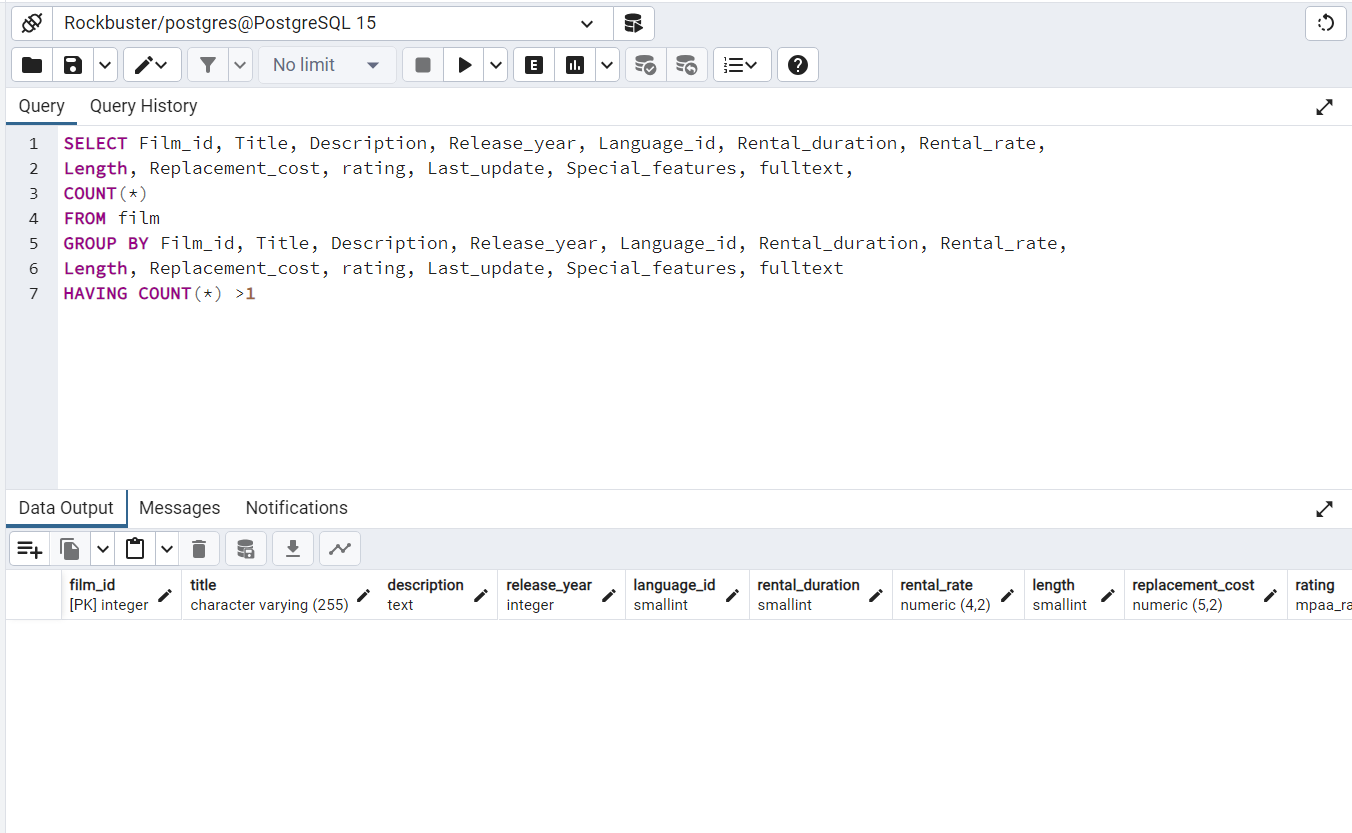
