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

# Team Members:

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

# Premise:

|  |
| --- |
| Looking at wildfires across the United States, is there an increase in data breaches? |

# Questions to Answer:

|  |
| --- |
| 1. Does the wildfire location increase the number of data breaches in that location? 2. How large are the fires being considered? |
| 1. How is data breach defined? |

# 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) |

# Task Breakdown:

|  |
| --- |
| 1. Define the answers in ‘Questions to Answer’ section. 2. Find, pull, and clean data. 3. Run analytics on both groups of datasets. 4. Using statistical modeling find if there is correlation between the two datasets. |