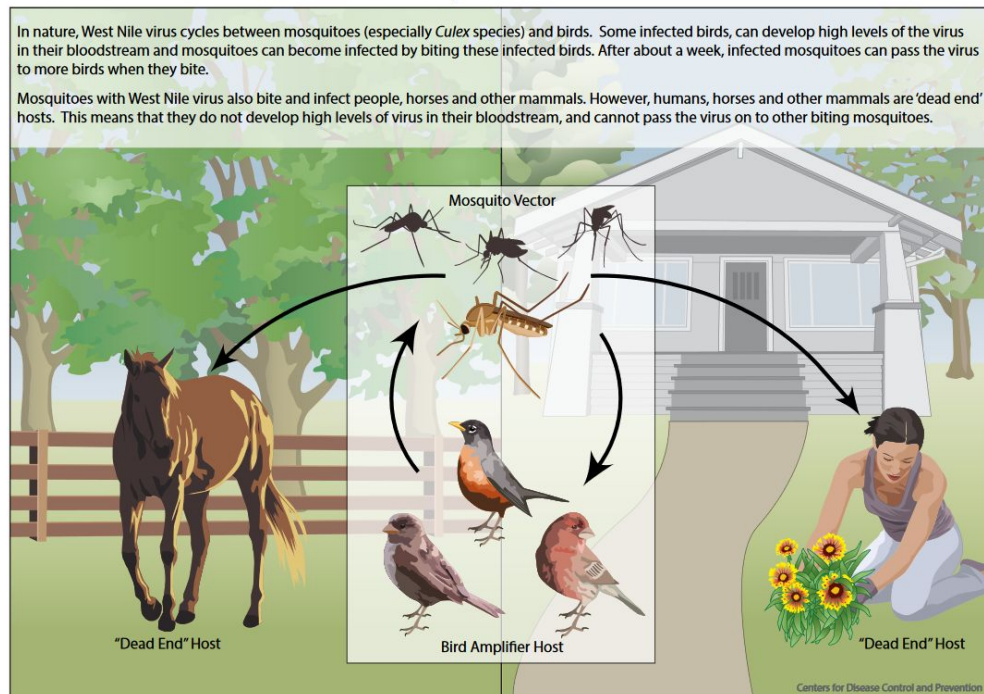# Keeping our city safe from West Nile virus

Mosquito Monitoring Team

West Nile virus (WNV) - the facts

- Most commonly spread to people by the bite of an infected mosquito.

- Cases of WNV occur during mosquito season - July and August.

- There are no vaccines to prevent or medications to treat WNV in people.

- Fortunately, most people infected with WNV do not feel sick.
    - About 1 in 5 people who are infected develop a fever and other symptoms.
    - About 1 out of 150 infected people develop a serious, sometimes fatal, illness.

West Nile Virus Transmission Cycle

In nature, West Nile virus cycles between mosquitoes (especially *Culex* species) and birds. Some infected birds, can develop high levels of the virus in their bloodstream and mosquitoes can become infected by biting these infected birds. After about a week, infected mosquitoes can pass the virus to more birds when they bite.

Mosquitoes with West Nile virus also bite and infect people, horses and other mammals. However, humans, horses and other mammals are 'dead end' hosts. This means that they do not develop high levels of virus in their bloodstream, and cannot pass the virus on to other biting mosquitoes.

Mosquito Vector

"Dead End" Host

Bird Amplifier Host

"Dead End" Host

Centers for Disease Control and Prevention

Problem statement

To develop a predictive model that will forecast the probability of WNV presence in 138 mosquito traps around Chicago over the course of a season. WNV is a communicable disease that is spread through its most common vector, mosquitoes

Information provided:

- Mosquito trap data: location of traps, how many mosquitos on which date, and whether they have the West Nile Virus
- Weather data: 2 Chicago weather station reports with many indicators (temperature, dew point, sunrise/sunset)
- Spray data: gis data for insecticide spraying by the city.

