West Nile Virus and Climate Change

Emma Hardin-Parker & Makenzie Hicks

Methodology

Library:

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v lubridate 1.9.3     v tibble    3.2.1
v purrr     1.0.2     v tidyr     1.3.1
-- Conflicts ---------------------------------------- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becor
Loading required package: carData


Attaching package: 'car'


The following object is masked from 'package:dplyr':

    recode


The following object is masked from 'package:purrr':

    some
```

Historic Data

Data Introduction:

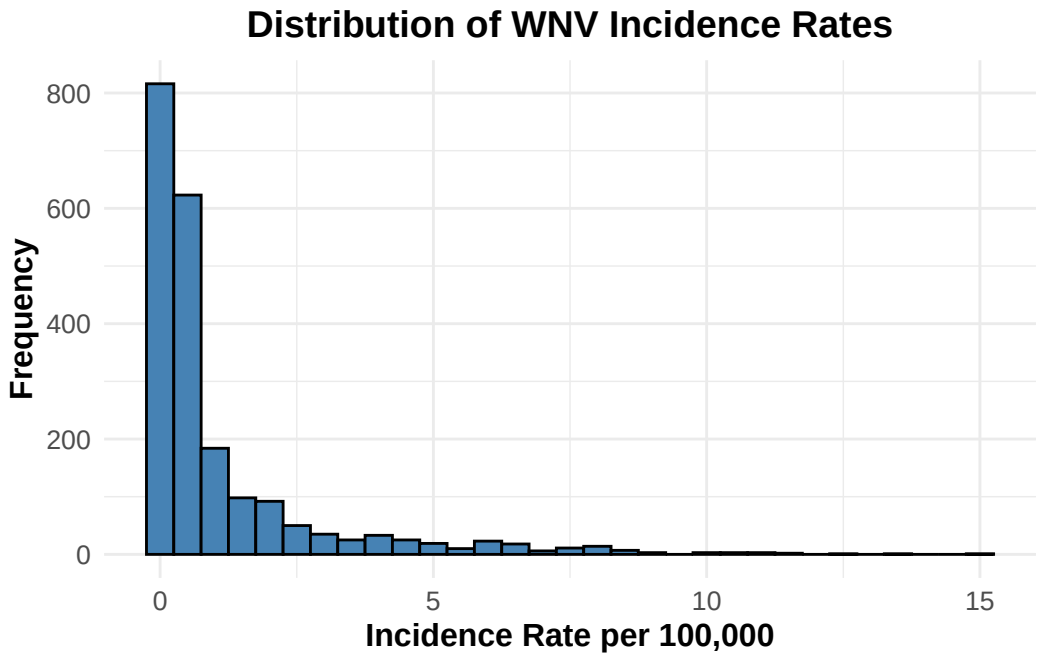
These data were accessed from the CDC Historic Data Page and were downloaded via .csv files. This data shows historic data of WNV incidence by county from 1993-2023. The goal with this data in particular are to identify both high-risk counties and high-risk states to move forward with our climate change indicator analysis. To get started, we calculated some descriptive statistics to become more comfortable with the data.

Basic Data Structure and Descriptive Statistics:

Descriptive Statistics for WNV Incidence and Population
Summary statistics and standard deviations

Statistic	Incidence	Population
Min	0.01	269.00
1st Qu.	0.15	13,871.75
Median	0.37	35,542.00
Mean	1.11	134,761.56
3rd Qu.	1.11	104,286.00
Max	14.87	9,931,394.00
Std. Dev.	1.83	383,630.78

Distribution of Incidence Rates:



Now that we have a general idea of the distribution of incidence rates from 1993-2023, we are going to move forward with both State and County summaries with accompanying figures.

State and County Distribution Summaries:

Average WNV Incidence and Population by State

Summary of West Nile Virus incidence rates across U.S. states

State	Mean Incidence Rate	Total Population
AK	0.01	298,806.00
AL	0.24	4,260,974.00
AR	0.58	2,596,134.00
AZ	0.74	6,556,344.00
CA	0.62	37,824,067.00
CO	2.49	5,082,308.00
CT	0.10	3,595,211.00
DC	0.48	635,737.00
DE	0.16	915,518.00
FL	0.20	17,829,381.00
GA	0.32	7,654,144.00
IA	0.81	2,864,440.00
ID	1.63	1,382,020.00
IL	0.43	12,608,424.00
IN	0.29	5,866,905.00
KS	1.91	2,570,120.00
KY	0.28	2,836,202.00

LA	0.98	4,542,745.00
MA	0.09	6,637,118.00
MD	0.17	5,711,775.00
ME	0.06	283,742.00
MI	0.28	9,203,646.00
MN	0.95	5,211,227.00
MO	0.42	5,181,446.00
MS	1.06	2,749,604.00
MT	2.92	899,488.00
NC	0.08	7,004,032.00
