

PRAKTIKUM IZ PROGRAMIRANJA 1

- domaći zadatak broj 5 -

Sastaviti program na programskom jeziku Python, kojim se vrši određena vrsta obrade nad tekstualnim datotekama. Glavni program treba da:

- učitava nazive jedne ili više datoteka sa kojima će se raditi,
- poziva odgovarajuće funkcije za čitanje datoteke i obradu,
- ispisuje sve dobijene rezultate na način tražen postavkom zadatka;

Zavisno od rednog broja problema, svaki student treba da sastavi jedan od programa koji su dati u prilogu ovog dokumenta. Programi se bave obradom podataka o filmovima.

Program treba da čita podatke uz vođenje računa o tipu podataka koji se čita. Svi programi čitaju tekstualnu datoteku u **Comma Separated Values (.CSV)** formatu koja sadrži podatke o filmovima u određenom periodu. Polja datoteke su razdvojena znakom tačka-zarez (;). Datoteka ima zaglavlje sa nazivima polja koje treba ignorisati prilikom obrade. Svaki red datoteke sadrži sledeće podatke o filmu: jedinstveni celobrojni identifikator filma (**id**), string koji sadrži naslov filma (**Movie Title**), string koji sadrži žanrove kojima film pripada odvojene znakom | (**Movie Genre**), datum izdavanja u formatu **dd.mm.yyyy**, gde **dd** predstavlja dan, **mm** mesec, a **yyyy** godinu izdavanja filma, string koji sadrži puno ime režisera (**Director**) i string koji sadrži podatke o zaradi filma u dolarima (**Revenue**), gde se znak \$ nalazi na početku stringa, a sledi ga realan broj. Kompletan primer ulaza je dat u datoteci **pp1_movies_2020.csv** u prilogu ovog dokumenta. Datoteka se može otvoriti za pregled u bilo kom editoru teksta ili pomoću *Microsoft Excel* alata. Za svaki problem su definisani i posebni, kraći ulazni primeri.

Voditi računa o dekompoziciji programa na potprograme prema gore navedenoj raspodeli i zahtevima pojedinačnih zadataka. Po potrebi, dozvoljeno je uvoditi i dodatne potprograme u rešenje. Student sam treba da definiše imena funkcija, kao i argumente potrebne za njihov rad. Voditi računa da se funkcijama dostave samo neophodni podaci. Funkcije ne smeju pristupati promenljivama glavnog programa direktno, već samo putem svojih argumenata.

Prilikom rada programa proveravati i obraditi moguće greške korišćenjem mehanizma izuzetaka. U slučaju greške pri radu sa datotekom, na standardnom izlazu ispisati poruku **DAT_GRESKA**. U slučaju greške prilikom konverzije podataka, na standardnom izlazu ispisati poruku **KONV_GRESKA**. U slučaju greške zbog nekorektno sadržaja nekog od polja (npr. polje u obliku datuma), na standardnom izlazu ispisati poruku **POLJE_GRESKA**. Zatim, korektno prekinuti izvršavanje programa.

Ako nešto u postavci zadatka nije dovoljno precizno definisano ili ako su neki od zahteva međusobno suprotstavljeni, usvojiti razumnu pretpostavku i rešiti zadatak korišćenjem te pretpostavke. Na samoj odbrani obavestiti demonstratora o usvojenoj pretpostavci ili pretpostavkama. Programski kod rešenja zadatka treba da bude uredno komentaran, tako da pri pregledu programa lako može biti uočeno šta radi bilo koja programska celina.

Radi boljeg testiranja programa, odabрати nekoliko dodatnih skupova podataka sa kojima će program biti testiran. Svaki primer treba da sadrži ulazne podatke i očekivani izlaz za te podatke.

Napomene:

1. Rok za predaju petog domaćeg zadatka je **utorak, 29.12.2020.** putem kursa predmeta na **Moodle** platformi za elektronsko učenje. Tačan termin za predaju će biti naknadno definisan za sve studente. Termin će biti ograničenog vremenskog trajanja.
2. Domaći zadaci će biti testirani i ocenjivani korišćenjem javnih i tajnih testova, kao i demonstracijom funkcionalnosti na odbrani.
3. Studentima će nekoliko dana pre roka za predaju biti dostupno okruženje za testiranje rešenja domaćeg zadatka na Moodle platformi za elektronsko učenje korišćenjem javnih testova.
4. Prilikom predaje domaćeg zadatka studenti će rešavati i kratak test znanja u vezi rešenja domaćeg zadatka i relevantnog gradiva iz oblasti programiranja na programskom jeziku Python koje obuhvata temu domaćeg zadatka.
5. Odbrana petog domaćeg zadatka će biti održana u **sredu, 30.12.2020.** Studenti mogu da se prijave za jedan od ponuđenih termina odbrane na Moodle platformi za elektronsko učenje.
6. Domaći zadaci se rešavaju **samostalno**. Predmetni nastavnici zadržavaju pravo da nakon predaje domaćih zadataka izvrše proveru sličnosti i preuzmu odgovarajuće disciplinske mere.
7. Svi drugi detalji oko predaje, ocenjivanja i odbrane domaćeg zadatka će biti blagovremeno objavljeni.
8. Formula za redni broj problema **i** koji treba rešavati je sledeća (R – redni broj indeksa, G – poslednje dve cifre godine upisa):

$$i = (R + G) \bmod 4$$

9. Kao rešenje domaćeg zadatka potrebno je predati sadržaj sledeće datoteke:

- **dz5.py**, koja sadrži izvorni tekst programa na programskom jeziku Python;

