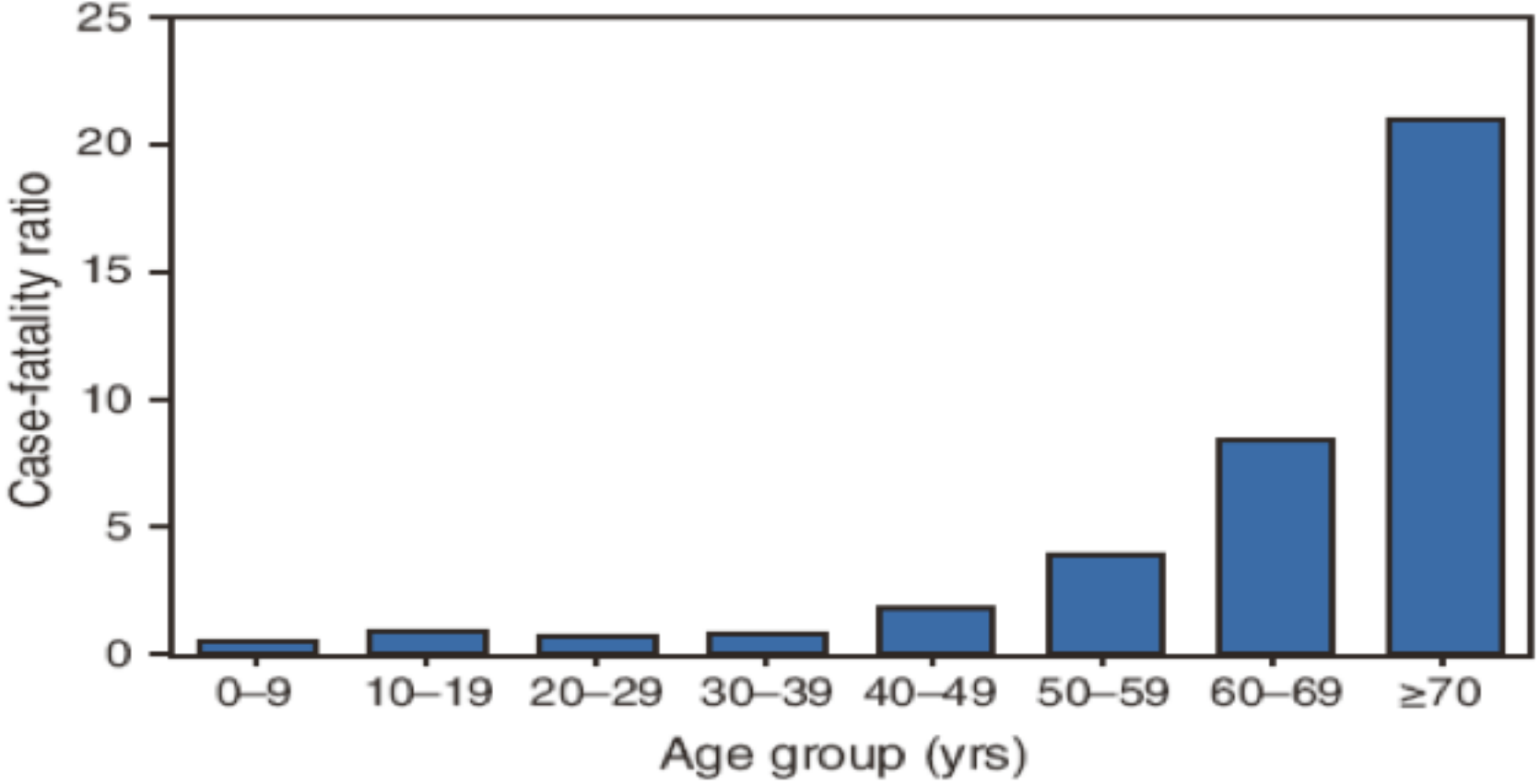
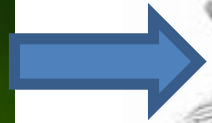
Maximizing effectiveness of mosquito spray in Chicago

By: Shane McGuckian, Ryan Lexow, and Paul Cyr

Case-Fatality Ratio



The “Illinois State Bird”



Problem Statement

Can we predict where to spray insecticide to reduce number of mosquitos that have West Nile?

West Nile Background

- Chicago had its first recorded death from West Nile in July 2018, this is likely to increase as more mosquitos become infected with West Nile

- Someone can contract West Nile with out showing symptoms, so we can not know the true number of people that are infected

- Since we are only catching a small percentage of the mosquito population in our traps it is hard to know where the infected mosquitos are most prevalent

- Focusing on other factors like high risk populations and high mosquito populations might be our best way to combat the problem with the information we have

Exploratory Data Analysis

EDA

- Drop columns: CodeSum, SnowFall, Water1 and Depth (values missing or very insignificant)
- Combined Station 1 and Station 2 for each day to a single row in order to reduce redundancy
- We dropped Depart, Sunrise, and Sunset from Station 2
- Converted the Date column to a date-time object and extracted date-time elements such as day of the week, week, month, year, and sunrise and sunset hours and minutes
- Filled missing values: changed M to 0 and T to 0.005
- Ensure train and test have same columns for modeling

EDA

Spray

- Combined spray data with train and test on a Date index column to show which dates the sprays occurred on