ND	4.35	687,041.00
NE	4.04	1,835,527.00
NH	0.04	987,245.00
NJ	0.13	8,698,668.00
NM	1.13	2,060,350.00
NV	1.14	2,705,322.00
NY	0.14	17,959,443.00
OH	0.27	10,934,731.00
OK	0.89	3,701,368.00
OR	0.71	1,912,453.00
PA	0.19	11,738,236.00
PR	0.01	379,547.00
RI	0.07	890,249.00
SC	0.15	4,065,910.00
SD	5.29	826,733.00
TN	0.20	5,060,253.00
TX	1.37	25,534,378.00
UT	0.50	2,658,954.00
VA	0.17	6,047,904.00
VT	0.20	343,780.00
WA	0.13	5,852,212.00
WI	0.24	4,896,381.00
WV	0.22	726,415.00
WY	2.27	503,178.00

Identifying High-Risk States:

High Risk States: WNV Incidence

Summary of High Risk States (Incidence > 1)

State	Mean Incidence	Population
SD	5.2919	826,733.0000
ND	4.3494	687,041.0000
NE	4.0398	1,835,527.0000
MT	2.9159	899,488.0000
CO	2.4929	5,082,308.0000
WY	2.2730	503,178.0000
KS	1.9127	2,570,120.0000
ID	1.6266	1,382,020.0000
TX	1.3745	25,534,378.0000
NV	1.1411	2,705,322.0000

NM	1.1345	2,060,350.0000
MS	1.0645	2,749,604.0000

After viewing the overall state summaries, we determined that our threshold for *high-risk* states were to be an Average Incidence greater than 1. After filtering out these *high-risk* states, we are left with 12 states to move forward with analysis. Additionally, we intend to conduct both state level and county level analyses; therefore we have elected to go through the *high-risk* process again with county-level data, with a threshold of Incidence > 5.

Identifying High-Risk Counties:

High Risk Counties of WNV Incidence

Summary of High Risk Counties (Incidence > 5)

County	**Type**	**Year**	**Population**	**Incidence**
TX, King	Neuroinvasive disease cases	1999-2023	269.0000	14.8700
SD, Dewey	Neuroinvasive disease cases	1999-2023	5,532.0000	13.7400
TX, Glasscock	Neuroinvasive disease cases	1999-2023	1,280.0000	12.5000
NE, Banner	Neuroinvasive disease cases	1999-2023	688.0000	11.6300
NE, Thomas	Neuroinvasive disease cases	1999-2023	700.0000	11.4300
MT, Treasure	Neuroinvasive disease cases	1999-2023	717.0000	11.1600
NE, Hooker	Neuroinvasive disease cases	1999-2023	718.0000	11.1400
ND, Mclean	Neuroinvasive disease cases	1999-2023	9,388.0000	11.0800
MT, Prairie	Neuroinvasive disease cases	1999-2023	1,141.0000	10.5200
SD, Potter	Neuroinvasive disease cases	1999-2023	2,340.0000	10.2600
WY, Goshen	Neuroinvasive disease cases	1999-2023	13,647.0000	10.2600
NE, Logan	Neuroinvasive disease cases	1999-2023	785.0000	10.1900