21.12.2020. godine

sa predmeta

0. Napisati program koji vrši određenu obradu nad tekstualnom datotekom u **Comma Separated Values (.CSV)** koja sadrži podatke o filmovima u određenom periodu prema prethodno opisanom formatu. Program treba da formira izlaznu datoteku koja će za svakog režisera odrediti do tri filma sa najvećom zaradom, kao i ukupnu zaradu u zadatom periodu. Jedan red izlazne datoteke treba da sadrži sledeća polja: string koji sadrži puno ime režisera (**Director**), listu sa naslovima do tri filma sa najvećom zaradom u zadatom vremenskom periodu, uređenu opadajuće po zaradi i razdvojenu znakom | (**Top3 Movies**) i ukupnu zaradu filmova u dolarima po istom formatu kao što je ulazni (**Top3 Movies Revenue**). Izlazna datoteka treba da bude u **csv** formatu. U izlaznu datoteku na početku upisati zaglavlje. U izlaznu datoteku ne upisivati podatke o režiserima koji nisu režirali ni jedan film u zadatom periodu. Izlaznu datoteku urediti nerastuće prema ukupnoj zaradi do tri filma sa najvećom zaradom jednog režisera u zadatom vremenskom periodu. Podatak o zaradi filma zaokružiti na dve decimale.

Program treba da:

- 1) Učita imena ulazne i izlazne datoteke sa standardnog ulaza.
- 2) Učita dva datuma u formatu **dd.mm.yyyy** koji predstavljaju vremenski period za koji se vrši obrada. Svaki datum se unosi u posebnoj liniji teksta.
- 3) Izvrši zahtevanu obradu prema tekstu zadatka.
- 4) Formira izlaznu datoteku prema tekstu zadatka.
- 5) Vodi računa i obradi moguće izuzetke koji mogu nastati prilikom rada programa.

Primeri

Primer 1	
Standardni ulaz	ulaz0.csv izlaz0_1.csv 01.01.2000 18.12.2020
Ulazna datoteka:	ulaz0.csv
Izlazna datoteka:	izlaz0_1.csv
Standardni izlaz:	
Primer 2	
Standardni ulaz	pp1_movies_2020.csv izlaz0_2.csv 01.01.2000 18.12.2020
Ulazna datoteka:	pp1_movies_2020.csv
Izlazna datoteka:	izlaz0_2.csv
Standardni izlaz:	
Primer 3	
Standardni ulaz	nepostoji.csv izlaz.csv 01.01.2000 18.12.2020
Ulazna datoteka:	-
Izlazna datoteka:	-
Standardni izlaz:	DAT_GRESKA

1. Napisati program koji vrši određenu obradu nad tekstualnom datotekom u **Comma Separated Values (.CSV)** koja sadrži podatke o filmovima u određenom periodu prema prethodno opisanom formatu. Program treba da formira izlaznu datoteku koja će sadržati podatke o najmanje popularnom filmskom žanru u svakoj godini zadanog perioda. Žanr se može smatrati najmanje popularnim ukoliko donosi najmanju zaradu posmatrane godine. Ukoliko ima više takvih, razdvojiti ih znakom |. Datoteka treba da bude uređena rastuće po godinama. Jedan red izlazne datoteke treba da sadrži sledeća polja: posmatranu godinu (**Year**), string koji sadrži ime žanra (**Movie Genre**) i ukupnu zaradu filmova najmanje popularnog žanra u dolarima po istom formatu kao što je ulazni (**Genre Revenue**). Izlazna datoteka treba da bude u **csv** formatu. U izlaznu datoteku na početku upisati zaglavlje. Ukoliko neke godine posmatranog perioda nisu izdavani filmovi, izostaviti je iz izlazne datoteke. Podatak o zaradi filma zaokružiti na dve decimale.

Program treba da:

- 1) Učita imena ulazne i izlazne datoteke sa standardnog ulaza.
- 2) Učita dva godine u formatu **yyyy** koje predstavljaju vremenski period za koji se vrši obrada. Svaka godina se unosi u posebnoj liniji teksta.
- 3) Izvrši zahtevanu obradu prema tekstu zadatka.
- 4) Formira izlaznu datoteku prema tekstu zadatka.
- 5) Vodi računa i obradi moguće izuzetke koji mogu nastati prilikom rada programa.

