A Bespoke Almanac for Monterey Event Planners

Jeff Trevino, 2019

Springboard, Capstone Project One

The Business Decision

- Event planners choose event dates based on weather forecasts.

The Data-Driven Solution

- An automatically generated, data-driven calendar that displays local predictions for
  - Temperature (degrees Fahrenheit)
  - Sky Obscuration (0-8 on the NOAA obscuration scale)
  - Humidity (%)
  - Relative Likelihood of Sky Obscuration

Wrangling

	STATION	DATE	REPORT_TYPE	SOURCE	AWND	BackupDirection	BackupDistance	BackupDistanceUnit	BackupElements	BackupElevation	Back
0	72491523259	2009-04-01T00:08:00	FM-16	7	NaN	NaN	NaN	NaN	NaN	NaN	
1	72491523259	2009-04-01T00:50:00	FM-16	7	NaN	NaN	NaN	NaN	NaN	NaN	
2	72491523259	2009-04-01T00:54:00	FM-15	7	NaN	NaN	NaN	NaN	NaN	NaN	
3	72491523259	2009-04-01T01:54:00	FM-15	7	NaN	NaN	NaN	NaN	NaN	NaN	
4	72491523259	2009-04-01T02:54:00	FM-15	7	NaN	NaN	NaN	NaN	NaN	NaN	

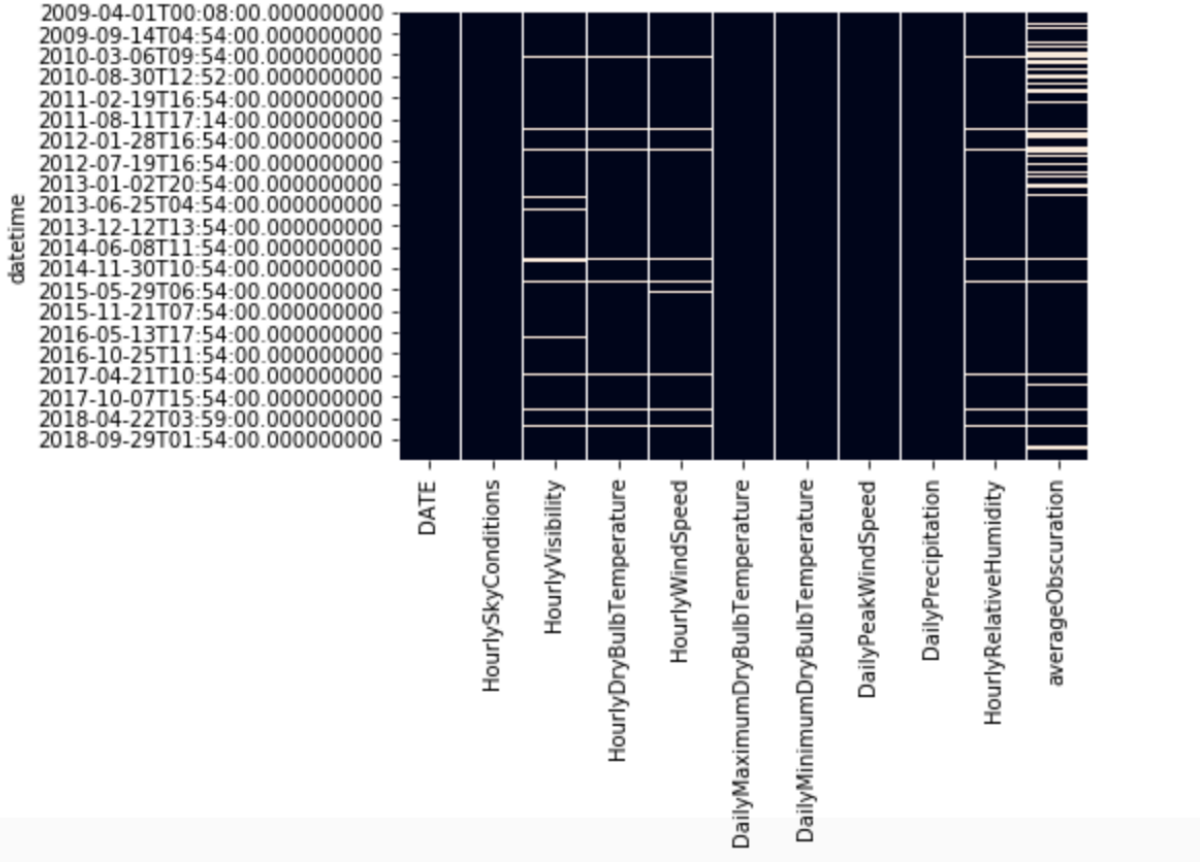
Wrangling

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 124599 entries, 2009-04-01 00:08:00 to 2019-03-31 23:59:00
Data columns (total 10 columns):
DATE                              124599 non-null object
HourlySkyConditions               102830 non-null object
