

## Ejercicio MapReduce/Hadoop

Considere los archivos de datos climáticos de NDCC disponibles en <https://www1.ncdc.noaa.gov/pub/data/noaa>.

El ejercicio propone implementar un algoritmo MapReduce en Hadoop que realice el procedimiento que se describe a continuación para la detección de días de calor:

- 1) Determine los registros *válidos* y *confiables* que indiquen que la temperatura del aire superó los 80F (temperatura cálida), los que superó los 90F (temperatura muy cálida) y los 100F (temperatura tórrida) y exprese su número total en valor absoluto y su porcentaje sobre el total de registros en el repositorio de datos.
- 2) Determine los años en los que la temperatura máxima superó los 115F (reportando el máximo de temperatura).
- 3) Determine las tres estaciones meteorológicas con más días tórridos en el siglo XX.
- 4) Cree en HBase una tabla con información de las estaciones meteorológicas disponible en <https://www1.ncdc.noaa.gov/pub/data/noaa/isd-history.txt>
- 5) Combine la información de ambas fuentes para dar como resultado una lista de registros con información de Uruguay.
- 6) Determine el mayor valor de temperatura reportado para Uruguay en el siglo XX.