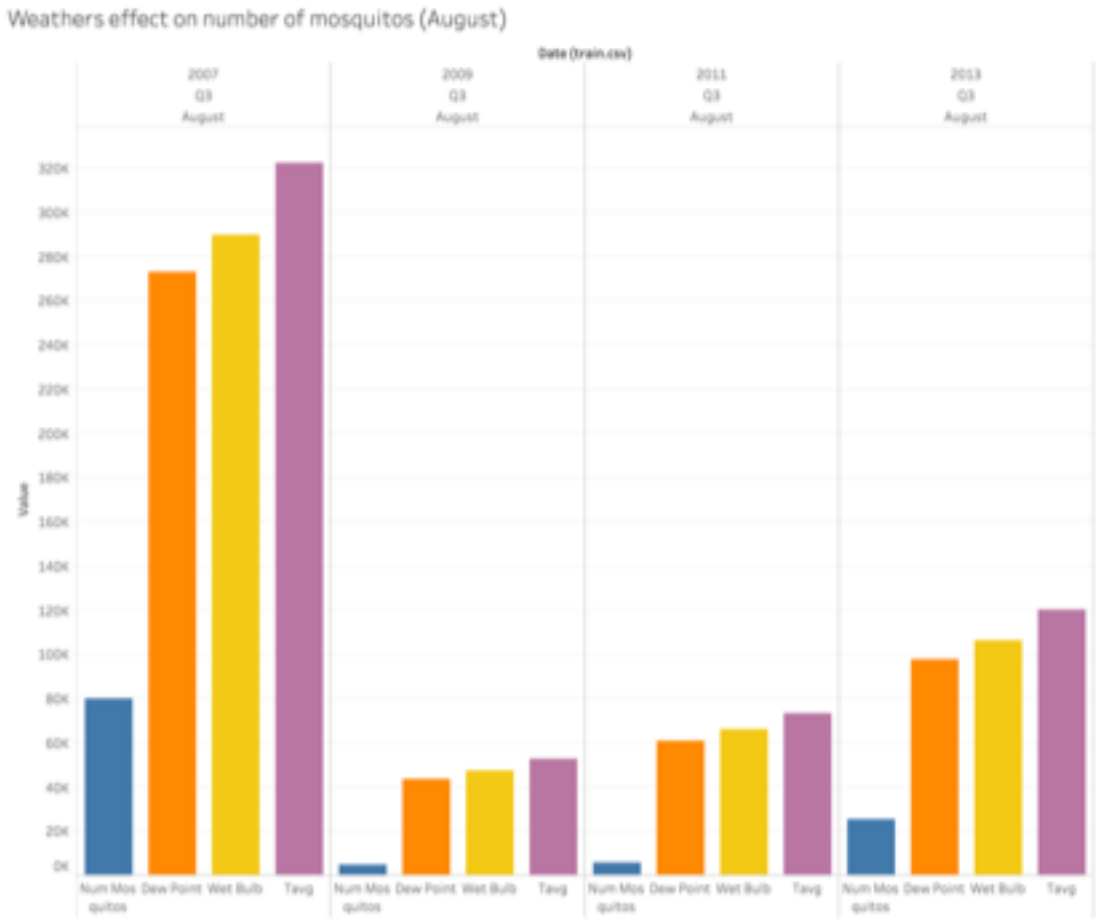
- Created a classifier for whether or not spraying occurred near a given trap within the past two weeks.

- Since there was no spray data for years in the test dataset we ended up not using spray data

Weather

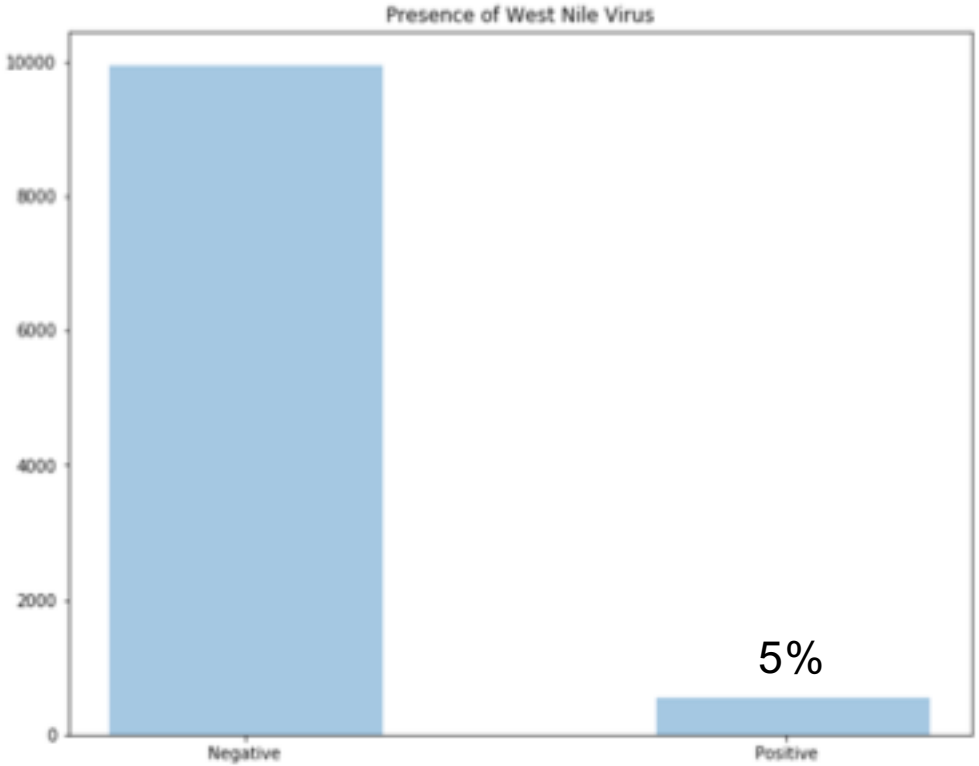
Most highly correlated features:

- Dew Point
- Wet Bulb
- Average Temp.



