Exploratory Data Visualization of Monterey Weather

Chaos determines weather, and the most powerful weather prediction systems depend on models of chaos. With this in mind, when asked to make a prediction for hypothetical event planners based on a decade of hourly weather data using a laptop, it's useful to frame the task of exploratory visualization with the question: from what perspectives is this data least chaotic?

With local knowledge of the bay's peculiar weather patterns, I first asked about each daily hour's relative obscuration rating, which reveals that 10 AM to 4 PM are notably clearer than all other hours on average. It's common for morning fog to burn off in the afternoon and for obscuration to return in the evening hours here. To see each hour's variation across the year, I plotted month-to-month averages for each daily hour's obscuration, which reveals narrower ranges for summer months. The summer months are more predictable, but, unfortunately, they tend to be overcast, and this doesn't help much with regard to predicting clear days.

I continued to plot a distribution of average daily obscuration. This revealed substantially more obscured days than clear days. I also considered the calendar day average across the decade, which compressed some of the extreme values toward the mean.

Next, I went looking for calendar days with clear average obscuration scores. There were 36 calendar days with decade average obscuration ratings of clear or scattered clouds. These days were concentrated in January and November, but a few days at the end of December had the clearest ratings.

After this, I investigated temperature trends. I plotted two-week moving averages for daily temperature maxima, minima, and means on one plot, to see temperature trends and ranges throughout the year, as this is a common technique for communicating weather time series data. To focus in on each month's predictability, I created box plots depicting each month's range of daily maximum and minimum temperature values over the decade. This revealed that the summer months have a compressed range of both maximum and minimum temperature values.

I looked at some correlations next. I plotted a hexagonal regression that shows a correlation between overcast weather and a particular range of temperatures between 60 and 70 degrees, and I plotted a scatterplot with a linear regression that shows the positive correlation between daily maximum wind speed and daily precipitation.

I looked at average annual precipitation, which had quite a wide range (from 4 inches one year to 22 inches another year).

I finished by plotting the 36 clear days in order of average daily precipitation and average maximum temperature.