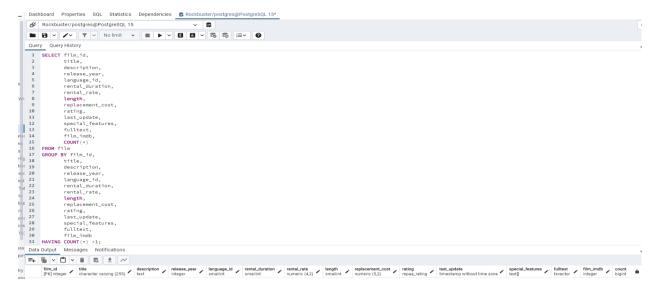
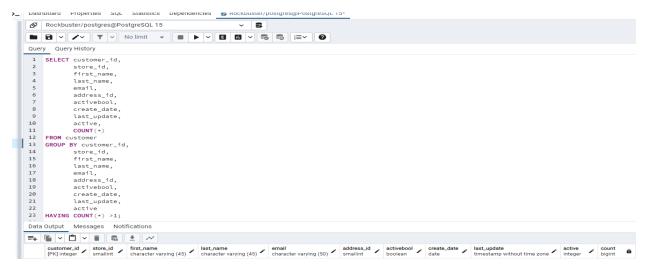
Summarizing & Cleaning – Rockbuster Stealth Data Analysis (SQL) Project

Ideal Kuci

- "Film" table searched for duplicates:

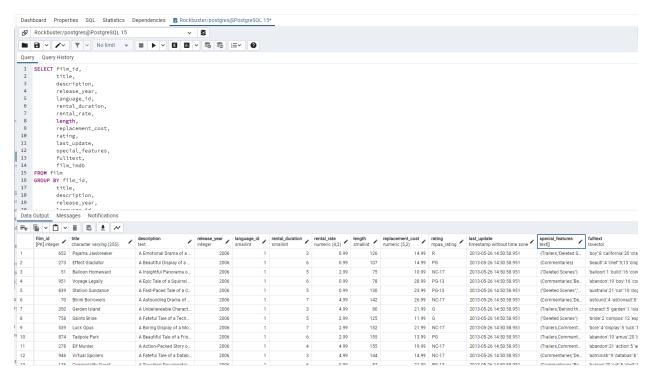


- "Customer" table searched for duplicates:

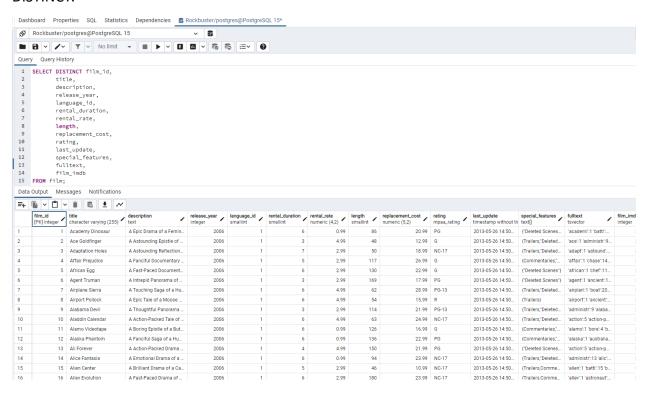


- "Film" table, query used for non – uniform check:

GROUP BY (I also used DISTINCT below for the sake of comparison):

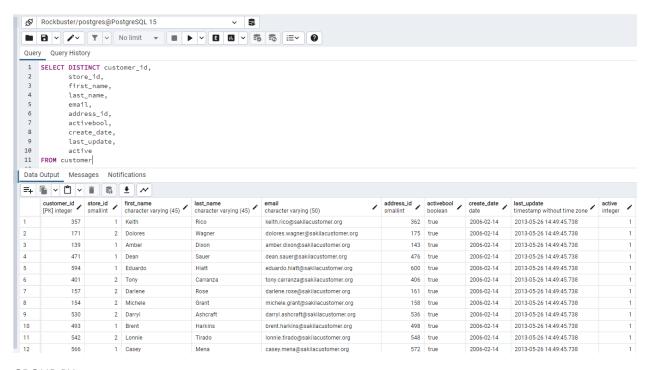


DISTINCT:

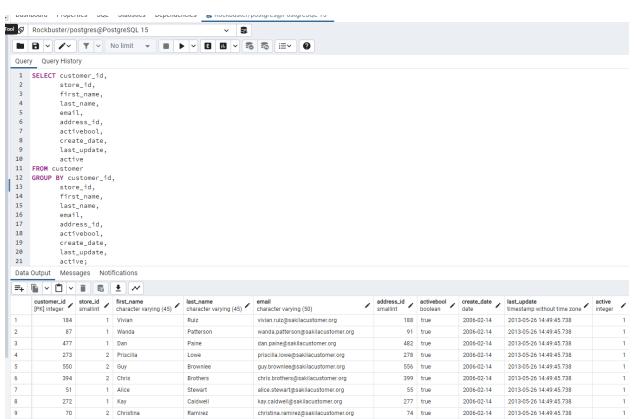


- "Customer" table, query used for *non* – *uniform* check:

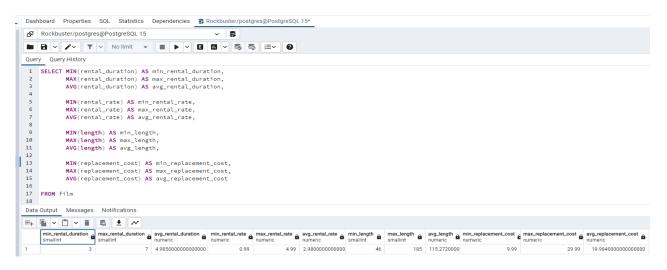
DISTINCT:



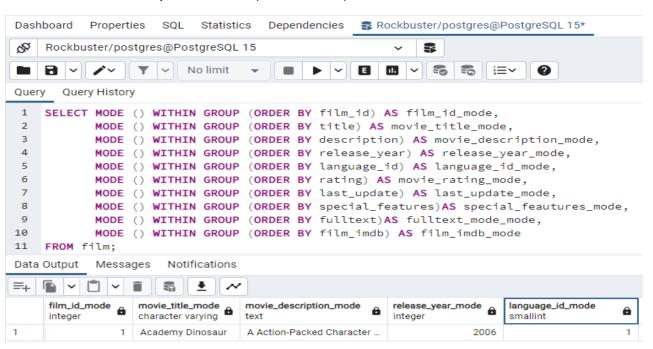
GROUP BY:



Film table descriptive statistics (numeric):



Film table descriptive statistics (non-numeric):



movie_rating_mode mpaa_rating	last_update_mode timestamp without time zone €	special_feautures_mode text[]	fulltext_mode_mode tsvector	film_imdb_mode integer
PG-13	2013-05-26 14:50:58.951	{Trailers,Commentaries,"Behind the Scenes"}	"baloon':19 'confront':14 'documentari':5 'feminist':8,11,16 'mile':2 'must':13 'spi':1 'thri	[null]

- Customer table descriptive statistics (numeric):

None.

- Customer table descriptive statistics (non-numeric):

