This reading describes the daily weather dataset used in the hands-on activities in this course.

The file **daily\_weather.csv** is a comma-separated file that contains weather data. This data comes from a weather station located in San Diego, California. The weather station is equipped with sensors that capture weather-related measurements such as air temperature, air pressure, and relative humidity. Data was collected for a period of three years, from September 2011 to September 2014, to ensure that sufficient data for different seasons and weather conditions is captured.

Sensor measurements from the weather station were captured at one-minute intervals. These measurements were then processed to generate values to describe daily weather. Since this dataset was created to classify low-humidity days vs. non-low-humidity days (that is, days with normal or high humidity), the variables included are weather measurements in the morning, with one measurement, namely relatively humidity, in the afternoon. The idea is to use the morning weather values to predict whether the day will be low-humidity or not based on the afternoon measurement of relatively humidity.

Each row in daily\_weather.csv captures weather data for a separate day. Each row, or sample, consists of the following variables:

|  |  |  |
| --- | --- | --- |
| **Variable** | **Description** | **Unit of Measure** |
| number | unique number for each row | NA |
| air\_pressure\_9am | air pressure averaged over a period from 8:55am to 9:04am | hectopascals |
| air\_temp\_9am | air temperature averaged over a period from 8:55am to 9:04am | degrees Fahrenheit |
| avg\_wind\_direction\_9am | wind direction averaged over a period from 8:55am to 9:04am | degrees, with 0 means coming from the North, and increasing clockwise |
| avg\_wind\_speed\_9am | wind speed averaged over a period from 8:55am to 9:04am | miles per hour |
| max\_wind\_direction\_9am | wind gust direction averaged over a period from 8:55am to 9:04am | degrees, with 0 being North and increasing clockwise |
| max\_wind\_speed\_9am | wind gust speed averaged over a period from 8:55am to 9:04am | miles per hour |
| rain\_accumulation\_9am | amount of rain accumulated in the 24 hours prior to 9am | millimeters |
| rain\_duration\_9am | amount of time rain was recorded in the 24 hours prior to 9am | seconds |
| relative\_humidity\_9am | relative humidity averaged over a period from 8:55am to 9:04am | percent |
| relative\_humidity\_3pm | relative humidity averaged over a period from 2:55pm to 3:04pm | percent |