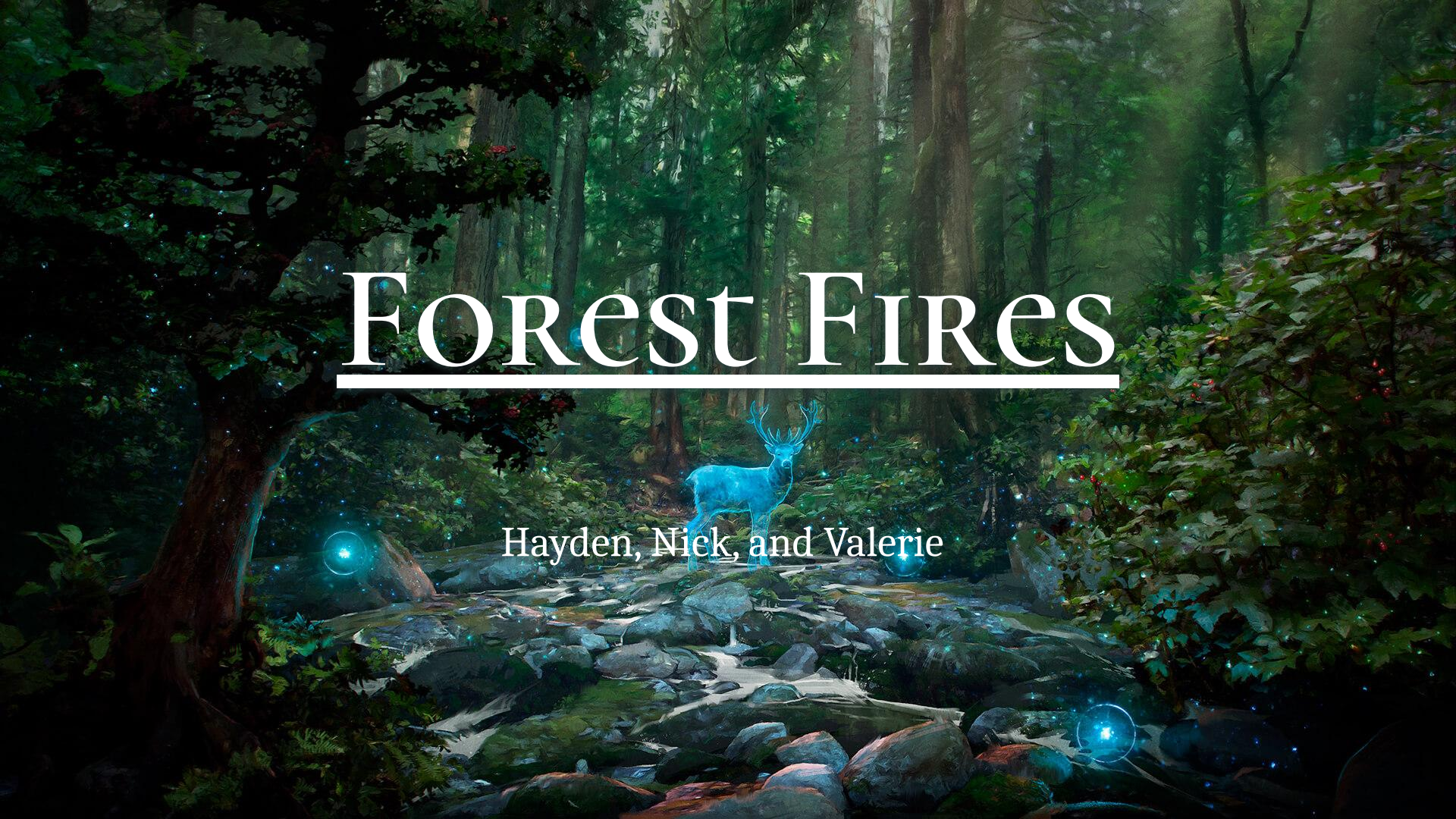
# Forest Fires



Hayden, Nick, and Valerie

# Our Mission

*Try and predict areas with high risk of forest fires based off of easy-to-gather statistics*