EDA

Mapped weather data across to training set

Final list of features selected -
- Latitude
- Longitude
- Tavg
- DewPoint
- Days since rain
- Tavg_rolling3day, Tavg_rolling7day, Tavg_rolling14day
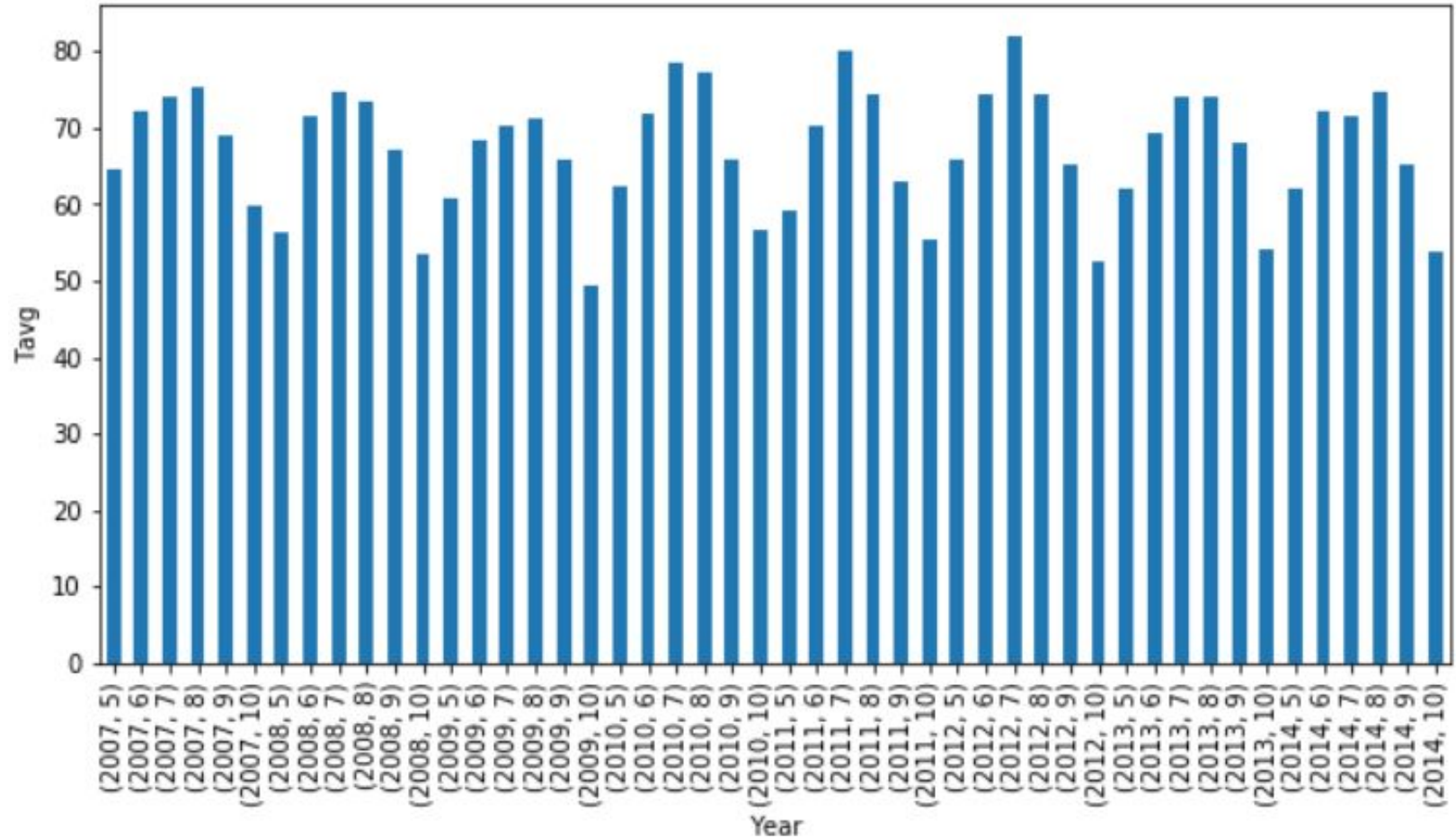- DewPoint_rolling3day, DewPoint_rolling7day, DewPoint_rolling14day