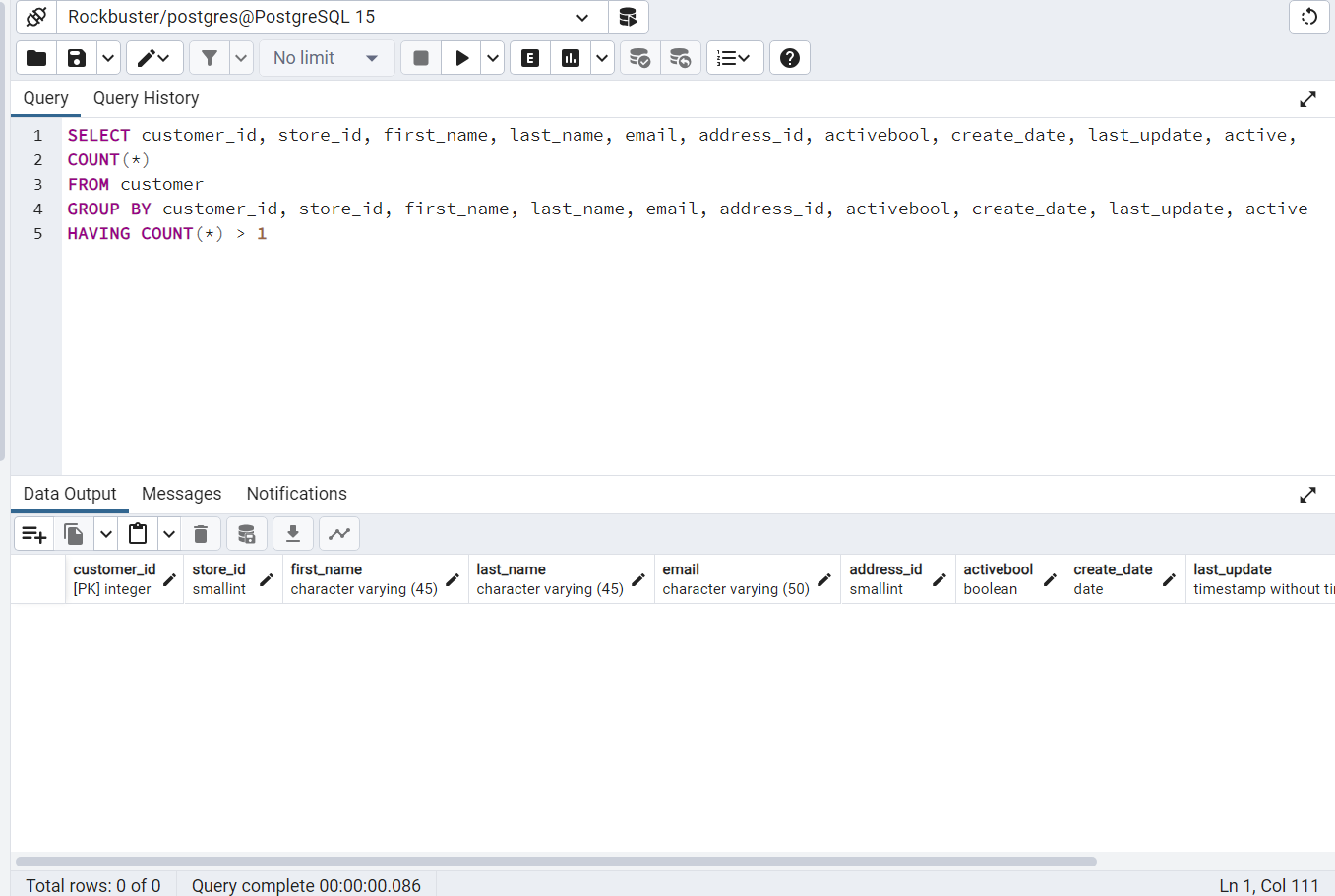
1.a Film Table