HourlyVisibility                  120764 non-null object
HourlyDryBulbTemperature          120422 non-null object
HourlyWindSpeed                   120617 non-null float64
DailyMaximumDryBulbTemperature    3601 non-null object
DailyMinimumDryBulbTemperature    3602 non-null object
DailyPeakWindSpeed                3611 non-null object
DailyPrecipitation                3610 non-null object
HourlyRelativeHumidity            120251 non-null object
dtypes: float64(1), object(9)
memory usage: 10.5+ MB
```

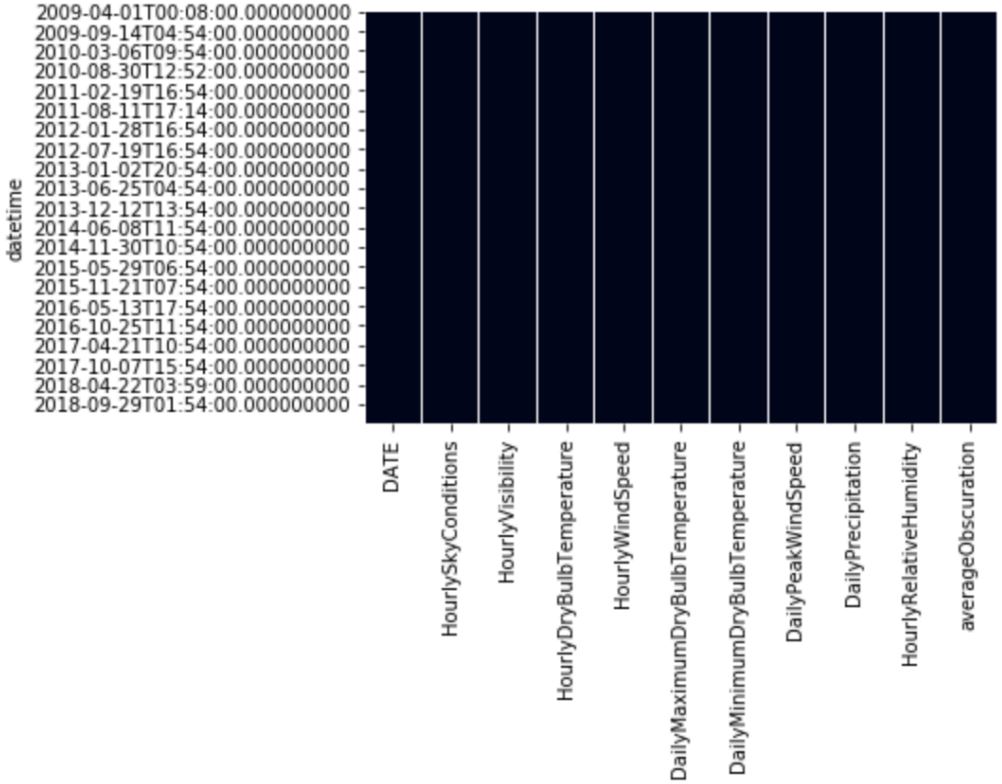
Wrangling

```
DATE                                 object