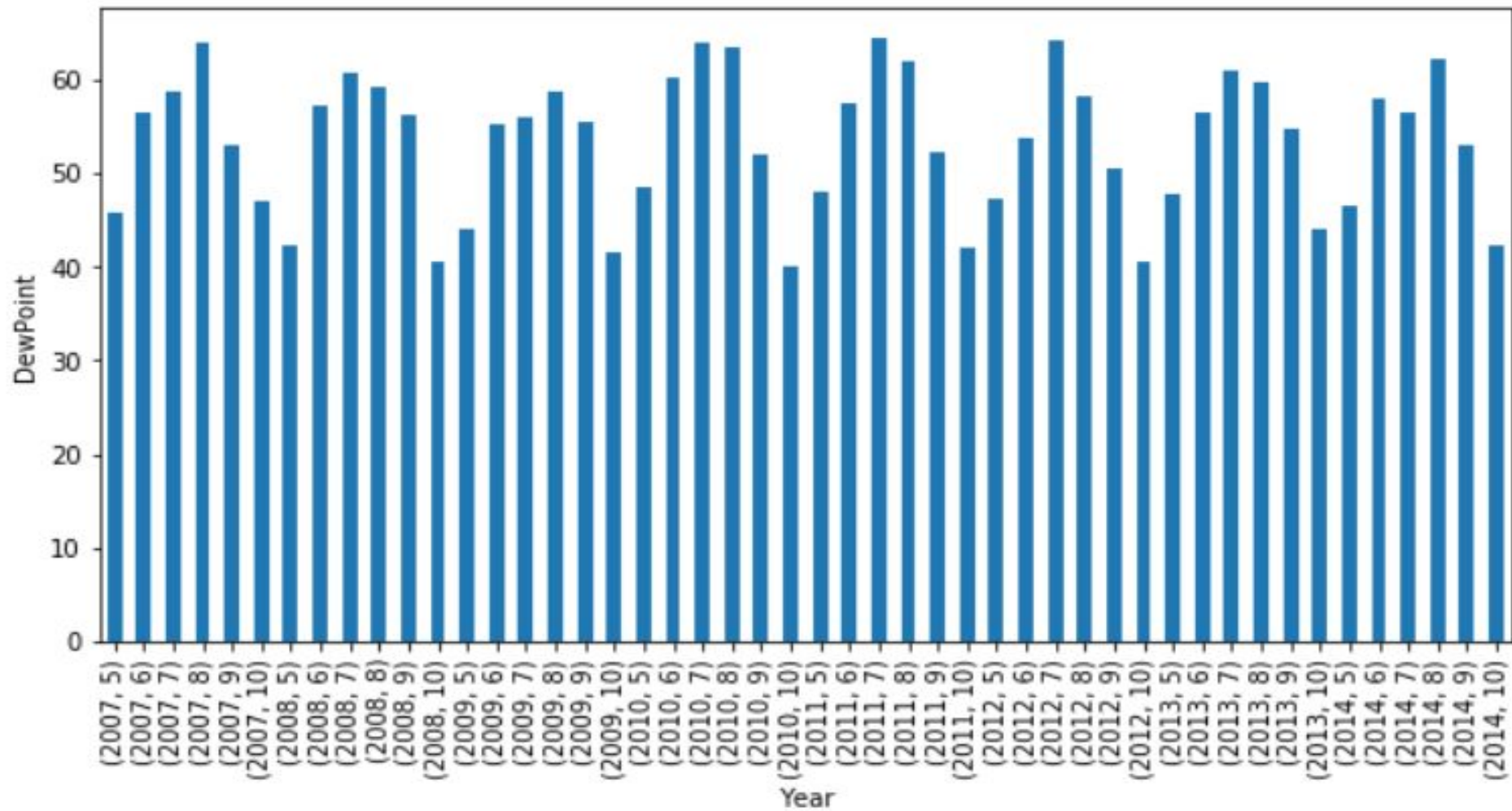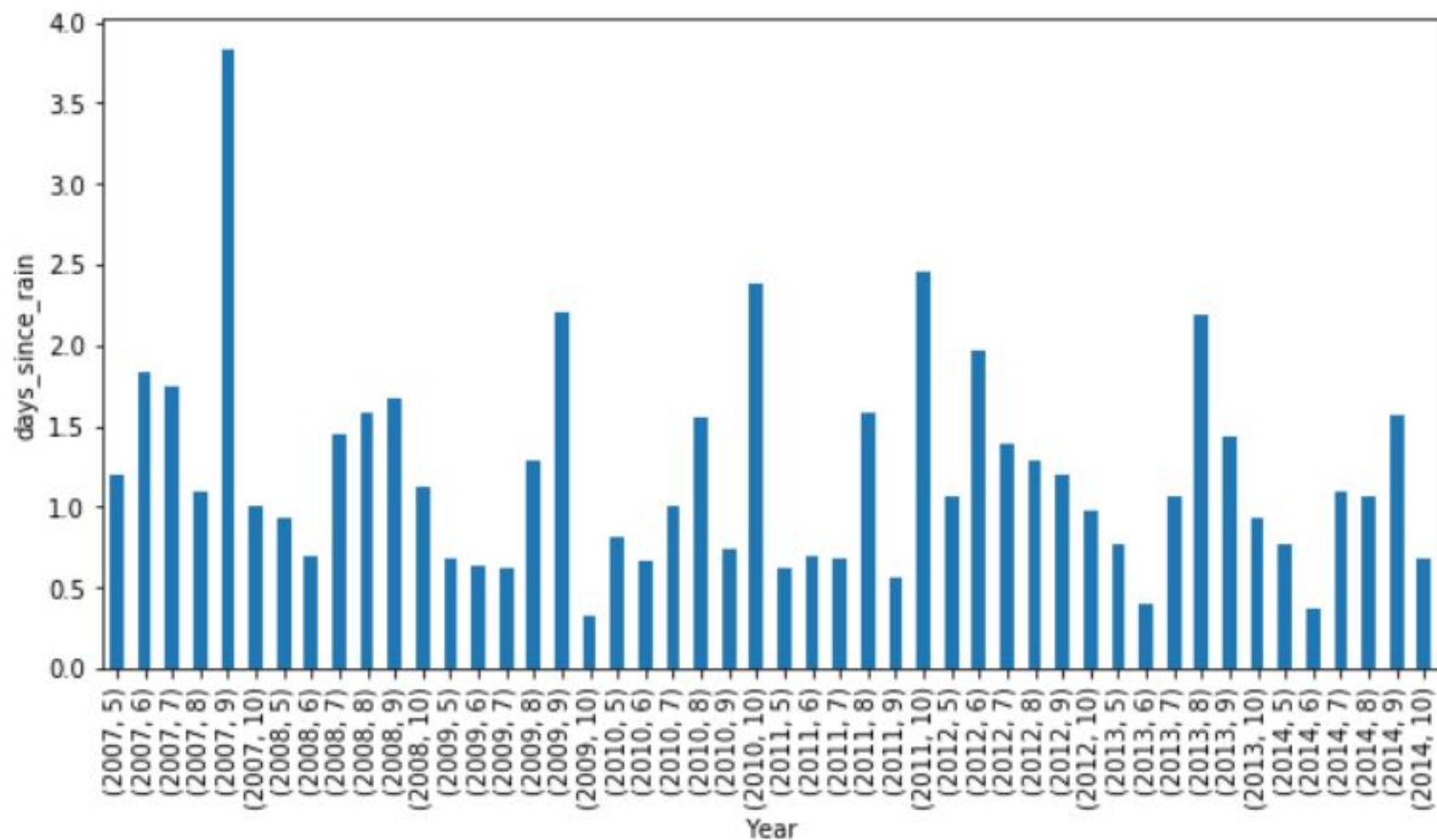
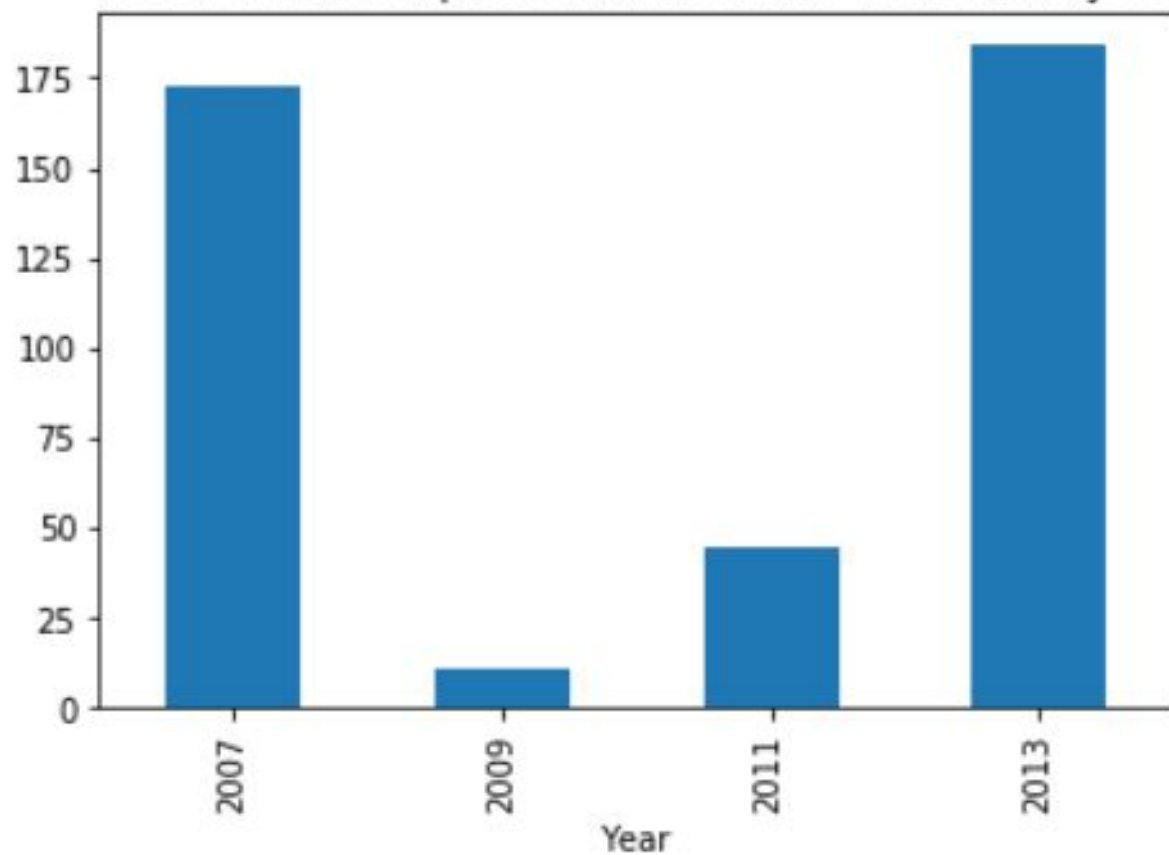# Visualising training data

