	0	

Convert weather type into a new “wet_dry” feature

CodeSum	Value
RA, TS, DZ, BR, UP, SH	1
All Others	0

sunrise	sunset	wetdry	preciptotal
0448	1849	0	0.00
0448	1849	0	0.00

Feature Engineering

Impute “station” feature based on distance of traps from the weather station 1 or 2

	date	species	trap	latitude	longitude
0	2007-05-29	2.0	T048	41.867108	-87.654224
1	2007-05-29	1.0	T048	41.867108	-87.654224
2	2007-05-29	1.0	T091	41.862292	-87.648860
3	2007-05-29	1.0	T049	41.896282	-87.655232
4	2007-05-29	1.0	T153	41.907645	-87.760886

wnvpresent	year	month	day	nummosquitos	station
0	2007	5	29	1	2.0
0	2007	5	29	2	2.0
0	2007	5	29	1	2.0
0	2007	5	29	1	2.0
0	2007	5	29	1	2.0

Merge train data with the weather data based on the nearest station

date	species	trap	latitude	longitude	wnvpresent
2007-05-29	2.0	T048	41.867108	-87.654224	0
2007-05-29	1.0	T048	41.867108	-87.654224	0
2007-05-29	1.0	T091	41.862292	-87.648860	0
2007-05-29	1.0	T049	41.896282	-87.655232	0
2007-05-29	1.0	T153	41.907645	-87.760886	0

nummosquitos	station	tmax	tmin	tavg	dewpoint	wetbulb
1	2.0	88.0	65.0	76.5	59.0	66.0
2	2.0	88.0	65.0	76.5	59.0	66.0
1	2.0	88.0	65.0	76.5	59.0	66.0
1	2.0	88.0	65.0	76.5	59.0	66.0
1	2.0	88.0	65.0	76.5	59.0	66.0

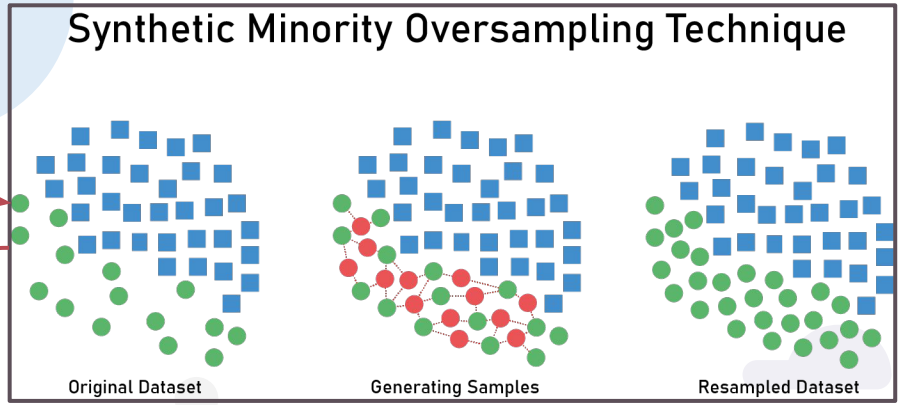
Preprocessing

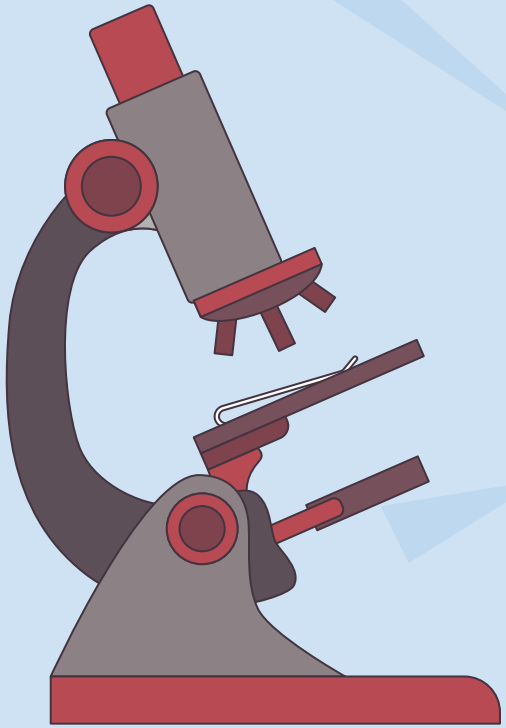
Dummify Trap feature and created and additional trap column for kaggle prediction

trap_T900	trap_T903	trap_T234
0	0	0
0	0	0

Data Imbalance	
Without WNV	0.95
With WNV	0.05

Numerical features: StandardScaler()





