Predicting West Nile

- Raj
    - Riordan
        - Mud

Which features are best at identifying the presence of West Nile?

Our Process

1. Question ✔
2. Data
3. Explore
4. Model
5. Communicate

Epidemiology

1. WNV introduced to the US in 1999 (NYC)

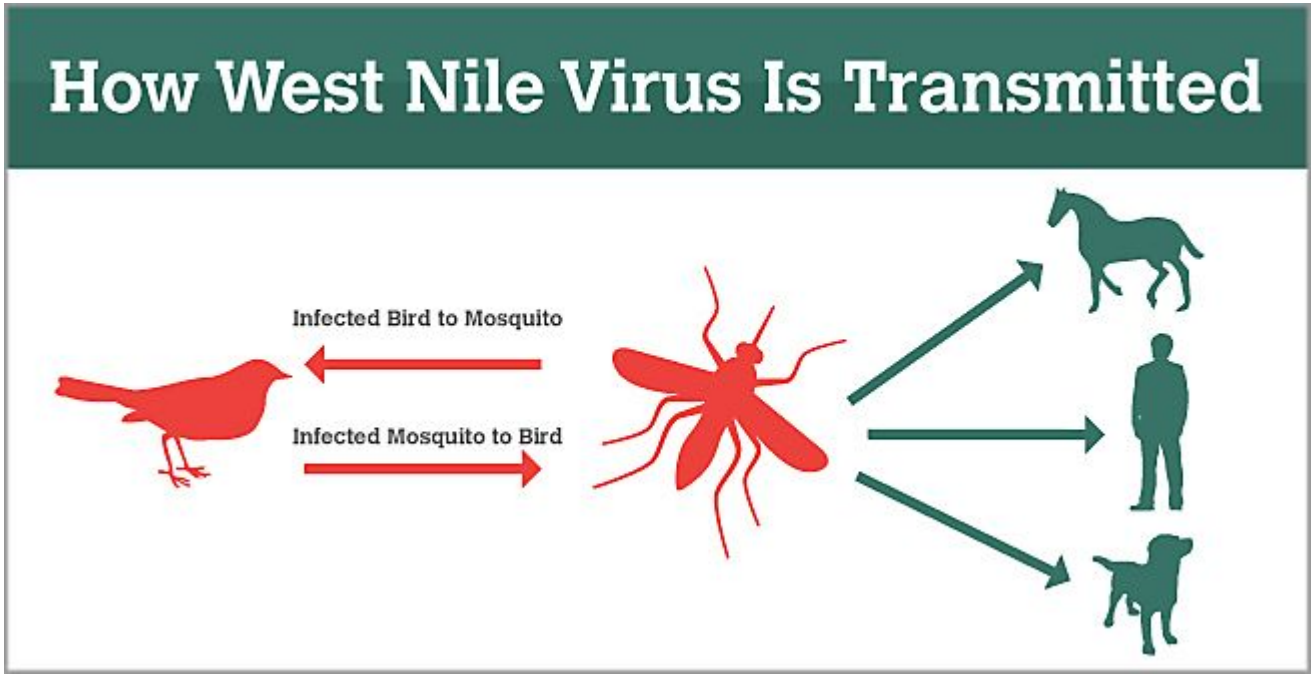
   Since then it has spread to the entire continental US

