Wildfires vs Data Breaches in the US between 2008 - 2014

...

Team Platypus

February 2021

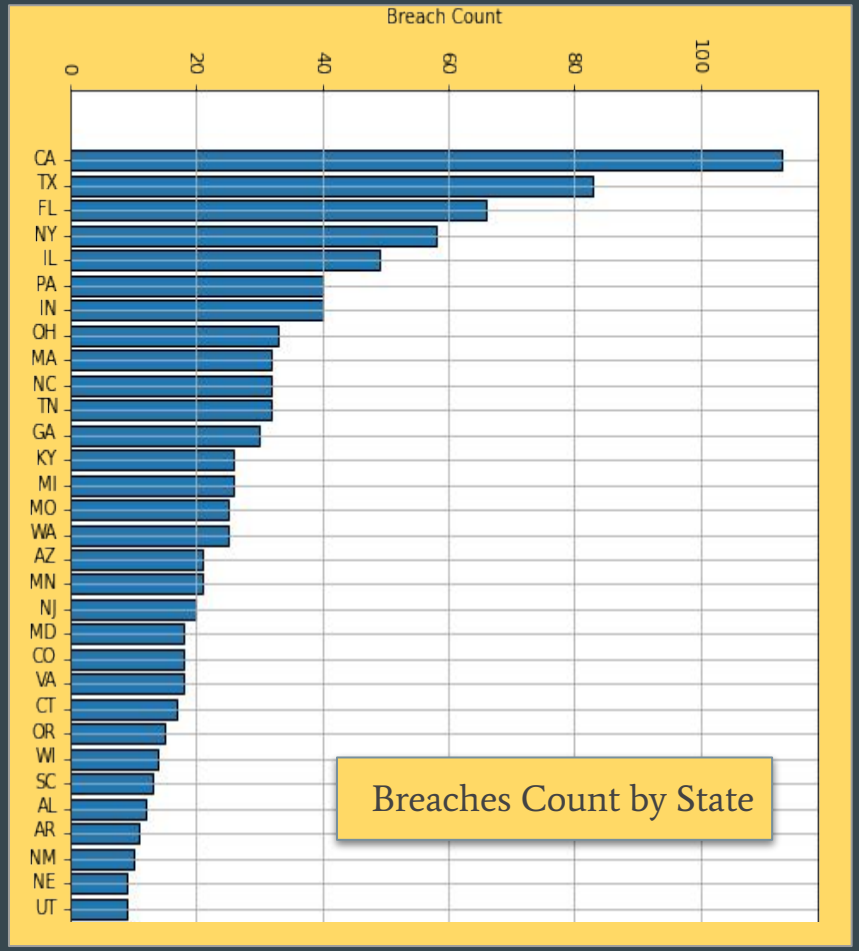
~ Understanding the problem ~

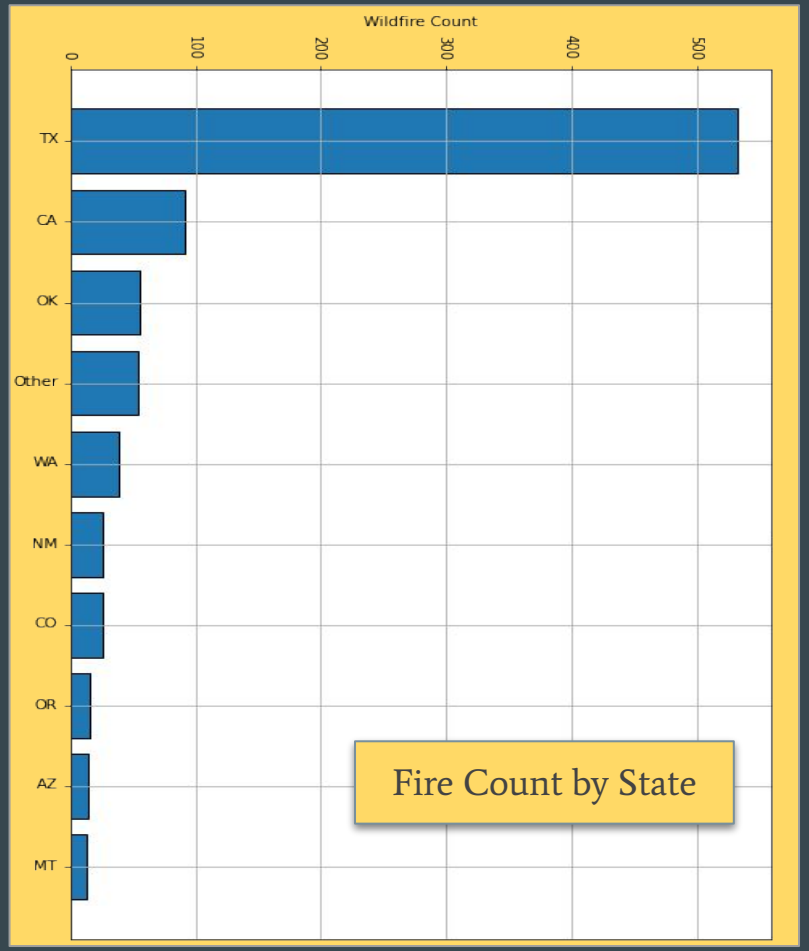
Null Hypothesis	Alternate Hypothesis
Looking at wildfires across the United States, between 2008 and 2014, there is no change in data breaches within that state.	Looking at wildfires across the United States, between 2008 and 2014, there is an increase in data breaches within that state.