EDA

Percentage of traps that tested positive for WNV



EDA

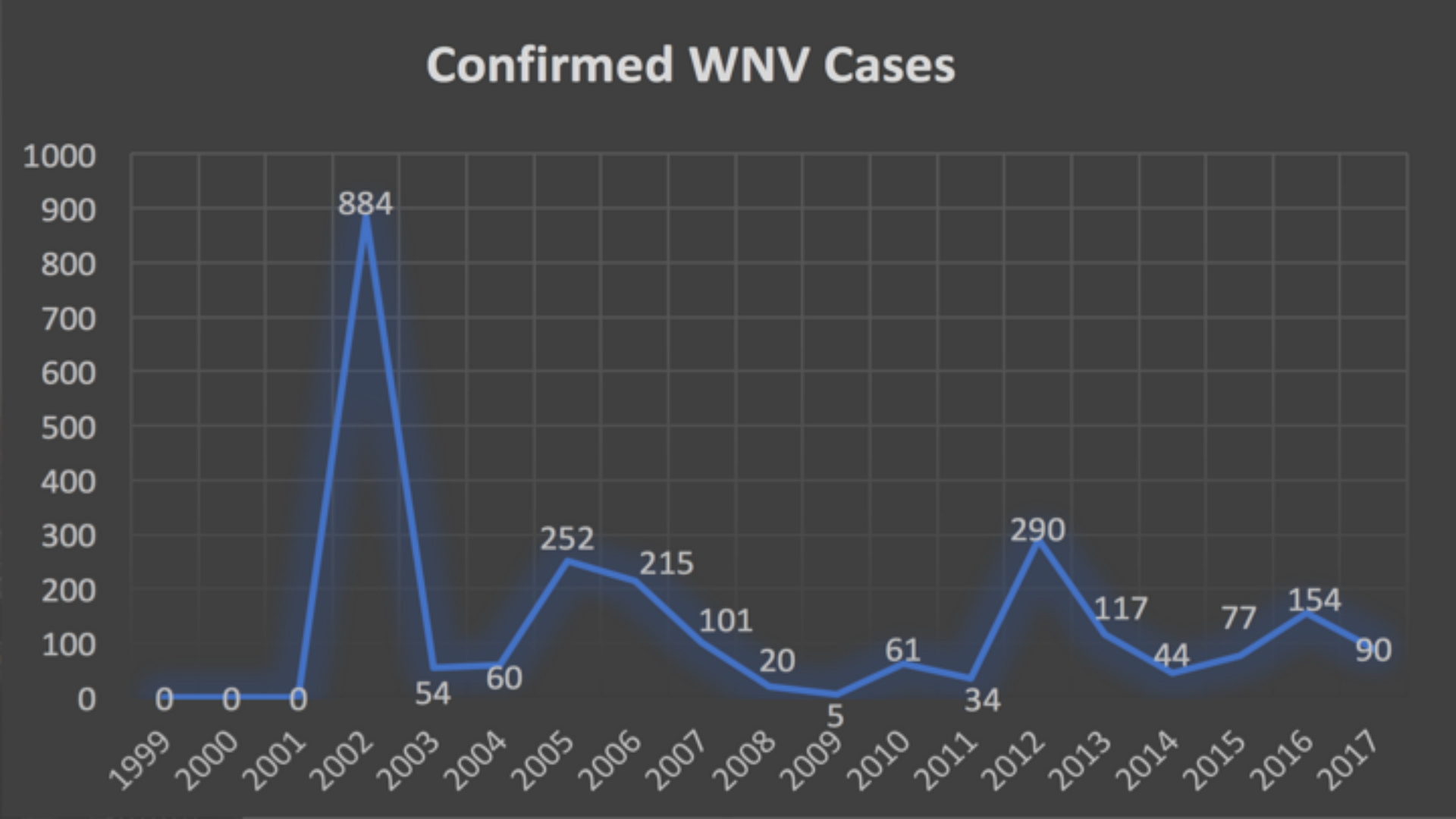
Preprocessing

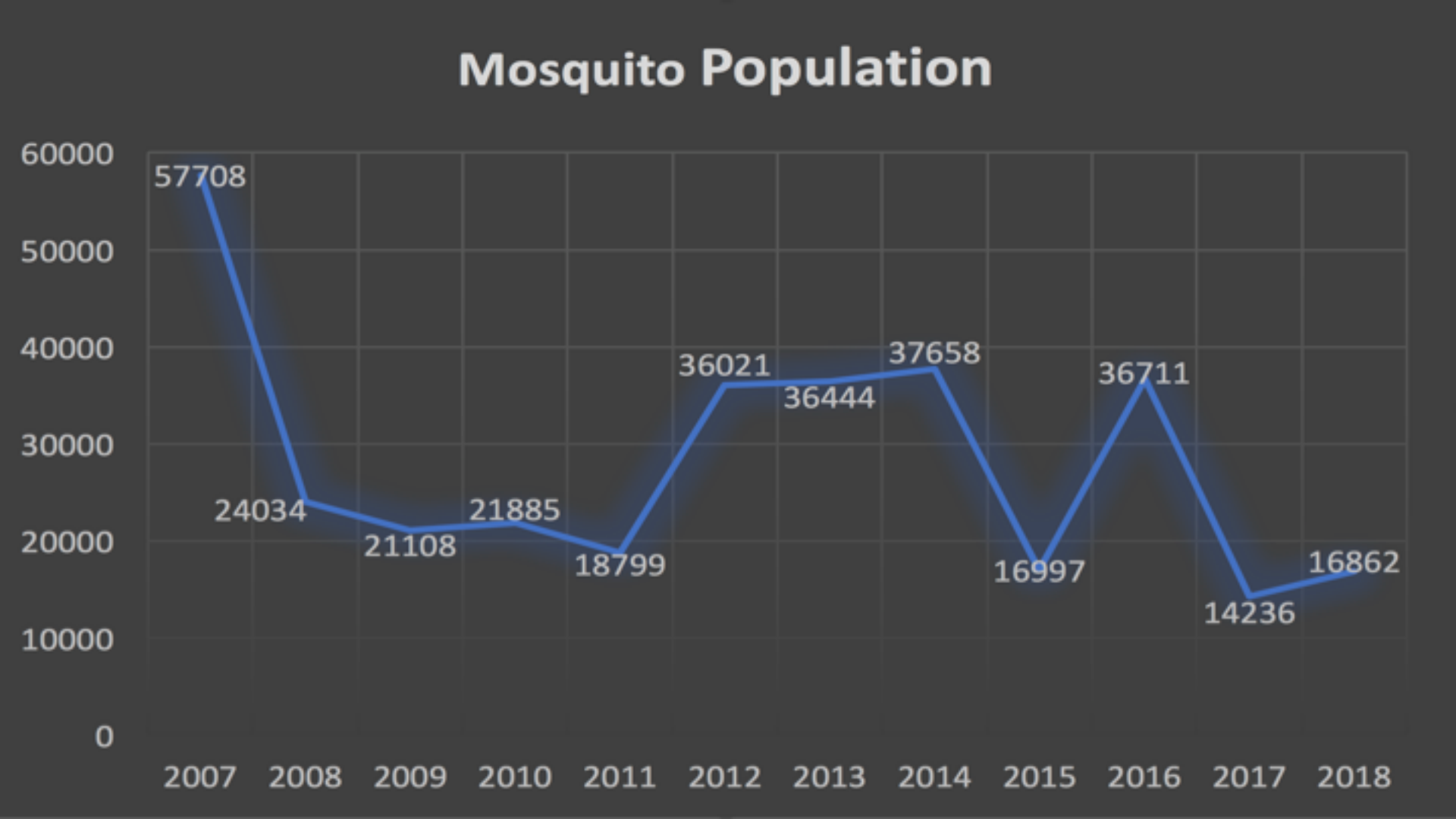
- Train, test, split

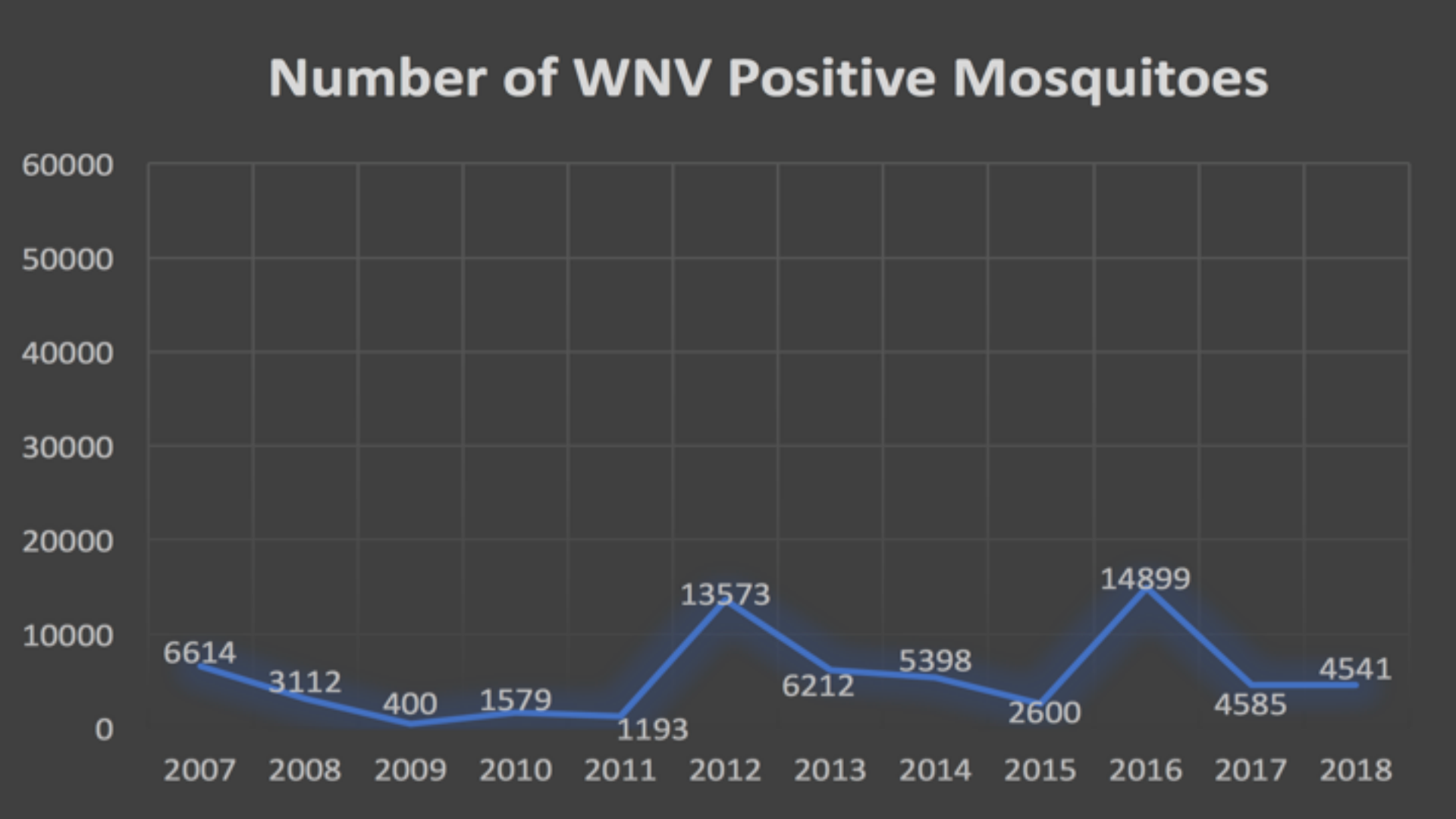
- Scaled the data

- Corrected imbalanced classes using SMOTE

- SMOTE creates slightly tweaked observation instead directly duplicating the minority class (bootstrapping)


Confirmed WNV Cases
1000
900
800
700
600
500
400
300
200
100
0
0
0
0
884
54
60
252
215
101
20
5
61
34