SD, Faulk	Neuroinvasive disease cases	1999-2023	2,355.0000	10.1900
SD, Buffalo	Neuroinvasive disease cases	1999-2023	2,016.0000	9.9200
SD, Harding	Neuroinvasive disease cases	1999-2023	1,304.0000	9.2000
MT, Sheridan	Neuroinvasive disease cases	1999-2023	3,549.0000	9.0200
SD, Walworth	Neuroinvasive disease cases	1999-2023	5,449.0000	8.8100
ND, Mcintosh	Neuroinvasive disease cases	1999-2023	2,750.0000	8.7300
ND, Billings	Neuroinvasive disease cases	1999-2023	921.0000	8.6900
SD, Sanborn	Neuroinvasive disease cases	1999-2023	2,318.0000	8.6300
CO, Cheyenne	Neuroinvasive disease cases	1999-2023	1,871.0000	8.5500
MT, Valley	Neuroinvasive disease cases	1999-2023	7,506.0000	8.5300
TX, Roberts	Neuroinvasive disease cases	1999-2023	938.0000	8.5300
SD, Sully	Neuroinvasive disease cases	1999-2023	1,429.0000	8.4000
NE, Harlan	Neuroinvasive disease cases	1999-2023	3,411.0000	8.2100
NE, Perkins	Neuroinvasive disease cases	1999-2023	2,929.0000	8.1900
TX, Armstrong	Neuroinvasive disease cases	1999-2023	1,953.0000	8.1900
SD, Douglas	Neuroinvasive disease cases	1999-2023	2,944.0000	8.1500
ND, Emmons	Neuroinvasive disease cases	1999-2023	3,471.0000	8.0700
SD, Jones	Neuroinvasive disease cases	1999-2023	991.0000	8.0700
NE, Garden	Neuroinvasive disease cases	1999-2023	1,984.0000	8.0600
KS, Gray	Neuroinvasive disease cases	1999-2023	5,963.0000	8.0500
KS, Wallace	Neuroinvasive disease cases	1999-2023	1,523.0000	7.8800
NE, Brown	Neuroinvasive disease cases	1999-2023	3,054.0000	7.8600

NE, Mcpherson	Neuroinvasive disease cases	1999-2023	509.0000	7.8600
SD, Mellette	Neuroinvasive disease cases	1999-2023	2,052.0000	7.8000
SD, Clark	Neuroinvasive disease cases	1999-2023	3,596.0000	7.7900
SD, Jerauld	Neuroinvasive disease cases	1999-2023	2,056.0000	7.7800
CO, Delta	Neuroinvasive disease cases	1999-2023	30,473.0000	7.7400
SD, Hutchinson	Neuroinvasive disease cases	1999-2023	7,257.0000	7.7200
NE, Gosper	Neuroinvasive disease cases	1999-2023	2,078.0000	7.7000
CO, Washington	Neuroinvasive disease cases	1999-2023	4,689.0000	7.6800
ND, Dickey	Neuroinvasive disease cases	1999-2023	5,219.0000	7.6600
NE, Polk	Neuroinvasive disease cases	1999-2023	5,265.0000	7.6000
TX, Floyd	Neuroinvasive disease cases	1999-2023	6,352.0000	7.5600
SD, Lyman	Neuroinvasive disease cases	1999-2023	3,775.0000	7.4200