How we got started

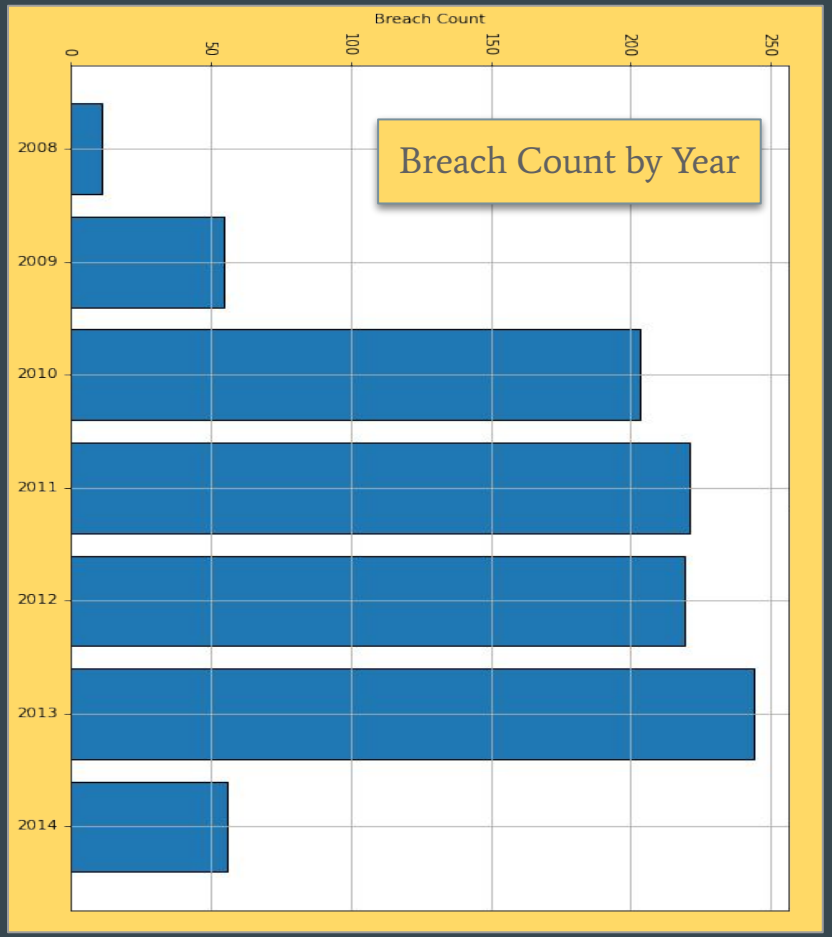
How did you decide on your hypothesis?	Background research, available data, measurability
What data did you use and how did you find it?	Kaggle - cyber breach data FEMA open API - wildfire data
What was your data clean up process?	Narrow scope, identify relevant info
How did you decide what information to focus on?	Compatibility, measurability