290
117
44
77
154
90
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017


Mosquito Population
60000
50000
40000
30000
20000
10000
0
57708
24034
21108
21885
18799
36021
36444
37658
16997
36711
14236
16862
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018

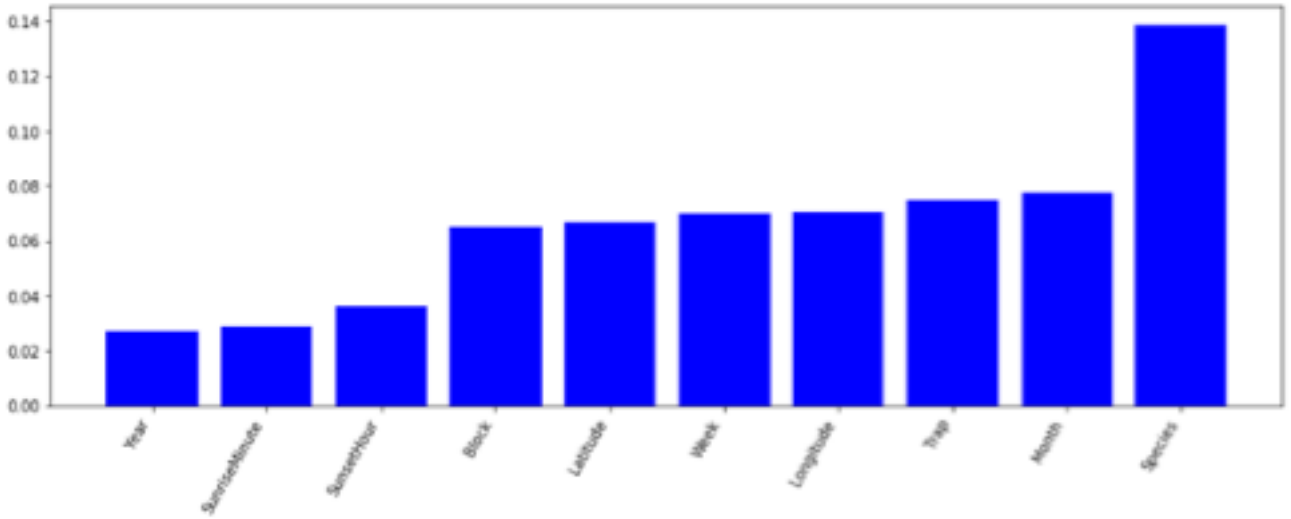

Number of WNV Positive Mosquitoes
60000
50000
40000
30000
20000
10000
0
6614
3112
400
1579
1193
13573
6212
5398
2600
14899
4585
4541
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018