2. Birds are main carriers of the virus

   99% of crows died in 1999

3. Transmitted to humans via *culex* mosquitoes

Mosquitoes, Birds, and West Nile



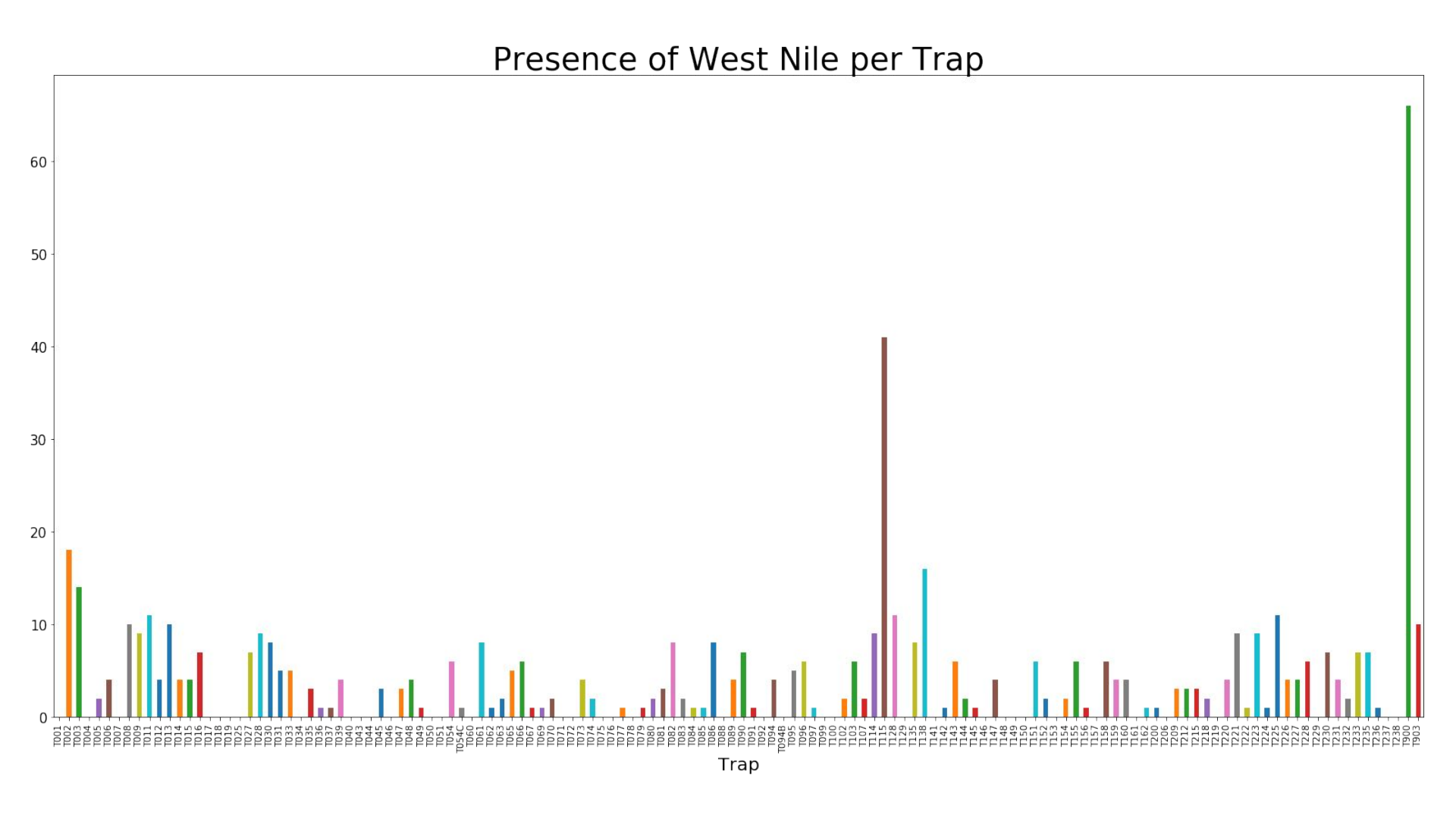
Our Process

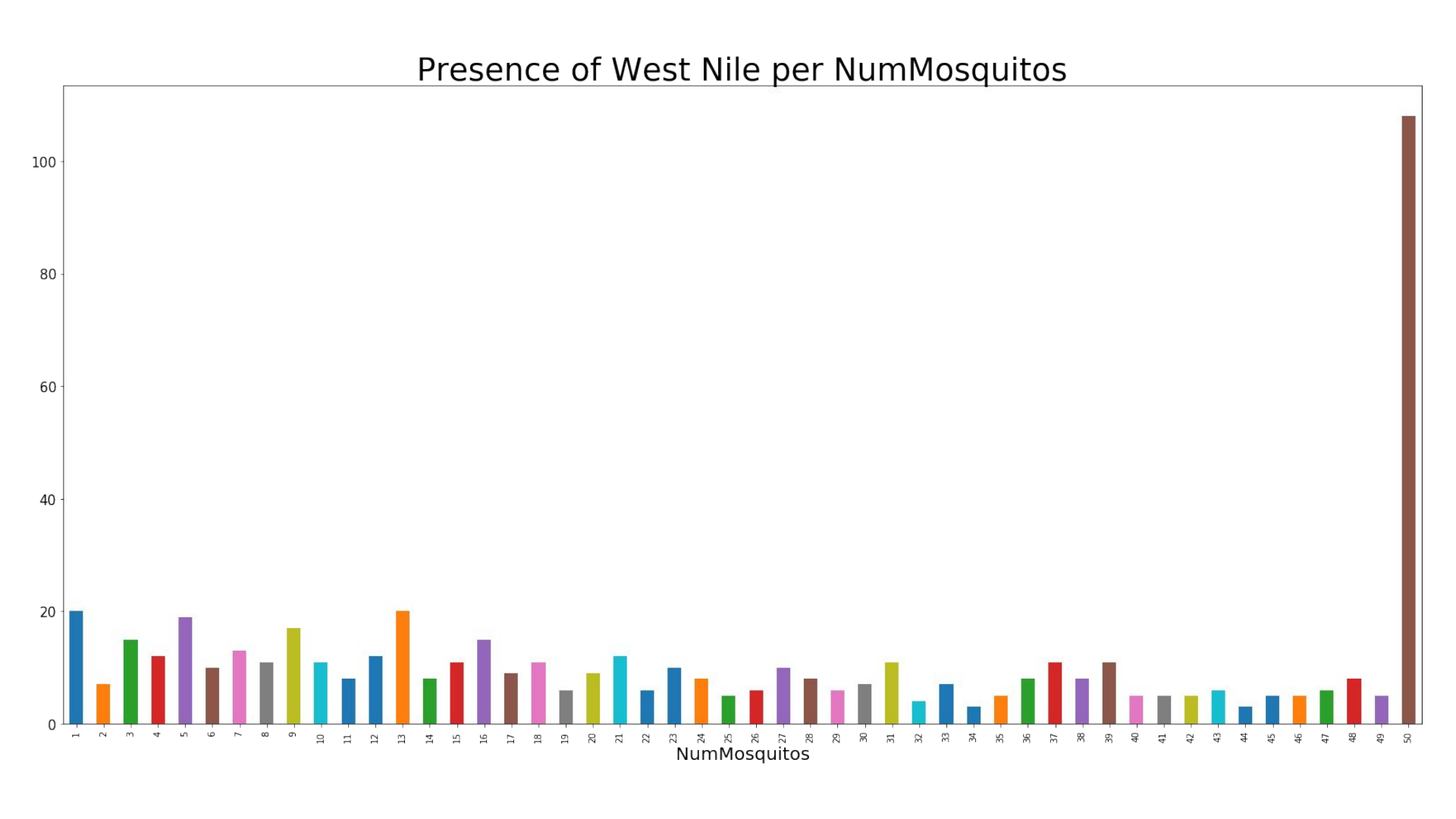
1. Question ✔
2. Data ✔
3. Explore
4. Model
5. Communicate

