

Python Minta Final (gyakorlati rész)

1. Írj egy függvényt, ami egy stringben kicseréli a „one” szavakat 1-esekre kis és nagybetűtől függetlenül! A visszatérési érték string legyen.
(1p)
2. Írj egy lambda függvényt tetszőleges számú paraméterrel, ami a reduce magasabb rendű függvény segítségével megszámolja a 3 karakternél hosszabb stringeket!
(Feltételezhetjük, hogy ez esetben csak stringeket adunk át a függvénynek)
(Segítség: functools modulban található a reduce függvény)
(1p)
3. Írj egy lambda függvényt, ami megszámolja, hogy hányszor szerepel egy stringben a „hop” szó!
(1p)
4. Írj komprehenzív listát, ami egy dictionary elemein végigiterálva kiválogatja azokat a kulcsokat, ahol a hozzátartozó érték nagyobb, mint 2!
(1p)
5. Adott az alábbi csv : (**spotify_yt.csv (minta kedvéért a beadandóból)**).
Olvasd be egy DataFrame-be a csv fájl tartalmát a pandas könyvtár segítségével!
(1p)
 - a. Írasd ki a beolvasott csv-re vonatkozó információkat, statisztikákat!
(1p)
 - b. Számold meg, hogy oszloponként és összesen hány null érték található az adataink között!
(1p)
 - c. Dobd el a nullértékeket tartalmazó sorokat a DataFrameből!
(0,5p)
 - d. Mentsd le a nullértékek nélküli DataFramet egy „sp_yt_clean.csv” nevű fájlba!
(0,5p)
 - e. Dobjuk el azokat a sorokat, ahol a „Danceability” érték 0,8 alatt van!
(1p)
 - f. Készíts egy új oszlopot „Popular” címmel, ami True értéket tartalmaz, ha a „Stream” oszlopban található érték nagyobb, mint 300 000 000 (egyébként False értéket adjunk meg)!
(1p)
 - g. Készítsünk egy új DataFramet, ahol az előadókat („Artist”) a „Stream” oszlop összege szerint csoportosítjuk! Csak azokat a zeneszámokat vegyük figyelembe, ahol a „Popular” oszlopban True található!
(1p)

- h. Rendezzük a g) feladatban lévő DataFramet értékek szerint, majd vizualizáljuk a 10 legnépszerűbb előadó értékeit oszlopdiagramon!
(1p)

Várható szorgalmi: OOP modellezés