Modeling

Random Forest

	Predicted Negative	Predicted Positive
Actual Negative	2374	110
Actual Positive	101	42

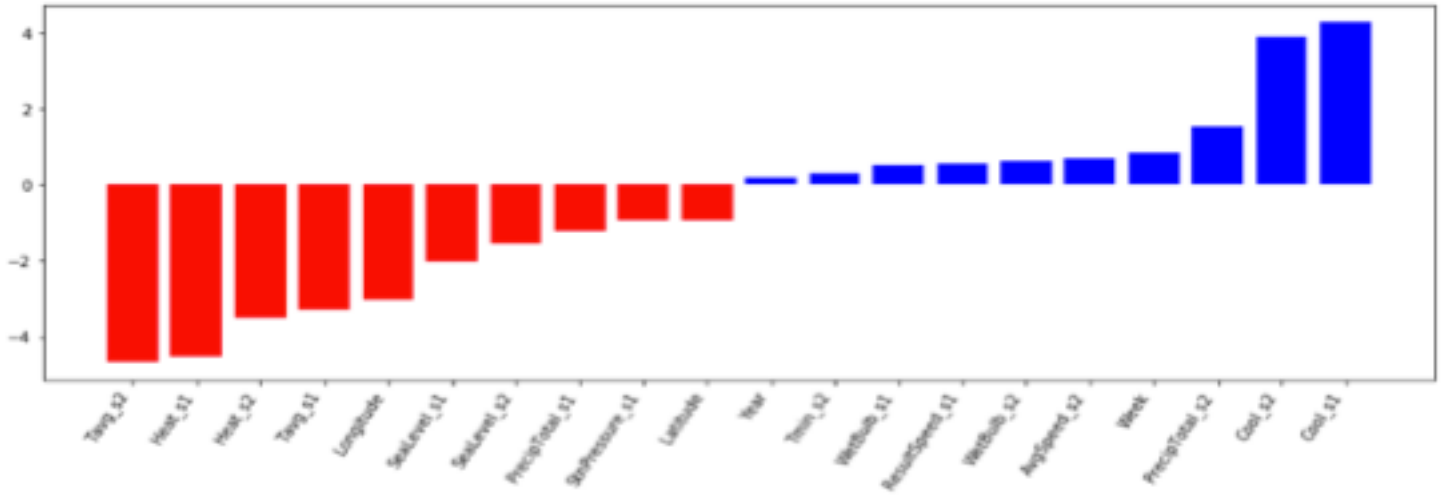
Metrics	
Accuracy	0.919680
Misclassifaction Rate	0.080320
Sensitivity/Recall	0.293706
Specificity	0.955717
Precision	0.276316



Logistic Regression

	Predicted Negative	Predicted Positive
Actual Negative	1728	756
Actual Positive	37	106

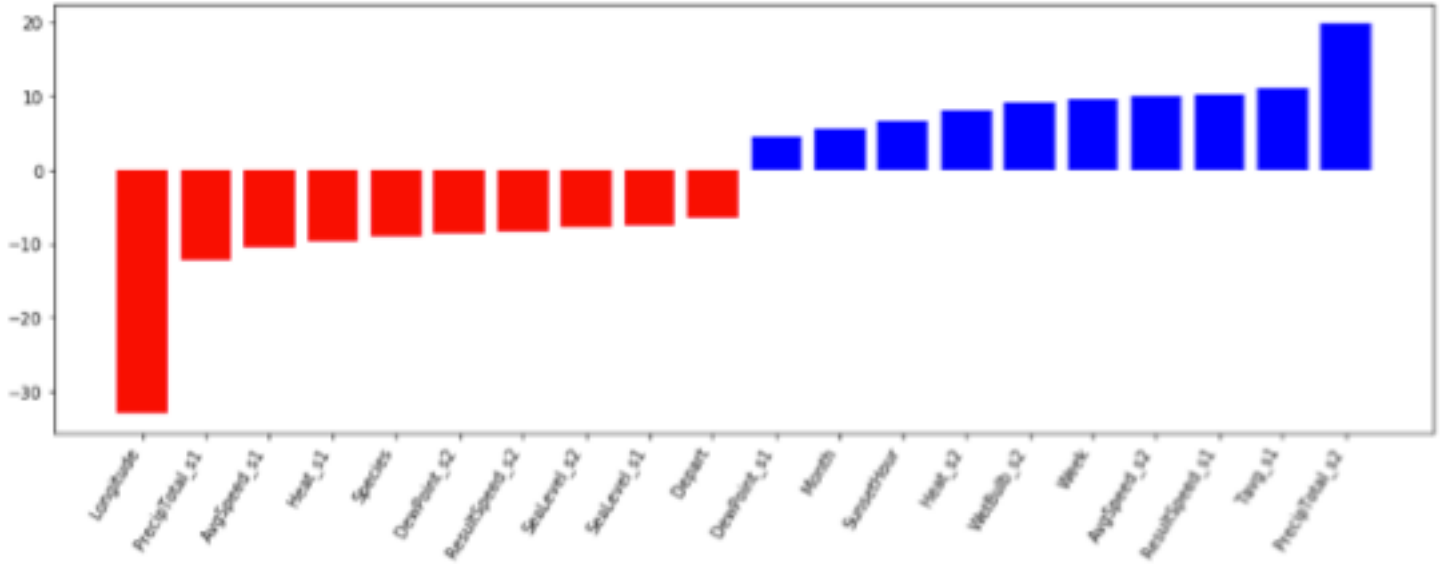
	Metrics
Accuracy	0.698135
Misclassifaction Rate	0.301865
Sensitivity/Recall	0.741259
Specificity	0.695652
Precision	0.122970