1.a Customer Table

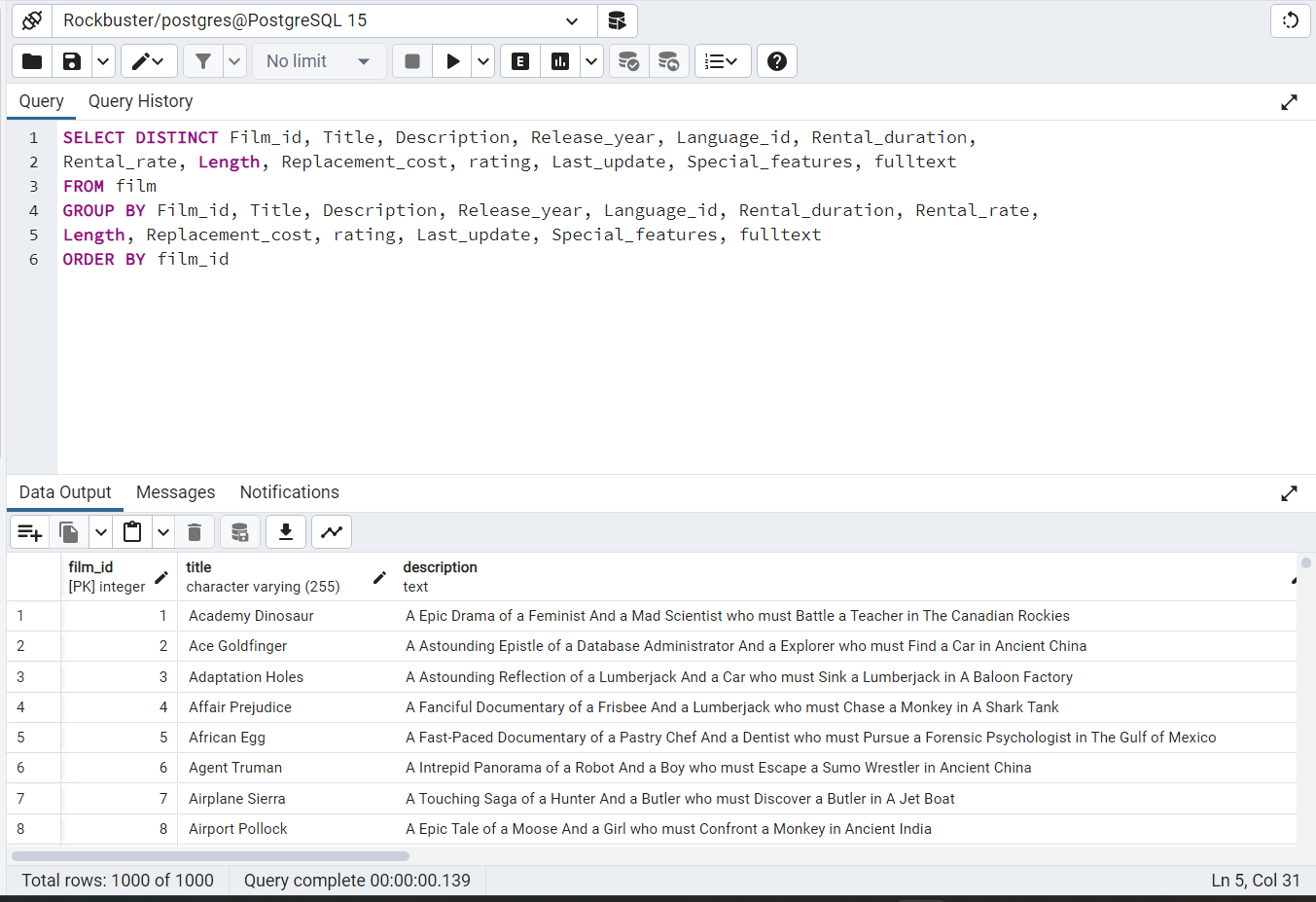


A screenshot of a computer

Description automatically generated

There are no duplicate records in either of the tables. Maintaining a viewable table with distinct records is deemed as a preferred approach. This is due to the fact that deleting records may lead to the loss of valuable information.