HourlySkyConditions                  object
HourlyVisibility                    float64
HourlyDryBulbTemperature            float64
HourlyWindSpeed                     float64
DailyMaximumDryBulbTemperature      float64
DailyMinimumDryBulbTemperature      float64
DailyPeakWindSpeed                  float64
DailyPrecipitation                  float64
HourlyRelativeHumidity              float64
dtype: object
```

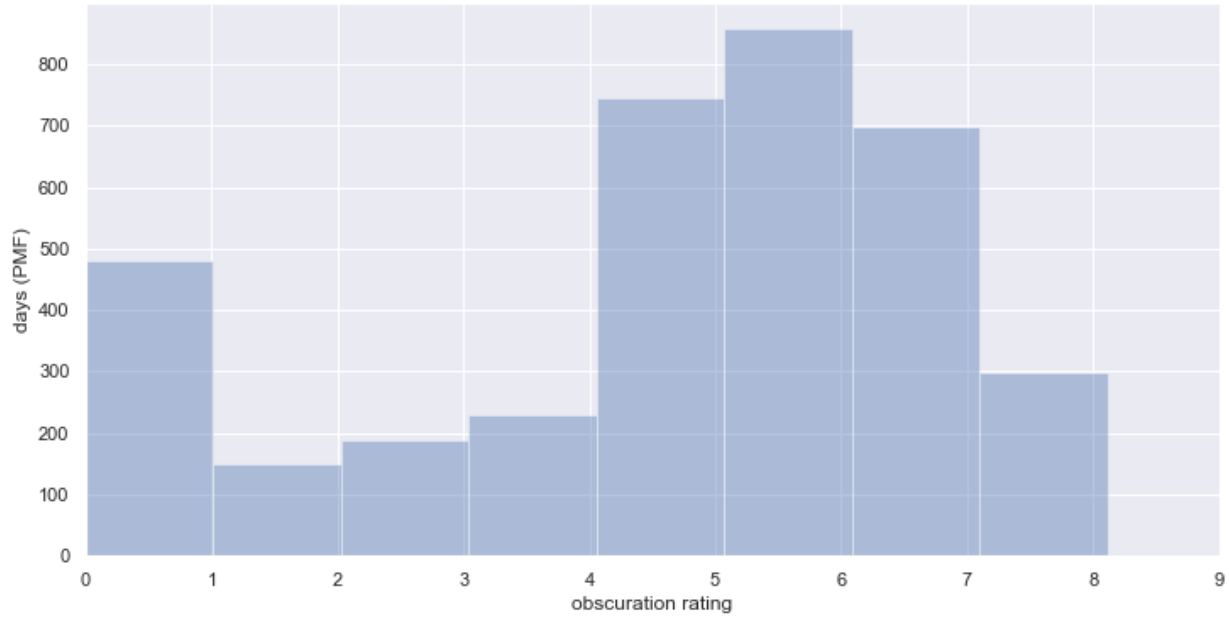
Wrangling – Heat Map of Missing Numeric Values



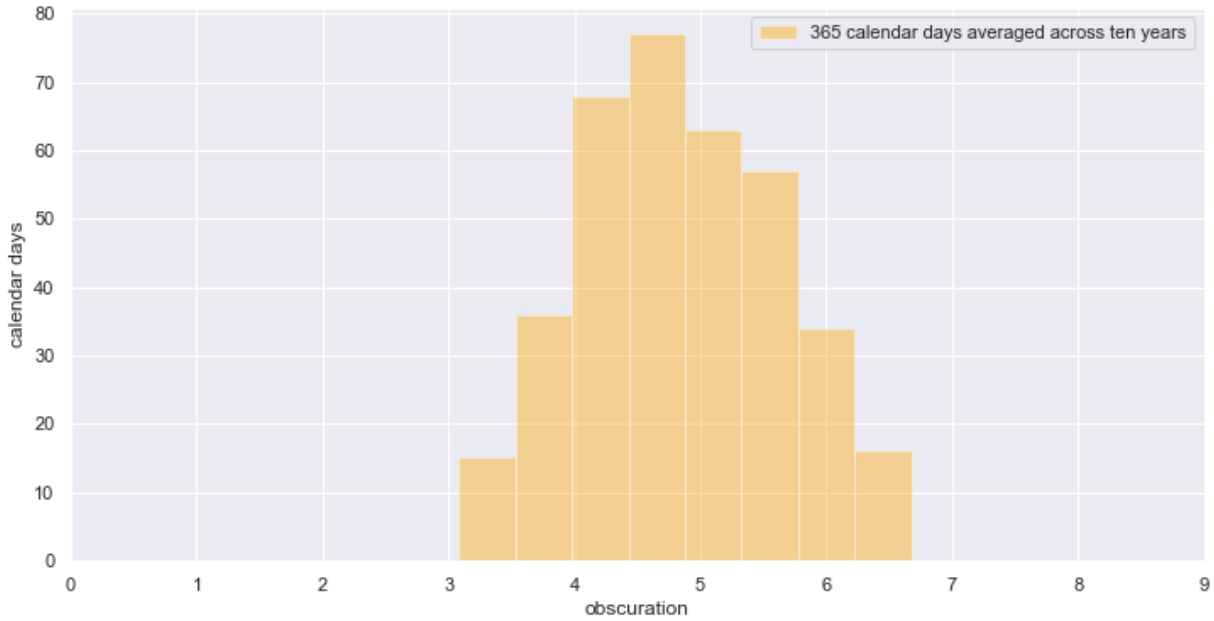
Wrangling – No Missing Values After Mean Imputation by Column