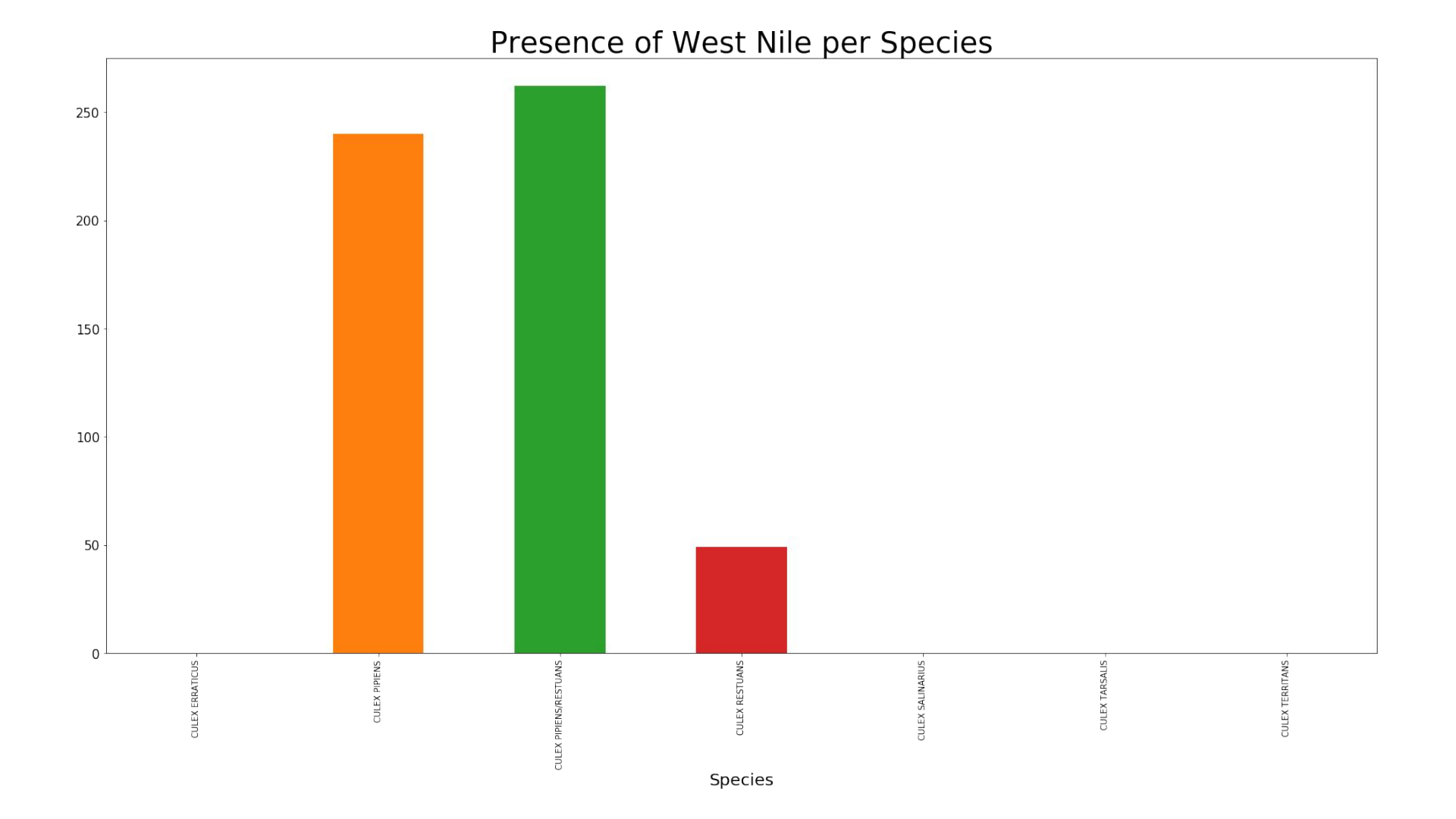
Explore - Traps, Spray, and Weather

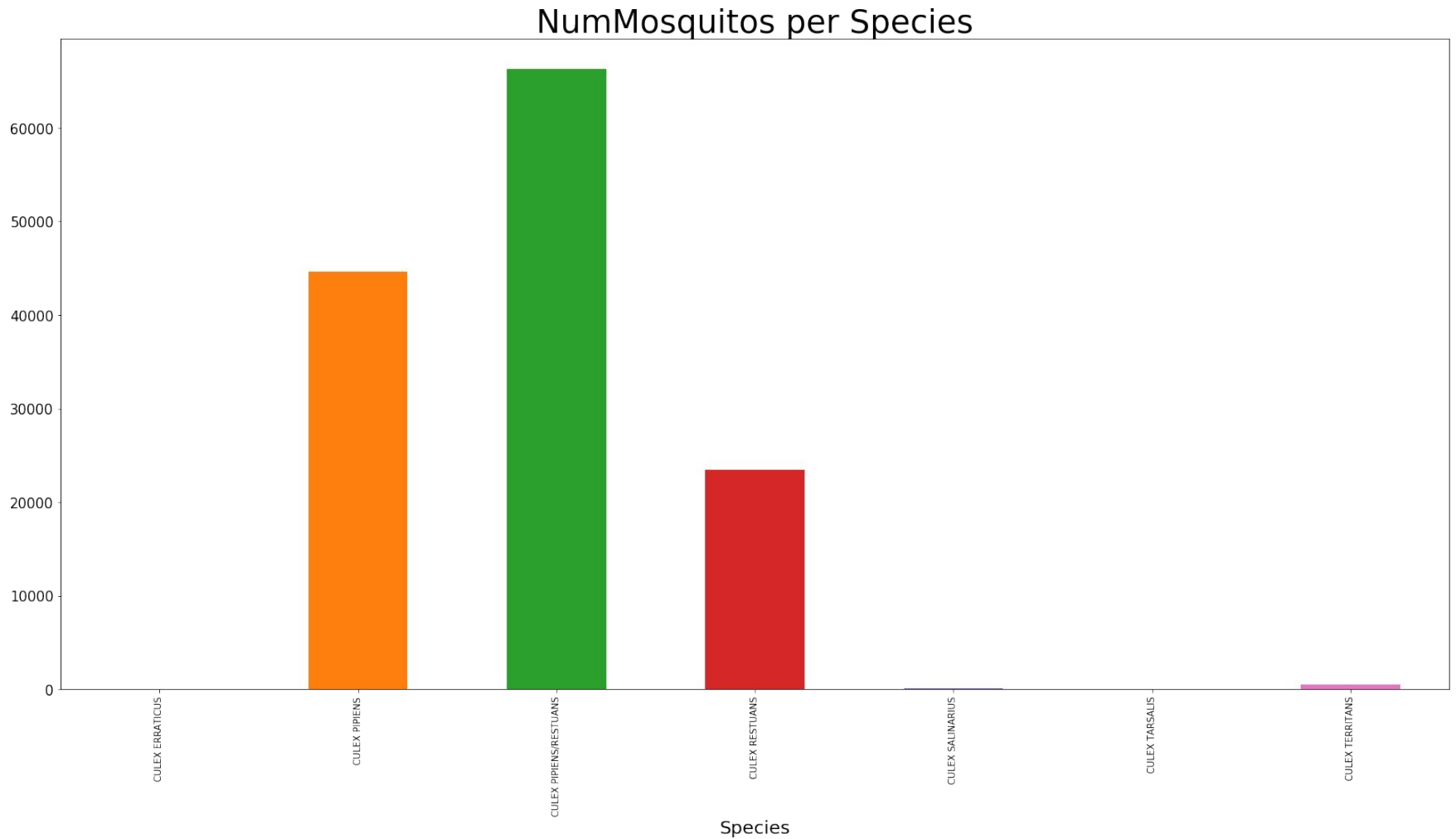
- Time
- Counts
- Location
- West Nile

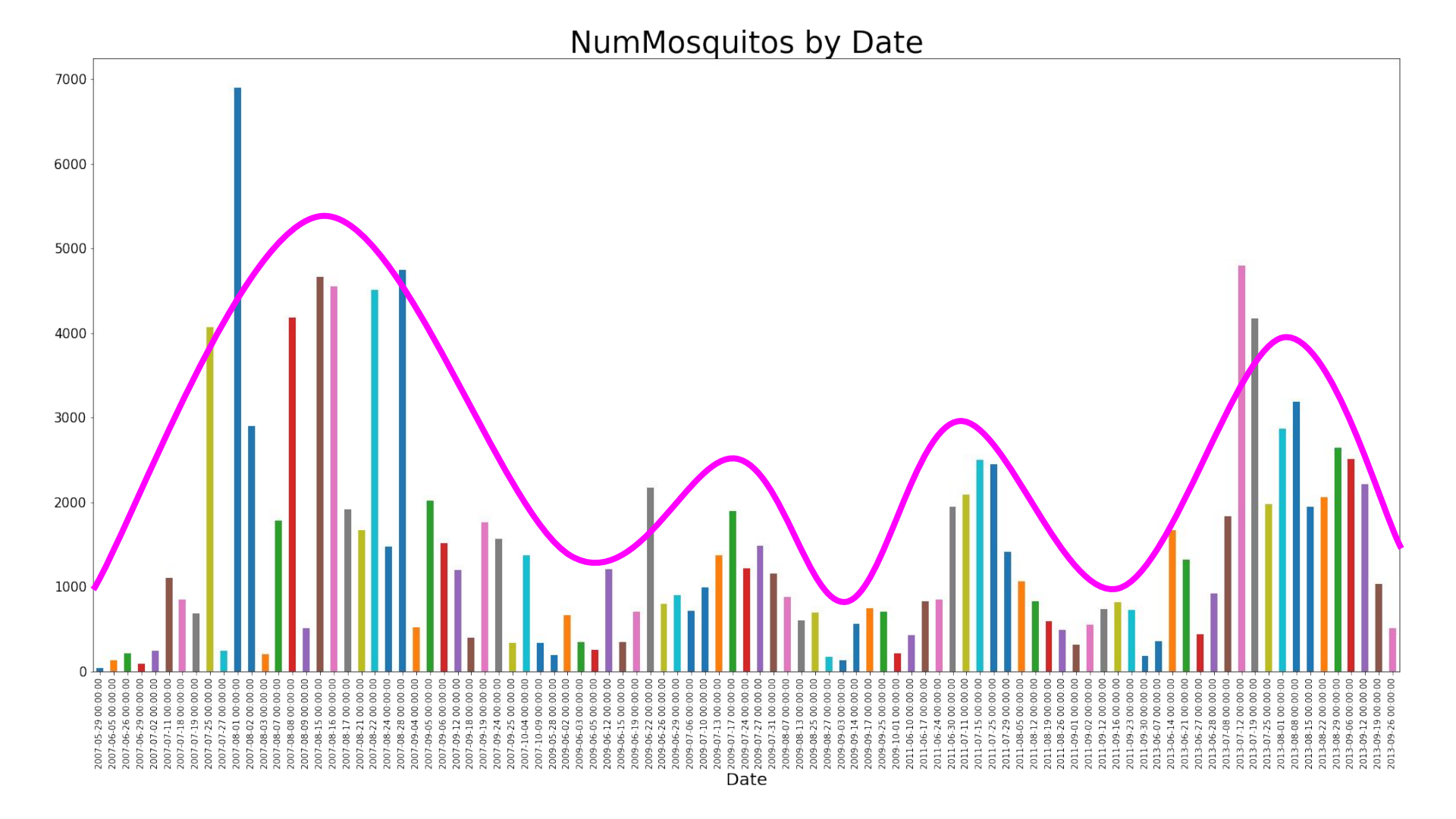
Any Trends?

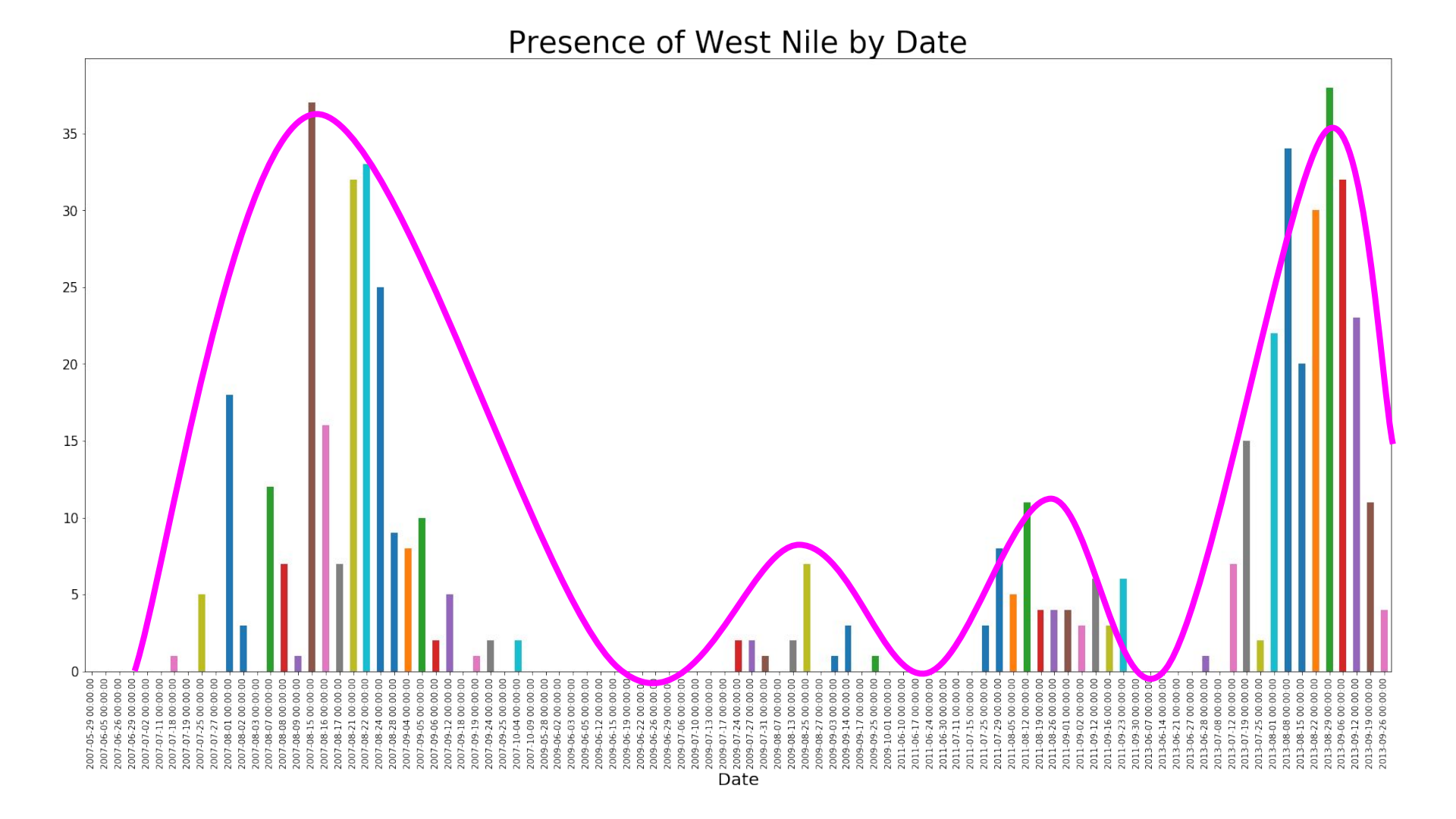

Presence of West Nile per Trap
60
50
40
30
20
10
0
T001 T002 T003 T004 T005 T006 T007 T008 T009 T011 T012 T013 T014 T015 T016 T017 T018 T019 T025 T027 T028 T030 T031 T033 T034 T035 T036 T037 T039 T040 T043 T044 T045 T046 T047 T048 T049 T050 T051 T054 T054C T060 T061 T062 T063 T065 T066 T067 T069 T070 T071 T072 T073 T074 T075 T076 T077 T078 T079 T080 T081 T082 T083 T084 T085 T086 T088 T089 T090 T091 T092 T094 T094B T095 T096 T097 T099 T100 T102 T103 T107 T114 T115 T128 T129 T135 T138 T141 T142 T143 T144 T145 T146 T147 T148 T149 T150 T151 T152 T153 T154 T155 T156 T157 T158 T159 T160 T161 T162 T200 T206 T209 T212 T215 T218 T219 T220 T221 T222 T223 T224 T225 T226 T227 T228 T229 T230 T231 T232 T233 T235 T236 T237 T238 T900 T903
Trap


