

**Part I.** (60 points) Do all calculations in  $\text{L}^{\text{T}}\text{E}^{\text{X}}$  + R + knitr. For this assignment, all R code should be well commented and be visible (`echo=TRUE`) in the document where you have written it. Every time you create or modify an object, please show the results with the appropriate function. Please do not display complete objects when they are large.

**Weather data:** This is an exercise in manipulating data. We will be using daily weather data from the Albuquerque International Airport (KABQ).

(30<sup>pts</sup>)

### 1. Read and manipulate

- (10 pts) Read in the data directly from the KABQ weather website: <http://www.wunderground.com/>, history, KABQ, date range, etc.
- (5 pts) Create a date column from the MST column (should be in `POSIXct` format). Use `as.Date()` or a function from the `lubridate` package.
- (10 pts) Create (numeric) month and year columns from the MST column. Use `?sub` and try of the examples there; also, google “regular expression R” for more info. Or try package `stringr` and function `str_split()` and `rbind()`.
- (5 pts) Replace all occurrences of T with 0 in the `PrecipitationIn` column make convert to a numeric variable using `as.numeric()`.

(30<sup>pts</sup>)

### 2. Subset and plot

- (10 pts) Use `subset` to keep only the columns related to the dates (including those created above), and mean precipitation, temperature, and wind.
- (10 pts) Use the `melt()` function to transform the subsetted object from wide form into long form. What are the `id.vars` here?
- (10 pts) Create a plot of the measured variables over time using `geom_point()` and `facet_wrap(, scales = )`. Organize your facets (small multiples) so they are easily comparable.
- (0 pts) EXTRA +5: Add a local trend line with `geom_smooth()`, and explain what this does, and improve other aspects of your plot. You’ll see that the the default smoother doesn’t capture the patterns that well. Improve by clicking through the help: follow `geom_smooth()` to `stat_smooth()` to the `method` used, to `loess()` and the `span` parameter.