

Tisztelt Hallgatók!

Az alábbi lista egy kis mankó a projektfeladat elvégzéséhez.

Kérem, hogy a munkájukat folyamatosan dokumentálják egy **jegyzőkönyv** keretében! Mit miért végeztek el, mi lett az eredménye!

Dokumentálják azt is, ha hibás útvonalon indultak el, és utána jöttek rá, hogy korrigálni kell, és másik úton mentek tovább!

Tapasztalni fogják, hogy a Python program készítését nem fogják tudni elkerülni.

A jegyzőkönyvben szerepeljen, hogy ki melyik rész elvégzésében vett részt, és a végén foglalják össze, hogy milyen eredményre jutottak az adatok elemzése során.

Ez nem sablon feladat! Nem csak egy megoldása létezik. Különböző utakon is lehet jó eredményre jutni. Legyenek kreatívak főleg az adattisztítás és a hiba azonosítás területén!

A végén kérni fogom a jegyzőkönyvet PDF formátumban, a teljes Python és SQL kódokat.

Az alábbi 12 pont a tananyag 12 fejezetével van összhangban, ezek segítenek elkészíteni a projekt feladatot:

1. A kapott adatokhoz milyen típusú adatbázist érdemes használni, és miért?
2. Tervezzék meg az adatbázist ER diagram segítségével. Lehet, hogy néhányszor újra fogják tervezni, ezeknek a miéértjét és eredményeit nyugodtan írják le.
3. Hozzák létre az adatbázist. A választott metódust indokolják meg.
4. Töltsék fel az adatokat a csv fájlokból a megfelelő táblákba. Ellenőrizték az adatok és az adattípusok helyességét!
5. DCL és TCL: SQLite-n belül nincs jogosultságkezelés.
 - a. Írják le, hogy melyik résztvevőnek milyen jogosultságra lett volna szüksége, a munkájához.
 - b. Vizsgálják meg, hogy van-e, és ha van, akkor milyen esetben tranzakciókezelésre.
 - c. 1 tranzakciót tegyenek bele akkor is, ha nem feltétlen szükséges csak a példa kedvéért.
6. Lekérdezések
 - a. Végezzenek el lekérdezéseket, melyek összecsatolnak táblákat (JOIN).
 - b. Használjanak aggregációs függvényeket.

- c. Írják le, hogy miért végezték el ezeket a műveleteket, és milyen következtetésekre jutottak.
7. Allekérdezések, UNION
- a. Szükség esetén készítsenek egymásba ágyazott lekérdezéseket
 - b. A részeredményeiket INSERT és SELECT segítségével tegyék ki külön táblákba
 - c. Vizsgálják meg, hogy mit lehetne kimutatni, vagy ellenőrizni az UNION utasításokkal
8. Teljesítmény optimalizálás
- a. Melyik mezőre milyen kulcsot, indexet alkalmaznának, és miért?
 - b. Milyen módszert javasolnának az adatműveletek gyorsítására, ha megkapnának 10 évnyi adatot és azt kellene feldolgozni?
9. Adattisztítás. Az adatok hibával terheltek.
- a. Milyen módszerrel, illetve logika alapján kerestek, szűrték hibás adatokat?
 - b. Milyen módszerrel kezelték azokat?
10. Milyen tudásuk alapján tudták azonosítani a hibákat?
- a. Mi alapján tekintették a hibás adatokat hibás adatnak? pl. Fizika órán tanulták, otthon tapasztalták, valamilyen logika alapján stb.
 - b. Adatbevitelhez milyen érvényesítési szabályokat, milyen kényszereket javasolnának
11. Az adatok egy ipari adatgyűjtő program írja az adatbázisba.
- a. Az adatok rögzítése és feldolgozása során milyen szerepköröket hoznának létre, és azokat milyen jogosultsággal látnák el?
12. Biztonsági mentésre milyen stratégiát választanának?

Javaslom, hogy azokat az anyagokat, melyeket már vettünk, annak álljanak neki!

Jó munkát!

Jó szerencsét!

Wizner Krisztián