SVC

	Predicted Negative	Predicted Positive
Actual Negative	1740	744
Actual Positive	33	110

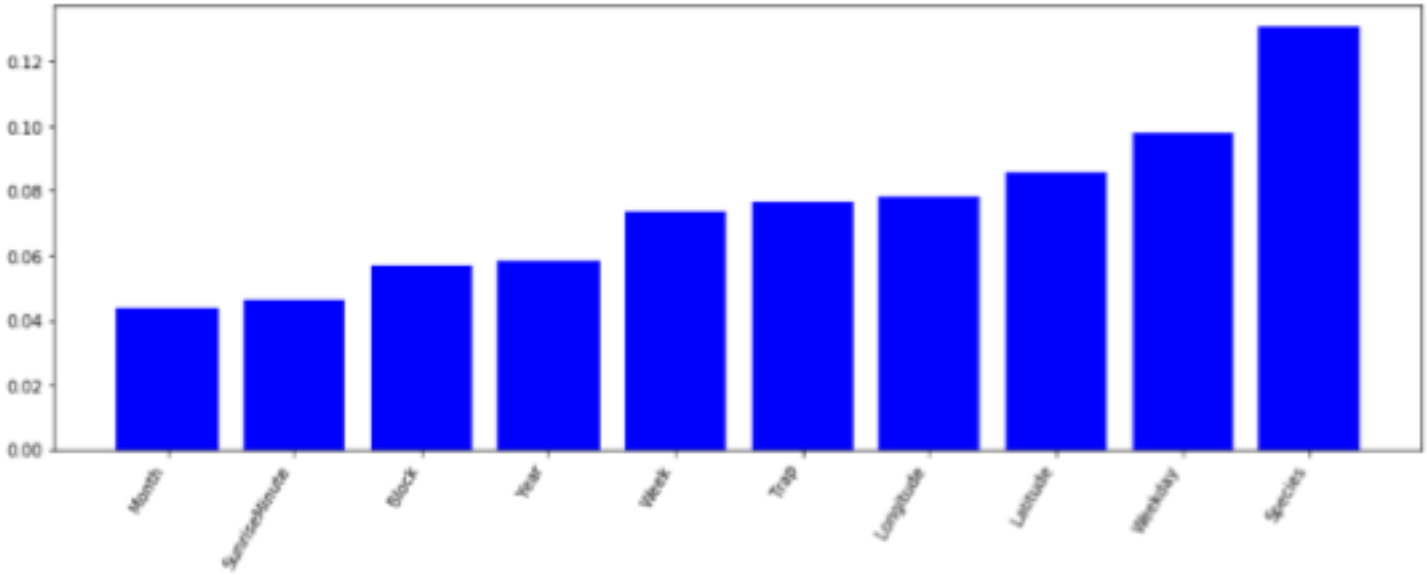
Metrics	
Accuracy	0.704225
Misclassifaction Rate	0.295775
Sensitivity/Recall	0.769231
Specificity	0.700483
Precision	0.128806



XGBoost

	Predicted Negative	Predicted Positive
Actual Negative	2166	318
Actual Positive	63	80

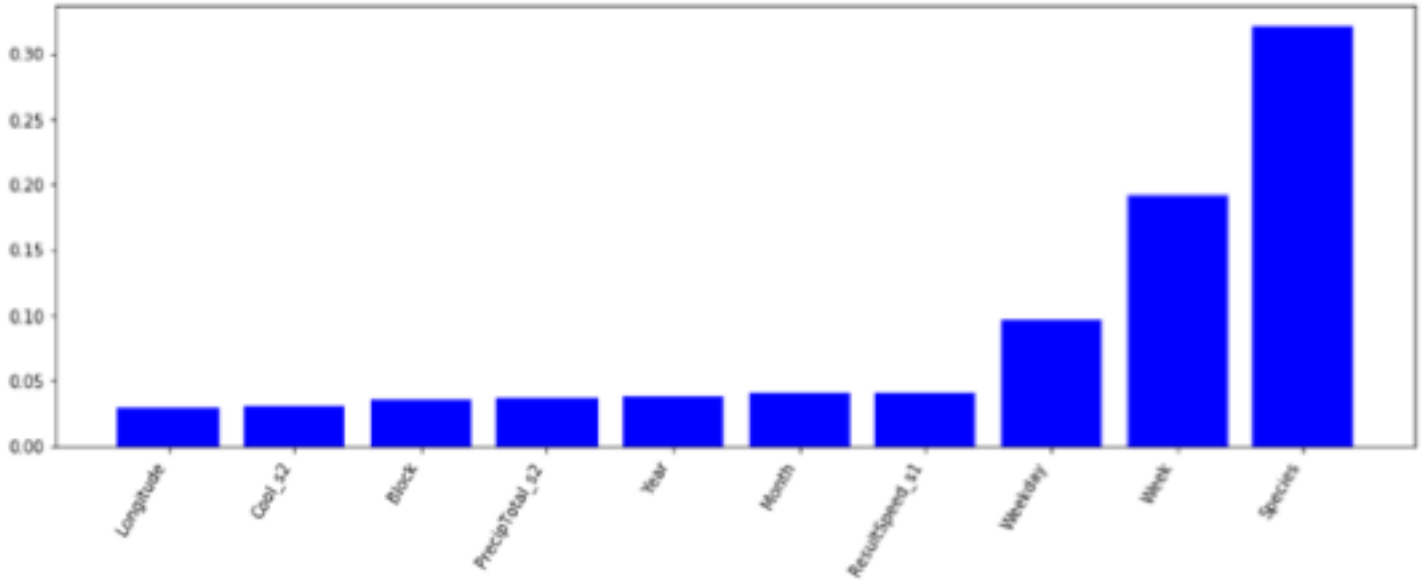
Metrics	
Accuracy	0.854968
Misclassifaction Rate	0.145032
Sensitivity/Recall	0.559441
Specificity	0.871981
Precision	0.201005