Presence of West Nile per NumMosquitos
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
NumMosquitos

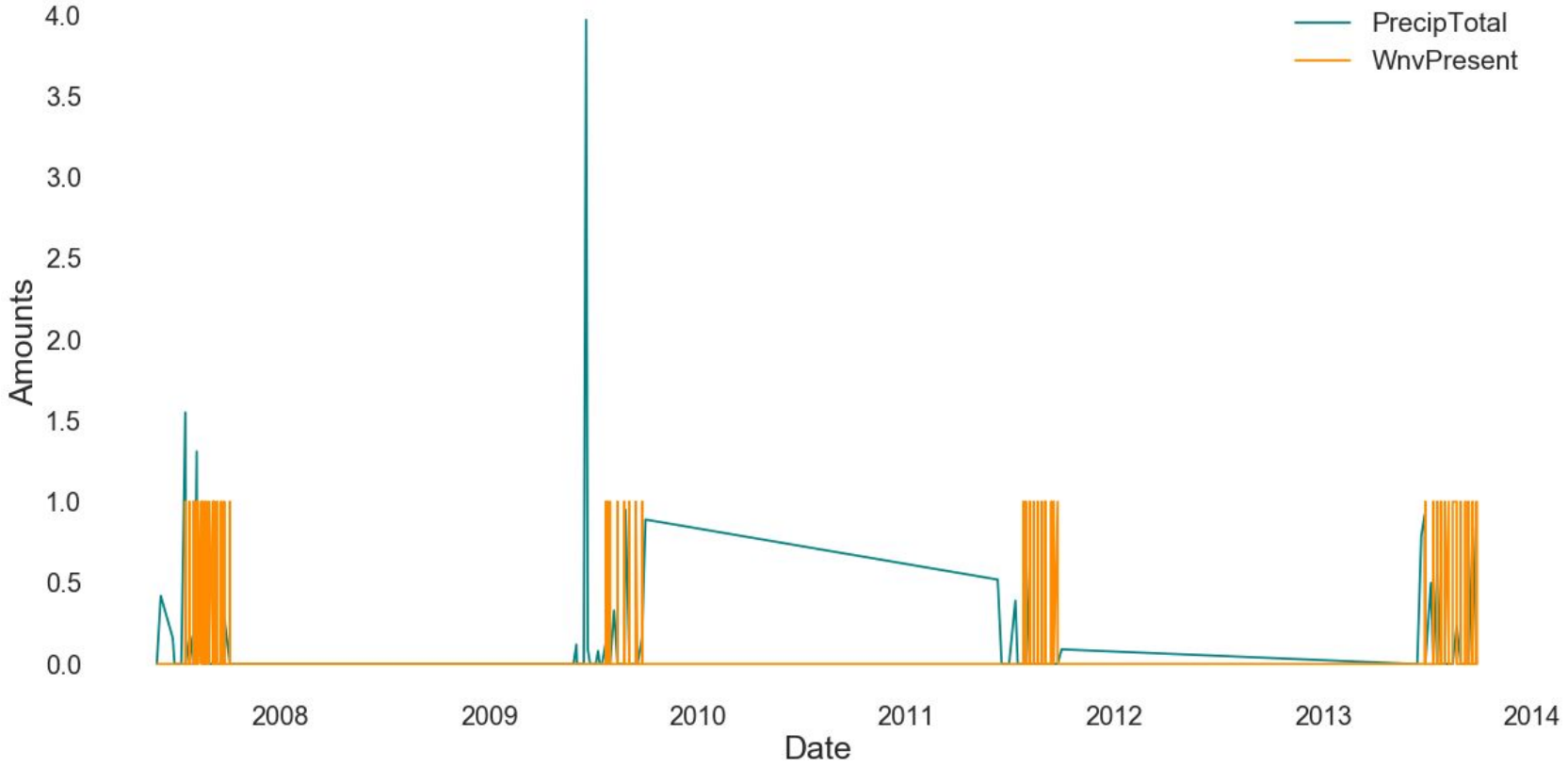

Presence of West Nile per Species
250
200
150
100
50
0
CULEX ERRATICUS
CULEX PIPIENS
CULEX PIPIENS/RESTUANS
CULEX RESTUANS
CULEX SALINARIUS
CULEX TARSALIS
CULEX TERRITANS
Species


NumMosquitos per Species
60000
50000
40000
30000
20000
10000
0
CULEX ERRATICUS
CULEX PIPIENS
CULEX PIPIENS/RESTUANS
CULEX RESTUANS
CULEX SALINARIUS
CULEX TARSALIS
CULEX TERRITANS
Species