NE, Kearney	Neuroinvasive disease cases	1999-2023	6,509.0000	7.3700
ND, Burke	Neuroinvasive disease cases	1999-2023	2,184.0000	7.3300
WY, Platte	Neuroinvasive disease cases	1999-2023	8,729.0000	7.3300
ND, Sargent	Neuroinvasive disease cases	1999-2023	3,895.0000	7.1900
ID, Washington	Neuroinvasive disease cases	1999-2023	10,028.0000	7.1800
SD, Hand	Neuroinvasive disease cases	1999-2023	3,377.0000	7.1100
MN, Traverse	Neuroinvasive disease cases	1999-2023	3,405.0000	7.0500
SD, Miner	Neuroinvasive disease cases	1999-2023	2,298.0000	6.9600
ND, Griggs	Neuroinvasive disease cases	1999-2023	2,345.0000	6.8200
TX, Sterling	Neuroinvasive disease cases	1999-2023	1,190.0000	6.7200
TX, Motley	Neuroinvasive disease cases	1999-2023	1,197.0000	6.6800
NE, Boone	Neuroinvasive disease cases	1999-2023	5,428.0000	6.6300
NE, Merrick	Neuroinvasive disease cases	1999-2023	7,848.0000	6.6300
TX, Crosby	Neuroinvasive disease cases	1999-2023	6,033.0000	6.6300
MT, Fallon	Neuroinvasive disease cases	1999-2023	3,025.0000	6.6100
NE, Cuming	Neuroinvasive disease cases	1999-2023	9,084.0000	6.6100
SD, Mcpherson	Neuroinvasive disease cases	1999-2023	2,428.0000	6.5900
MT, Blaine	Neuroinvasive disease cases	1999-2023	6,683.0000	6.5800
ND, Nelson	Neuroinvasive disease cases	1999-2023	3,055.0000	6.5500
SD, Spink	Neuroinvasive disease cases	1999-2023	6,736.0000	6.5300
NE, Morrill	Neuroinvasive disease cases	1999-2023	4,936.0000	6.4800
OK, Beaver	Neuroinvasive disease cases	1999-2023	5,592.0000	6.4400
NE, Grant	Neuroinvasive disease cases	1999-2023	626.0000	6.3900
SD, Brown	Neuroinvasive disease cases	1999-2023	37,578.0000	6.3900
SD, Fall River	Neuroinvasive disease cases	1999-2023	6,974.0000	6.3100
NE, Fillmore	Neuroinvasive disease cases	1999-2023	5,710.0000	6.3000
SD, Jackson	Neuroinvasive disease cases	1999-2023	3,174.0000	6.3000
SD, Haakon	Neuroinvasive disease cases	1999-2023	1,922.0000	6.2400
NE, Hamilton	Neuroinvasive disease cases	1999-2023	9,042.0000	6.1900
ND, Logan	Neuroinvasive disease cases	1999-2023	1,941.0000	6.1800
ND, Steele	Neuroinvasive disease cases	1999-2023	1,953.0000	6.1400
TX, Foard	Neuroinvasive disease cases	1999-2023	1,309.0000	6.1100
CO, Montrose	Neuroinvasive disease cases	1999-2023	40,675.0000	6.1000
NE, Sioux	Neuroinvasive disease cases	1999-2023	1,319.0000	6.0700
NE, York	Neuroinvasive disease cases	1999-2023	13,839.0000	6.0700
