Nome: _____ Cognome: _____ Matricola: _____

Laurea Magistrale in Ingegneria Informatica — **Corso di Big Data**

Tempo a disposizione: un'ora

Prima domanda

Si vuole realizzare un sistema di prenotazione on-line su scala mondiale di auto private simile ad Airbnb con il quale sia possibile conoscere disponibilità di auto in varie città del mondo e noleggiarle per un breve periodo. **Disegnare** e descrivere sinteticamente una possibile architettura per la **gestione e l'analisi di questi dati** con indicazione delle scelte tecnologiche fatte tenendo conto dei seguenti requisiti: (a) l'applicazione deve scalare su **centinaia di milioni** di record; (b) per le prenotazioni si vuole garantire **disponibilità e resistenza alle partizioni** (sistema AP) con un sistema **p2p "eventually consistent"**; (c) si vogliono effettuare **analisi off-line** dei dati per individuare fenomeni ricorrenti. La risposta deve essere contenuta nello spazio a disposizione.

Seconda domanda

Considerando il seguente possibile tracciato di record per un noleggio relativo all'applicazione descritta nella prima domanda:

RENTALS [RentalId: i, NameOfUser: n, CarId: c, ModelOfCar: m, StartDate: d1, EndDate, d2, Price: p]

Scrivere: (a) un job **MapReduce** descritto in pseudo-codice (qualunque) in grado di calcolare la media dei prezzi di noleggio per ciascun modello di auto e (b) una query **Hive** in grado di calcolare il minimo e il massimo prezzo di noleggio per ciascun modello di auto. La risposta deve essere contenuta nello spazio a disposizione.