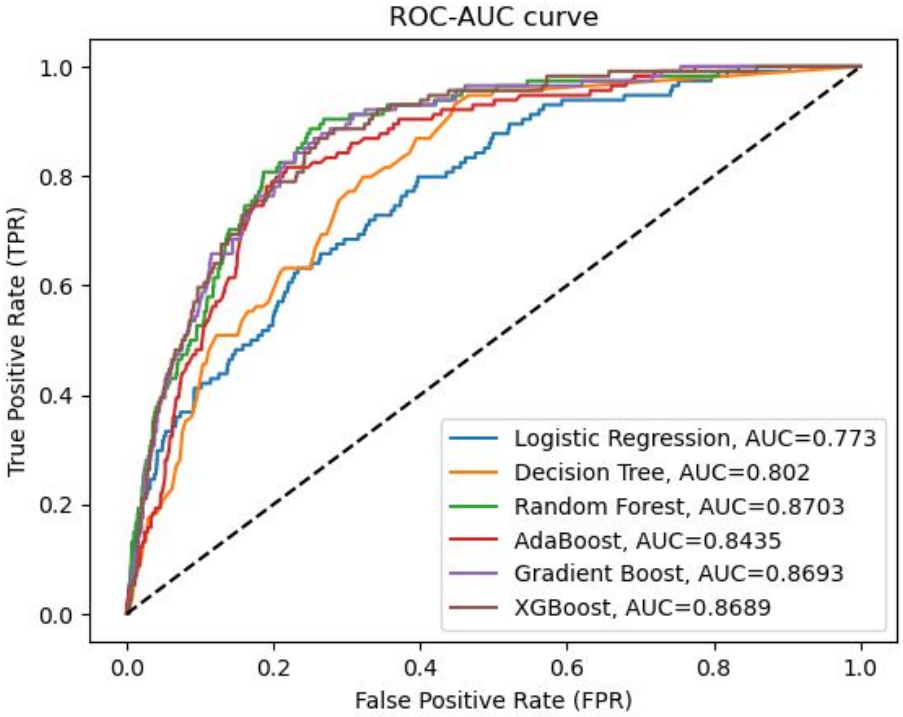
04

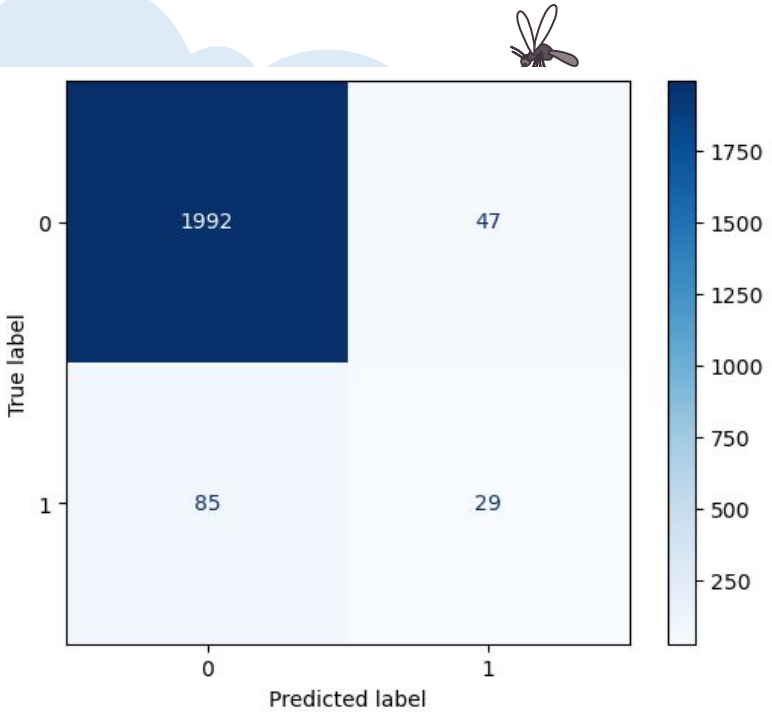
Model Evaluation

Model Evaluation

Classifier	Train score	Test score	Generalisation	Recall	F1 score	ROC AUC	Runtime
Logistic Regression	0.81	0.74	9.19%	0.64	0.21	0.77	13 min 33 s
Decision Tree	0.92	0.89	2.13%	0.33	0.25	0.80	11.2s
Random Forest	0.96	0.94	2.70%	0.25	0.31	0.87	32 min 42 s
Ada Boost	0.89	0.85	4.56%	0.58	0.30	0.84	1 min 33 s
Gradient Boost	0.97	0.94	3.21%	0.12	0.18	0.87	5 min 26 s
XGBoost	0.97	0.94	2.98%	0.088	0.13	0.87	8 min 37 s

