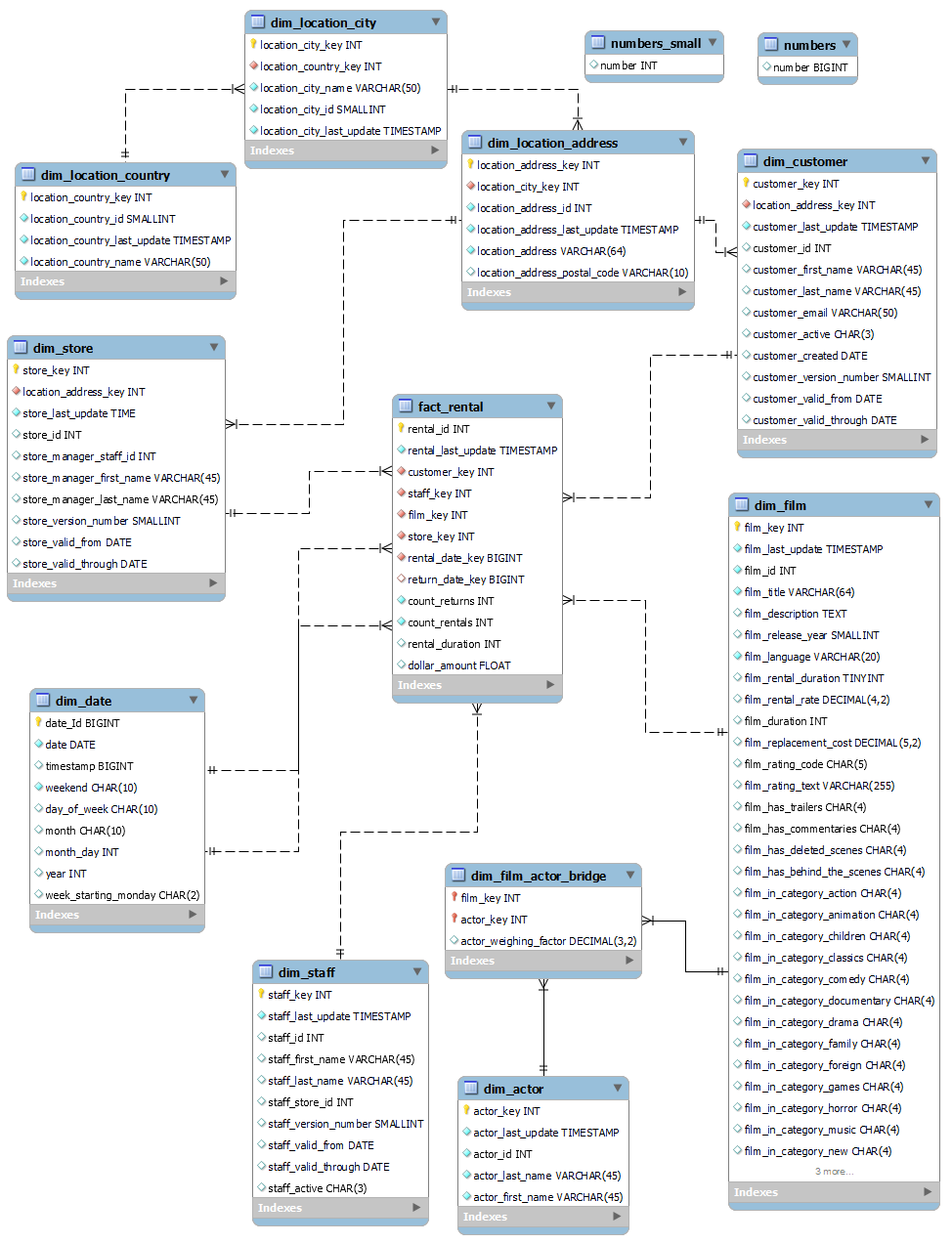