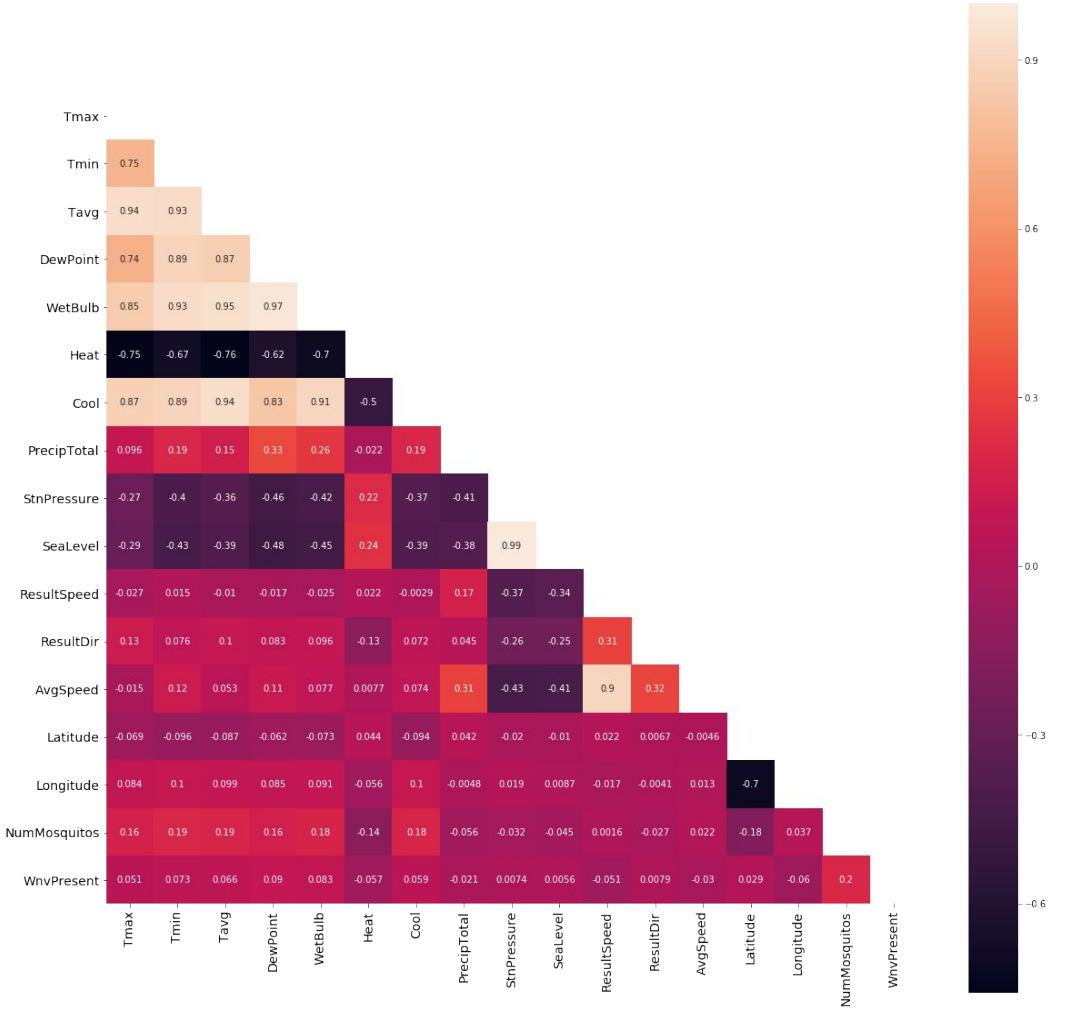

NumMosquitos by Date
7000
6000
5000
4000
3000
2000
1000
0
2007-05-29 00:00:00
2007-06-05 00:00:00
2007-06-26 00:00:00
2007-06-29 00:00:00
2007-07-02 00:00:00
2007-07-11 00:00:00
2007-07-18 00:00:00
2007-07-19 00:00:00
2007-07-25 00:00:00
2007-07-27 00:00:00
2007-08-01 00:00:00
2007-08-02 00:00:00
2007-08-03 00:00:00
2007-08-07 00:00:00
2007-08-08 00:00:00
2007-08-09 00:00:00
2007-08-15 00:00:00
2007-08-16 00:00:00
2007-08-17 00:00:00
2007-08-21 00:00:00
2007-08-22 00:00:00
2007-08-24 00:00:00
2007-08-28 00:00:00
2007-09-04 00:00:00
2007-09-05 00:00:00
2007-09-06 00:00:00
2007-09-12 00:00:00
2007-09-18 00:00:00
2007-09-19 00:00:00
2007-09-24 00:00:00
2007-09-25 00:00:00
2007-10-04 00:00:00
2007-10-09 00:00:00
2009-05-28 00:00:00
2009-06-02 00:00:00
2009-06-03 00:00:00
2009-06-05 00:00:00
2009-06-12 00:00:00
2009-06-15 00:00:00
2009-06-19 00:00:00
2009-06-22 00:00:00
2009-06-26 00:00:00
2009-06-29 00:00:00
2009-07-06 00:00:00
2009-07-10 00:00:00
2009-07-13 00:00:00
2009-07-17 00:00:00
2009-07-24 00:00:00
2009-07-27 00:00:00
2009-07-31 00:00:00
2009-08-07 00:00:00
2009-08-13 00:00:00
2009-08-25 00:00:00
2009-08-27 00:00:00
2009-09-03 00:00:00
2009-09-14 00:00:00
2009-09-17 00:00:00
2009-09-25 00:00:00
2009-10-01 00:00:00
2011-06-10 00:00:00
2011-06-17 00:00:00
2011-06-24 00:00:00
2011-06-30 00:00:00
2011-07-11 00:00:00
2011-07-15 00:00:00
2011-07-25 00:00:00
2011-07-29 00:00:00
2011-08-05 00:00:00
2011-08-12 00:00:00
2011-08-19 00:00:00
2011-08-26 00:00:00
2011-09-01 00:00:00
2011-09-02 00:00:00
2011-09-12 00:00:00
2011-09-16 00:00:00
2011-09-23 00:00:00
2011-09-30 00:00:00
2013-06-07 00:00:00
2013-06-14 00:00:00
2013-06-21 00:00:00
2013-06-27 00:00:00
2013-06-28 00:00:00
2013-07-08 00:00:00
2013-07-12 00:00:00
2013-07-19 00:00:00
2013-07-25 00:00:00
2013-08-01 00:00:00
2013-08-08 00:00:00
2013-08-15 00:00:00
2013-08-22 00:00:00
2013-08-29 00:00:00
2013-09-06 00:00:00
2013-09-12 00:00:00
2013-09-19 00:00:00
2013-09-26 00:00:00
Date