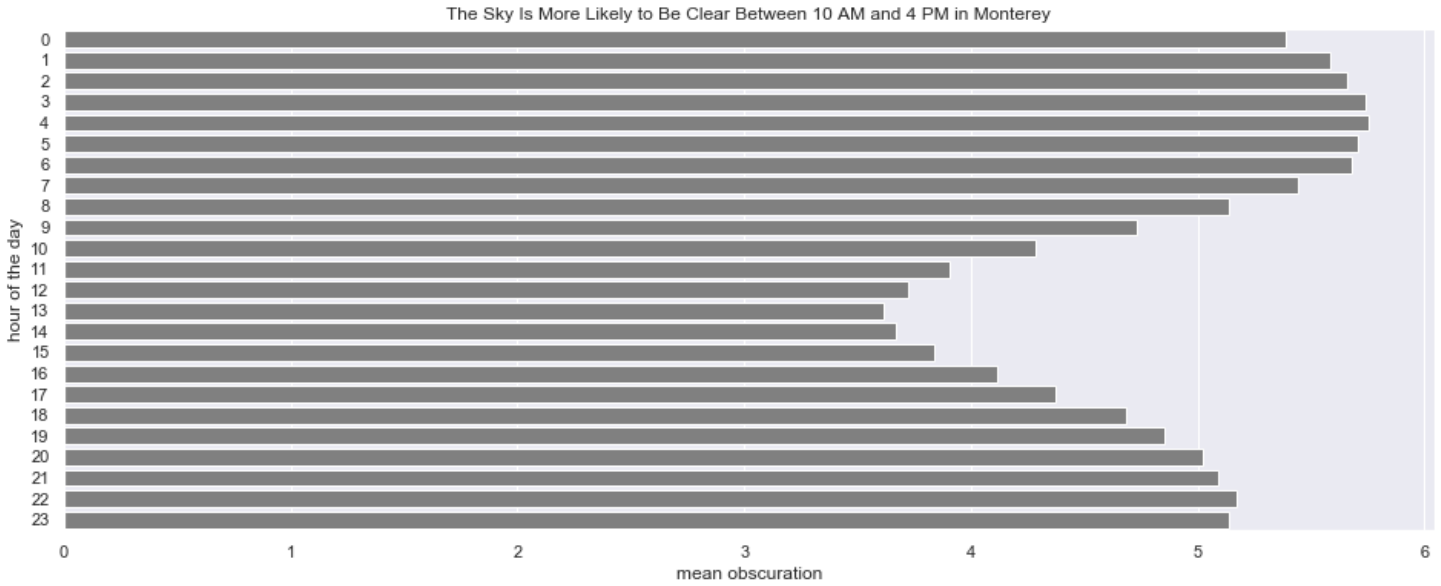
Exploration – Obscuration Distribution



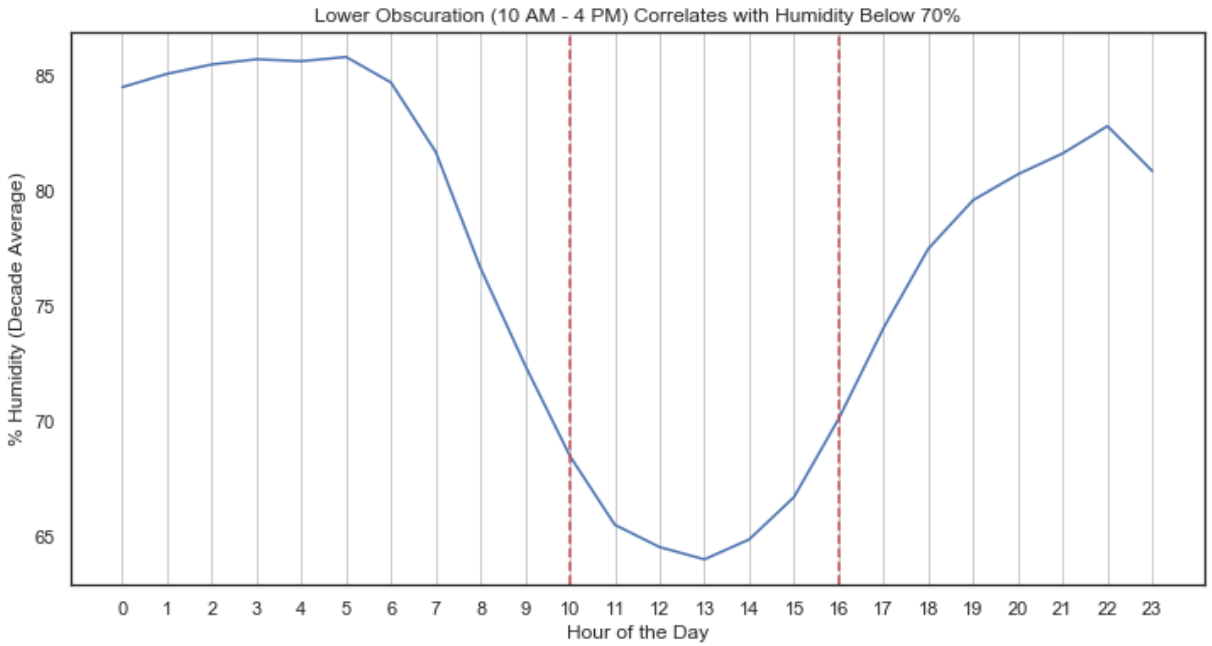
Exploration – Obscuration Distribution (Calendar Days Averaged Over Decade)



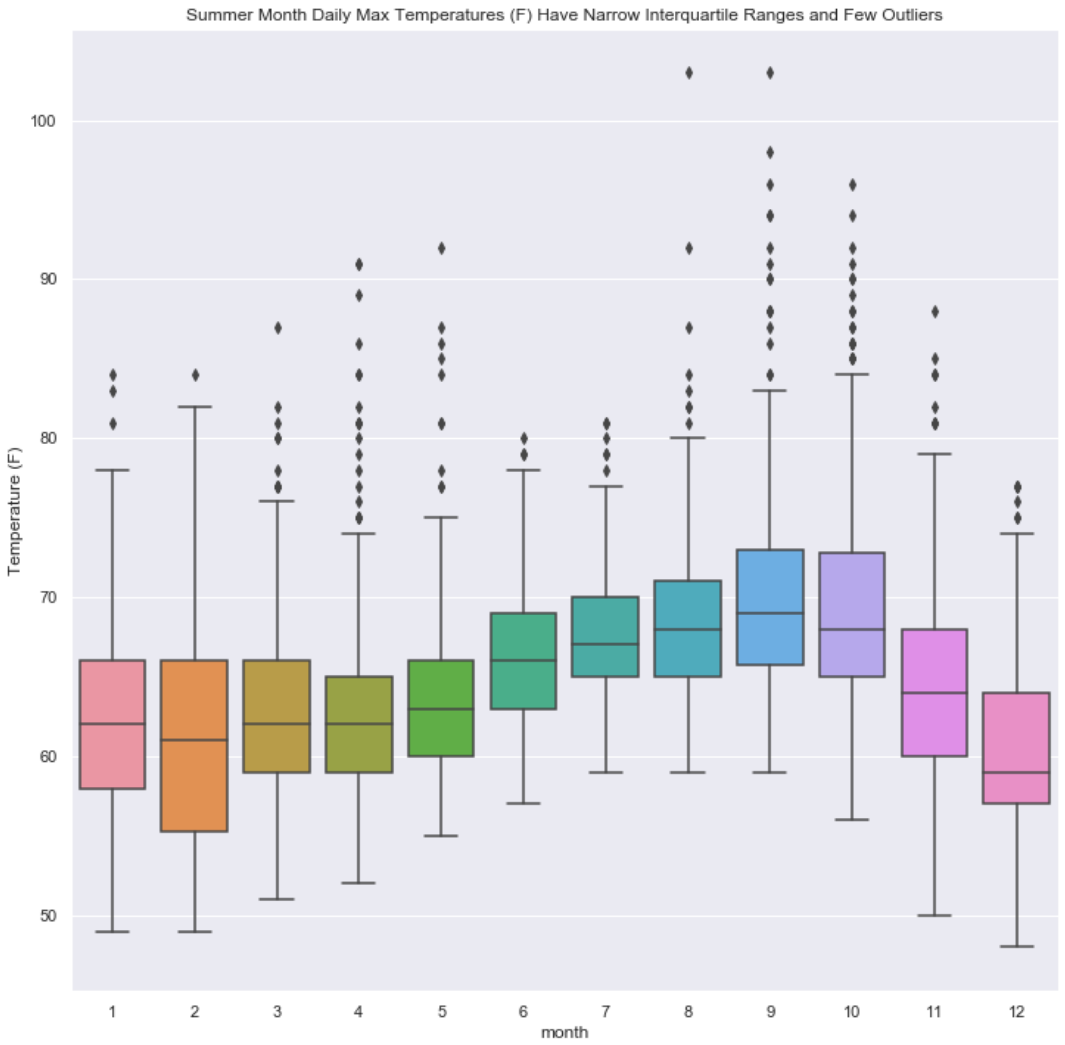
Exploration – Daily Obscuration Pattern



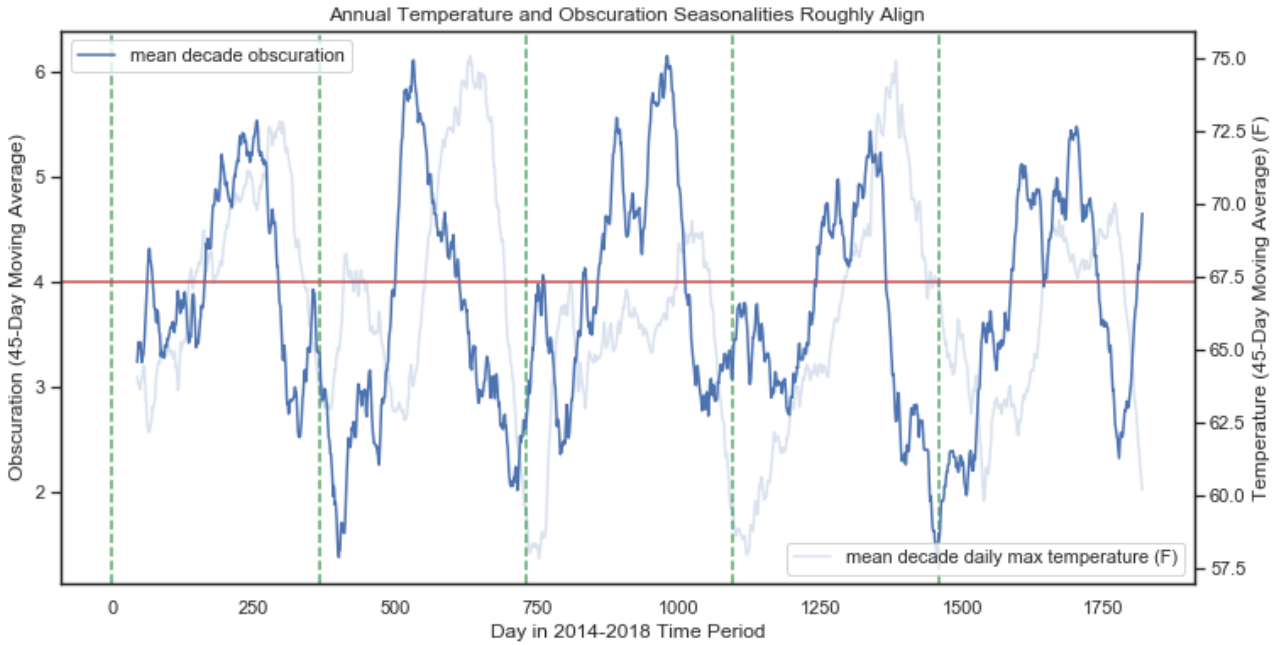
Exploration – Daily Humidity Pattern



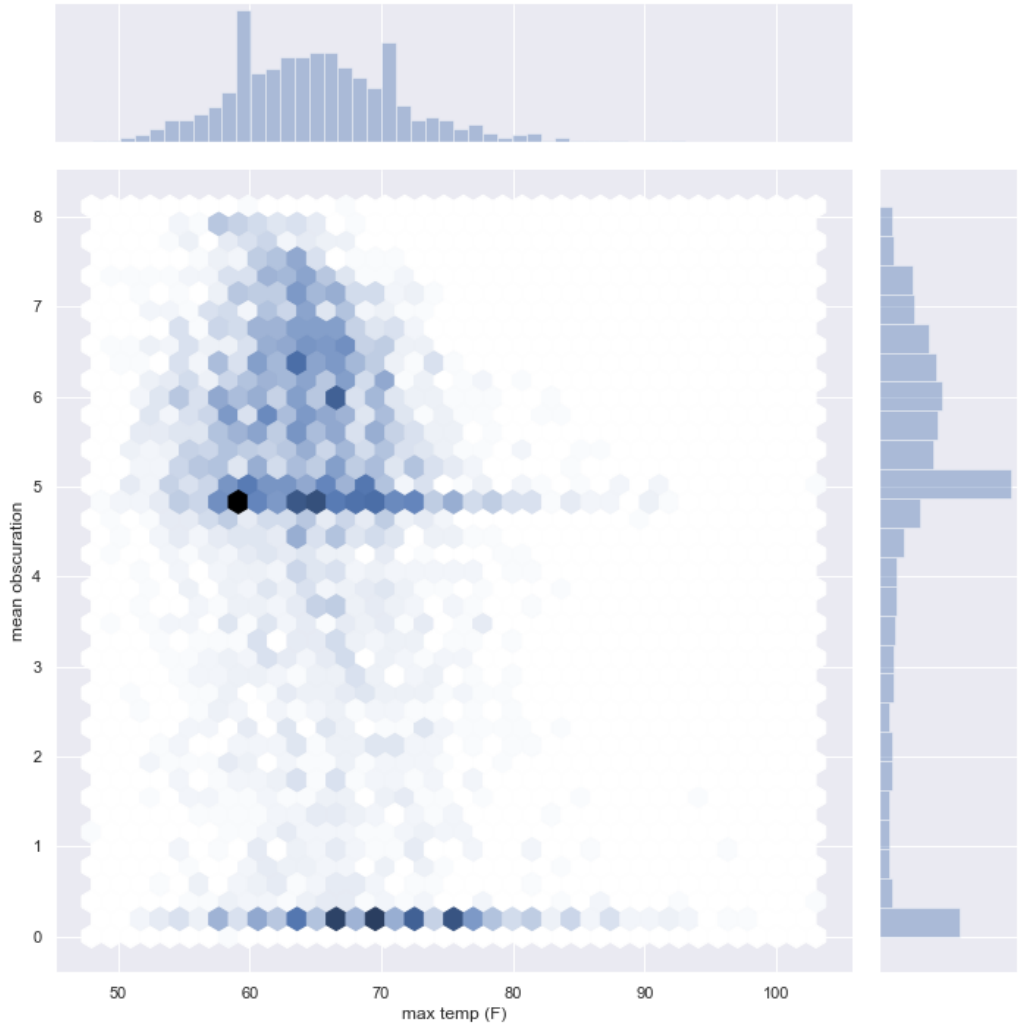
Exploration – Monthly Temperature Distributions



Exploration – Annual Temperature and Obscuration Seasonalities



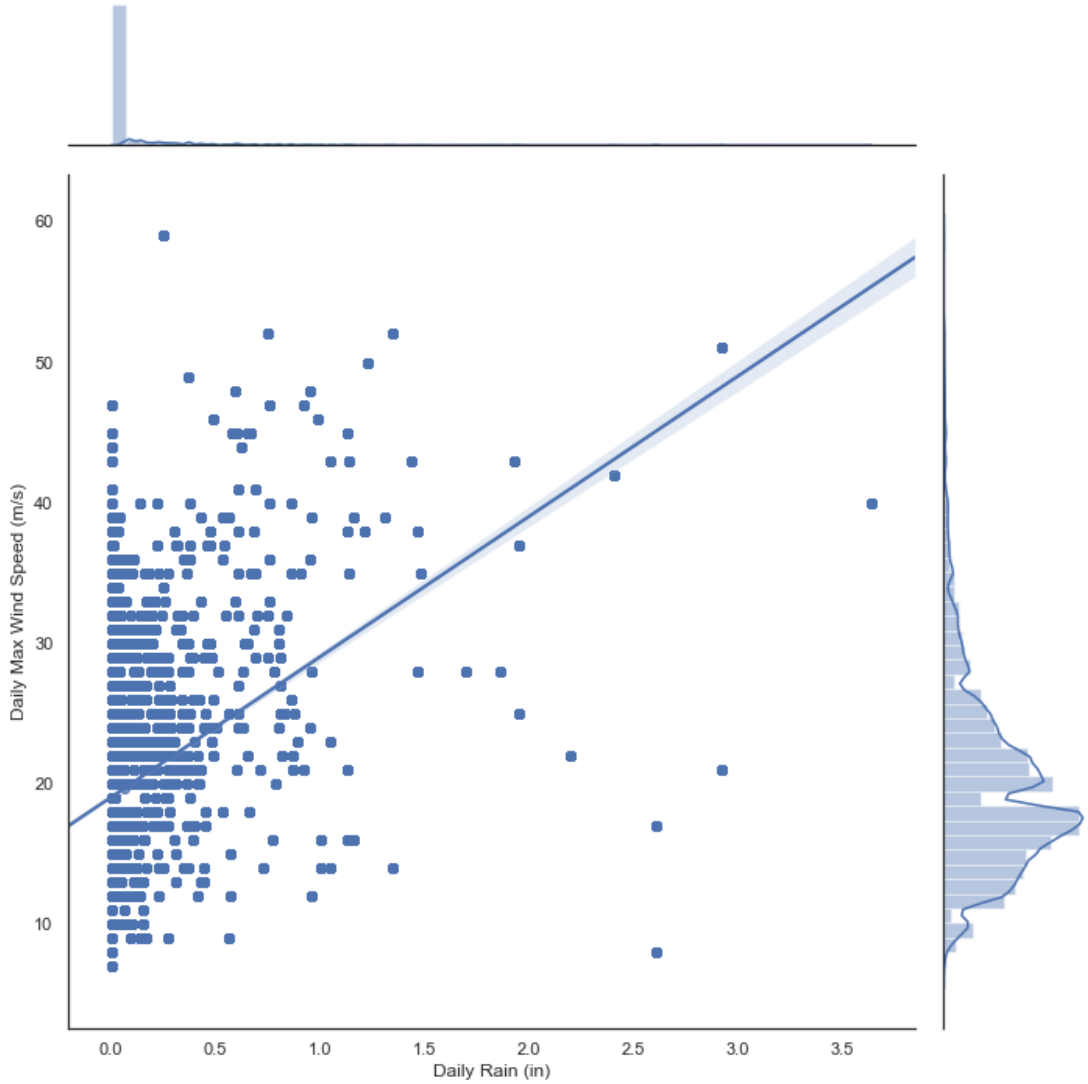
Exploration – Annual Temperature and Obscuration Correlation



