Amrich Csongor (Adatbázisok G01) beadandó dokumentáció

A kiinduló adatbázis: <https://www.kaggle.com/datasets/zusmani/pakistans-largest-ecommerce-dataset>

Server name: f7fwli.database.windows.net

login: hallgato

jelszó: Password123

Excelben módosítottam az adatkészletet, több oszlop nevét átírtam és kitöröltem a nem használatos oszlopokat (created\_at, price, total\_amount sales\_commission\_code, discount\_amount, BI Status, MV, Year, Month, M-Y, FY, és a 4 üres oszlop). Továbbá kibővítettem még saját magam által generált Customer\_name oszloppal, nem igaziak a nevek (török név generátort találtam csak, szóval a nevek többsége nem pakisztáni, de ez részlet kérdés). Bővítettem még pár ID oszloppal excel függvények segítségével illetve a price és a total amount oszlopok értékeit szerkesztenem kellett párszor mert nem feleltek meg az adatok, így nem reprezentálnak valós értékeket, de lehet használni őket.

Beimportáltam a tisztított állományt a Management Studioba, megfelelő adattípusokkal ellátva.

Elkezdtem szétszedni az importált állományt, létrehoztam SELECT INTO-val az 5 adattáblát. Elején használtam pár temporary táblát utána már csak select distinct-el szedtem szét őket mert rájöttem, hogy úgy egyszerűbb és hasonlóan működik. Ahol kellett ott még hozzáadtam egy új oszlopot és ID-t rendeltem hozzá (review\_id). A végén töröltem az összes temporary táblát és teszteltem, hogy működnek-e.

Miután összeálltak a táblák, utána elkezdtem a constrainteket rájuk szabni, létrehoztam az elsődleges kulcsokat. Ezután a feladatnak megfelelően rendeltem constrainteket minden típusból különböző adatokhoz logikusan. A végén pedig megadtam az idegen kulcsokat is, ahova kellett.

Végül létrehoztam a scripteket amiket fel kell tölteni (data only, schema only, data and schema továbbá a query amiben manipuláltam az adatokat (ez nem tudom hogy kell-e de feltöltöttem))

A képen diagram, kör, tervezés látható

Automatikusan generált leírásER-Diagram: