



UNIVERSITÀ
DEGLI STUDI
DI BERGAMO

Tecnologie Cloud & Mobile

HOMework

TEDx & AWS GLUE

mauro.pelucchi@gmail.com

Mauro Pelucchi

Homework

- Aggiungere al servizio “MyTEDx” la possibilità di navigare attraverso i talk suggeriti
- Descrivete come aggiungere e trattare il dataset `watch_next`
- Modificare il job pyspark per ricaricare il nostro DWH
- Come potremmo utilizzare questo dato?
 - Che tipo di funzionalità offrite ai vostri utenti?
- Implementate (opzionale) qualche funzionalità della vostra board (sempre dal punto di vista dati)

Homework

- Prepare una presentazione (max 5 slides) che descriva:
 - Il job pypspark che avete realizzato
 - I dati che avete trattato e lo schema finale che avete riempito
 - Criticità tecniche
 - Possibili evoluzioni

Modalità

- La presentazione (formato PDF) dovrà essere inviata via mail
 - Oppure, è possibile caricare anche l'elaborato su GitHub
- L'esercizio potrà essere svolto anche in gruppi da **MASSIMO** tre persone (in tal caso mandate una sola email indicando il vostro gruppo)
- Durante la prossima lezione presenterete i lavori più interessanti (1 o 2 presentazioni - 15 minuti a presentazione)
- NOTA BENE: da questa consegna sarà possibile partire per il terzo homework