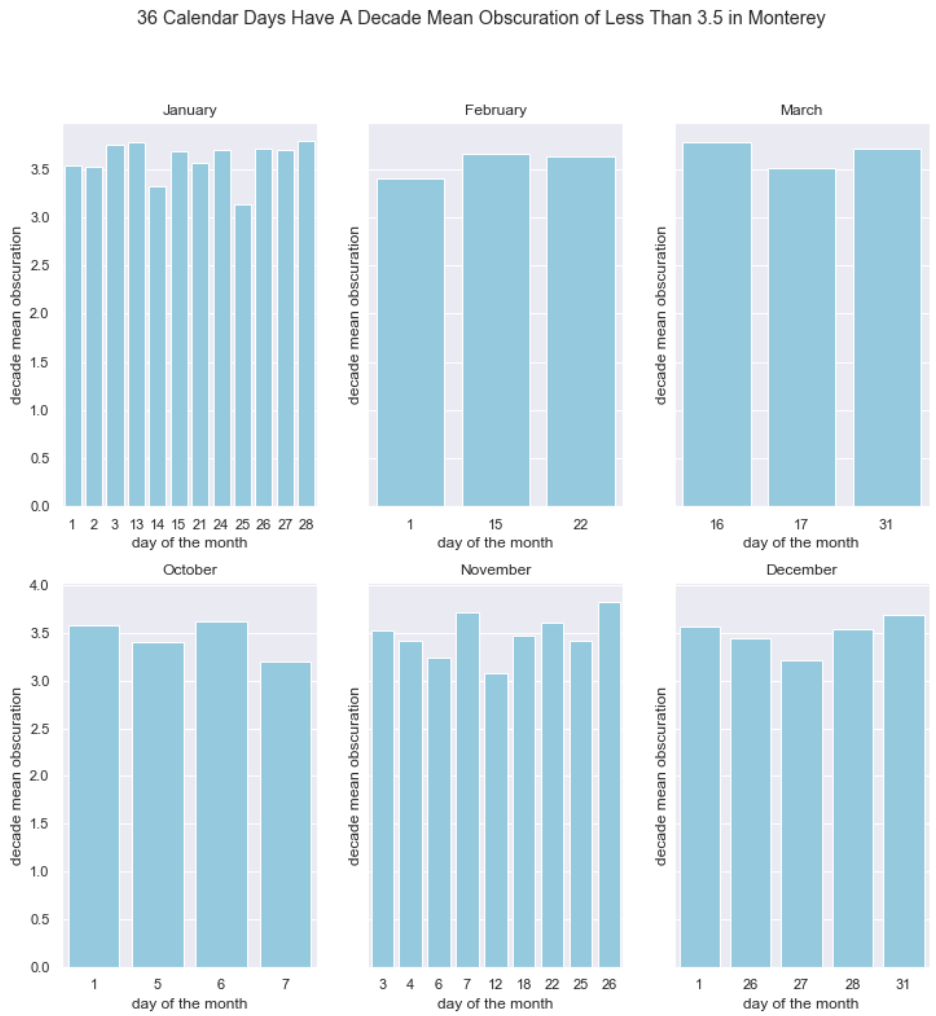
Exploration – Annual Rainfall

year	Days Without Rain
2010	265
2011	295
2012	303
2013	341
2014	303
2015	324
2016	295
2017	286
2018	312

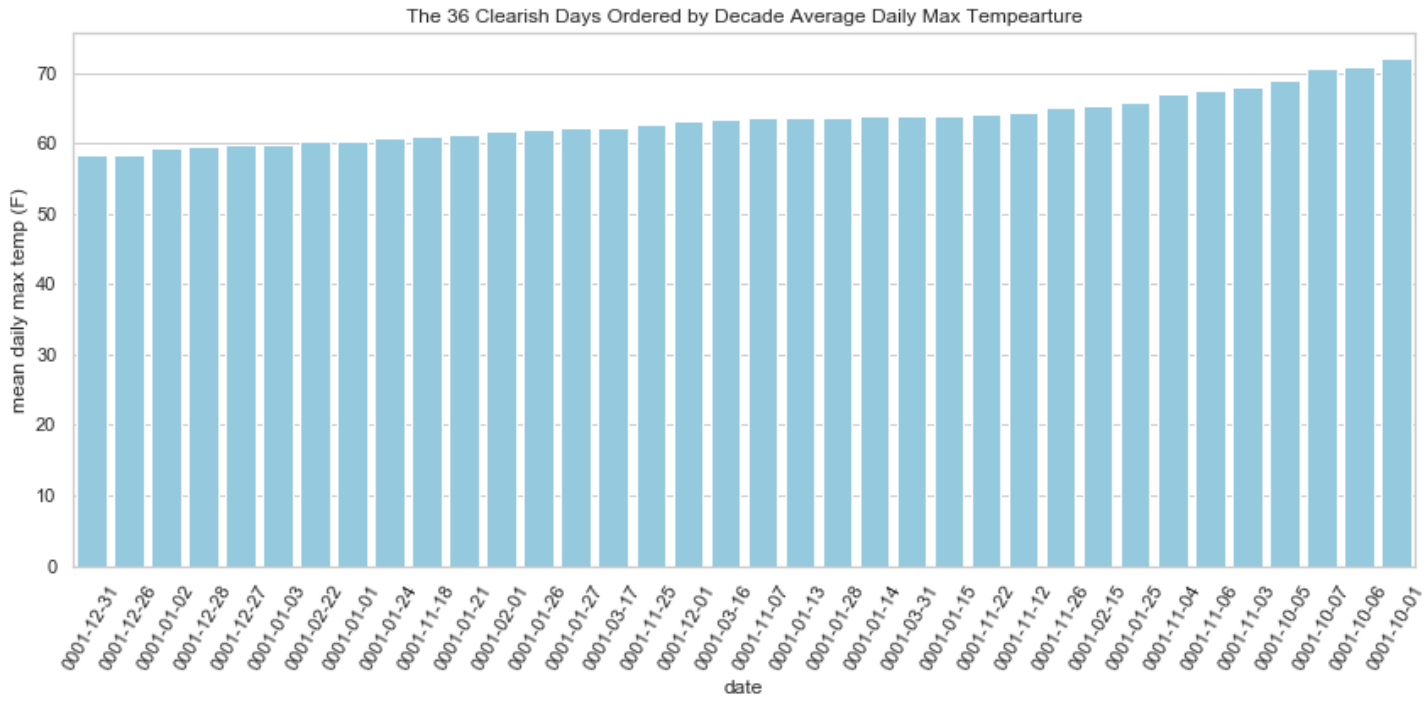
Exploration – Correlation Between Wind and Rain



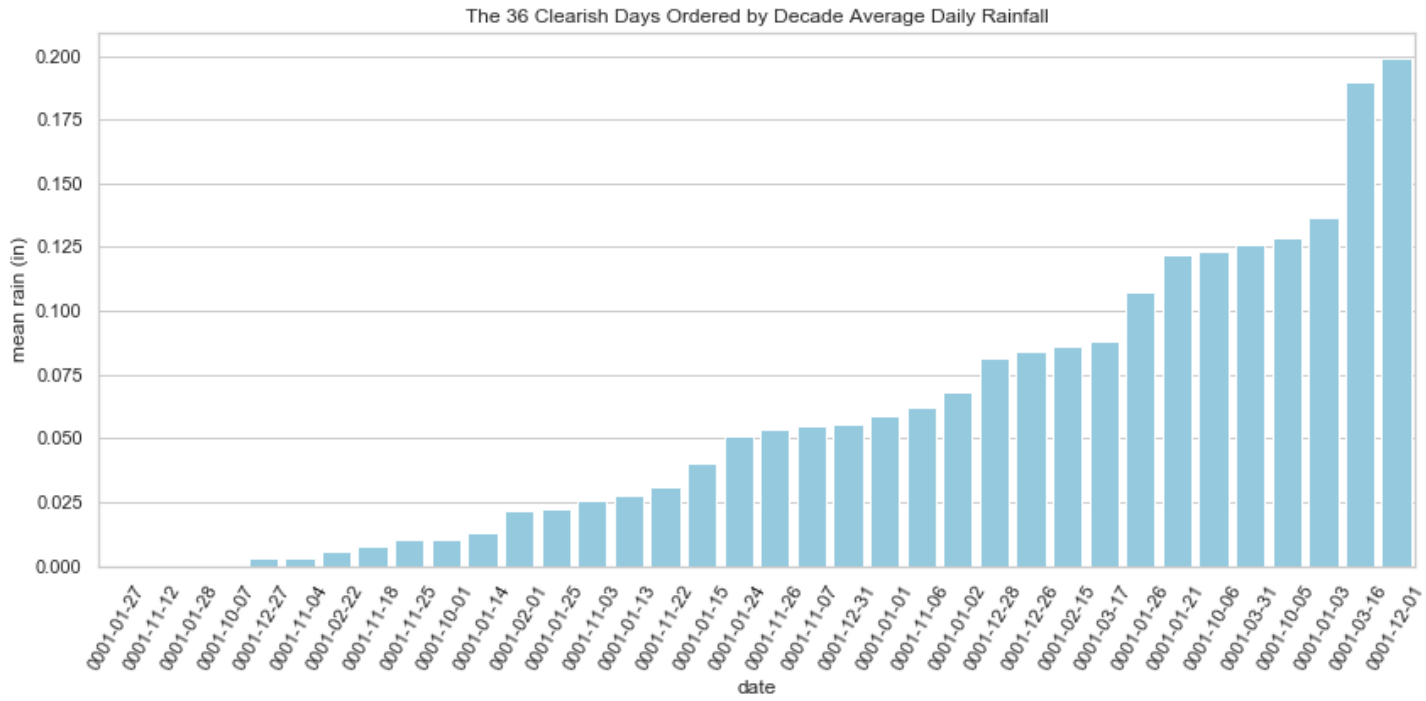
Exploration – Clearish Days



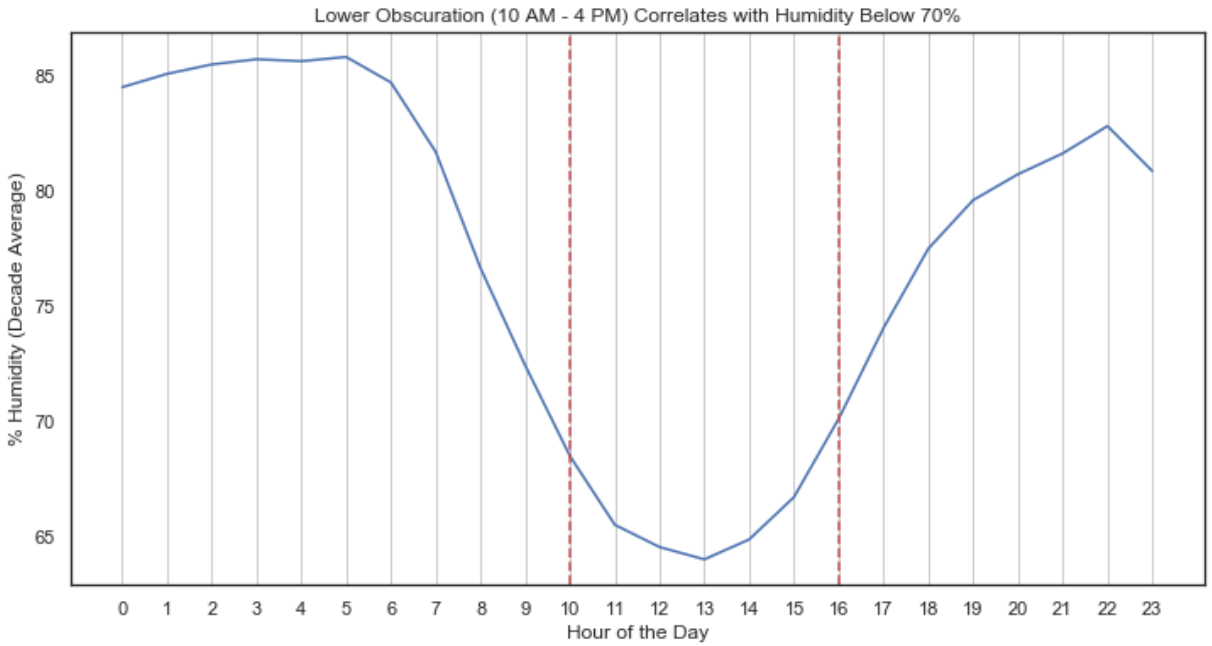
Exploration – Clearish Days by Temperature



Exploration – Clearish Days by Rainfall



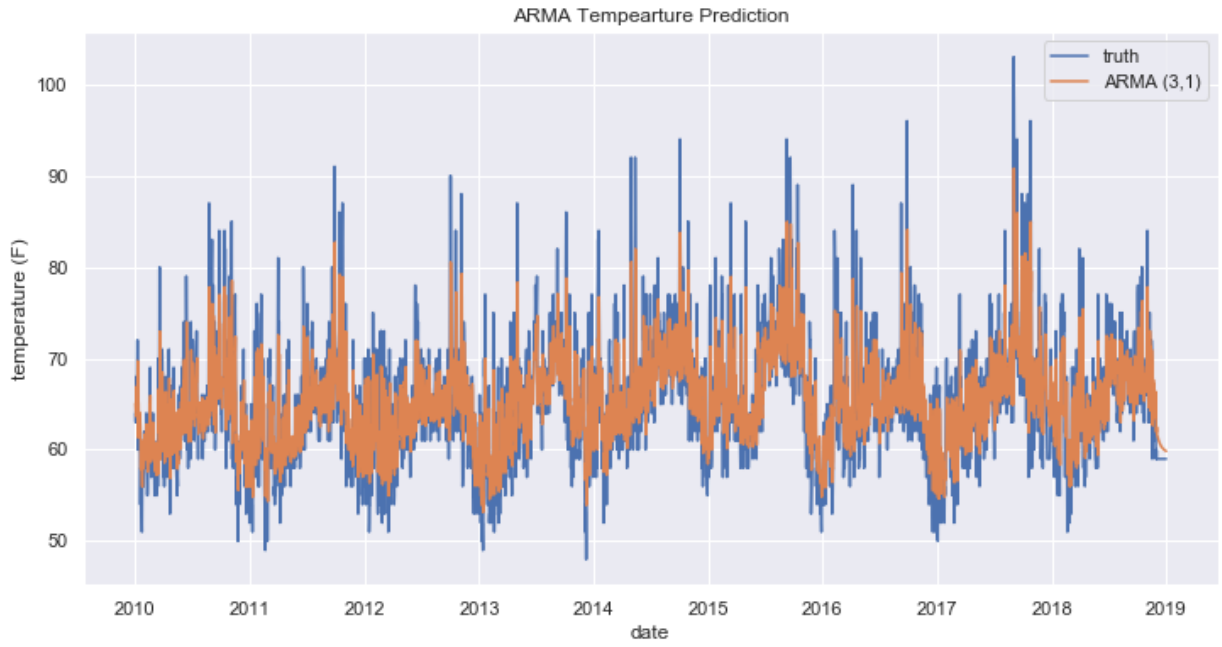
Exploration – Daily Obscuration Pattern



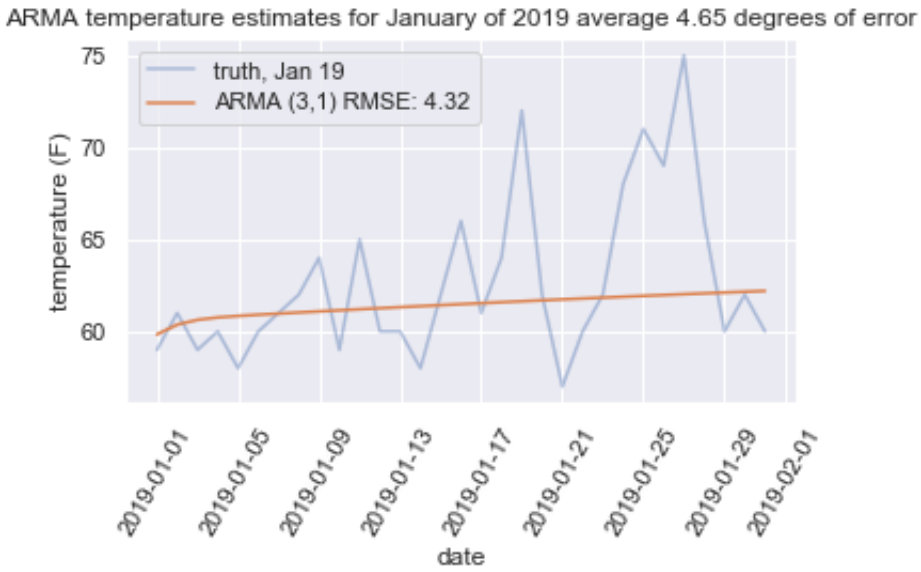
Modeling