NE, Dundy	Neuroinvasive disease cases	1999-2023	1,979.0000	6.0600
SD, Brule	Neuroinvasive disease cases	1999-2023	5,292.0000	6.0500
SD, Hamlin	Neuroinvasive disease cases	1999-2023	5,946.0000	6.0500
NE, Sheridan	Neuroinvasive disease cases	1999-2023	5,330.0000	6.0000
SD, Edmunds	Neuroinvasive disease cases	1999-2023	4,013.0000	5.9800
SD, Marshall	Neuroinvasive disease cases	1999-2023	4,694.0000	5.9700

SD, Hanson	Neuroinvasive disease cases	1999-2023	3,381.0000	5.9200
ND, La Moure	Neuroinvasive disease cases	1999-2023	4,085.0000	5.8800
TX, Oldham	Neuroinvasive disease cases	1999-2023	2,046.0000	5.8700
CO, Crowley	Neuroinvasive disease cases	1999-2023	5,468.0000	5.8500
NE, Keith	Neuroinvasive disease cases	1999-2023	8,234.0000	5.8300
NE, Butler	Neuroinvasive disease cases	1999-2023	8,250.0000	5.8200
ND, Barnes	Neuroinvasive disease cases	1999-2023	11,009.0000	5.8100
SD, Aurora	Neuroinvasive disease cases	1999-2023	2,763.0000	5.7900
NE, Holt	Neuroinvasive disease cases	1999-2023	10,378.0000	5.7800
NE, Furnas	Neuroinvasive disease cases	1999-2023	4,884.0000	5.7300
SD, Campbell	Neuroinvasive disease cases	1999-2023	1,406.0000	5.6900
KS, Clark	Neuroinvasive disease cases	1999-2023	2,166.0000	5.5400
ND, Sioux	Neuroinvasive disease cases	1999-2023	4,331.0000	5.5400
ND, Mchenry	Neuroinvasive disease cases	1999-2023	5,798.0000	5.5200
KS, Meade	Neuroinvasive disease cases	1999-2023	4,378.0000	5.4800
SD, Oglala Lakota County	Neuroinvasive disease cases	1999-2023	14,041.0000	5.4100
SD, Kingsbury	Neuroinvasive disease cases	1999-2023	5,229.0000	5.3500
TX, Sherman	Neuroinvasive disease cases	1999-2023	3,037.0000	5.2700
ID, Owyhee	Neuroinvasive disease cases	1999-2023	11,411.0000	5.2600
SD, Charles Mix	Neuroinvasive disease cases	1999-2023	9,192.0000	5.2200
CO, Logan	Neuroinvasive disease cases	1999-2023	22,360.0000	5.1900
ND, Adams	Neuroinvasive disease cases	1999-2023	2,334.0000	5.1400
SD, Bon Homme	Neuroinvasive disease cases	1999-2023	7,028.0000	5.1200
ND, Grant	Neuroinvasive disease cases	1999-2023	2,348.0000	5.1100
CO, Sedgwick	Neuroinvasive disease cases	1999-2023	2,377.0000	5.0500
ND, Dunn	Neuroinvasive disease cases	1999-2023	3,957.0000	5.0500
TX, Bailey	Neuroinvasive disease cases	1999-2023	7,139.0000	5.0400

