

Quiz 3+4 (7+7 ქულა)

მინი პროექტი

ლექტორი: ლიკა სვანაძე

საგანი: მონაცემთა ანალიტიკა (Python)

დავალბა:

მოცემულია „Significant Earthquakes 1965-2016“ მონაცემთა ბაზა csv ფაილის სახით. დეტალური ინფორმაცია ბაზის შესახებ შეგიძლიათ იხილოთ ლინკზე: <https://www.kaggle.com/datasets/usgs/earthquake-database> ქვემოთ მოცემული დავალებები შეგიძლიათ შეასრულოთ აღნიშნულ მონაცემებზე ან ნებისმიერ სხვა მონაცემზე (ბაზების ფაილი შეგიძლიათ მოიძიოთ Kaggle-ის პლატფორმაზე. გაითვალისწინეთ, უნდა აარჩიოთ ისეთი მონაცემთა ბაზა, სადაც გექნებათ თარიღის/დროის სვეტი). მონაცემთა ბაზის ფაილში უნდა იყოს მინიმუმ 7 სვეტი და 100 სტრიქონი. შეასრულეთ შემდეგი დავალებები:

კოდის ფაილში თითოეული დავალბას დაუწერეთ აღწერა.

1. (L6 – 3 ქულა) გამოიყენეთ მონაცემთა წმედის რომელიმე ფუნქციონალი (მაგ. რომელიმე სვეტის მონაცემები გადაანაწილეთ რამდენიმე ახალ სვეტად, წაშალეთ რომელიმე სვეტი, ცარიელი მნიშვნელობებისთვის გამოიყენეთ სასურველი ფორმატი, ა.შ.)
2. (L7 – 3 ქულა) რიცხვითი ტიპის ველის მონაცემები დაყავით კატეგორიებად cut() ფუნქციის გამოყენებით (მაგ. ასაკის ველი 3 კატეგორიად „Children“, „Adolescent“, „Old“). შესაბამისად დაამატეთ კატეგორიის სვეტი.
3. (L9 – 3 ქულა) გამოიყენეთ მონაცემთა “Reshaping”-ის ჩამოთვლილი ფუნქციებიდან რომელიმე 2 მათგანი (stack(), unstack(), melt(), pivot(), explode(), pivot_table(), Crosstab())
4. (L10 – 3 ქულა) ააგეთ თქვენთვის სასურველი 4 განსხვავებული დიაგრამა შემდეგი ტიპებიდან (ხაზოვანი დიაგრამა, სვეტოვანი დიაგრამა, გაფანტულობის დიაგრამა, წრიული დიაგრამა, პისტოგრამა).
5. (L11 – 2 ქულა) გამოიყენეთ თარიღებთან სამუშაო მინიმუმ 2 განსხვავებული ფუნქცია/ტრანსფორმაცია სურვილისამებრ (მაგ. resample ფუნქცია).

ატვირთეთ პროექტის ფაილი და csv ფაილი Github-ზე (ან სხვა მსგავს პლატფორმაზე). README ფაილში დეტალურად აღწერეთ რას წარმოადგენს თქვენი პროექტი და როგორ მუშაობს.

დავალბების გაგზავნა:

6. დავალბების გაგზავნის დედლაინია: **27 დეკემბერი (23:00 საათი)**
7. დავალბების ფაილები ატვირთეთ Github-ზე და ლინკი მიუთითეთ Google Classroom-ში Quiz 3&4 ლინკზე