# Our Data

## Burn Severity Data

**Key Variables:**

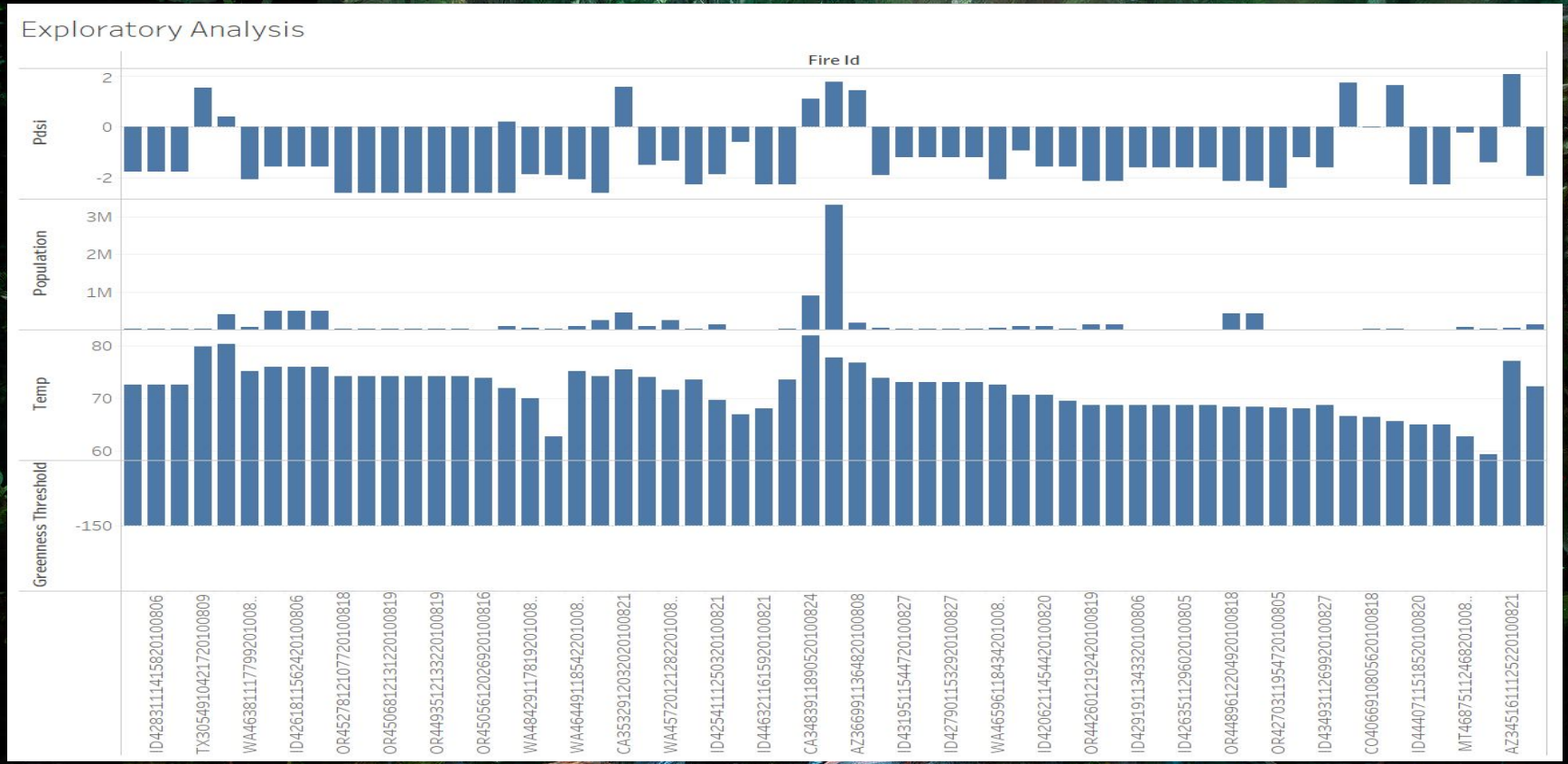
- Area
- Low Threshold
- Moderate Threshold
- High threshold
- Greenness threshold