Presence of West Nile by Date
0
5
10
15
20
25
30
35
2007-05-29 00:00:00
2007-06-05 00:00:00
2007-06-26 00:00:00
2007-06-29 00:00:00
2007-07-02 00:00:00
2007-07-11 00:00:00
2007-07-18 00:00:00
2007-07-19 00:00:00
2007-07-25 00:00:00
2007-07-27 00:00:00
2007-08-01 00:00:00
2007-08-02 00:00:00
2007-08-03 00:00:00
2007-08-07 00:00:00
2007-08-08 00:00:00
2007-08-09 00:00:00
2007-08-15 00:00:00
2007-08-16 00:00:00
2007-08-17 00:00:00
2007-08-21 00:00:00
2007-08-22 00:00:00
2007-08-24 00:00:00
2007-08-28 00:00:00
2007-09-04 00:00:00
2007-09-05 00:00:00
2007-09-06 00:00:00
2007-09-12 00:00:00
2007-09-18 00:00:00
2007-09-19 00:00:00
2007-09-24 00:00:00
2007-09-25 00:00:00
2007-10-04 00:00:00
2007-10-09 00:00:00
2009-05-28 00:00:00
2009-06-02 00:00:00
2009-06-03 00:00:00
2009-06-05 00:00:00
2009-06-12 00:00:00
2009-06-15 00:00:00
2009-06-19 00:00:00
2009-06-22 00:00:00
2009-06-26 00:00:00
2009-06-29 00:00:00
2009-07-06 00:00:00
2009-07-10 00:00:00
2009-07-13 00:00:00
2009-07-17 00:00:00
2009-07-24 00:00:00
2009-07-27 00:00:00
2009-07-31 00:00:00
2009-08-07 00:00:00
2009-08-13 00:00:00
2009-08-25 00:00:00
2009-08-27 00:00:00
2009-09-03 00:00:00
2009-09-14 00:00:00
2009-09-17 00:00:00
2009-09-25 00:00:00
2009-10-01 00:00:00
2011-06-10 00:00:00
2011-06-17 00:00:00
2011-06-24 00:00:00
2011-06-30 00:00:00
2011-07-11 00:00:00
2011-07-15 00:00:00
2011-07-25 00:00:00
2011-07-29 00:00:00
2011-08-05 00:00:00
2011-08-12 00:00:00
2011-08-19 00:00:00
2011-08-26 00:00:00
2011-09-01 00:00:00
2011-09-02 00:00:00
2011-09-12 00:00:00
2011-09-16 00:00:00
2011-09-23 00:00:00
2011-09-30 00:00:00
2013-06-07 00:00:00
2013-06-14 00:00:00
2013-06-21 00:00:00
2013-06-27 00:00:00
2013-06-28 00:00:00
2013-07-08 00:00:00
2013-07-12 00:00:00
2013-07-19 00:00:00
2013-07-25 00:00:00
2013-08-01 00:00:00
2013-08-08 00:00:00
2013-08-15 00:00:00
2013-08-22 00:00:00
2013-08-29 00:00:00
2013-09-06 00:00:00
2013-09-12 00:00:00
2013-09-19 00:00:00
2013-09-26 00:00:00
Date

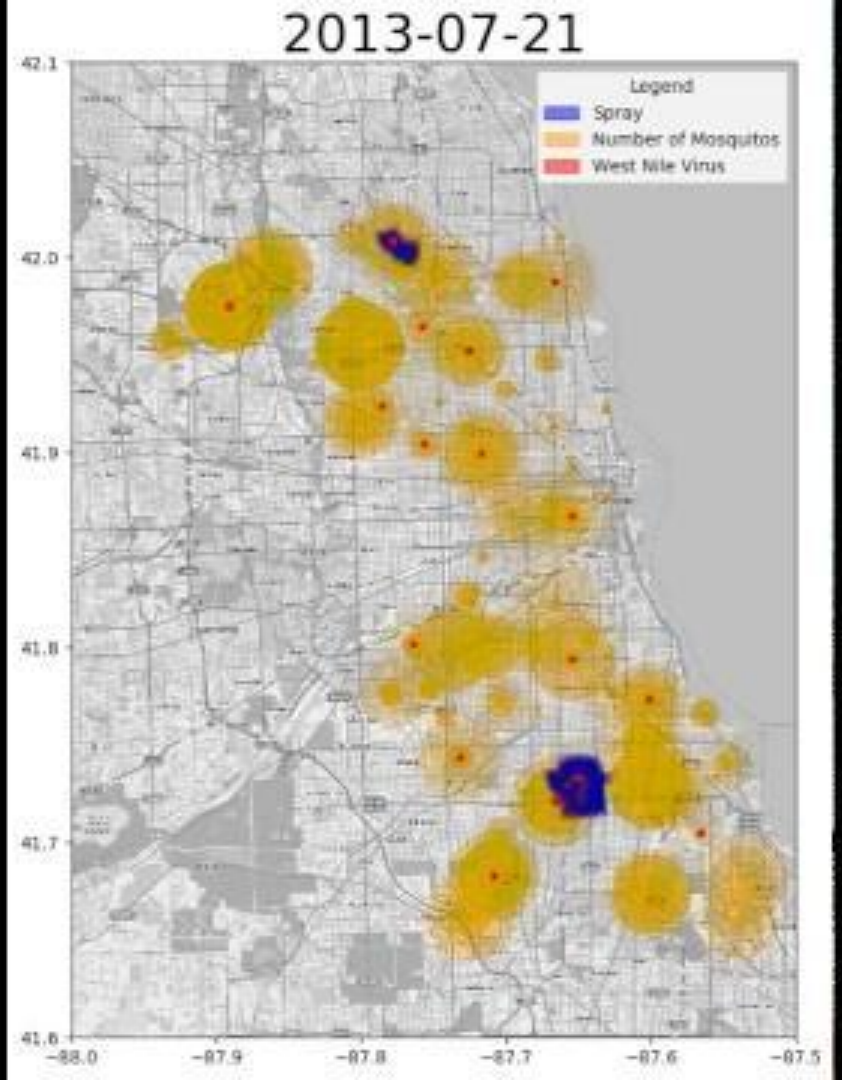
How does rain correlate to West Nile?



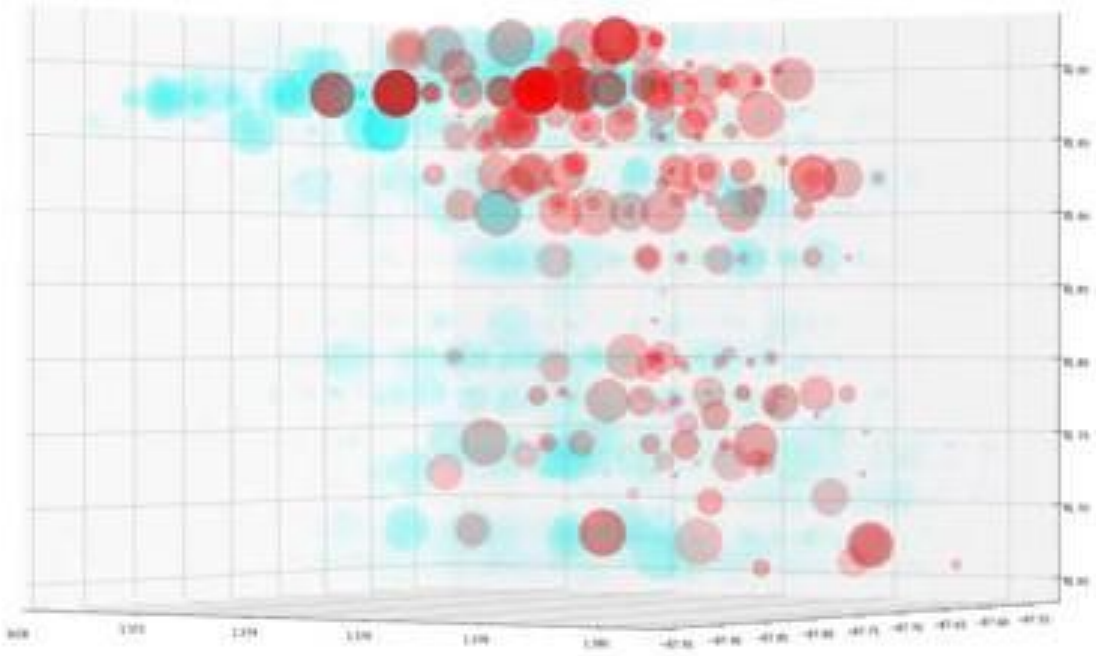
How do different parts of the weather data relate?