1.b Film Table



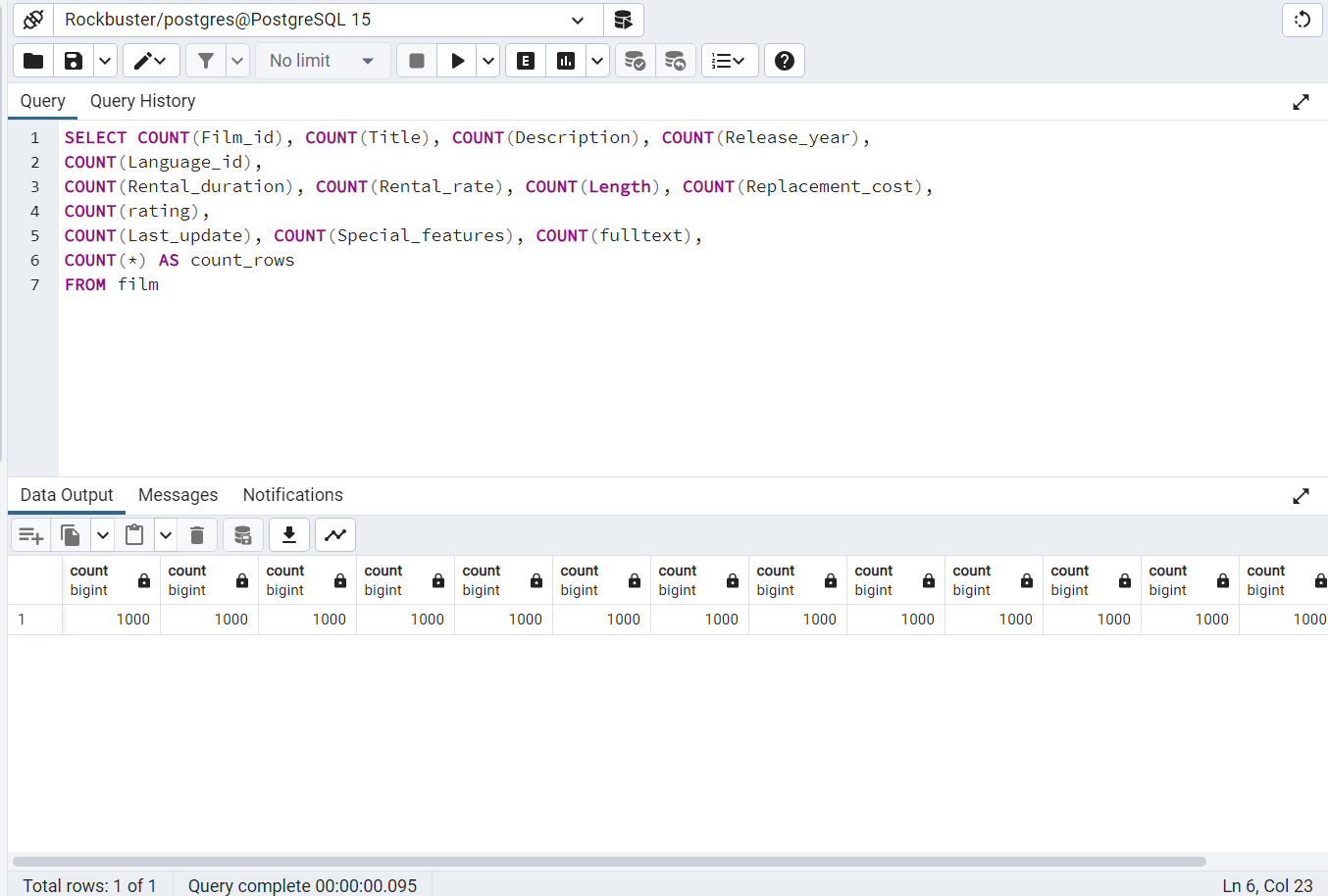
1.b Customer Table

A screenshot of a computer

Description automatically generated

The data appears uniform for both tables. If the data was not uniform, I would use the Update function along with the Where and Set functions to resolve this problem.