After filtering out counties with an Incidence > 5, we are left with 114 counties for further analysis. We plan to cross-reference climate change indicator data (precipitation, temperature, and drought) to determine relationships and significant effects of these indicators on incidence rates in these areas.

Further Analysis

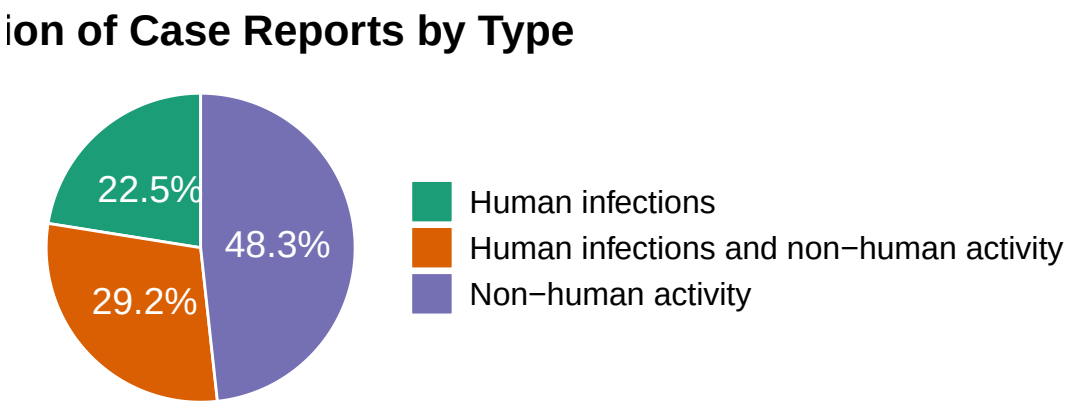
Reported Cases by Year:

These data were individually downloaded from the CDC Historic Data. Each individual data set contained information on the number of cases reported, their activity type, case type, and corresponding county. Activity type here is defined by the following: **Non-human activity**: indicates that veterinary disease cases or infections in mosquitoes, birds, or sentinel animals have been reported to the CDC, **Human infections**: indicates that human disease cases or infections in blood donors have been reported to the CDC, and **Human infections and non-human activity**: indicates that both human infections and non-human infections have been reported to the CDC. These data sets were individually aggregated and written into a .csv file for future use.

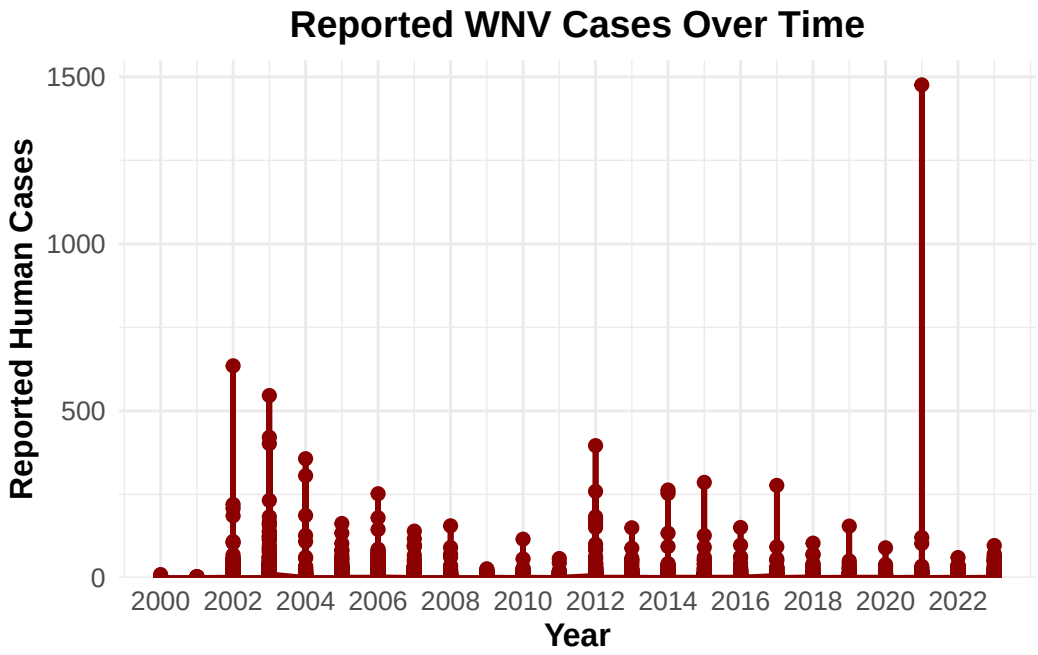
```
[1] "/Users/ehardinparker/Desktop/Completed/CPH_Spring24/BIOS8060E/emmahardinparker-MADA-por
```

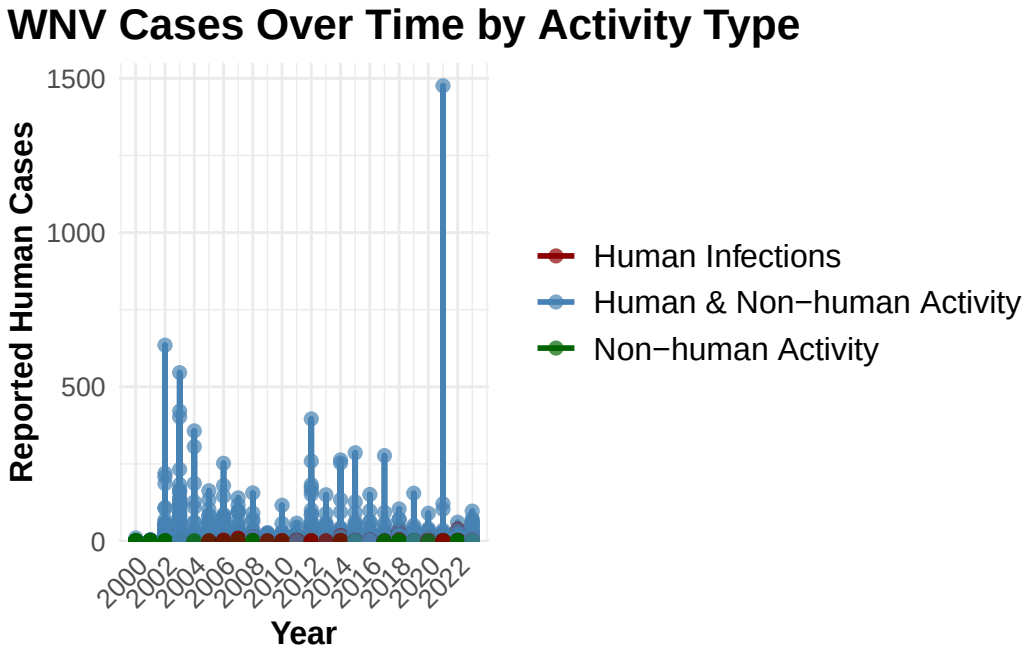
```
     FullGeoName Year County                                Activity
1  CT, Fairfield 2000   9001 Human infections and non-human activity
2   CT, Hartford 2000   9003                     Non-human activity
3 CT, Litchfield 2000   9005                     Non-human activity
4  CT, Middlesex 2000   9007                     Non-human activity
5  CT, New Haven 2000   9009                     Non-human activity
6 CT, New London 2000   9011                     Non-human activity
  Reported.human.cases Neuroinvasive.disease.cases
1                    1                           0
2                    0                           0
3                    0                           0
4                    0                           0
5                    0                           0
6                    0                           0
  Identified.by.Blood.Donor.Screening
1                                   0
2                                   0
3                                   0
4                                   0
5                                   0
6                                   0
```

Descriptive Statistics:




