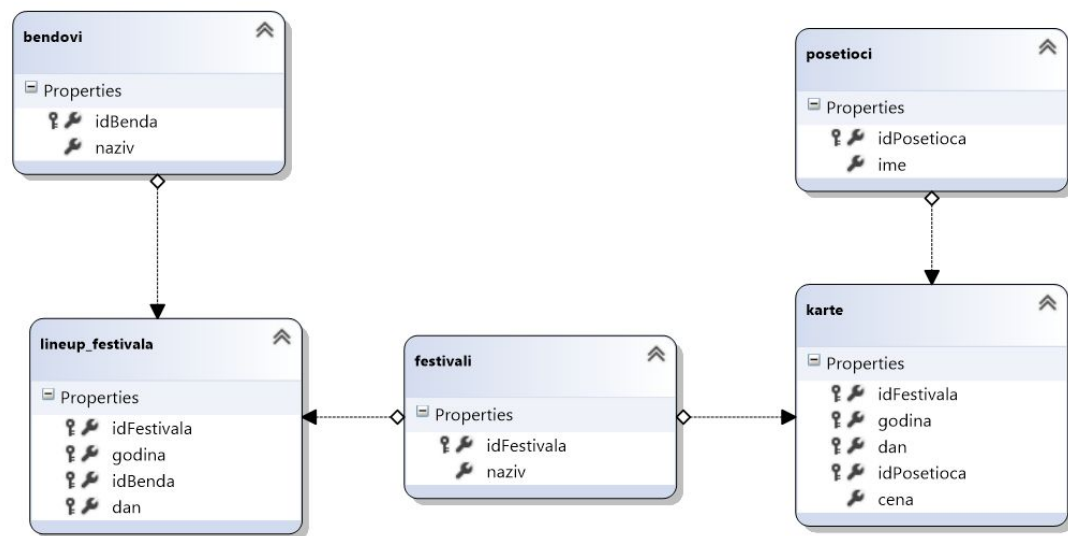


Dati su sledeći CSV fajlovi:

- **bendovi.csv** (*idBenda, naziv*) - Spisak svih bendova.
- **festivali.csv** (*idFestivala, naziv*) - Spisak svih festivala.
- **lineup_festivala.csv** (*idFestivala, godina, idBenda, dan*) - Spisak bendova koji su učestvovali na festivalima, po godinama i danima (redni broj dana festivala).
- **pocetioci.csv** (*idPosetioca, ime*) - Spisak svih posetilaca.
- **karte.csv** (*idFestivala, godina, dan, idPosetioca, cena*) - Spisak svih kupljenih karata za određene dane i godine festivala.



Parametarski posao (9 poena)

- Odrediti festivale koji svakim novim održavanjem ostvaruju veći profit u odnosu na prethodno održavanje (jedan festival se može održati najviše jednom godišnje). Potrebno je prikazati nazive festivala.

Rešenje dati u vidu parametarskog posla gde je potrebno kreirati:

1. Bash skriptu za pokretanje parametarskog posla (*pokreni.sh*), kao i integrisanje rezultata u fajl *resenje.txt*.
2. Odgovarajuće *PBS* komandne fajlove.
3. Programske kodove za obradu podataka i kreiranje rešenja.

Apache Spark

- **(3 poena)** Odrediti pesetioce koji su potrošili najviše novca na kupovinu karata.
- **(6 poena)** Odrediti festivale koji svakim novim održavanjem ostvaruju veći profit u odnosu na prethodno održavanje (jedan festival se može održati najviše jednom godišnje). Potrebno je prikazati nazive festivala.

Rešenje dati u vidu jednog *Python* fajla pod nazivom *festivali.py*.

Vreme izrade kolokvijuma je 150 minuta.