פרוייקט חלק א' – עיבוד מקדים של הנתונים מגישים: איתי בכר 206948218 אלעד וינר 209639855

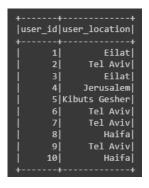
הנחות

סוגי השאילתות המופיעים בטבלת queries הם הסוגים היחידים הקיימים ועל סמך סוגי השאילתות הנ"ל נעבד את הנתונים. (לדוגמה ניתן לתשאל רק על סמך שנה מסוימת שממנה והלאה הסרטים יהיו בפלט ולכן יהיה ניתן לחלץ אך ורק את השנה שהסרטים יצאו לאור ולא להציג את התאריך המלא).

USERS – טבלה אשר מכילה מידע על המשתמשים.

.user id, user location נשאיר את הטבלה כמו שהיא עם העמודות

לבסוף הטבלה תראה באופן הבא:



MOVIES – טבלה אשר מכילה מידע על הסרטים.

נבצע כמה שינויים:

- overview, revenue, tagline מחיקת עמודות לא רלוונטיות
- ניתן לתשאל רק על סמך שנה מסוימת שממנה והלאה הסרטים יהיו בפלט ולכן נחלץ מעמודה release date
- עמודות genres, production_companies, production_countries, spoken_languages,cities עמודות ולכן את הנתונים מהאתר הוא היה בפורמט כזה) ולכן (ככה"נ כשחילצו את הנתונים מהאתר הוא היה בפורמט כזה) ולכן נרצה "לייפה" את העמודות הנ"ל, נשים לב לפי טבלת השאילתות נרצה רק חלק מהנתונים אשר נמצאים בעמודות הנ"ל.
 - genres: ניצור רשימה של השמות של הז'אנרים של הסרט (id לא רלוונטי).
 - production_companies: ניצור רשימה של השמות של חברות ההפקה של הסרט (id) לא רלוונטי). production_companies: ניצור רשימה של השמות של המדינות שבהן הופק הסרט (קידוד לא רלוונטי). spoken_languages: ניצור רשימה של השפות שהסרט זמין בהן (קידוד לא רלוונטי).
 - cities: ניצור רשימה של הערים שהסרט זמין בהם.
 - " ","?????", "??????" במחוק מהעמודה spoken_languages ערכי
 - כשים לב כי יש שפה בשם No Language אך זה לא ערך זבל מכיוון שיש סרטים אילמים ללא
 שפה כלשהי (רובם מאוד ישנים) ולכן לא נמחוק את הערכים הנ"ל, נשים לב שיש משתמשים
 אשר מתעניינים ומחפשים סרטים מהסוג הנ"ל.

לבסוף הטבלה תראה באופן הבא:

movie id genres production companies production countries	release date spoken languages title cities
	terease_uate spoken_ianguages title titles
862 [Animation, Comed [Pixar Animation [United States of	1995 [English] Toy Story [Eilat, Tel Aviv,
8844 [Adventure, Fanta [TriStar Pictures [United States of	1995 [English, Franāsais] Jumanji [Jerusalem, Tiber
15602 [Romance, Comedy] [Warner Bros., La [United States of	1995 [English] Grumpier Old Men [Eilat, Haifa, Te
31357 Comedy, Drama, R Twentieth Centur United States of	1995 [English] Waiting to Exhale [Tel Aviv, Tiberi]
11862 [Comedy] [Sandollar Produc [United States of	1995 [English] Father of the Bri [Haifa, Jerusalem
949 [Action, Crime, D [Regency Enterpri [United States of	1995 [English, Español] Heat [Tiberias, Jerusa
11860 [Comedy, Romance] [Paramount Pictur [Germany, United	1995 [FranASais, English] Sabrina [Eilat, Jerusalem
45325 [Action, Adventur [Walt Disney Pict [United States of	1995 [English, Deutsch] Tom and Huck [Haifa, Jerusalem
9091 [Action, Adventur [Universal Pictur [United States of	1995 [English] Sudden Death [Tiberias, Haifa,
710 [Adventure, Actio [United Artists, [United Kingdom,	1995 [English, PÑĐNĐNĐ GoldenEve [Eilat, Tiberias
9087 [Comedy, Drama, R [Columbia Picture [United States of	1995 [English] The American Pres [Haifa, Tel Aviv,
12110 [Comedy, Horror] [Columbia Picture [France, United S	1995 [English, Deutsch] Dracula: Dead and [Tiberias, Jerusa
21032 [Family, Animatio [Universal Pictur [United States of	1995 [English] Balto [Jerusalem, Haifa
10858 [History, Drama] [Hollywood Pictur [United States of	1995 [English] Nixon [Eilat, Haifa, Ti