```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```





From the above pie chart we can see that nearly half of all case reports were from non-human activity, with a combination of human infections and non-human activity following at 29.2%.

Results

Descriptive Statistics

Strengths and Limitations:

It should be noted that the CDC itself reported limitations of using ArboNET surveillance data. The following limitations were accessed from the CDC Historic Data Page and are reported here for inclusion in our own limitation section. Firstly, the CDC reports that under-reporting is a common limitation to all surveillance systems that solely rely on healthcare providers to consider the disease (WNV) as a possible diagnosis, obtain the proper laboratory tests, and report confirmed diagnoses to public health authorities. Secondly, mild illness (non-neuroinvasive disease) are more likely to be under-reported compared to more severe (neuroinvasive) cases. Thirdly, data are reported using the county of residence of the individual, not necessarily the county or state of exposure, possibly skewing any state and county level data. Finally, non-human surveillance is variably conducted across the country and the absence of non-human activity data should not be interpreted as no risk.

Our original data sources for climate change indicators were unfortunately affected by the impact of Hurricane Helene in Asheville, North Carolina. We were unable to access some data we were intending to use in our analyses, which resulted in the use of other data sources and analysis techniques. While we find our analyses to be thorough and possess tha chance of having long-term impact on climate change policy and recognition of the impact of climate change on vector-borne diseases, we find it appropriate to mention.