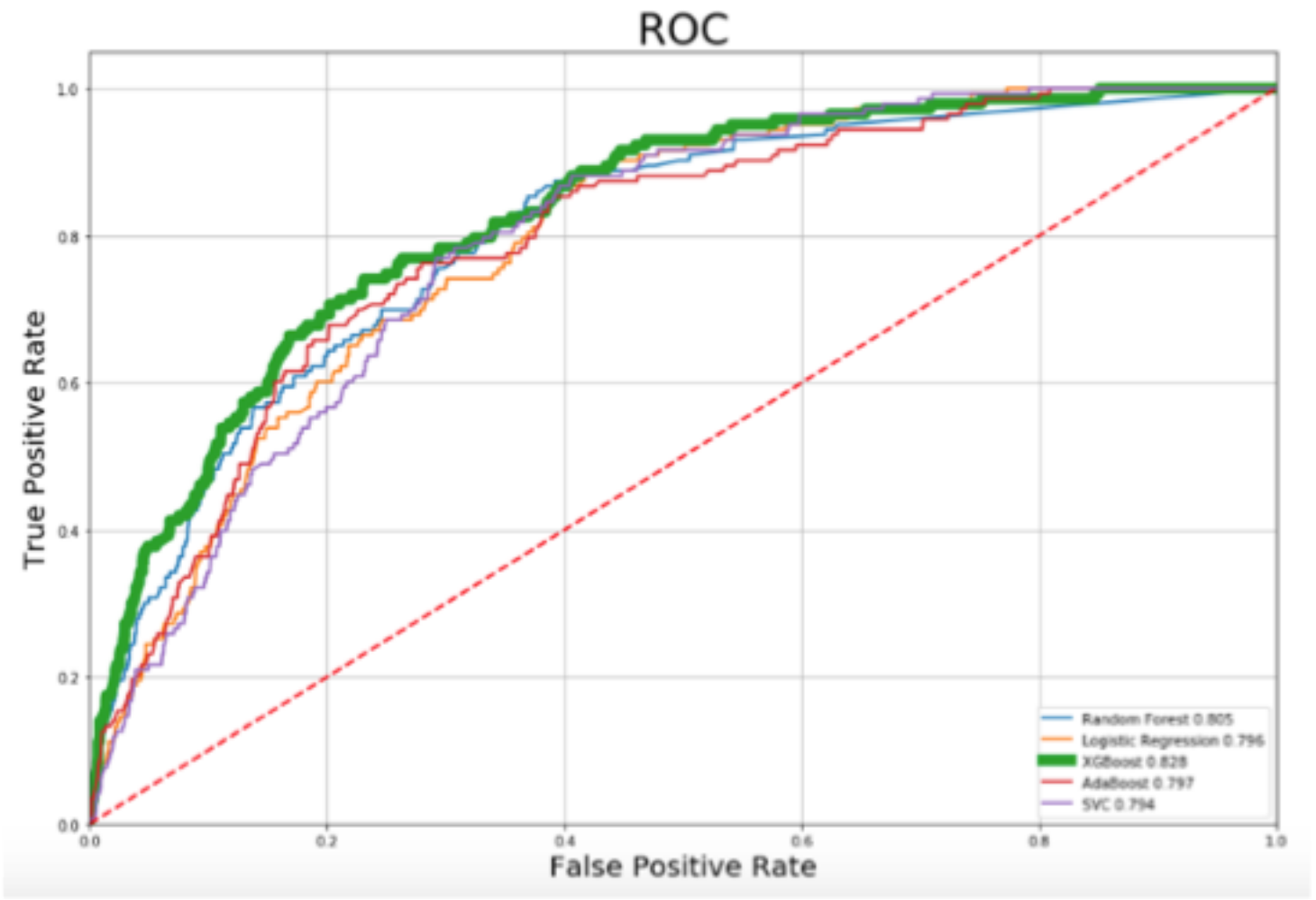


AdaBoost

	Predicted Negative	Predicted Positive
Actual Negative	2076	408
Actual Positive	55	88

	Metrics
Accuracy	0.854968
Misclassifaction Rate	0.145032
Sensitivity/Recall	0.559441
Specificity	0.871981
Precision	0.201005




ROC
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
Random Forest 0.805
Logistic Regression 0.796
XGBoost 0.828
AdaBoost 0.797
SVC 0.794

Cost-benefit analysis

Cost-benefit analysis

Is spray effective?

- No effective treatment exists for WNV; prevention of disease relies on management of mosquitoes through various control tactics.
- Spaying decreased WNV mosquitoes (*Culex pipiens* and *Cx. Tarsalis*) by 57.5%, compared with the pre-spray population in the treated area.
- A decrease in WNV infection rates in mosquitoes to 3.9/1,000 for trapped females in the treated areas, compared with 6.7/1,000 in the untreated areas.
- In treated areas, no incidents of human cases were found, compared with 18 cases in the untreated area.
- Consequently, the emergency aerial spray seemed to effectively reduce both mosquito populations and human WNV cases.
- This is evidence of planned spraying being cost effective.

Cost-benefit analysis

Economics savings from spraying (Sacramento study)

- Costs of inpatient and outpatient medical care, productivity loss, the state's public health department, and vector control were ≈$20.14 million for the 329 cases, including $9.2 million for mosquito control and public health agency costs.
- Total cost of the 2005 Sacramento County WNV epidemic was ≈$2,979,037. Costs for treating WNND patients alone exceeded costs of emergency vector control by $1,438,619, a ratio of 3:1.
- This difference suggests that for the benefits of the vector control to outweigh the cost of the epidemic, the spray event would need to prevent only 15 WNND cases.
- https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3322011/

Cost-benefit analysis

Where should be sprayed?

