**Red Team Planning Document**

Alex Zadel, Jason Mallet, and Teddy Pender

**Problem Statement:**

*The city of Chicago has battled with West Nile Virus since the human infection was first identified in 2002. As part of the public health effort to track the disease, traps have been placed throughout the city and checked regularly to track instances of mosquitos carrying the disease.*

We have been tasked with using a subset of this data to predict the outcomes of any particular trap on a given day.

Furthermore, we have been asked to prepare an effective plan for deploying pesticides throughout the city to neutralize the threat of infected mosquitos while remaining cost conscious and efficient.

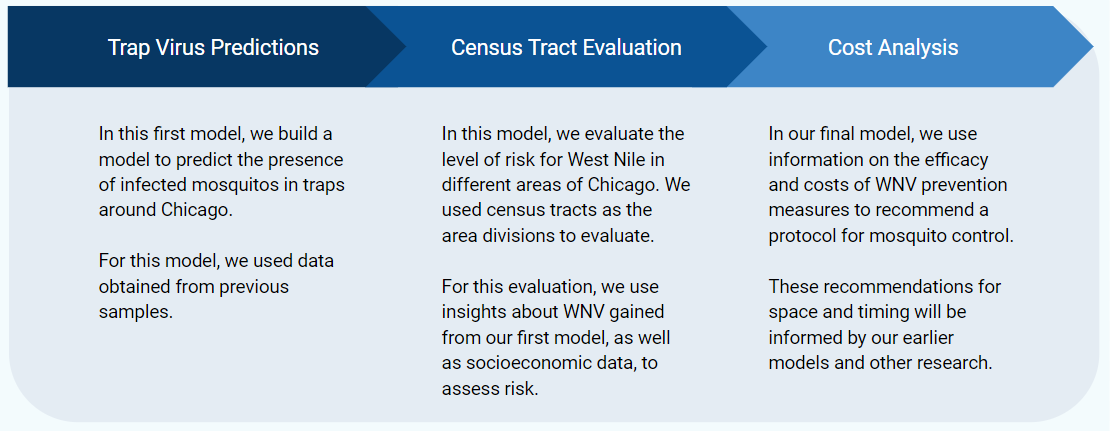
**Relevant Background for Analysis:**

* Only certain species of mosquitos carry the disease but there are more than those in our database. We can sort on that since the non-carriers won’t test positive.
* Different species have different life cycles and therefore we would expect them to have different patterns of blooming over time given the same weather conditions.
* That said, ideal weather conditions are rain—which creates standing water locations preferred for larva—followed by warm dry weather. If it rains too soon after, the larva can be washed out of their space before they fully develop.
* We should look for those weather and species patterns in the data.\
* West Nile is more likely to be contracted by vulnerable populations, who are themselves most at-risk of serious complications.

**Central Questions**

* What are the costs and benefits of preventative measures such as spraying?
* How can we target preventative measures to make our policy most cost efficient?
* What are the costs of treating WNV? Our goal is complete prevention of human cases, but how much is too much to spend?

**Analysis Process**



**Phase by Phase Goals:**

1. Create a model that will predict whether a certain trap will test positive for WNV given other conditions such as recent weather patterns.
2. Translate those model results to information about the risk of WNV in different areas of Chicago. Use other information about those areas to evaluate their risk profile or other reasons for priority.
3. Use risk information found in Part 2 to create a cost estimate for prevention spraying in targeted regions.

**First Steps**

Jason: Data search, exploration, and cleaning of weather data.

Alex: Obtain census data, GIS shapefiles, geographic exploration

Teddy: WNV research, early EDA

All: General EDA, research on WNV