No. of samples w/ WNV in a given year
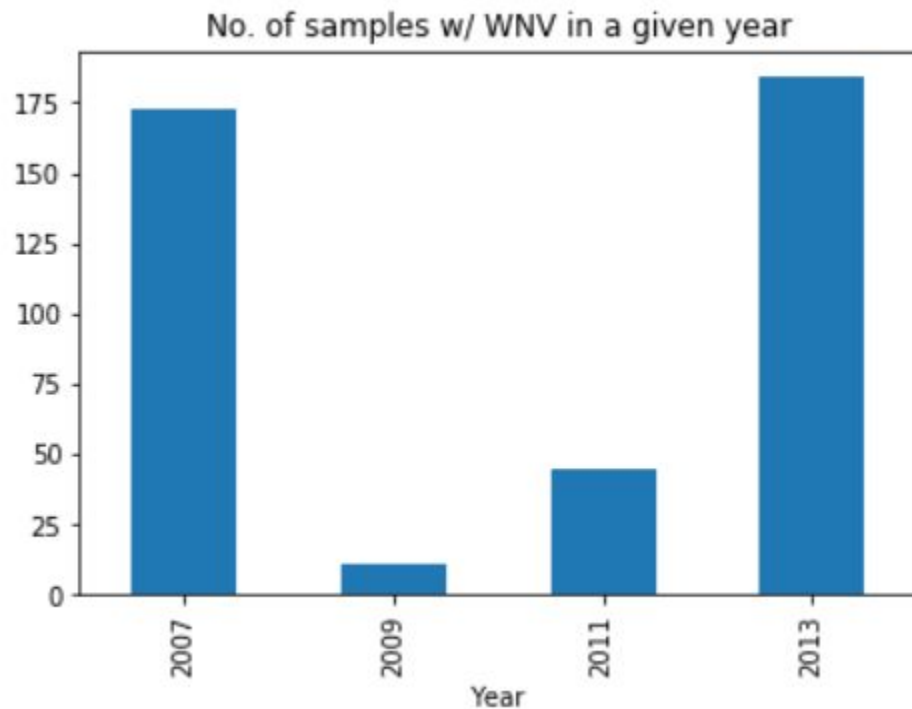
No. of samples w/ WNV in a given (year, month)
