Weekend Weather Prediction for Outdoor Event Scheduling in Monterey

Background + The Client's Decision (The Problem)
To decide when to schedule events, Monterey peninsula event planners for outdoor weddings and cultural festivals, such as the Monterey Jazz Festival, must often guess whether or not a given weekend will be sunny, sometimes up to a year in advance. But the area's climate can be notoriously volatile and counterintuitive: sunny weather ends sometime in May or June and begins again sometime in September, after a few months of fog throughout the summer season, with gaps of sun between

The Action
I propose to develop a predictive model based on weather data for the city of Monterey for the last ten years, to assign a sunniness probability for each calendar weekend. I will need to align calendar weekends across years, i.e., "the first weekend in March," and create an average sunniness score for a given weekend that takes into account three component days (Friday, Saturday, and Sunday) and averages across ten different corresponding weekends through time; a "calendar weekend" data structure will include data for thirty different calendar days (Friday, Saturday, and Sunday for ten years into the past). The dataset will be ordered from [the National Center for Environmental Information](#), and will include several daily summary fields: total daily precipitation, average wind speed, and (most importantly) daily weather type. The data requested spans a period of May 14th, 2008 to May 14th 2019.

The project requires several considered modeling choices. The reality of climate change problematizes the question of timeframe: how many past years of data really reflect a weather trend relevant to the immediate future, rather than pre-global-warming patterns that no longer apply? But also how many years of data will be necessary to avoid overfitting to this small amount of data? One challenge of interpretation will be the quantification of qualitative, categorical data: how should a "partly cloudy" or "partly sunny" day impact a weekend's overall sunniness score? The model will also require binarization, as the goal is to determine a "sunny/not sunny" binary value for each calendar weekend — should this binarization occur for each day, for each weekend, for each day throughout history, or for each weekend throughout history?

Deliverable
Because the aim of the project is to improve event planners' future event scheduling decisions, I will create a calendar that displays sunniness probabilities for each weekend, an interface that will allow planners to make more informed choices when scheduling outdoor events in the area.

The Value of an Improved Decision
While this won't enable us to predict the future, it should allow event and festival planners in the Monterey area to make more informed event planning choices in the future. This might be especially useful to those planners who have recently moved to the area and aren't yet familiar

with the area's peculiar weather patterns, and to smaller businesses or individual entrepreneurs who can't afford data consultants.