

Laboratorio 3. (2° giorno - mattina)

Obiettivi

- Fare esercizi sui dataframe usando il notebook Zeppelin
- In particolare utilizzare i metodi `select`, `selectExpr`, `filter`, `orderBy`, `where`, `distinct`, `count`.

Esercizi

Esercizio.1

Creare un notebook vuoto e caricare il file *earthquake.csv*. Visualizzare le prime righe del file in forma tabellare.

Esercizio.2

Quanti record ci sono in tutto?

In molti record l'intensità (richter) è zero. In quanti esattamente?

Discorso analogo per le città. Molti terremoti non hanno una città (ad es. sono avvenuti in mare). Quanti sono i terremoti con intensità maggiore di 3.0 e che hanno colpito qualche città?

Esercizio.3

Qual è la città dove si è verificato il terremoto più grave? Visualizzare sulla mappa i 10 terremoti più gravi.

Esercizio.4

Creare un dataframe derivato che contiene solo città, data, intensità e profondità. Visualizzare il dataframe in forma tabellare.

Esercizio.5

Creare un dataframe derivato che abbia la colonna della data di tipo `date`. Visualizzare il grafico che mostra l'andamento dell'intensità dei terremoti rispetto al tempo.