## US Counties Data

**Key Variables:**

- Temperature
- Palmer Drought severity index (PDSI)
- Population

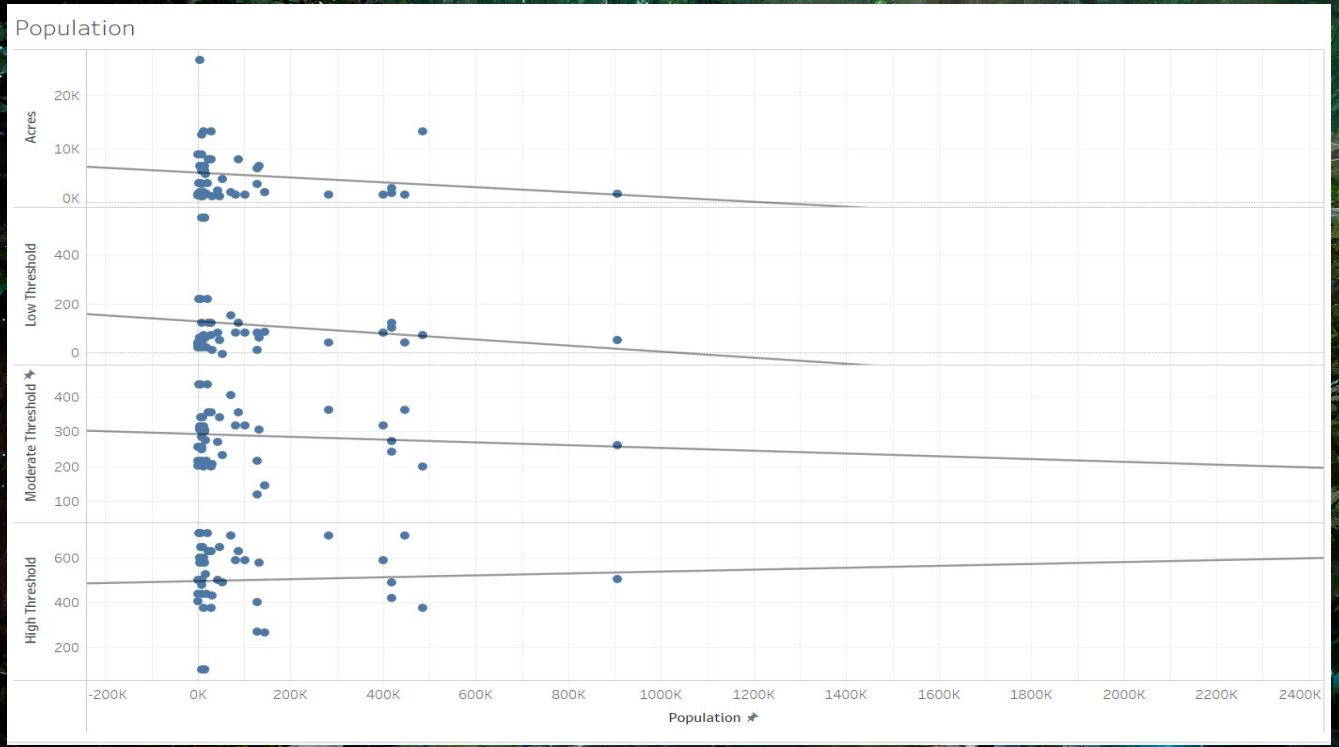
# Exploratory Data Analysis



# Levels of burn

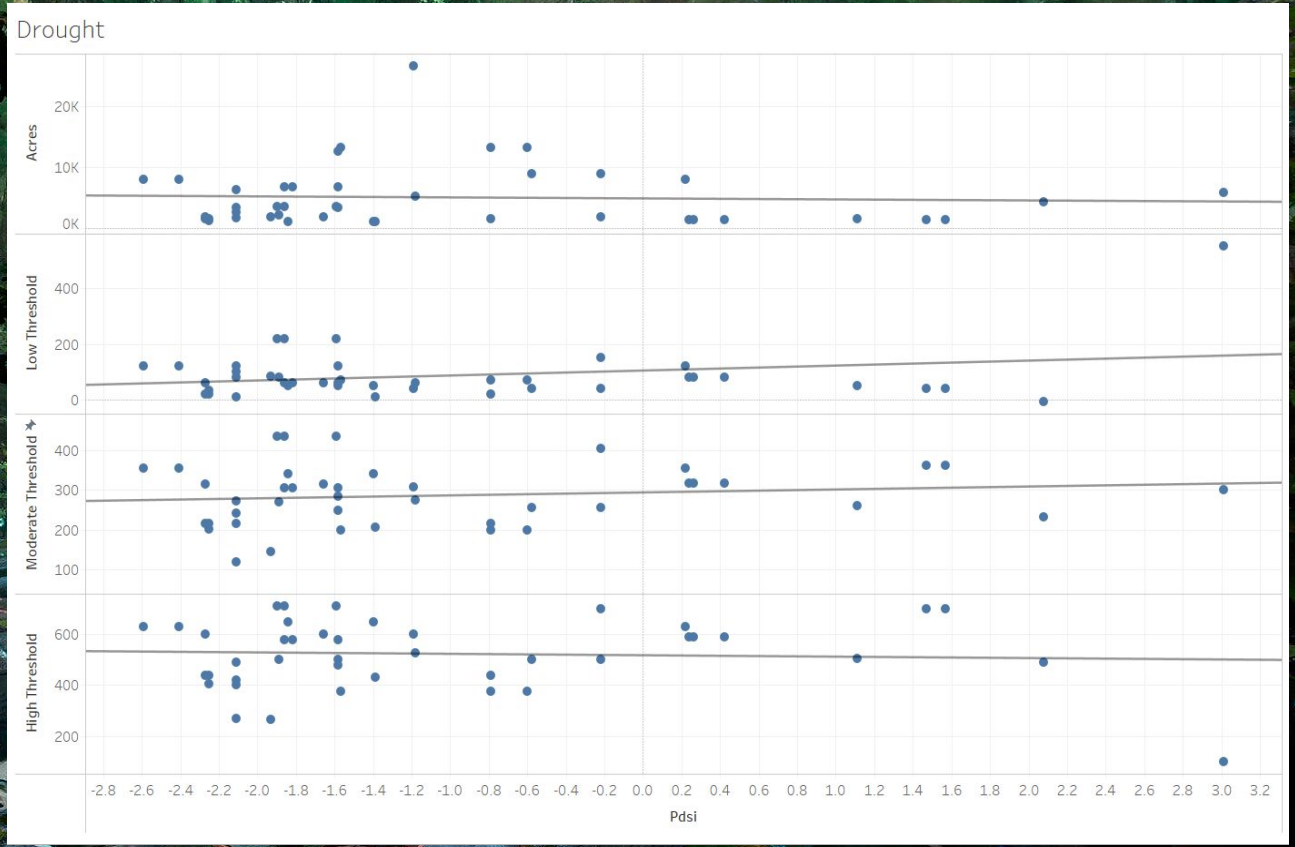
- Data separated into 3 levels of burn severity
- Most severed areas being the high threshold and least burned areas means lower threshold