- Compared ARMA, Holt-Winters, and GAN Models
- GAN model (fbprophet) used for prediction:
  - best performance
  - best interface (plots and parameters)

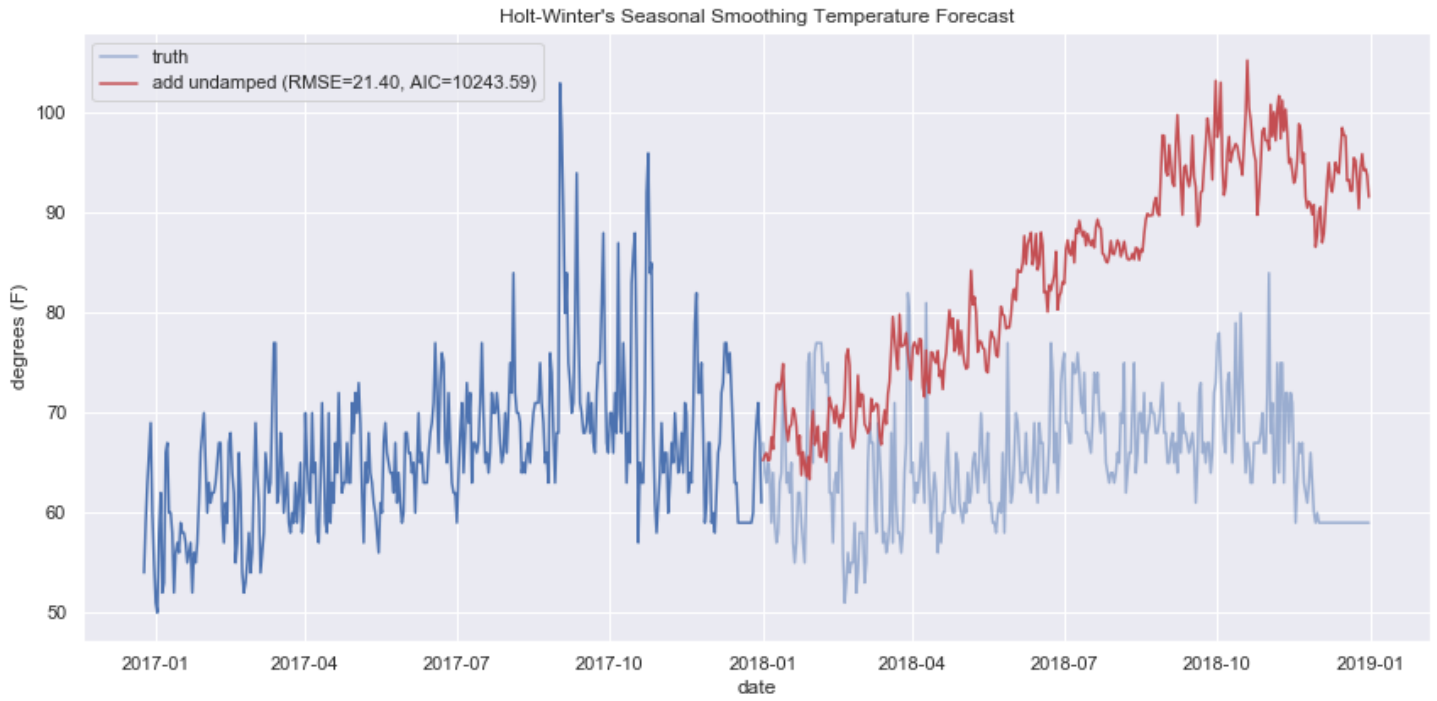
Modeling – ARMA



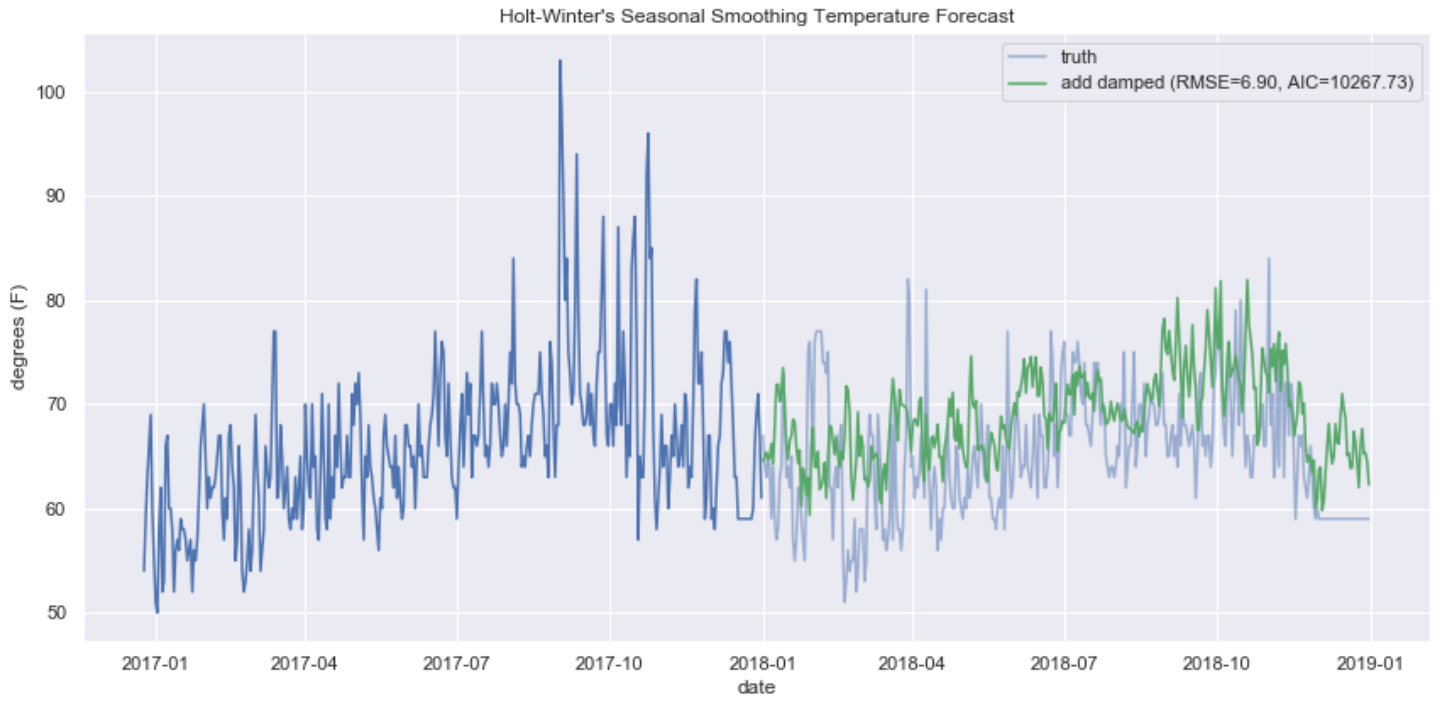
Modeling – ARMA



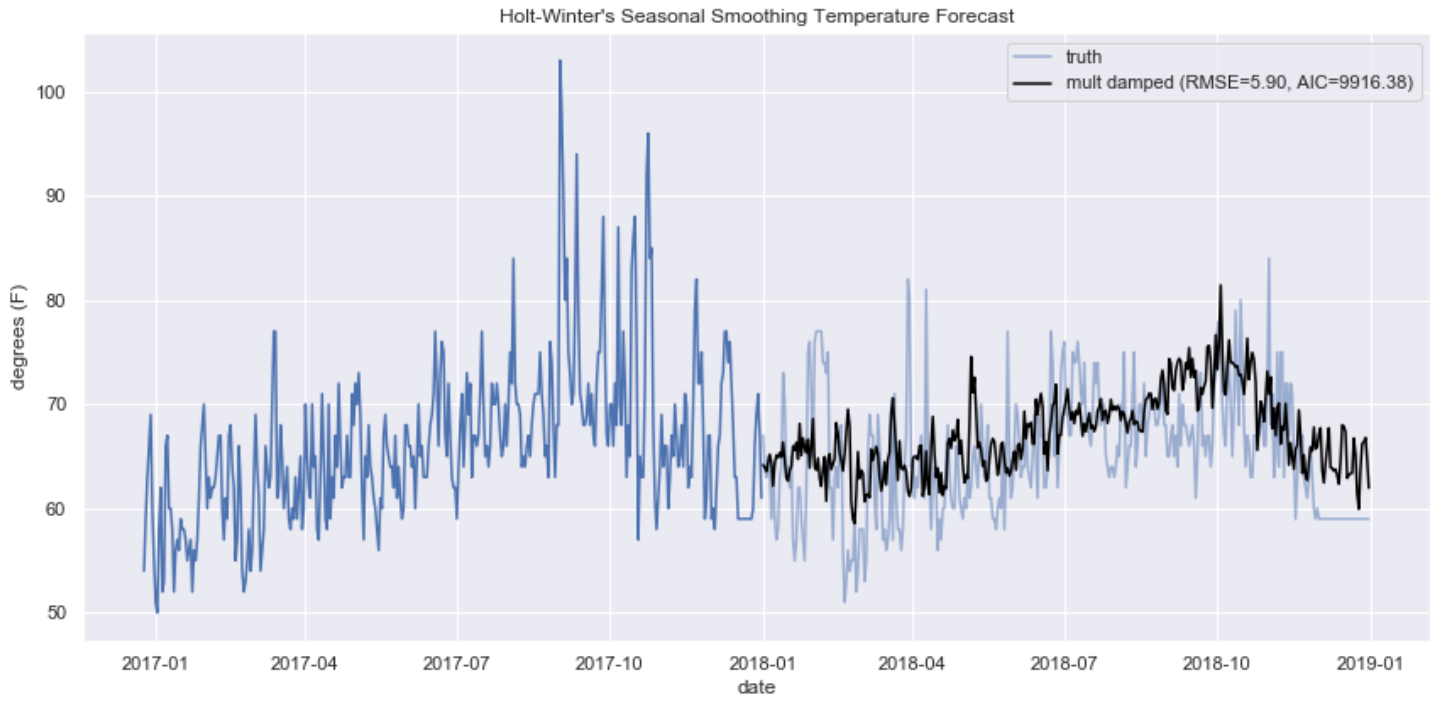
Modeling – Exponential Smoothing



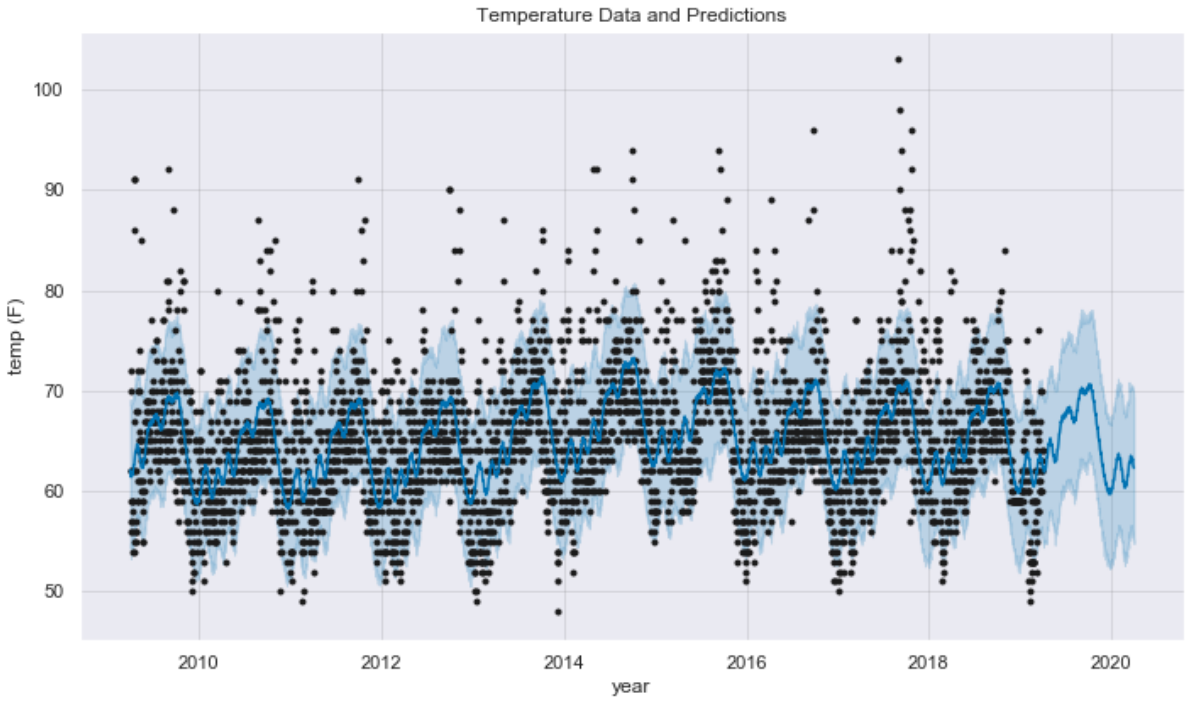
Modeling – Exponential Smoothing

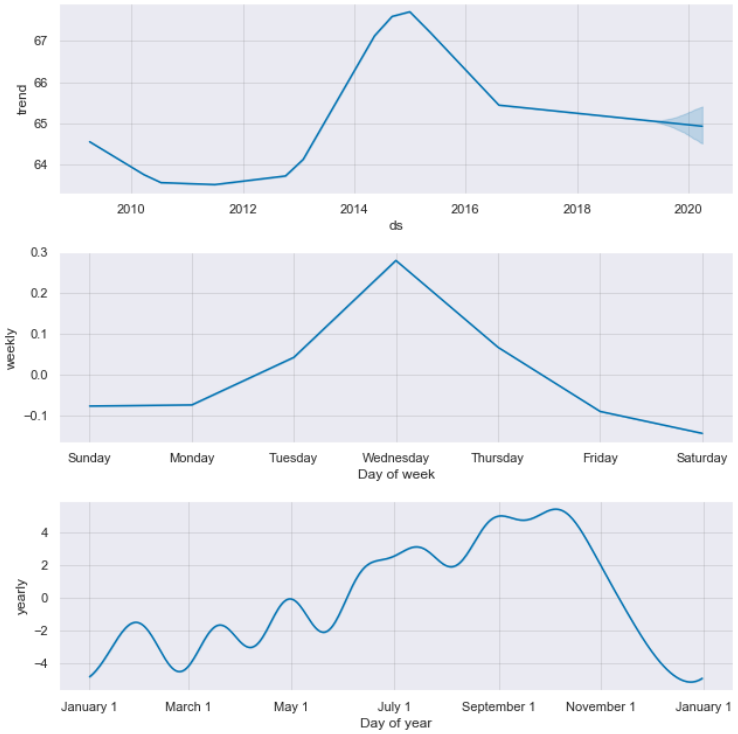


Modeling – Exponential Smoothing

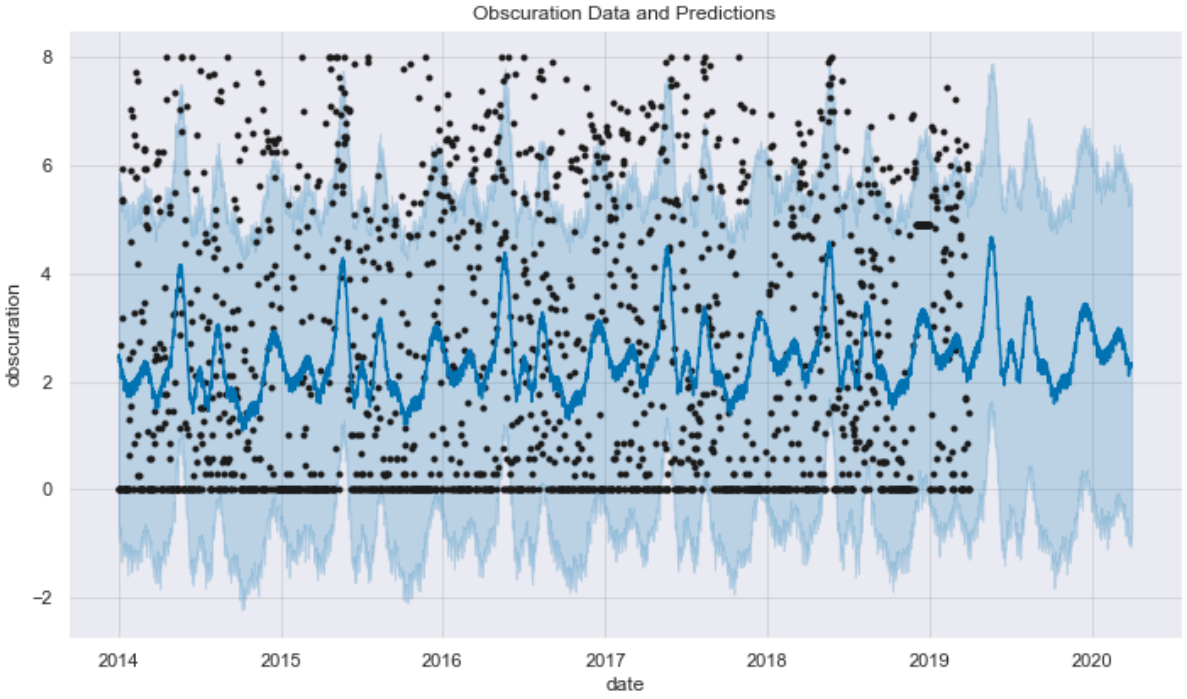


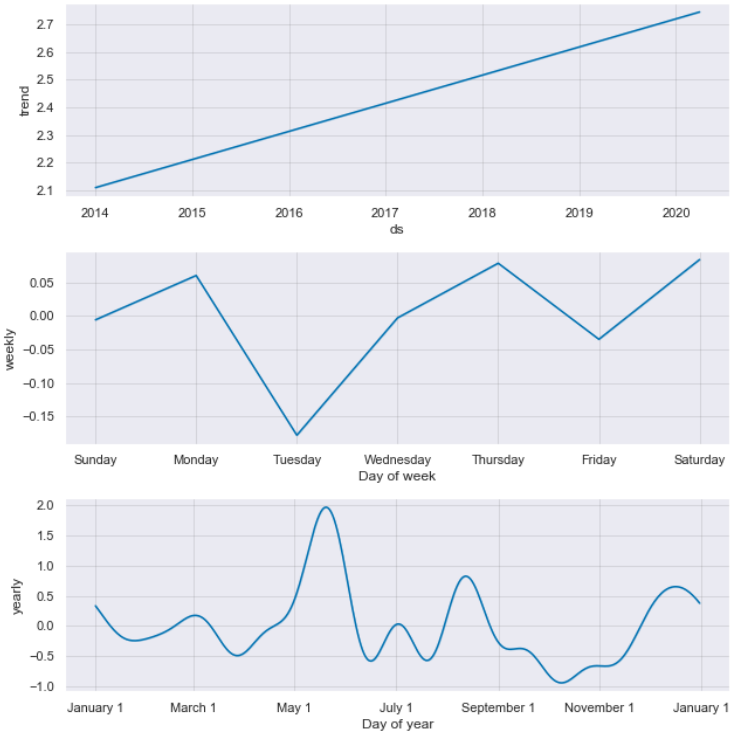
Modeling – Generalized Additive Model