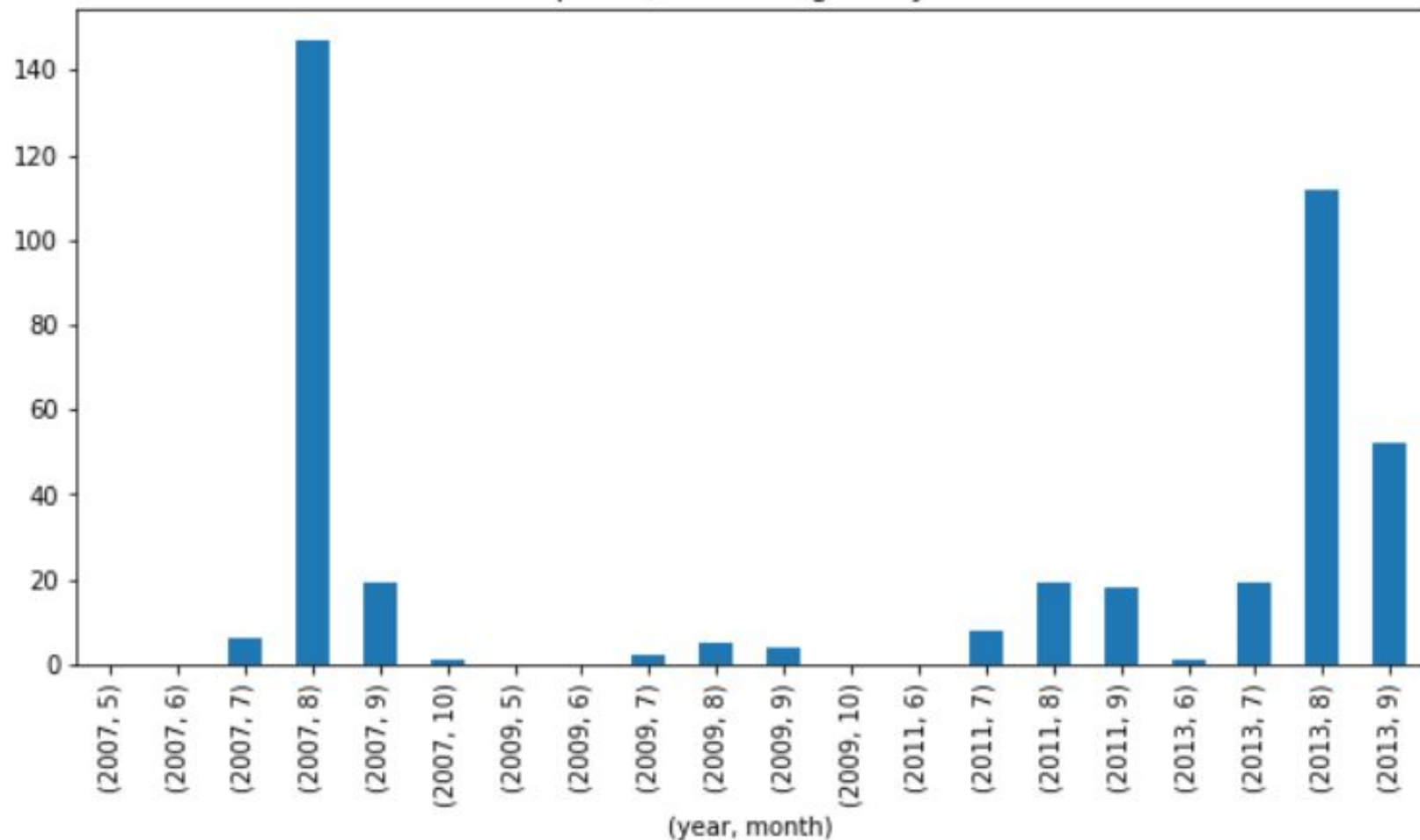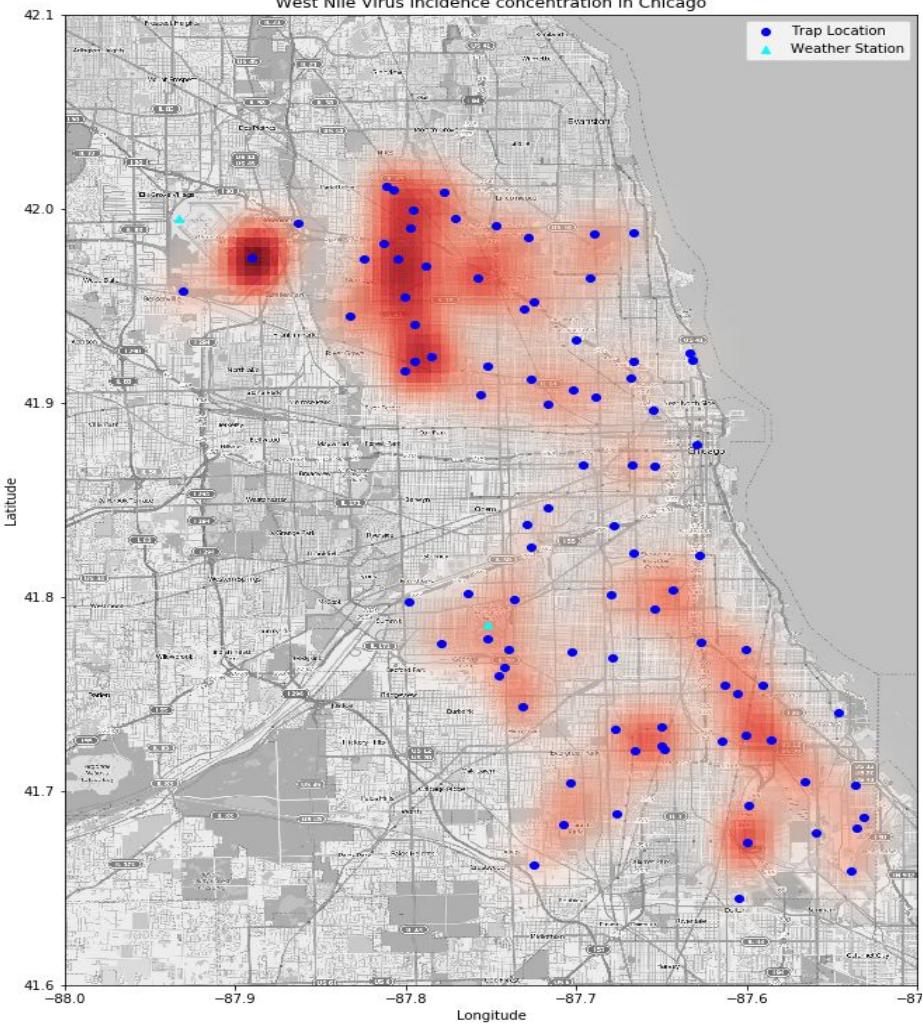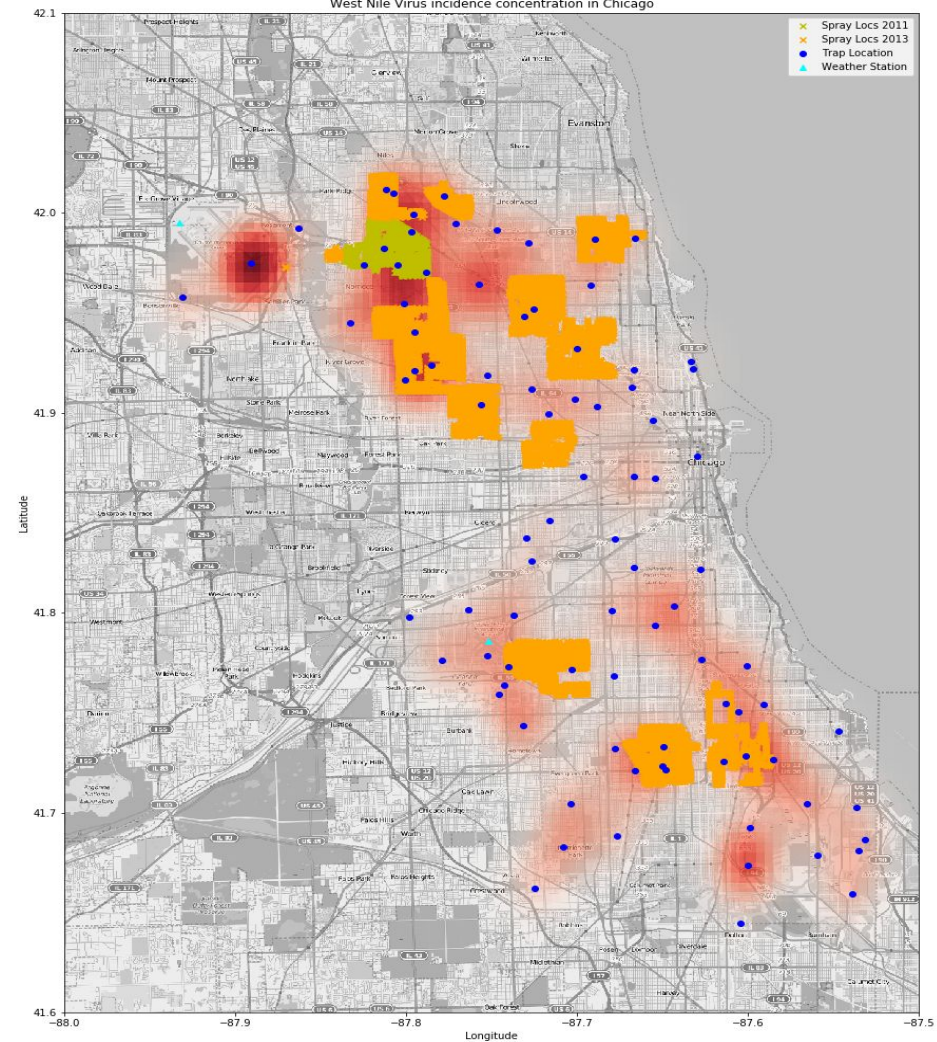
