

Cognome e Nome _____

Matricola _____

Big data – Graph Mining

A.A. 2025/2026

Esonero del 18 dicembre 2025

Esercizio [8 punti]

Si consideri un gestore di telefonia mobile.

Siano date in input:

- Delle tuple ("**chiamata**", **id_sim_chiamante**, **id_sim_destinatario**, **durata_in_secondi**, **costo**, **tempo**) che rappresentano delle chiamate telefoniche (**tempo** rappresenta l'istante di inizio della chiamata).
- Delle tuple ("**ricarica**", **id_sim**, **importo**, **tempo**) che rappresentano le ricariche effettuate sulle varie sim (**tempo** rappresenta l'istante di esecuzione della ricarica).

Si assuma che costi e importi siano espressi in Euro, e che **id_sim** (e anche **id_sim_chiamante** e **id_sim_destinatario**) vari tra 0 e $n-1$. Si assuma altresì che ogni sim effettua e riceve un numero di chiamate costante, non dipendente quindi da n , e che ugualmente costante sia il numero di ricariche per ogni sim.

1. **[4 punti]** Si progettino (eventualmente utilizzando più step di computazione tra loro concatenati) le funzioni *Map* e *Reduce* per dare in output il **massimo valore del credito** associato a una sim (ottenuto sottraendo il totale dei costi delle chiamate effettuate al totale degli importi delle ricariche effettuate), utilizzando spazio locale $O(n^{1/3})$. Si discuta il *reducer size* di tutti gli step individuati.

Suggerimento: calcolare prima le coppie (*id_sim*, *credito*) per poi procedere al calcolo del massimo credito.

[Esercizio alternativo, valutazione 2 punti]: Si assuma di partire dalle coppie (*id_sim*, *credito*) date direttamente in input.

2. **[4 punti]** Si progettino (eventualmente utilizzando più step di computazione tra loro concatenati) le funzioni *Map* e *Reduce* per dare in output **uno degli id_sim che ha ricevuto chiamate per il maggior tempo** (maggiore somma dei valori di *durata_in_secondi* di tutte le chiamate ricevute), utilizzando spazio locale $O(n^{1/2})$. Si discuta il *reducer size* di tutti gli step individuati.

Suggerimento: calcolare prima le coppie (*id_sim*, *durata_chiamate_ricevute*) per poi procedere al calcolo dell'*id_sim* che ha il massimo valore di *durata_chiamate_ricevute*.

[Esercizio alternativo, valutazione 2 punti]: Si assuma di partire dalle coppie (*id_sim*, *durata_chiamate_ricevute*) date direttamente in input.

Regole per lo svolgimento della prova scritta:

- Per svolgere il compito si hanno a disposizione **60** minuti.
- Scrivere **subito** nome, cognome, matricola su OGNI FOGLIO (**compreso questo**).
- Durante la prova scritta **non** è possibile abbandonare l'aula.
- Non è ammesso **per nessun motivo** comunicare in qualsiasi modo con altre persone
- Non è possibile consultare appunti, libri e dispense.
- Qualsiasi strumento elettronico di calcolo o comunicazione (telefoni cellulari, calcolatrici, palmari, computer, etc...) deve essere **completamente disattivato** e **depositato in vista sulla cattedra**
- Mettere in vista sul banco un valido documento di identità.