CREDITS – טבלה אשר מכילה מידע על הצוות והשחקנים של הסרטים.

נבצע כמה שינויים:

עמודות cast,crew היו מיוצגות ע"י סוג של פורמט JSON (ככה"נ כשחילצו את הנתונים מהאתר הוא היה בפורמט כזה) ולכן נרצה "לייפה" את העמודות הנ"ל, נשים לב לפי טבלת השאילתות נרצה רק חלק מהנתונים אשר נמצאים בעמודות הנ"ל.

cast: ניצור רשימה של השמות של השחקנים של הסרט (כל שאר הנתונים לא רלוונטים). crew: ניצור רשימה של השמות של הבמאיים של הסרט (כל שאר הנתונים לא רלוונטים).

לבסוף הטבלה תראה באופן הבא:



QUERIES - טבלה אשר מכילה מידע על השאילתות של המשתמשים. נבצע שינוי:

עמודה from_release_date מיוצגת ע"י שנה אחת בלבד בתוך מערך, לכן נחלץ רק את השנה עצמה.

לבסוף הטבלה תראה באופן הבא:

user_id	genres	lang	actors	director	cities	country	from_realese_date	production_company
	['Western', 'Myst ['Action', 'Weste			Nae Caranfil'] 1	Haifa', 'Tiberi Tel Aviv'			'Katakuri-ke no 'Clavius Base'
4095		['English', '⊋₹ ['	Kenneth Alton',	[] ['Philip Dunne']		Belgium', 'Mold	1995 ['P	eter Carsten P 'Centrul de Prod
	['Action', 'Docum ['TV Movie']	['English']			['Tel Aviv', 'Jer		2018 ['Vortex Words Pi ['Northern Lights']
	['Fantasy', 'Adve ['Drama', 'Animat	['English'] ['English', '泙	ia i	10 10		['Guatemala', 'Ky 'Qatar', 'Israel']	1996 [1999	'Warner Bros', ' ['Chill']
	['Crime', 'Romance'] ['Drama', 'Animat	['English'] ['English']		[] ['Michael Tiddes']	['Eilat'] ['Haifa']	['Switzerland', '		'Monogram Pictur 'Joel Production
8260	['Family']	['English']	["Terry O'Quinn",	ii di	['Tiberias']	['Namibia', 'Moro	1998 ['TROS Bridge Rig

TICKETS – טבלה אשר מכילה מידע על ההזמנות של המשתמשים. נבצע שינוי:

.cinema_id - מחיקת עמודה לא רלוונטית

לבסוף הטבלה תראה באופן הבא:

++		+	+
user_id r	novie_id numbe	r_of_tickets	city
++	+	+	+
1	110	1	Eilat
1	147	1	Tiberias
1	858	2	Tel Aviv
1	1246	1	Eilat
1	1968	1	Eilat
1	2762	1	Tiberias
1	2959	1	Tel Aviv
1	4226	1	Eilat
1	54503	3	Eilat
1	58559	1	Haifa
1	96821	2	Haifa
11	110	2	Eilat
11	2959	1	Eilat
11	4226	1	Jerusalem