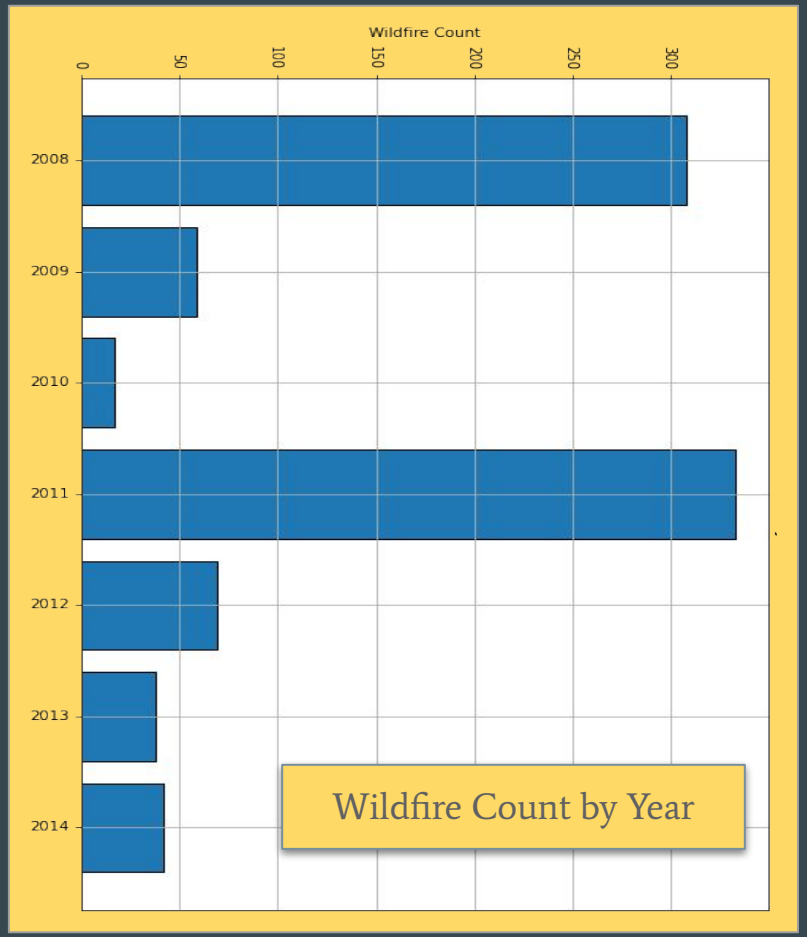
Total Breach & Fire count by State



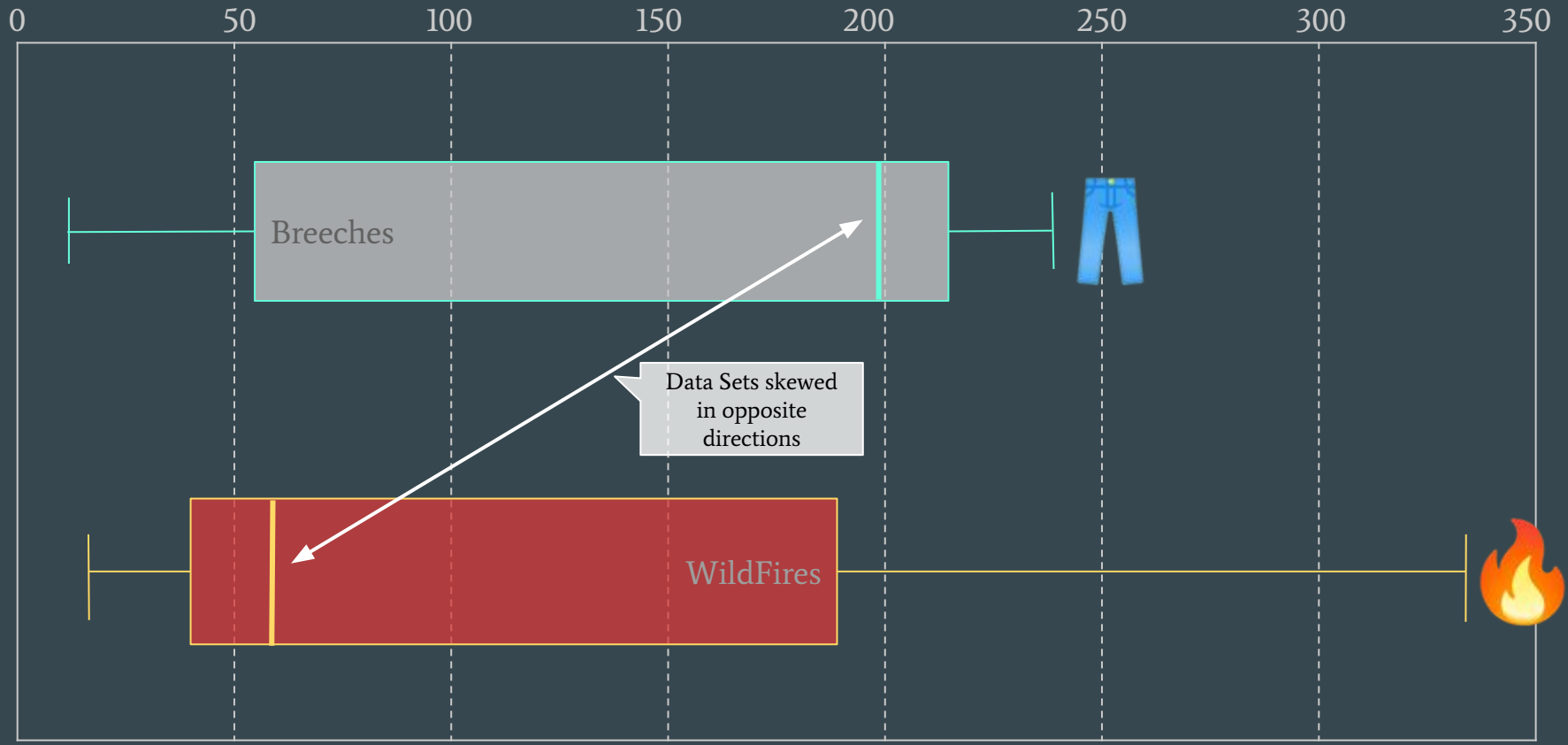