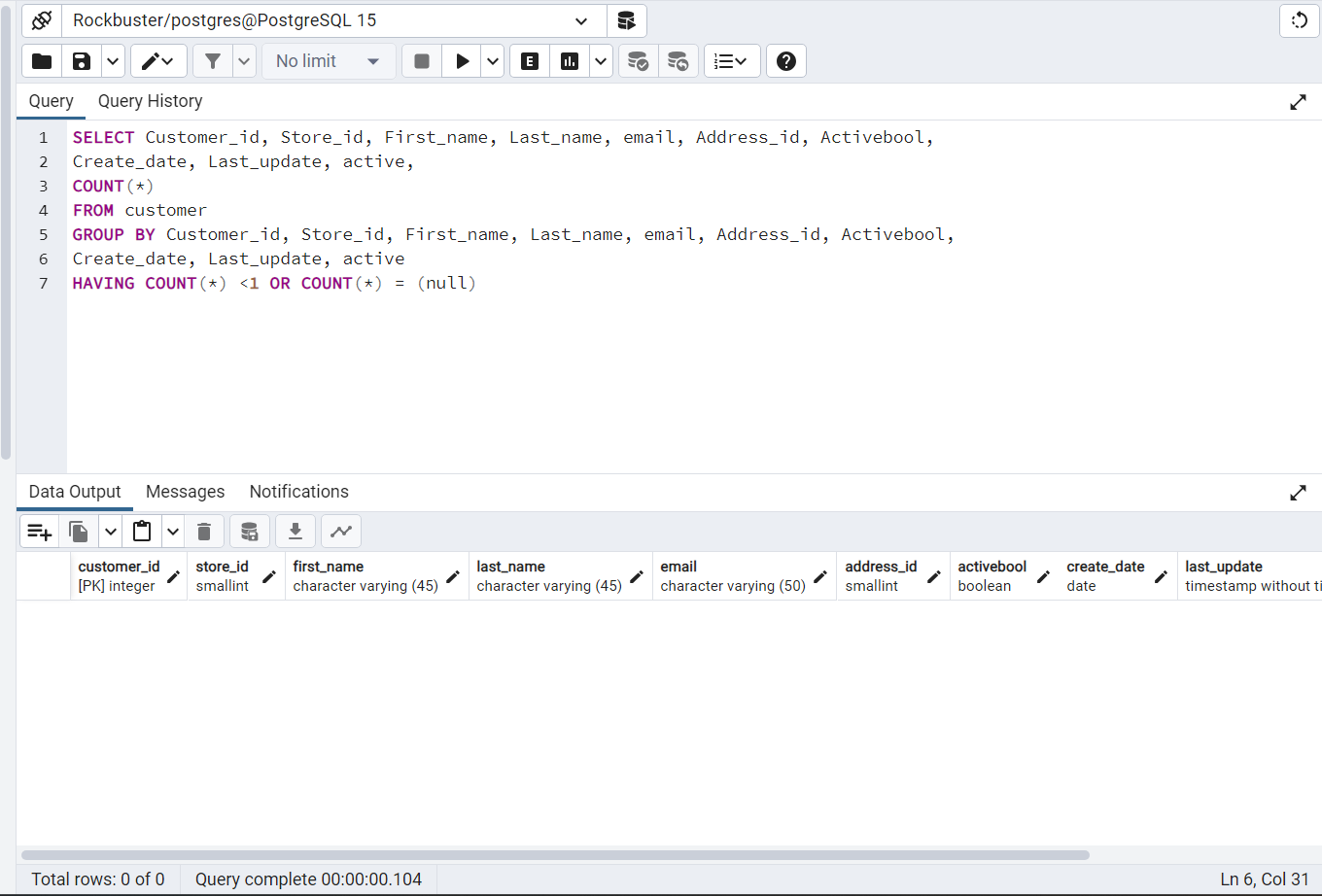
1.c Film table



A screenshot of a computer

Description automatically generated

1.c Customer Table

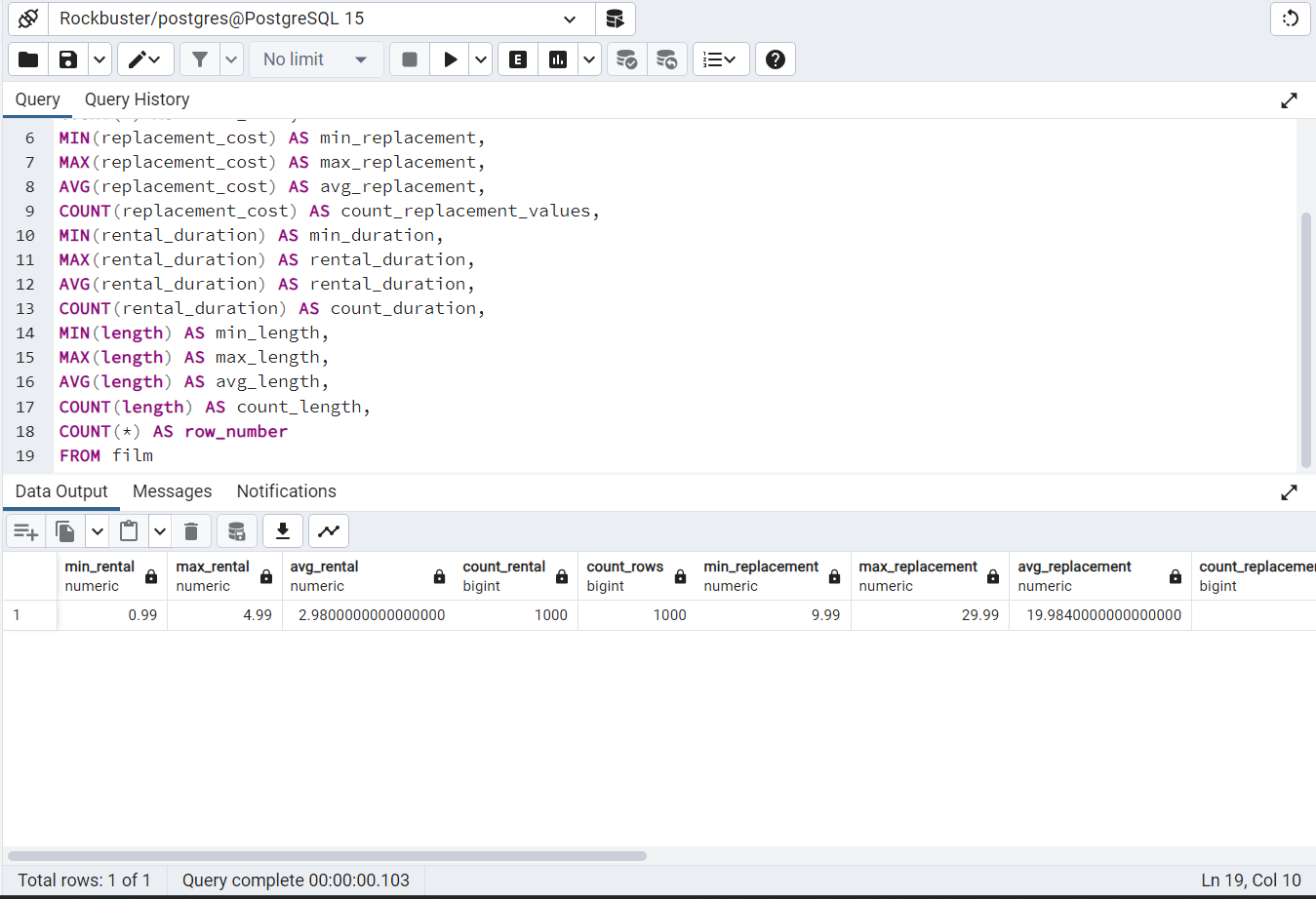


A screenshot of a computer

Description automatically generated

There appears to be no missing data. If there were missing data you could use the Update, Where, Set functions could be used to impute the average value for a small amount of missing data. However, you can also disregard a column or row that has a significant proportion of missing data.

2.a Film Table Numerical



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| min\_rental | max\_rental | avg\_rental | count\_rental | count\_rows | min\_replacement | max\_replacement |
| 0.99 | 4.99 | 2.98 | 1000 | 1000 | 9.99 | 29.99 |
| avg\_replacement | count\_replacement\_values | min\_duration | rental\_duration | rental\_duration-2 | count\_duration | min\_length |
| 19.984 | 1000 | 3 | 7 | 4.985 | 1000 | 46 |
| max\_length | avg\_length | count\_length | row\_number |
| 185 | 115.272 | 1000 | 1000 |