Primeri

Primer 1	
Standardni ulaz	ulaz1.csv izlaz1_1.csv 2000 2020
Ulazna datoteka:	ulaz1.csv
Izlazna datoteka:	izlaz1_1.csv
Standardni izlaz:	
Primer 2	
Standardni ulaz	pp1_movies_2020.csv izlaz1_2.csv 2000 2020
Ulazna datoteka:	pp1_movies_2020.csv
Izlazna datoteka:	izlaz1_2.csv
Standardni izlaz:	
Primer 3	
Standardni ulaz	nepostoji.csv izlaz.csv 2000 2020
Ulazna datoteka:	-
Izlazna datoteka:	-
Standardni izlaz:	DAT_GRESKA

2. Napisati program koji vrši određenu obradu nad tekstualnom datotekom u **Comma Separated Values (.CSV)** koja sadrži podatke o filmovima u određenom periodu prema prethodno opisanom formatu. Program treba da formira izlaznu datoteku koja će za svakog režisera sadržati podatke o broju filmova odgovarajućih žanrova koje je režirao. Žanrovi od interesa se unose iz jedne linije sa standardnog ulaza odvojeni zarezom (,). Jedan red izlazne datoteke treba da sadrži sledeća polja: string koji sadrži puno ime režisera (**Director**), listu odgovarajućih žanrova i broja režiranih filmova po formatu **žanr : broj filmova**, odvojenih znakom | i uređenih leksikografski rastuće prema nazivu žanra (**Movie Genre Count**). Izlazna datoteka treba da bude u csv formatu. U izlaznu datoteku na početku upisati zaglavlje. Poredak režisara u izlaznoj datoteci treba da bude isti kao u ulaznoj datoteci.

Program treba da:

- 1) Učita imena ulazne i izlazne datoteke sa standardnog ulaza.
- 2) Učita string koji sadrže žanrove za obradu razdvojene zarezima.
- 3) Izvrši zahtevanu obradu prema tekstu zadatka.
- 4) Formira izlaznu datoteku prema tekstu zadatka.
- 5) Vodi računa i obradi moguće izuzetke koji mogu nastati prilikom rada programa.

Primeri

Primer 1	
Standardni ulaz	ulaz2.csv izlaz2_1.csv Romance,Drama,Mystery
Ulazna datoteka:	ulaz2.csv
Izlazna datoteka:	izlaz2_1.csv
Standardni izlaz:	
Primer 2	
Standardni ulaz	pp1_movies_2020.csv izlaz2_2.csv Romance,Drama,Mystery
Ulazna datoteka:	pp1_movies_2020.csv
Izlazna datoteka:	izlaz2_2.csv
Standardni izlaz:	
Primer 3	
Standardni ulaz	nepostoji.csv izlaz.csv Romance,Drama,Mystery
Ulazna datoteka:	-
Izlazna datoteka:	-
Standardni izlaz:	DAT_GRESKA

3. Napisati program koji vrši određenu obradu nad tekstualnom datotekom u **Comma Separated Values (.CSV)** koja sadrži podatke o filmovima u određenom periodu prema prethodno opisanom formatu. Program treba da formira izlaznu datoteku koja će za svakog reditelja sadržati podatke o filmovima odgovarajućeg žanra sa najmanjom i najvećom zaradom koje je režirao u zadatom periodu. Žanr od interesa se unosi sa standardnog ulaza. Jedan red izlazne datoteke treba da sadrži sledeća polja: string koji sadrži puno ime režisera (**Director**), naslov i zaradu filma sa najmanjom zaradom po formatu **naslov filma : zarada (Movie Title Min Revenue)**, naslov i zaradu filma sa najvećom zaradom po formatu **naslov filma : zarada (Movie Title Max Revenue)**. Smatrati da ne postoje dva filma sa istom zaradom. Izlazna datoteka treba da bude u **csv** formatu. U izlaznu datoteku na početku upisati zaglavlje. Izlazna datoteka treba da bude sortirana leksikografski rastuće po punom imenu režisera.

Program treba da:

- 1) Učita imena ulazne i izlazne datoteke sa standardnog ulaza.
- 2) Učita dva datuma u formatu **dd.mm.yyyy** koji predstavljaju vremenski period za koji se vrši obrada. Svaki datum se unosi u posebnoj liniji teksta.
- 3) Učita string koji sadrži žanr za obradu.
- 4) Izvrši zahtevanu obradu prema tekstu zadatka.
- 5) Formira izlaznu datoteku prema tekstu zadatka.
- 6) Vodi računa i obradi moguće izuzetke koji mogu nastati prilikom rada programa.

Primeri

Primer 1	
Standardni ulaz	ulaz3.csv izlaz3_1.csv 01.01.2000 18.12.2020 Drama
Ulazna datoteka:	ulaz3.csv
Izlazna datoteka:	izlaz3_1.csv
Standardni izlaz:	
Primer 2	
Standardni ulaz	pp1_movies_2020.csv izlaz3_2.csv 01.01.2000 18.12.2020 Drama
Ulazna datoteka:	pp1_movies_2020.csv
Izlazna datoteka:	izlaz3_2.csv
Standardni izlaz:	
Primer 3	
Standardni ulaz	nepostoji.csv izlaz.csv 01.01.2000 18.12.2020 Drama
Ulazna datoteka:	-
Izlazna datoteka:	-
Standardni izlaz:	DAT_GRESKA