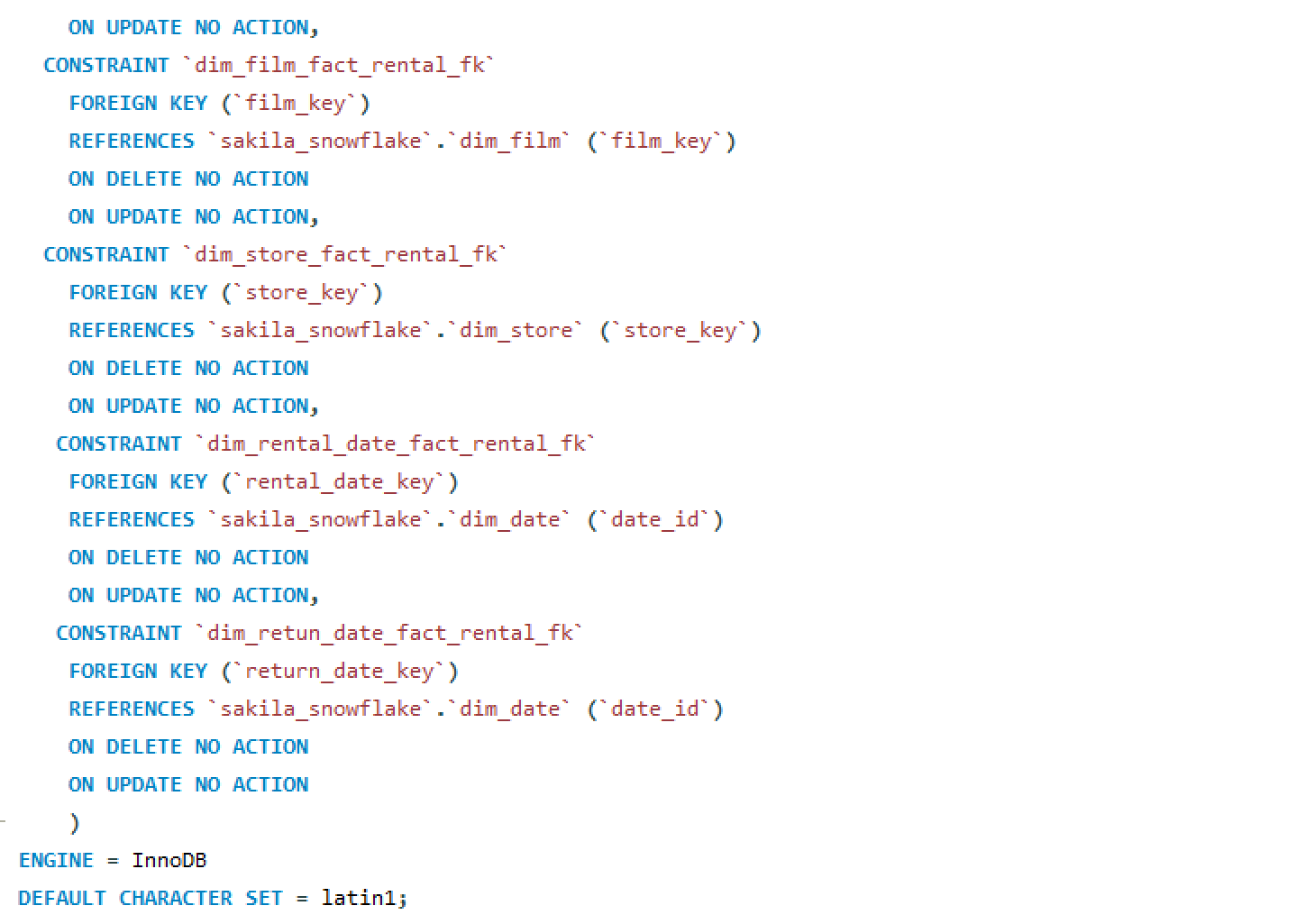
**Assignment 3**