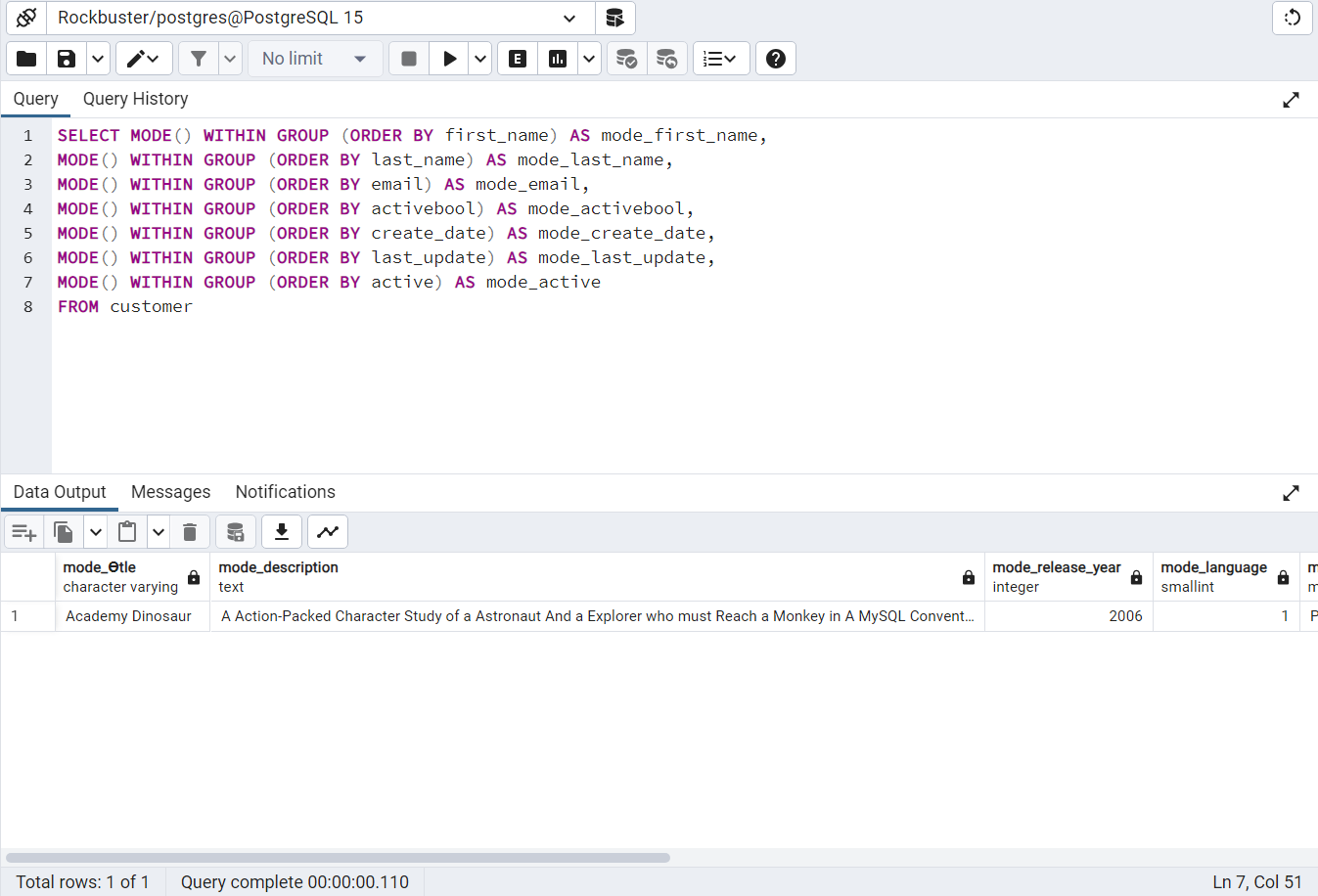
2.a Film Table Non-Numerical

A screenshot of a computer

Description automatically generated

|  |  |  |  |
| --- | --- | --- | --- |
| mode\_ÆŸtle | mode\_description | mode\_release\_year | mode\_language |
| Academy Dinosaur | A Action-Packed Character Study of a Astronaut And a Explorer who must Reach a Monkey in A MySQL Convention | 2006 | 1 |
| mode\_rating | mode\_last\_update | mode\_special\_features | mode\_fulltext |
| PG-13 | 50:59.0 | {Trailers,Commentaries,"Behind the Scenes"} | 'baloon':19 'confront':14 'documentari':5 'feminist':8,11,16 'mile':2 'must':13 'spi':1 'thrill':4 |

2.b Customer Table Non-Numerical



|  |  |  |  |
| --- | --- | --- | --- |
| mode\_ÆŸtle | mode\_description | mode\_release\_year | mode\_language |
| Academy Dinosaur | A Action-Packed Character Study of a Astronaut And a Explorer who must Reach a Monkey in A MySQL Convention | 2006 | 1 |
| mode\_rating | mode\_last\_update | mode\_special\_features | mode\_fulltext |
| PG-13 | 50:59.0 | {Trailers,Commentaries,"Behind the Scenes"} | 'baloon':19 'confront':14 'documentari':5 'feminist':8,11,16 'mile':2 'must':13 'spi':1 'thrill':4 |

3.a I find it is easier to use SQL as I continue profile data. Excel requires multiple steps including pivot tables and equations that you have to enter individually to get the same results as SQL which only takes you entering the equation in one solid flow versus multiple times. Also, SQL allows you to check multiple statistics of your data without having to use a pivot table. I believe (especially with a large data set like a data base) SQL is the way to go for efficiency purposes.