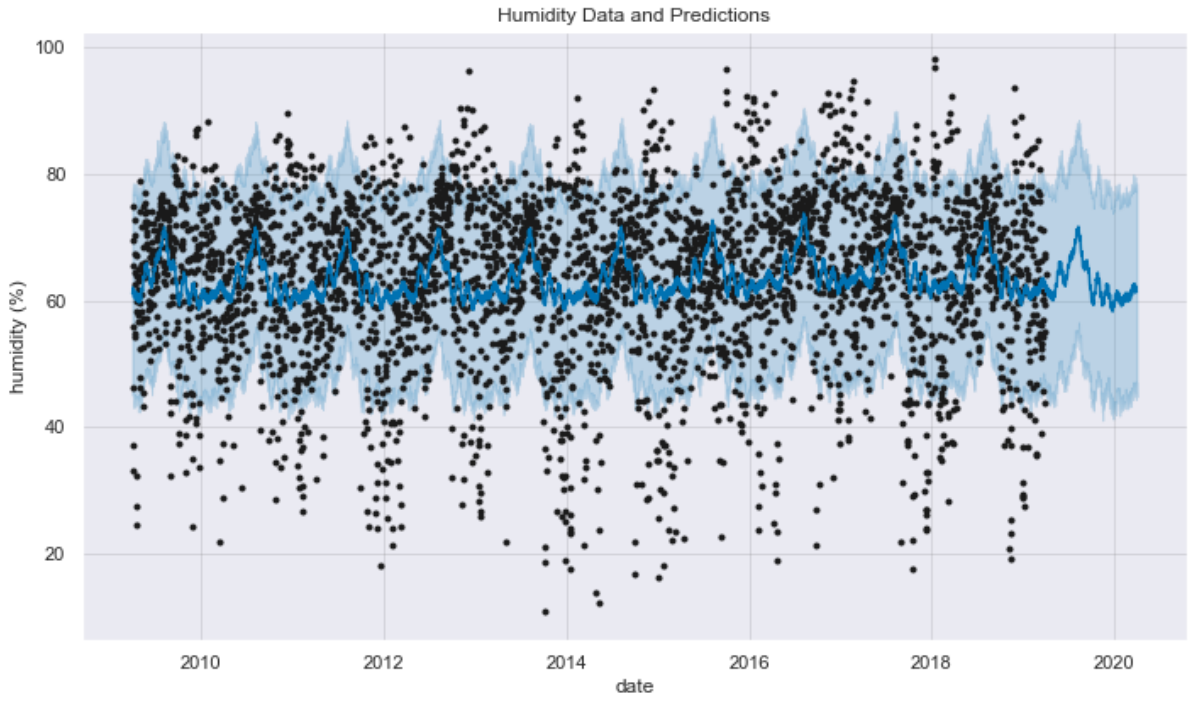


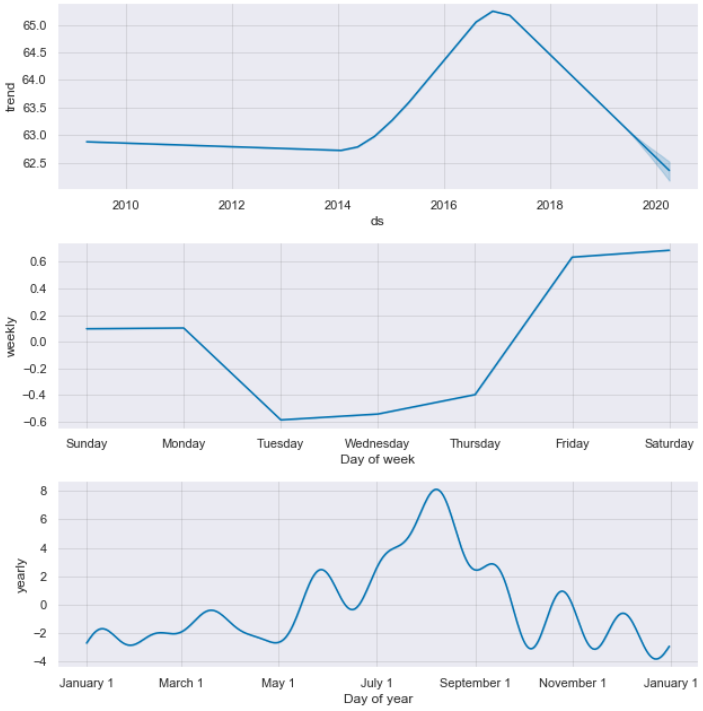
Modeling – Generalized Additive Model





Modeling – Generalized Additive Model





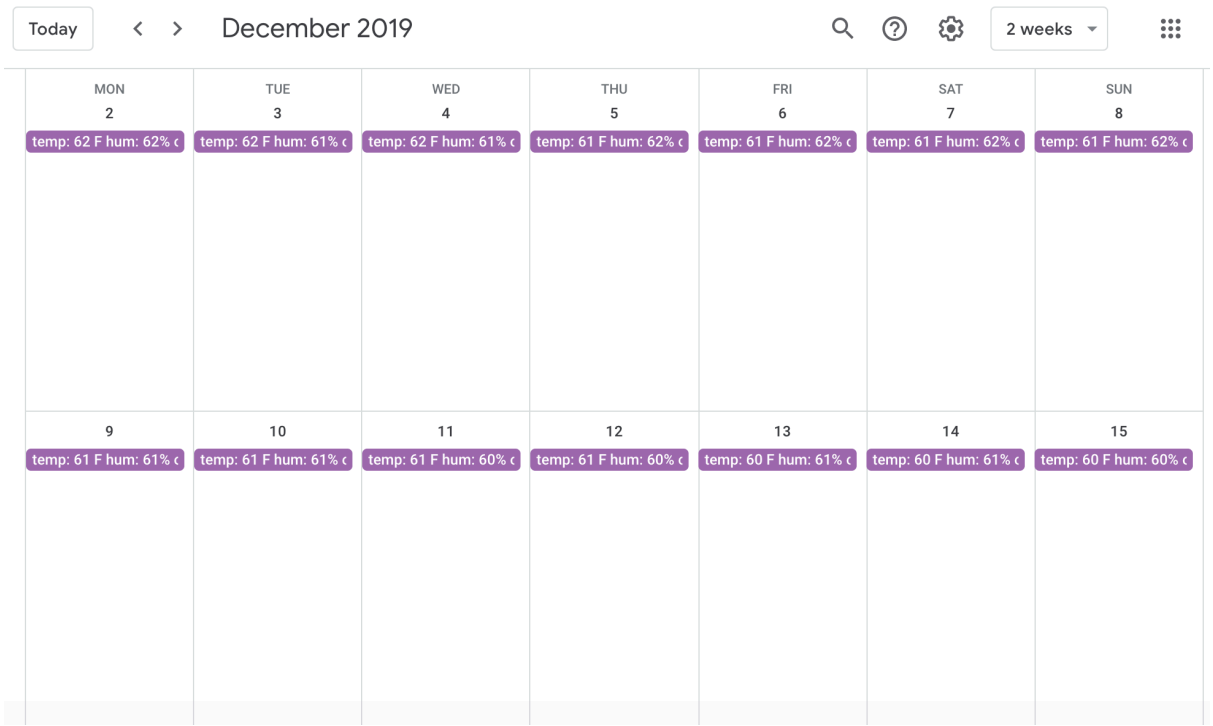
The Calendar

- Publicly Visible Calendar
- Python + Google Calendar API
- Displays prediction for remaining days in 2019 for:
  - Temperature
  - Obscuration
  - Humidity

The Calendar – Obscuration Risk Ratio

- Ratio of probabilities
  - Numerator: calendar day's average obscuration risk over decade
  - Denominator: the same for all other days in that month over decade

The Calendar



The Calendar