

**Architetture Dati**  
**Prof. Andrea Maurino**  
**Corso di Laurea Magistrale in Informatica**  
**Anno Accademico 2020/2021**  
**2° Appello 10 Febbraio 2021**

**Tempo massimo 45 minuti**

**Domanda 1**

Come l'architettura lambda può gestire l'acquisizione e analisi dei dati veloci.

**Domanda 3**

Cosa succede durante il two phase commit se il transaction manager muore prima della global decision

**Esercizio 1**

Diversi dataset rappresentano le informazioni relative agli alberi presenti nella città di New York, nei diversi censimenti fatti nel corso dell'anno 2015. Dopo una modellazione seguendo la strategia per embedded la struttura è data come in seguito:

- **\_id** : ovvero l'identificativo interno per ogni documento;
- **tree\_id** : identificativo del csv originale;
- **block\_id** : secondo identificativo del csv originale;
- **status**: indica lo stato di vita. I valori possono essere due: Alive o Dead;
- **spc\_latin** : indica il nome in latino della specie;
- **spc\_common** : indica il nome di uso comune dell'albero;
- **address** : indirizzo in cui si trova l'albero;
- **zipcode** : il "cap" italiano;
- **zip\_city** : la città a cui fa riferimento il cap;
- **borocode** : codice relativo al distretto in cui si trova;
- **boroname** : nome relativo al distretto in cui si trova;
- **cncldist, st\_assem, st\_senate** : informazioni burocratiche relative al distretto;
- **nta** : codice relativo al quartiere;
- **nta\_name** : nome del quartiere;
- **boro\_ct** : informazione relativa al borgo di appartenenza;
- **state** : informazione relativa allo stato di appartenenza (New York);
- **latitude, longitude** : dati numerici che riguardano la latitudine e la longitudine dell'albero;
- **spc\_info** : sotto-documento Porta informazioni relative all'albero:
  - **form** : indica la forma dell'albero;
  - **growth\_rate** : indicatore di crescita della specie;
  - **fall\_color** : colore delle foglie nel periodo in cui cadono dai rami;

Si richiede di eseguire le seguenti interrogazioni:

1. Ricerca di tutti gli alberi con status=dead e proiettando status, indirizzo, boroname, latitude e longitude ma non l'id;
2. Ricerca nome latino degli alberi , il distretto, lo zip\_city e indirizzo di alberi vivi per un lo zipcode 11204;
3. Ricerca di tutti gli alberi con colore delle foglie che cadono "Yellow", forma "Pyramidal" or "Upright" e percentuale di crescita "Slow";
4. Ricerca di tutti gli alberi con latitudine compresa tra due valori: