

Adatbányászat a Gyakorlatban

Kuknyó Dániel

2024/25/1

1. ZH: Dash

A feladatok véletlenszerűen sorsolódnak ki az 5. gyakorlaton. A feladatok beadása e-mail-en keresztül történik, ahol csak egy linket kell beküldeni a daniel.kuknyo@mailbox.org címre, ami a feladatot megvalósító Git tárhelyre mutat. **Késő beadás nem lehetséges**, csak a beadási dátum előtt Gitre feltöltött anyagok számítanak bele a jegyre. minden további információ a tantárgyi útmutatóban található.

A beadás tartalma:

- Dash alkalmazás és segédprogramjai
- Adatforrások (100Mb-nál nagyobb állomány esetén nyílt link az adatforrásra)
- `requirements.txt` a projekthez tartozóan
(Anaconda prompt-ban: `pip freeze > requirements.txt`)

Beadási határidő:

- 2024.10.27. 24:00

Tantárgyi útmutató, anyagok:

- <https://github.com/basictask/Adatbanyaszat>
- <https://github.com/basictask/Adatbanyaszat/tree/main/feladatok/1zh>

1 Emberi fejlettség (50p)



1. Tölts be és transzformálja az elemzéshez az `1_emberi_fejlettseg.csv` állományt a `feladatok/1zh` mappából. A beolvasáshoz ne használjon keménykódolt, fix elérési utat. (3p)
2. Hozzon létre Plotly Dash alkalmazást a megfelelő struktúrával. Az alkalmazás tartalmazzon egy főcímet és egy lapolt konténert a projekt információkkal, markdown formátumban. A lapok a következők legyenek: (5p)
 - Az Ön adatai: Név, Neptun-kód, E-mail...
 - A projekt adatai: cél, megvalósítás módja, adathalmaz információk, stb...
3. Hozzon létre egy legördülő listát, ahol a legetséges fejlettségi szintek közül lehet választani (**Human Development Groups**). Ha a felhasználó kiválaszt egy fejlettségi szintet, mutassa meg az ebben szereplő országokat (**Country**) és a helyezésüket a 2021-es HDI rangsorban (**HDI rank**) egymás alatt, felsorolásban, alfabetikusan rendezve. (5p)
4. Hozzon létre egy legördülő listát, ahol országot lehet választani. Ha a felhasználó kiválaszt egy országot, mutassa meg a megfelelő változók értékeit az adott évek szerint egy táblázatban. (7p)
5. Hozzon létre egy többváltozós legördülő listát, ahol az adatkészletben szereplő országok közül lehet választani. Ha a felhasználó kiválaszt egy országot, mutassa meg diagramon a fejlettségi szintet (**Human Development Index**) egy időszorban, évek szerint. Ha a felhasználó több országot választ ki, ugyanazon a diagramon egyszerre több ország adatai legyenek láthatóak. (7p)
6. Hozzon létre egy csúszkát, amin évet lehet választani, és mellette jobb oldalt egy legördülő listát, amin változót. Ha a felhasználó kiválaszt egy évet és egy országot, mutassa meg gyakorisági diagramon az országok adatainak eloszlását a választott változó szerint. A gyakorisági osztályközök száma legyen a felhasználó által szabályozható komponens segítségével. (7p)
7. Hozzon létre egy legördülő listát, amiben a lehetséges időszoros változók szerepelnek, és egy tematikus térképet, amiben az országok. Ha a felhasználó kiválaszt egy változót, mutassa meg a változó alakulását évek szerint a tematikus térképen dinamikusan lejátszható módon. (7p)
8. Hozzon létre egy legördülő listát, amin országot lehet választani, és egy legördülő listát amiben a változók közül. Ha a felhasználó kiválaszt egy országot és egy változót, mutassa meg a kiválasztott változó és a GDP/fő alakulását évek szerint diagramon. A diagramhoz adjon hozzá regressziós modellt, amit a megfelelő formában jelenít meg. A regressziós modell legyen paraméterezhető és futtatható felhasználói komponensek segítségével. (7p)
9. Adjon hozzá egységes, egyedi háttérszínt az alkalmazáshoz, minden diagramhoz és üres diagramhoz. Adjon elválasztást és alcímeket a diagramoknak. Adjon hozzá hibakezelést a visszahívó függvényekhez, ahol szükséges. minden diagramon ahol szükséges, szerepeljen megfelelő mennyiségű információ, tengelyfelirat, tengelycím, jelmagyarázat, marker, marker felirat stb... (2p)

