

[Luogo di discussione](#)

[Descrizione dataset](#)

[Record di esempio per un solo inquinante](#)

[Schema dati](#)

[Elaborazioni da effettuare sui dataset originari](#)

[Elenco grezzo](#)

Luogo di discussione

E' stato creato un [repository github](#), dove tutti gli interessati potranno collaborare al progetto. Il luogo più semplice per discutere è quello delle "issue", che come vedrete non è nulla di diverso da un forum. Per partecipare l'unica cosa necessaria è crearsi un [account](#) su github.

Descrizione dataset

I dataset originari di output hanno queste caratteristiche:

- formato file -> CSV;
- encoding dei caratteri -> us-ascii;
- separatore di celle -> ";"
- separatore dei decimali -> ".";
- 6 righe di intestazione;
- il carattere " " per delimitare tutte le celle.

Qui un esempio:

https://raw.githubusercontent.com/SiciliaHub/arpasicilia_od/gh-pages/data/aria/exautoparco/exautoparco%2001092015.CSV

Record di esempio per un solo inquinante

Datetime	Aver	SWrd	MinT	Min	MaxT	Max	Sigma	Samp
05/09/2015 01:00	0.44376185	80400000	00:08	0.39558147	00:23	0.49235716	0.03427188	4

Schema dati

Campo	Descrizione	Unità di misura	Formato (data, intero, decimale, testo, ora, data-ora, ecc.)	Note
DateTime			data-ora	Questa intestazione in realtà non esiste nel dato originario. E' quindi da creare
Aver			decimale	il separatore è il “.”
SWrd				
MinT				
Min			decimale	il separatore è il “.”
MaxT				
Max			decimale	il separatore è il “.”
Sigma				
Samp				Probabilment e è un campo da rimuovere

Elaborazioni da effettuare sui dataset originari

Elenco grezzo

- avere soltanto una riga di intestazione;
- convertire il formato data da "05/09/2015 01:00" a "2015-09-05 01:00";
-