Model Evaluation





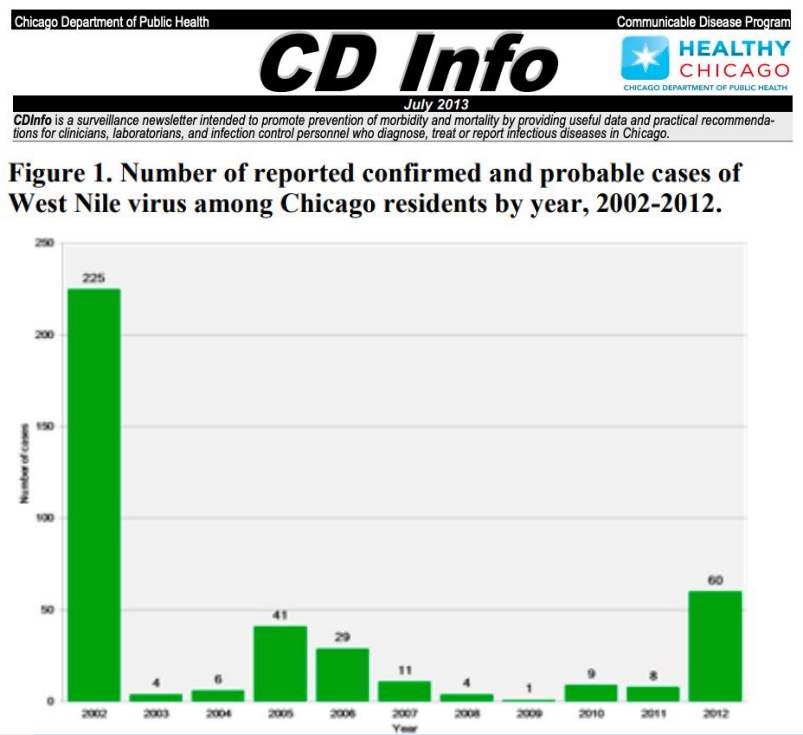
Final Model + HyperParameter

Final Model:	Random Forest Classifier
Reason:	Has good generalisation, with best F1 score and good auc_roc score
Best parameters:	rf__max_depth : None, rf__min_samples_leaf: 5, rf__min_samples_split: 10, rf__n_estimators: 1500

05

Cost Benefit Analysis

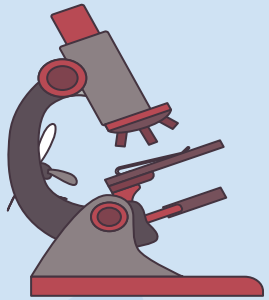
Cases on WNV among Chicago residents



Hospitalisation

Cases between 2007-2012:	1- 60
Average Initial Cost (USD)	~ $7,501/ person
Average Long term Cost (USD)	~ $7,015/ person
Total (USD / Annum)	$14,516 - $870,960

Cost Benefit Analysis



Adulticides (Zenivex E20)

	Whole of Chicago	Target Traps
Traps	N/A	27
Area (acres)	145,300	69,120
Cost (USD / acre)	67¢	
Spray Frequency	11#	
Total (USD)	$1,070,861	$509,414.40

Cost savings:
USD $561,446.60

~ 52%

*Each square mile = 640 acres. Up to 4sq mile = 2560 acres/ trap
#Based on the surveillance newsletter from CDPH

06

Conclusion, Recommendation

Conclusions

Our model has identified 27 traps for targeted action and with the cost-benefit analysis performed, significant cost savings could be achieved.

However, the current analysis is oversimplified and generalised which may not be an accurate reflection of the projected spendings/ cost incurred. In addition, other factors such as population density is not considered to evaluate the cost and effort of spraying in low density neighbourhood.

Recommendation

Explore other efforts such as:

1. Introducing male wolbachia mosquitoes to mate with virus carrying female Culex mosquitoes to prevent eggs from hatching hence reducing the likelihood of new mosquitoes carrying West Nile Virus.
2. Looking into larval control in stagnant water bodies that could potentially be a favourable mosquitoes breeding ground to then reduce the number of adult mosquitoes.
3. Educating people on preventive methods to reduce the breeding of mosquitoes

Thanks

Do you have any questions?

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