2 CO₂ kibocsájtás (50p)



1. Tölts be és transzformálja az elemzéshez az `2_co2_kibocsajtas.csv` állományt a **feladatok/1zh** mappából. A beolvasáshoz ne használjon keménykódolt, fix elérési utat. (3p)
2. Hozzon létre Plotly Dash alkalmazást a megfelelő struktúrával. Az alkalmazás tartalmazzon egy főcímet és egy lapolt konténert a projekt információkkal, markdown formátumban. A lapok a következők legyenek: (5p)
 - Az Ön adatai: Név, Neptun-kód, E-mail...
 - A projekt adatai: cél, megvalósítás módja, adathalmaz információk, stb...
3. Hozzon létre egy legördülő listát, amin országot lehet választani. Ha a felhasználó kiválaszt egy országot, jelenítse meg táblázatos formában az adott ország nettó energiamérlegét évekre lebontva (az energiamérleg az összes felhasználás és az összes termelés összege). (5p)
4. Hozzon létre egy csúszkát, amin a felhasználó kiválaszthatja egy ország (**Country**) kumulált energiafogyasztását (**Energy_consumption**). Ha a felhasználó választott egy értéket, jelenítse meg felsorolásban azokat az országokat, amelyek a választott értéktől több energiát használtak 1980 óta. (7p)
5. Hozzon létre egy többváltozós legördülő listát, ahol országot lehet választani. Ha a felhasználó kiválaszt egy országot, mutassa meg diagramon az egyes erőforrások termelésének (**Energy_production**) alakulását évek szerint megfelelő diagram típuson. Ha több, mint egy ország van kiválasztva, országonként külön diagramon jelenjenek meg az adatok. (7p)
6. Hozzon létre egy csúszkát, amin évet (**Year**) lehet választani a lehetséges értékek közül. Ha a felhasználó választ egy értéket, mutassa meg az országok adott évre jellemző CO₂ kibocsájtását (**CO2_emission**) gyakorisági diagramon. Az osztályközök száma legyen dinamikusan szabályozható felhasználói komponensek segítségével. A szabályozó a csúszka mellett, jobb oldalt szerepeljen. (7p)
7. Hozzon létre legördülő listát, amin energiatípust (**Energy_type**) lehet választani, és egy tematikus térképet. Ha a felhasználó kiválaszt egy értéket, mutassa meg a változó alakulását évek szerint a tematikus térképen, dinamikusan lejátszható módon. (7p)
8. Hozzon létre egy legördülő listát, amin országot lehet választani. Ha a felhasználó kiválaszt egy országot, mutassa meg a CO₂ kibocsájtás alakulását évek szerint a diagramon. A diagramhoz adjon hozzá regressziós modellt, aminenek a predikciót a megfelelő módon jeleníti meg. A regressziós modell legyen paraméterezhető és futtatható felhasználói komponensek segítségével. (7p)
9. Adjon hozzá egységes, egyedi háttérszínt az alkalmazáshoz, minden diagramhoz és üres diagramhoz. Adjon elválasztást és alcímeket a diagramoknak. Adjon hozzá hibakezelést a visszahívó függvényekhez, ahol szükséges. minden diagramon ahol szükséges, szerepeljen megfelelő mennyiségű információ, tengelyfelirat, tengelycím, jelmagyarázat, marker, marker felirat stb... (2p)

3 Várható élettartam (50p)



1. Tölts be és transzformálja az elemzéshez az `3_varhato_elettartam.csv` állományt a `feladatok/1zh` mappából. A beolvasáshoz ne használjon keménykódolt, fix elérési utat. (3p)
2. Hozzon létre Plotly Dash alkalmazást a megfelelő struktúrával. Az alkalmazás tartalmazzon egy főcímet és egy lapolt konténert a projekt információkkal, markdown formátumban. A lapok a következők legyenek: (5p)
 - Az Ön adatai: Név, Neptun-kód, E-mail...
 - A projekt adatai: cél, megvalósítás módja, adathalmaz információk, stb...
3. Hozzon létre egy legördülő listát, amin országot (**Country**) lehet választani. Ha a felhasználó kiválaszt egy országot, jelenítse meg az ország egy főre jutó GDP-jét évekre lebontva. (5p)
4. Hozzon létre egy csúszkát, amin népesség (**Population**) és nemzeti össztermék (**GDP**) tartományt lehet választani az adathalmazban szereplő értékeknek megfelelően. Ha a felhasználó értéket választ minden két komponensen, jelenítse meg azokat az országokat és a fejlettségi szintjeiket, amelyek a választott tartományba esnek egy felsorolásban, egymás alatt. (7p)
5. Hozzon létre többváltozós legördülő listát, ahol országot lehet választani. Ha a felhasználó országot választ, mutassa meg a várható élettartam (**Life expectancy**) alakulását az adott országban évek szerint. Ha több ország van kiválasztva, ugyanazon a diagramon jelenjenek meg. (7p)
6. Hozzon létre egy csúszkát, amin évet lehet választani és jobbra mellette egy másikat, amin változót. Ha a felhasználó minden két komponensen választott értéket, jelenítse meg a változó értékeit gyakorisági diagramon. A gyakoriság osztályközei legyenek dinamikusan szabályozhatóak felhasználói komponensek segítsével. (7p)
7. Hozzon létre egy legördülő listát, amin változót lehet választani az adathalmazból, és egy tematikus térképet. Ha a felhasználó értéket választ, jelenítse meg a változó alakulását évek szerint tematikus térképen, dinamikusan lejátszható módon. (7p)
8. Hozzon létre egy legördülő listát, amin évet lehet választani. Ha a felhasználó értéket választ, mutassa meg a várható élettartam alakulását egy főre jutó GDP szerint. A diagramhoz adjon hozzá regressziós modellt, aminek a predikciót megfelelő módon jeleníti meg a diagramon. A regressziós modell legyen paraméterezhető és futtatható felhasználói komponensek segítségével. (7p)
9. Adjon hozzá egységes, egyedi háttérszínt az alkalmazáshoz, minden diagramhoz és üres diagramhoz. Adjon elválasztást és alcímeket a diagramoknak. Adjon hozzá hibakezelést a visszahívó függvényekhez, ahol szükséges. minden diagramon ahol szükséges, szerepeljen megfelelő mennyiségű információ, tengelyfelirat, tengelycím, jelmagyarázat, marker, marker felirat stb... (2p)