Visualising weather data
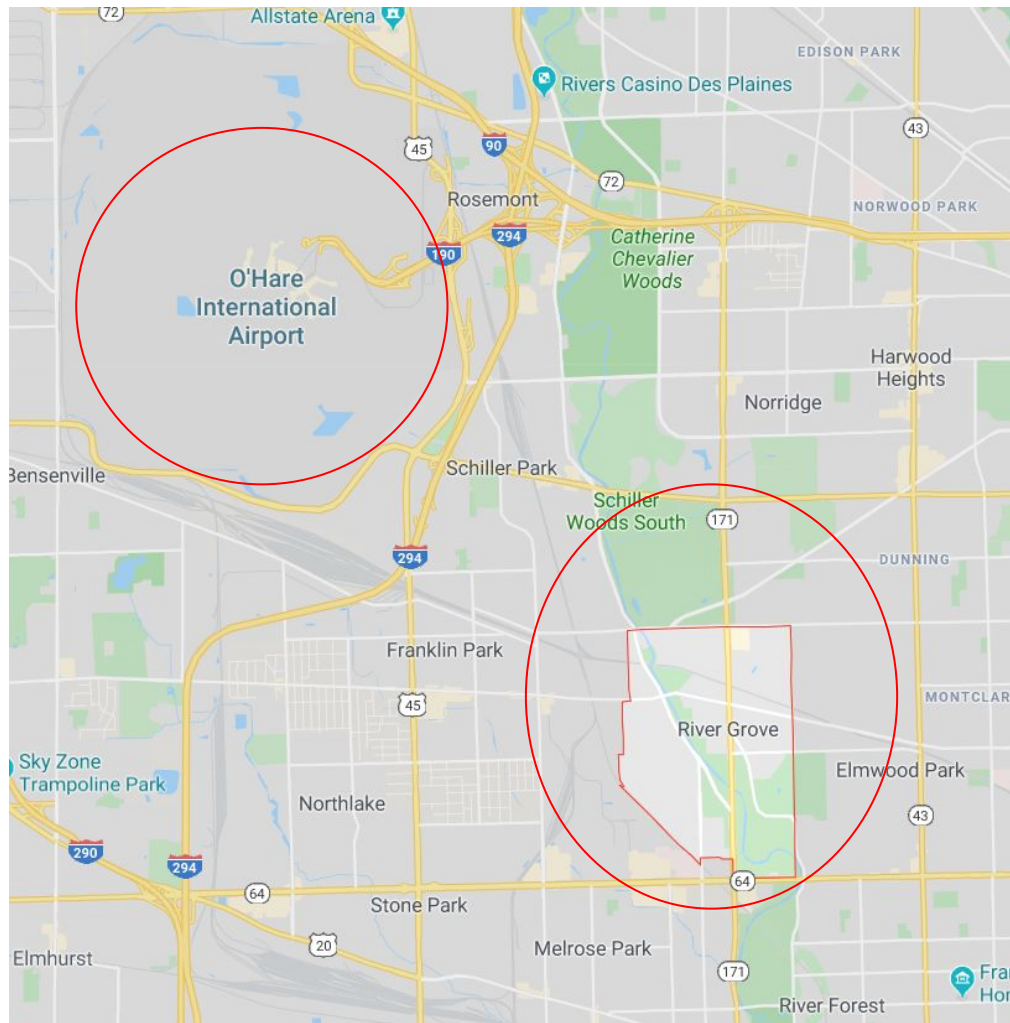
Plot of no. of samples found with WNV in a certain year