Total Breach & Fire Counts per Year





Breach & Wildfires Across All Years

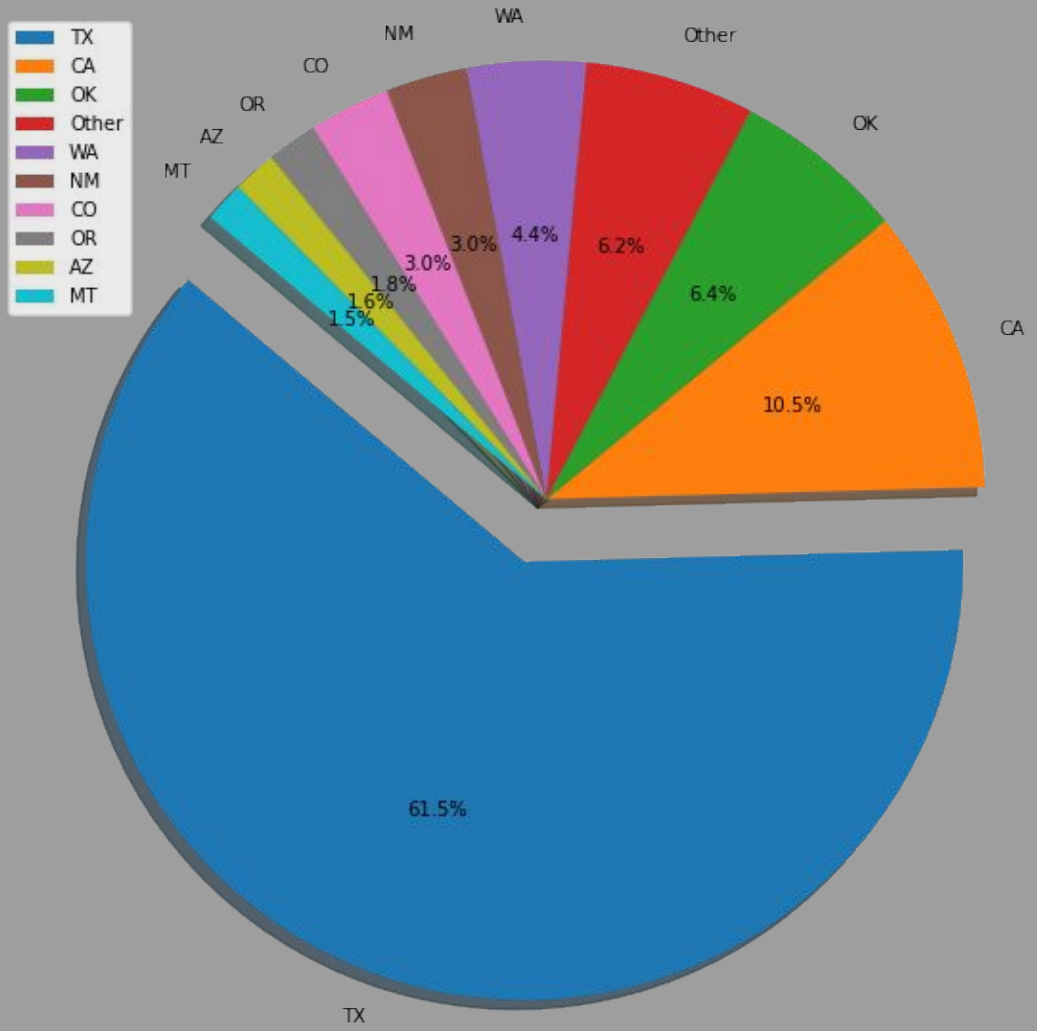


What else was interesting?

