Milestone overview

The first milestone consists in extracting meaningful information from the dataset. The methods to implement live in the **Extraction.scala** file.

You are given several .csv files containing two kinds of data:

- Weather station's locations (**stations.csv** file) ;

- Temperature records for a year (files **1975.csv**, **1976.csv**, etc.).

Your goal is to merge the data from these sources to get a series of information of the form date × location × temperature.

You can monitor your progress by submitting your work at any time during the development of this milestone. Your submission token and the list of your graded submissions is available on this page.

Data files

The data you will use comes from the National Center for Environmental Information of the United States.

Our files use the comma-separated values format: each line contains an information record, which is itself made of several columns.

Stations

The **stations.csv** file contains one row per weather station, with the following columns:

STN identifier	WBAN identifier	Latitude	Longitude

You might have noticed that there are two identifiers. Indeed, weather stations are uniquely identified by the compound key