Simulating Mosquito Population Growth in Chicago


2013-07-21
Legend
Spray
Number of Mosquitos
West Nile Virus
42.1
42.0
41.9
41.8
41.7
41.6
−88.0
−87.9
−87.8
−87.7
−87.6
−87.5

2013



EDA Takeaways

1. Only 3 observed species carry West Nile.

2. July-September have the highest rate of the virus.

3. Spraying has little to no effect on the number of mosquitoes.

Our Process

1. Question ✔
2. Data ✔
3. Explore ✔
4. Model
5. Communicate

Modelling Time - Round 1 & 2

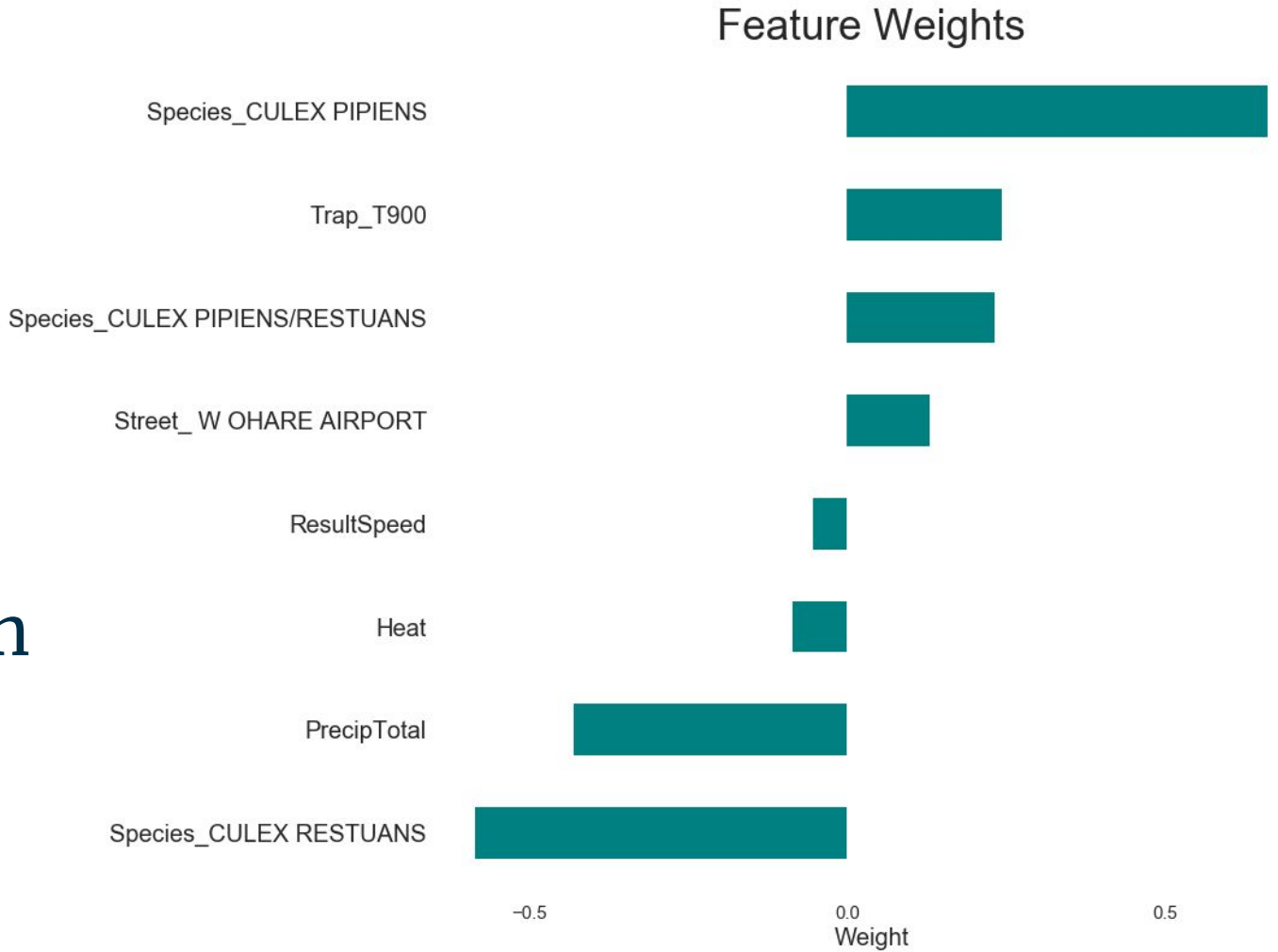
- No Spray Data
- Test set doesn't have Num Mosquitoes
- Round 1: Let the computer weight features

Extra Trees ROC AUC Score: 53%

Logistic Regression ROC AUC Score: 65%

Round 2:

Logistic Regression



Modelling Time - Round 3

- Group Areas Together
- Group by Dates, Months
- Label Encode Species, Streets, Traps

Random Forest ROC AUC Score: 71.9%

Heavy Weights: Labelled, Block, Area

Modelling Time - Round 4

- Maximize Weather: *It's Cyclical!*
  - Let the **MEANS** *ROLL*
- Binary Species!
- Feature Extraction [PCA]

… And?

Modelling Time - Round 4

Random Forest ROC AUC Score: 72%

Rolling Means:

- ResultSpeed: 21 (3 wks)
- PrecipTotal: 15 (2 wks)
- DewPoint: 16 (2 wks)
- AvgSpeed: 19 (3 wks)
- Heat: 28 (4 wks)
- Tmax: 4 & Tmin: 8

Modelling Time - Round 5

- **Bootstrap?**
  - *Didn't Help*
- **Predict the Number of Mosquitoes?**
  - Didn't Help

Our Process

1. Question ✔
2. Data ✔
3. Explore ✔
4. Model ✔
5. Communicate

Recommendations

We need to rethink our approach to spraying:

- Begin spraying earlier in the season
- Focus on O’Hare area
- Impact of Birds [Crows & Sparrows]
  - If they’re dead, collect and test them

Recommendations

Public Awareness Campaigns in Early Spring

- Report dead birds
- **Know the dangers**
- **NO STANDING WATER**

Bioengineering

Imperial College of London research:

- Complete eradication of a subspecies within 8 generations

$100m budget

Production model in 5-10 years

Alternative Solution

- Time is running out
- Resources are depleting
- West Nile is **NOT** being prevented

Look at previous attempts...? **Malaria**

- $100M dealt with mosquitoes before
- Chicago is smaller scale

EASIER. FASTER. SAFER.