

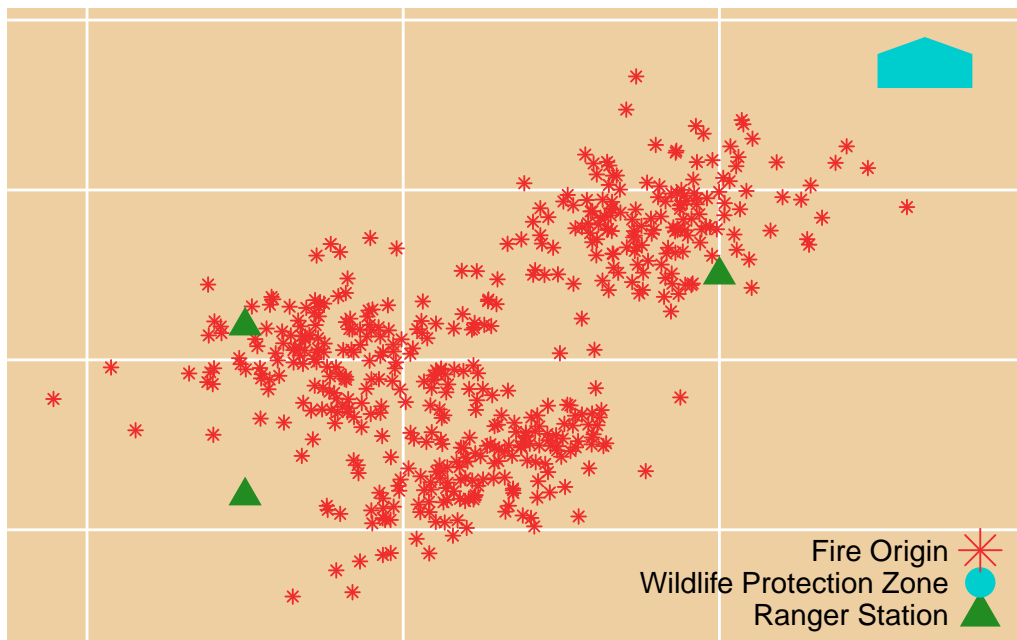
# Wildfires

## Simulated Dataset

4/15/23

### Description

This data describes 500 wildfires that started within a large national park. The origin of each fire is shown below in red.



### Data

The dataset contains observations on 500 fires started in a large national park. For each fire, the data contain 15 variables, in addition there are two outcomes ('burned' and 'wlf'). These variables are:

- **x**, and **y**: the (x, y) coordinates of where the fire started.
- **temp**: air temperature when the fire started
- **humidity**: air humidity when the fire started
- **windspd**: wind speed when the fire started
- **winddir**: wind direction when the fire started
- **rain**: rainfall in the week preceding the fire
- **days**: number of days since the last fire
- **vulnerable**: amount of vulnerable (unburned) foliage near the fire
- **other**: indicates if another fire is ongoing when the fire starts
- **ranger**: indicates if the fire started near a ranger station, and that ranger station was manned. Note that even if a fire started near a ranger station, that station may not have been manned at the time.
- **pre1950**: indicates if the fire occurred before 1950, which it became more common to use planes and helicopters to fight forest fires.
- **heli**: indicates if a helicopter was available to fight the fire
- **resources**: composite score for the resources available to fight the fire, including manpower and available air support
- **traffic**: indicates the level of foot traffic in the national park when the fire started
- **burned**: number of hectares burned by the fire
- **wlf**: indicates fire reached the wildlife protection zone