|  |  |  |
| --- | --- | --- |
| US Wildfires and Data Breaches | | |
|  |  |  |
| Project 1 – Team Platypus | | |

# Team Members:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1. Pramod | 1. Lindsey | 1. Juan | 1. Emerald | 1. Nader |

# Premise:

|  |
| --- |
| Ho: Looking at wildfires across the United States, between xx and xx, there is no change in data breaches within that state.  Ha: Looking at wildfires across the United States, between xx and xx, there is an increase in data breaches within that state. |

# Questions to Answer:

|  |
| --- |
| 1. How large are the fires being considered? 2. How is data breach defined? 3. Where did the most & least fires occur? 4. Where were the largest fires? 5. Where did the most & least data breaches occur? 6. What is the ‘normal’ ratio of data breaches? 7. How long before and/or after a fire will we be considering as acceptable for indicators of increased data breaches? 8. How many data breaches occur outside of this range? 9. How many data breaches occur inside of this range? 10. Do the number of breaches go up or down? |

# Datasets to Be Used:

|  |
| --- |
| [Cyber Security Breaches Data (kaggle)](https://www.kaggle.com/alukosayoenoch/cyber-security-breaches-data?select=Cyber+Security+Breaches.csv) |
| [~~Data for currently active wildfires across the U.S (aerisweather.com)~~](https://www.aerisweather.com/support/docs/api/reference/endpoints/fires/#properties)  <https://www.fema.gov/about/openfema/api> (Fema) |

# Task Breakdown:

|  |
| --- |
| 1. Clean data breach file – Lindsey 2. Pull fire data – Pramod 3. Put fire data into DF – Nader 4. Statistical analysis - Emerald 5. Create plots of data for comparison – Juan 6. Written analysis – Group |
|  |