

מכללת הדסה, החוג למדעי המחשב

אנליזה של ביג דאטה / חננאל פרל

סמסטר ב', תשפ"ד

תרגיל 1 / 4.04.2024

תאריך הגשה:

שבוע מהיום - יום חמישי 11.04.2024 בשעה 23:59

מטרות התרגיל:

תרגול שפת SQL בסיסית

הוראות:

הורד את הדאטה בייס:

<https://drive.google.com/file/d/1C4MNtoLAAwGWAjBHOUGL1fTn0zo-rkHL/view?usp=sharing>
לצורך משחק ובדיקות אפשר לפתוח אותו בקלינט הזה (כמובן שיש גם אחרים):

<https://sqlitebrowser.org/>

פתור התרגילים.

ההגשה צריכה להכיל:

- קובץ טקסט אחד (.py) שמכיל תוכנית פייטון שניתן להריצה.
- קובץ טקסט אחד (.txt) של ההדפסות של הרצה שאתם עשיתם אצלכם.

התוכנית פייטון צריכה לפתוח את הדאטהבייס מקובץ לוקלי בסיפרייה נוכחית (ללא תתי תיקיות וכד'), ולהריץ מולו את השאילתות שעונות על השאלות.
עבור כל שאלה יש להדפיס את:

(1) מספר התרגיל.

(2) את השאילתא עצמה בעימוד לפי הקונבנציה שלמדנו.

(3) יש להדפיס גם את התוצאה באופן הבא:

- יש להדפיס את כמות השורות שיצאו סה"כ (לכתוב את מספר השורות, לא כולל כותרות כמובן)
- יש להדפיס 5 שורות ראשונות וגם 5 שורות אחרונות של התוצאה (לא להדפיס את כל השורות, רק ראשונות ואחרונות).

הדגשה: השאילתא עצמה צריכה להחזיר את כל השורות לפי התרגיל. רק בהדפסה למסך בפייטון צריך להדפיס כאמור את השורות הראשונות והאחרונות.
השאילתא צריכה להיות כתובה בקונבנציה שלמדנו (בכמה שורות, אותיות גדולות איפה שצריך וכו') בפייטון אפשר בנוחות לכתוב מחרוזת פרוסה על פני כמה שורות עם שלוש גרשיים "" (ואפשר באופן כזה במחרוזת לכתוב גם " ' ועוד). ככה:

```
query = ""  
SELECT *  
FROM City  
""
```

יש לשים שורה ריקה וקו מפריד באורך 55 '=' (התו =) בין כל שאלה ושאלה.
הערה: אין לכתוב עברית בשאילתות ובקוד.

שאלה שאתם לא יודעים את התשובה עליה, עדיין לכתוב את הקו המפריד ומספר השאלה, אך להשאיר את השאילתא והתוצאה ריקים.

=====

Question: 1

The query:

```
SELECT *
FROM City
```

Num of rows: 4079

The results:

	ID	Name	CountryCode	District	Population
0	1	Kabul	AFG	Kabul	1780000
1	2	Qandahar	AFG	Qandahar	237500
2	3	Herat	AFG	Herat	186800
3	4	Mazar-e-Sharif	AFG	Balkh	127800
4	5	Amsterdam	NLD	Noord-Holland	731200
	ID	Name	CountryCode	District	Population
4074	4075	Khan Yunis	PSE	Khan Yunis	123175
4075	4076	Hebron	PSE	Hebron	119401
4076	4077	Jabaliya	PSE	North Gaza	113901
4077	4078	Nablus	PSE	Nablus	100231
4078	4079	Rafah	PSE	Rafah	92020

=====

Question: 2

The query:

=====

Question: 3

..

..

השאלות לא בהכרח ממוינות לפי סדר קושי..

בהצלחה!

השאלות SQL 1:

1. להציג את כל הערים
2. להציג את כל המדינות
3. להציג את הערים במדינה שהקוד שלה NLD
4. הצג את המדינות שהקוד שלהם הוא LBR או IOT או TKL
5. להציג את כל הערים עם אוכלוסיה מעל 4 מיליון
6. להציג את כל הערים עם אוכלוסיה מעל 3 מיליון וגם בברזיל (קוד BRA)
7. להציג את כל הערים עם אוכלוסיה בין 150,000 ל 170,000
8. להציג את המדינות ששנת עצמאות 1970 או 1980 או 1990
9. להציג את המדינות ששנת עצמאות 1980 וגם 1990
10. להציג את המדינות ששנת עצמאות בין 1980 ל 1990
11. להציג את המדינות באפריקה וגם שנת העצמאות 1964
12. להציג את המדינות באפריקה או שנת העצמאות 1964
13. להציג את המדינות שהם באסיה
14. להציג את המדינות שהם לא באסיה
15. להציג את המדינות שהם לא באסיה ולא באירופה
16. להציג את הערים שמתחילים באות H
17. להציג את הערים שאין בהם את האות e
18. להציג את כל השפות (פעם אחת כל שפה) ממוינות
19. להציג מדינות ממוינות לפי שנת עצמאות, ואם זהה אז לפי א"ב
20. להציג מדינות ממוינות לפי שנות חיים סדר יורד
21. להציג 10 מדינות תמג (GNP) הכי גבוה
22. להציג את ה 10 מדינות הבאות אחרי ה 10 התמג (GNP) הכי גבוה
23. להציג 10 מדינות עם התמג (GNP) הכי נמוך
24. להציג את "שם המדינה" בעמודה ראשונה ואז בעמודה השניה "היבשת והאזור בסוגריים" ממוין לפי העמודה הראשונה
25. להציג כל המדינות שמתחילות באות Z, הראה את אוכלסייתם במיליונים (ז"א אם יש מיליון וחצי יראה 1.5)
26. להציג מדינות שהתמג (GNP) עלה ביותר מפי 2
27. להציג המדינות והצפיפות (אוכלוסיה/שטח) שלהם, ממוין לפי צפיפות מהגדול לקטן
28. להציג את שם המדינה והצפיפות של מדינות שהצפיפות מעל 2000, ממוין לפי קוד המדינה
29. להציג את כל המדינות ממוינות לפי היבשת בסדר עולה, ובתוך כל יבשת ממוין לפי שטח מהגדול לקטן
30. הצג את כל המדינות שמסתיימות באות N ושם המדינה פחות מ-10 תוים
31. הצג כל המדינות שלא ידוע שנת העצמאות שלהם