

תרגיל קריאה ושמירת קבצים בפייתון וטיפול בשגיאות

מטרת התרגיל: תרגול שימוש בקריאה מקבצים וטיפול בחריגות.

דרישות התרגיל:

- בתרגיל זה עליכם לכתוב תוכנית אשר תקרא ותנתח קובץ של מסד נתונים רפואי אשר כתוב בפורמט של פקודות SQL ותמירו לפורמט CVS.
- אם יהיו שגיאות בקובץ, עליכם לטפל בהם בעזרת try, except finally and raise. עליכם ליצור לפחות 3 raise ולפחות 1 try and except בהתאם לשיקול דעתכם בצורה הגיונית שתשמור על הקוד מאורגן, ברור ולא מאולץ.
- הקוד צריך להיות מאורגן בפונקציות קצרות ויעילות.
- הפונקציה הראשונה תקבל כפרמטר את הנתיב ממנו לקרוא את הקובץ. יש להגיש הכול בקובץ py אחד.
- הפונקציה הראשונה צריכה להיות מסוג generator וכל קריאה לnext שלה, תיצור את הטבלה הבאה.

הסבר התרגיל:

- SQL זהו סוג של מסד נתונים, זאת אומרת הוא שומר נתונים רבים בצורה יעילה ומאורגנת בתוך טבלאות.
- קובץ פקודות SQL הוא קובץ המכיל אוסף של פקודות SQL לביצוע במסד הנתונים. קובץ זה מכיל את הפקודות וההוראות הנדרשות כדי ליצור, לעדכן, למחוק או לשאול מידע ממסד הנתונים.
- קובץ CSV מכיל סידרת נתונים אשר הופרדה על ידי סימני פסיק כאשר שורה חדשה מייצגת טור נתונים חדש. ניתן לייצא את הקבצים הללו לתוך תאים על ידי תוכנת גליון אלקטרוני, כמו אקסל, או להכניסם כמו שהם לתוך מסד נתונים. הם כלי נפוץ ופופולרי להעברת אינפורמציות ונתונים.

מה אתם מקבלים מאיתנו:

אתם מקבלים מאיתנו את הקובץ demo.sql (מופיע במודל מתחת לתרגיל) קובץ זה הוא קובץ פקודות של SQL שבונה את מסד הנתונים. מסד הנתונים מכיל הרבה טבלאות. בכל טבלה יש את השדות שבה (העמודות בטבלה) ואת הנתונים שנכנסים אליה.

הפקודה: CREATE TABLE יוצרת טבלה חדשה. ובשורות שאחר כך יש את הפירוט של השדות השונים.

למשל, בדוגמא הבאה מוגדרת טבלה בשם active_list_type ולה עשרה שדות: active_list_type_id, name, description, creator, date_created, retired, retired_by, date_retired, retire_reason, uuid

```

CREATE TABLE `active_list_type` (

`active_list_type_id` int(11) NOT NULL AUTO_INCREMENT,

`name` varchar(50) NOT NULL,

`description` varchar(255) DEFAULT NULL,

`creator` int(11) NOT NULL,

`date_created` datetime NOT NULL,

`retired` tinyint(1) NOT NULL DEFAULT '0',

`retired_by` int(11) DEFAULT NULL,

`date_retired` datetime DEFAULT NULL,

`retire_reason` varchar(255) DEFAULT NULL,

`uuid` char(38) NOT NULL,

```

נתעלם מטיפוס השדות והמפתחות.

מיד לאחר הגדרת הטבלה, ישנה סדרת שורות שמתחילה במילים INSERT INTO אשר מורה על הכנסת ערכים לתוך הטבלה.

פקודות הINSERT INTO יכולות להופיע כל אחת בשורה נפרדת, או משורשרות באותה שורה ומופרדות על ידי פסיק. להלן דוגמא לשתי האפשרויות השונות.

```

INSERT INTO `active_list_type` VALUES (1,'Allergy','An Allergy the Patient may have',1,'2010-05-28 00:00:00',0,NULL,NULL,NULL,'96f4f603-6a99-11df-a648-37a07f9c90fb')

```

```

INSERT INTO `active_list_type` VALUES (2,'Problem','A Problem the Patient may have',1,'2010-05-28 00:00:00',0,NULL,NULL,NULL,'a0c7422b-6a99-11df-a64837a07f9c90fb');

```

או

```

INSERT INTO `active_list_type` VALUES (1,'Allergy','An Allergy the Patient may have',1,'2010-05-28 00:00:00',0,NULL,NULL,NULL,'96f4f603-6a99-11df-a648-37a07f9c90fb'),(2,'Problem','A Problem the Patient may have',1,'2010-05-28 00:00:00',0,NULL,NULL,NULL,'a0c7422b-6a99-11df-a648-37a07f9c90fb')

```

עליכם ליצור מתוך הפקודות הנ"ל את הקובץ CSV ששמו יהיה כשם הטבלה: active_list_type והמבנה שלו יהיה:

active_list_type_id,name,description,creator,date_created,retired,retired_by,date_retired,retire_reason,uuid














































```
1,'Allergy','An Allergy the Patient may have',1,'2010-05-28
00:00:00',0,NULL,NULL,NULL,'96f4f603-6a99-11df-a648-37a07f9c90fb'

2,'Problem','A Problem the Patient may have',1,'2010-05-28
00:00:00',0,NULL,NULL,NULL,'a0c7422b-6a99-11df-a648-37a07f9c90fb'
```

אם נייצא אותו לאקסל, נראה:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	active_list_name	description	creator	date_created	retired	retired_by	date_retired	retire_reason	uuid					
2	1	'Allergy'	'An Allergy the Patient may have'	1	'2010-05-28 00:00:00'	0	NULL	NULL	NULL	'96f4f603-6a99-11df-a648-37a07f9c90fb'				
3	2	'Problem'	'A Problem the Patient may have'	1	'2010-05-28 00:00:00'	0	NULL	NULL	NULL	'a0c7422b-6a99-11df-a648-37a07f9c90fb'				
4														
5														
6														
7														
8														

ניתן להסתמך על מבנה הקובץ, כלומר: אם תופיע שורת INSERT INTO היא תמיד תופיע אחרי ה CREATE TABLE יכול להיות שיהיו טבלאות ריקות, כלומר רק שורת הצהרה על טבלה של CREATE TABLE ללא שורה\ות INSERT INTO אחריה ואז אין ליצור את הטבלה. יש לייצא כל טבלה לקובץ נפרד. סה"כ הפלט של שאלה זו צריך להכיל את רשימת הקבצים הבאה:

 active_list_type.csv	 encounter_role.csv	 patient_identifier_type.csv
 care_setting.csv	 encounter_type.csv	 person.csv
 concept.csv	 field.csv	 person_address.csv
 concept_answer.csv	 field_type.csv	 person_attribute.csv
 concept_class.csv	 form.csv	 person_attribute_type.csv
 concept_datatype.csv	 form_field.csv	 person_name.csv
 concept_description.csv	 global_property.csv	 privilege.csv
 concept_map_type.csv	 hl7_source.csv	 provider.csv
 concept_name.csv	 liquibasechangelog.csv	 relationship_type.csv
 concept_numeric.csv	 liquibasechangeloglock.csv	 role.csv
 concept_set.csv	 location.csv	 role_privilege.csv
 concept_stop_word.csv	 obs.csv	 scheduler_task_config.csv
 drug.csv	 order_type.csv	 user_property.csv
 encounter.csv	 patient.csv	 user_role.csv
 encounter_provider.csv	 patient_identifier.csv	 users.csv

בהצלחה!