Springboard Capstone 2 Project Proposal:
The Phenology of Flowering Plants on the Carrizo Plain

Description: The Carrizo Plain National Monument is an area of native and exotic grassland approximately 30 miles long and 10 miles wide. It is located east of San Louis Obispo, California, between the Caliente and Temblor mountain ranges. The Carrizo Plain is known for spectacular spring wildflower blooms that can vary in intensity from year to year, depending on rainfall. Total amount of rainfall between October and May and temperature influences blooms. It is unclear how late (March and April) versus early (November through February) influence bloom times of species with differing phenologies, the cyclic or seasonal patterns seen in the lifecycle of living organisms. Some flowering species bloom early in the season while others don't bloom until late spring or early summer, when conditions have already become dry. A model will be created that uses climatological data and iNaturalist observations to predict bloom date for flower species on the Carrizo Plain National Monument and explore the impact of early versus late rainfall.

Context: How can iNaturalist observations and NOAA climate data be used to predict when certain flower species will bloom based on temperature and rainfall patterns from previous years and recorded climate data for the current year?

Criteria for Success: The model will predict bloom start and end within a week of the actual observed bloom time for any year in the dataset.

Scope of Solution Space: The model will focus on predicting first bloom and last bloom across a range of flowering plant species for an upcoming or current spring and summer season.

Constraints: Some weather stations have large amounts of missing data while others have 100 percent coverage. The tool may not work in areas with low daily data coverage or in areas where the nearest weather station is in a separate micro-climate. Flower observations in iNaturalist do not always indicate if the plant is in bloom. This project may require a method for filtering observations of plants that are not in bloom. Climate change has altered weather patterns in recent years and this could impact the ability to predict current bloom year with a model that includes historic data. Some constraint on number of years of past data used should be included in the model. The tool will work best when a large part of the current year rainy season has already occurred.

Stakeholders: Stakeholders are clients who may use the tool to plan outings, inform guided activities, provide tourism support and more. These clients may be individuals, hikers, park employees, biologists, or anyone interested in flower phenology.

Key Data Sources: Research grade observation data will come from iNaturalist through the iNaturalist API. These data will include flower species seen in the area of interest and be filtered for plants that are in bloom. Climate data will come from NOAA using their API to gather data

from the nearest climate station to the park of interest. Specifically, daily high and low temperatures and daily rainfall will be acquired.

Methodology: Daily climate data including high temperature, low temperature, precipitation, and precipitation type will be used in exploratory analyses to visualize patterns in the years of interest. Bloom observations will be plotted with climate variables to look for relationships between climate attributes and bloom time. Additional model variables will be created and may include: time between first rainfall and first bloom, time between last freeze and bloom time, time between last rainfall and last bloom, rainfall by week, rainfall by month, and any other measurement which seems relevant based on exploratory analyses. A separate model will be optimized for each species of interest using a pipeline, identifying the variables that influence that species most. These influential variables will be reported as interesting species attributes. The model will report the predicted first bloom and last bloom for the species of interest, producing a list of species that might be seen on a given date. Predictions can be based on recent or current rainy season data or input manually as a positive or negative difference from the average for each week or each month.

Products and Deliverables:
The model and analysis will be used to produce a project report and a slide deck outlining the methodology and resulting prediction tool.