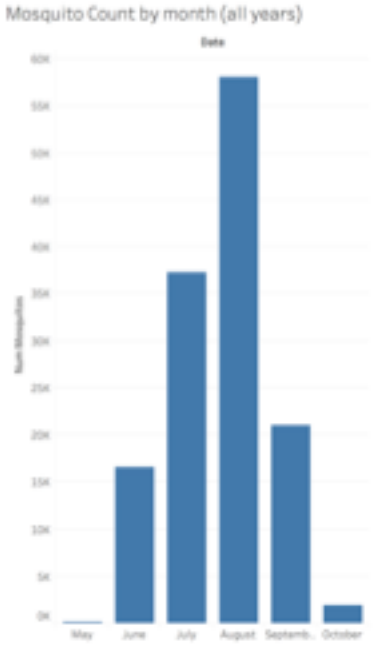
- They estimated an average cost per case of approximately $27,000.
- The total societal costs of WNV hospitalized cases and deaths as reported to CDC for 1999–2012 was estimated to be roughly $778 million dollars or ~$56 million per year.
- Illinois spent $5.3 million on mosquito abatement in 2015.
- However, the annual cost of WNV disease varies substantially as the number of WNV disease cases has ranged from 21 cases in 2000 to 9,862 cases in 2003
- http://www.ajtmh.org/content/journals/10.4269/ajtmh.13-0206#html_fulltext

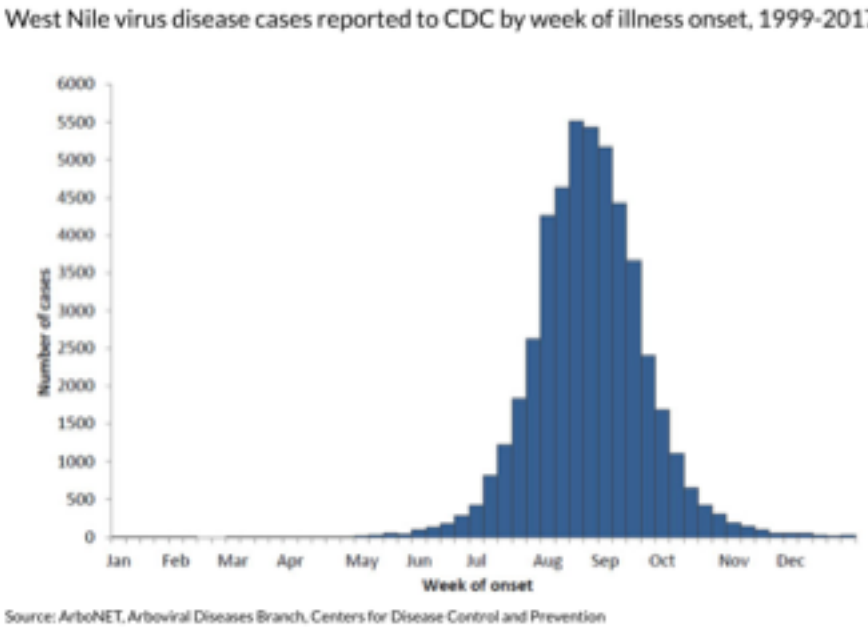
Conclusions

Recommendations

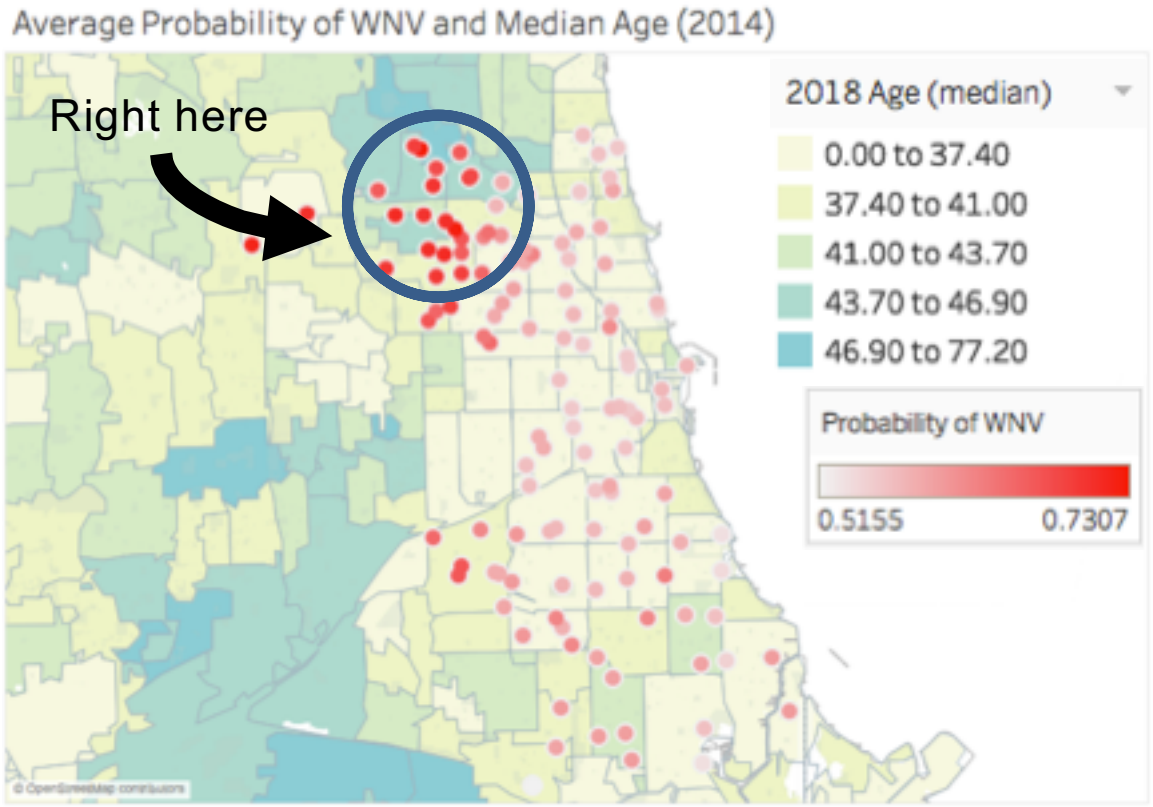
When to spray?

- In August WNV cases are increasing the most and when it would be most effective to spray





Where should we spray?



Recommendations

Alternatives

- Use CRISPR to make mosquitoes disease-resistant or kill all mosquitos world wide
- Vaccinations

https://www.vox.com/science-and-health/2018/5/31/17344406/crispr-mosquito-malaria-gene-drive-editing-target-africa-regulation-gmo
https://www.smithsonianmag.com/innovation/kill-all-mosquitos-180959069/

Additional research

- How does weather effect mosquitos life span/reproduction cycle
- Bird and mosquito WNV relationship
- Landscape (elevation, vegetation)
- Housing (year built)
- Age demographics
- Spray vs. Larvicide

https://ij-healthgeographics.biomedcentral.com/articles/10.1186/1476-072X-6-10