West Nile Virus incidence concentration in Chicago

Cost-Benefit Analysis

- Spraying not really worth it
- USD ~500 to spray 2000 m^2
- O'hare Airport would cost USD ~6,525,000
- Killing adult mosquitoes is not really effective in reducing WNV incidences
  - Instead, larvae should be targetted

Modelling

Model	AUC value for ROC curve	Kaggle private score	Kaggle public score
Logistic Regression	0.724	0.6095	0.6360
Decision Tree	0.788	0.6212	0.6544
Random Forest	0.770	0.6150	0.6432
ADA Boost	0.797	0.6586	0.6776
Gradient Boost	0.801	0.6540	0.6809
XG Boost	0.807	0.6737	0.7045

Selected model - XG Boost Classifier

```
############################
XG Boost Classifier model
############################

Best params: {'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 100}
Train score: 0.7681766273315569, Test score 0.73734722634665
```

	actual WnV Not Present	actual WnV Present
Predict WnV Not Present	1921	568
Predict WnV Present	41	97

```
** Accuracy %: 76.82% **
Misclassification %: 23.18&
Precision %: 0.1459
Recall: 0.7029
Specificity: 0.7718
```


ROC Curve with AUC = 0.807
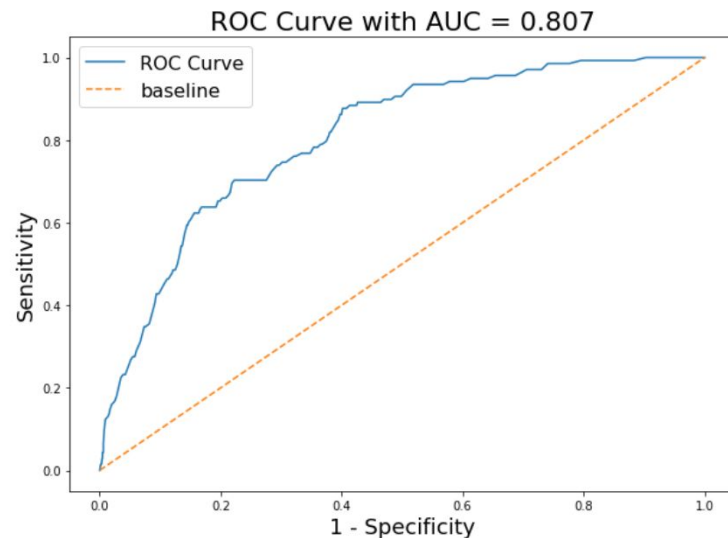
Conclusion

- We would recommend against spraying on a large scale
  - Maybe areas with really high density of mosquitoes should be sprayed though
  - Instead, educate population on how to prevent mosquito breeding
- O'hare Airport and River Grove areas should be focused on
- Maybe usage of weather forecasts might be useful

Limitations

- No bird data (birds are WNV carriers as well)

Further work

- Try out neural network (possible increase in performance)