# Population, Amount of Area burned, burn severity

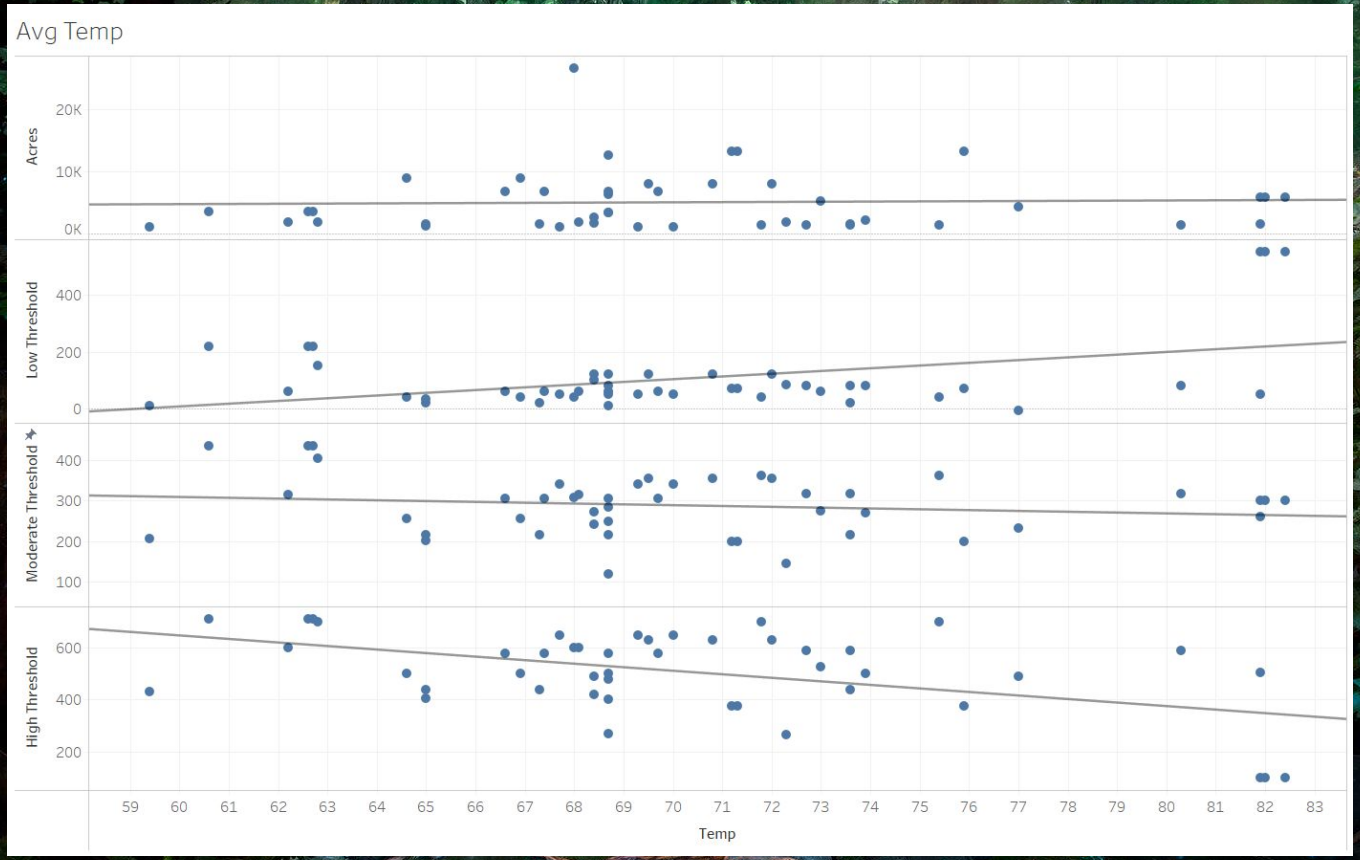


# PDSI Levels

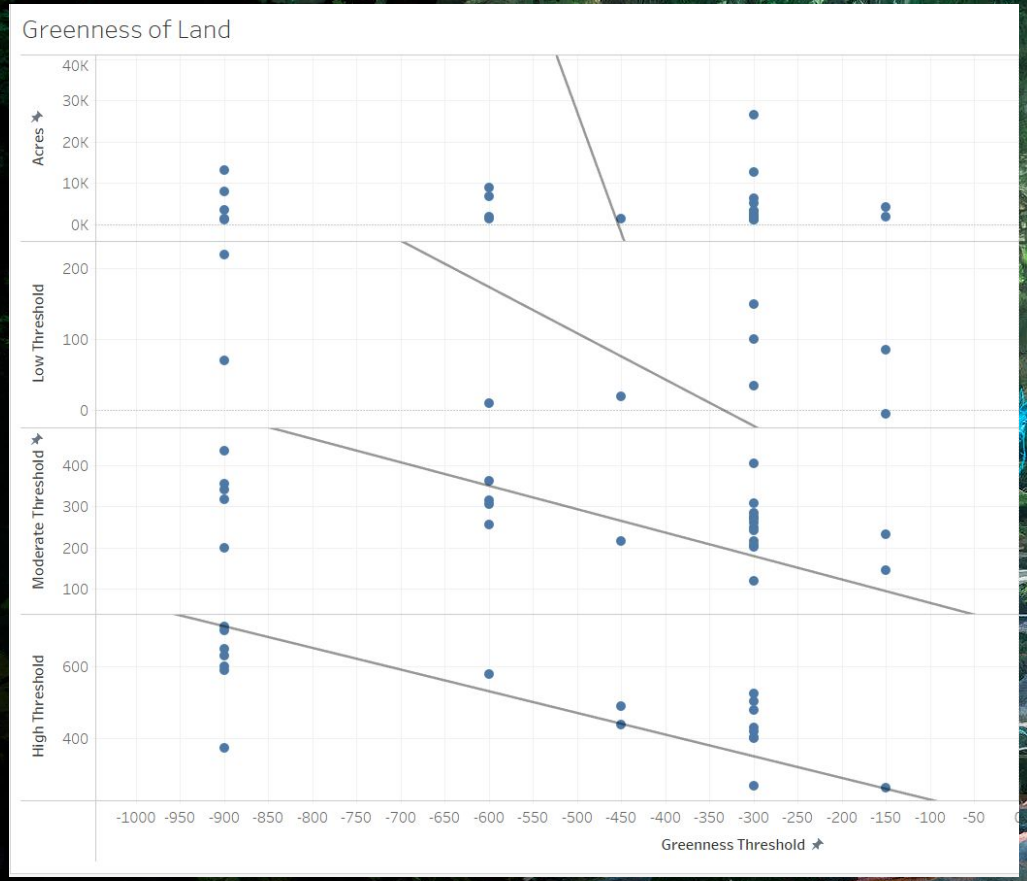
PDSI shows how dry an area is, negative values mean drier and positive means more wet



# Average Temperature



# Greenness of Land at Time of Fire




Greenness of an area has a negative relationship here. The greener the area is means that the burn was less severe. This data is categorical unlike our other models that we used previously.

# CONCLUSIONS

- Results were not exactly as we expected them to be
  - Temperature had an inverse relationship with burn severity
  - P-Values and R-Squared values very poor
- Success when it came to finding important factors that lead to the severity of forest fires
- Had a lot of trouble when it came to scraping and gathering the data itself
  - Only could use August 2010 data
  - Octoparse kept crashing on us
- Maybe more data would have led to better overall predictive performance