Surprising results on the pie chart

Findings

We expected to see CA with the most fires but TX had a shocking majority.

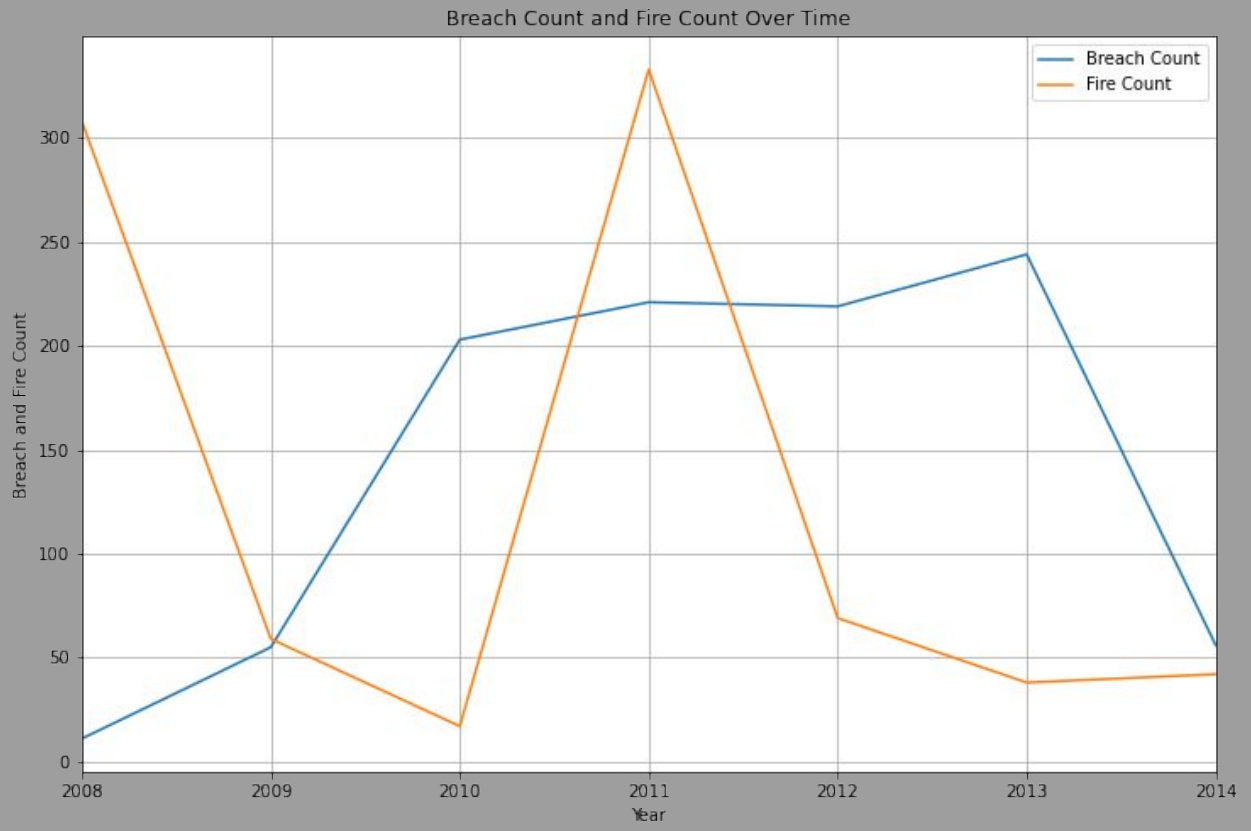


Surprising results on the line chart

Findings

It appears that our two data sets have almost opposite behaviors.

With more time and resources we would like to have tried to find a correlation between Breaches & Fires over time.



Limitations

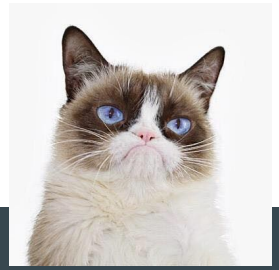
With more time & resources we could have . . .

- More data on breaches
- Cross country data breaches from wildfires
- Other trends hidden in the data

The Team

Lindsey

Story Teller
(Data Clean Up)

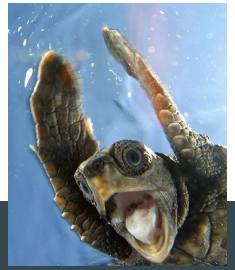


Nader

Bob the Data Builder
(API DF Bilder)

Emerald

Data scissorhands
(Isolating Data)



Pramod

Data Yoderler
(API Call)

Juan

Bob Ross
(Plotting)