**Individual:**

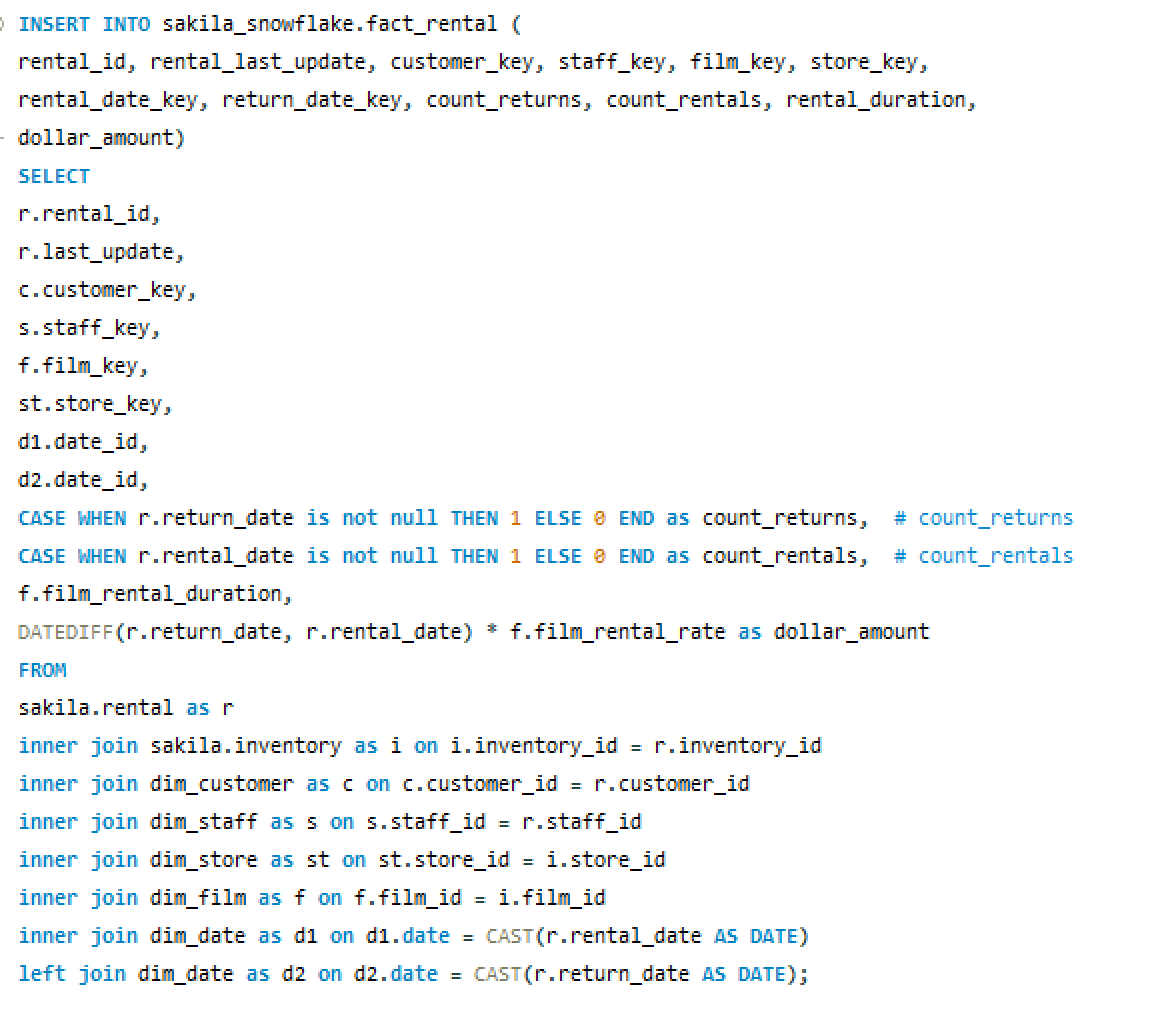
* **Screenshot of the EER diagram**



* **SQL (DDL/DML) queries used along with any assumptions & outputs (first few rows) for each question**
* **DDL:**



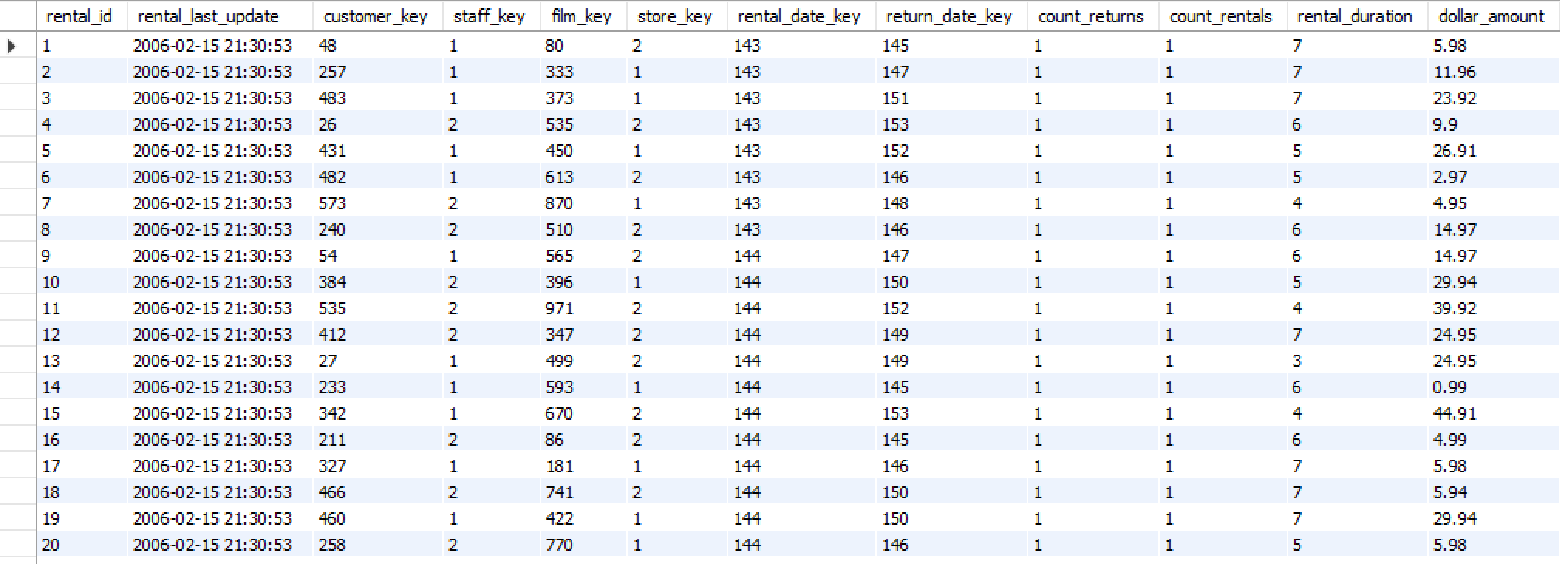
* **DML:**



* **Assumptions:**

For the dollar\_amount, if the return\_date is null in the original rental table, then the dollar\_amount would also be null. (use the datediff function and it would return null)

* **Output:**



* **Identify any scripts errors and areas of improvement in the data model.**

1. In the DML script, the actor information is inserted twice into the dim\_actor table.

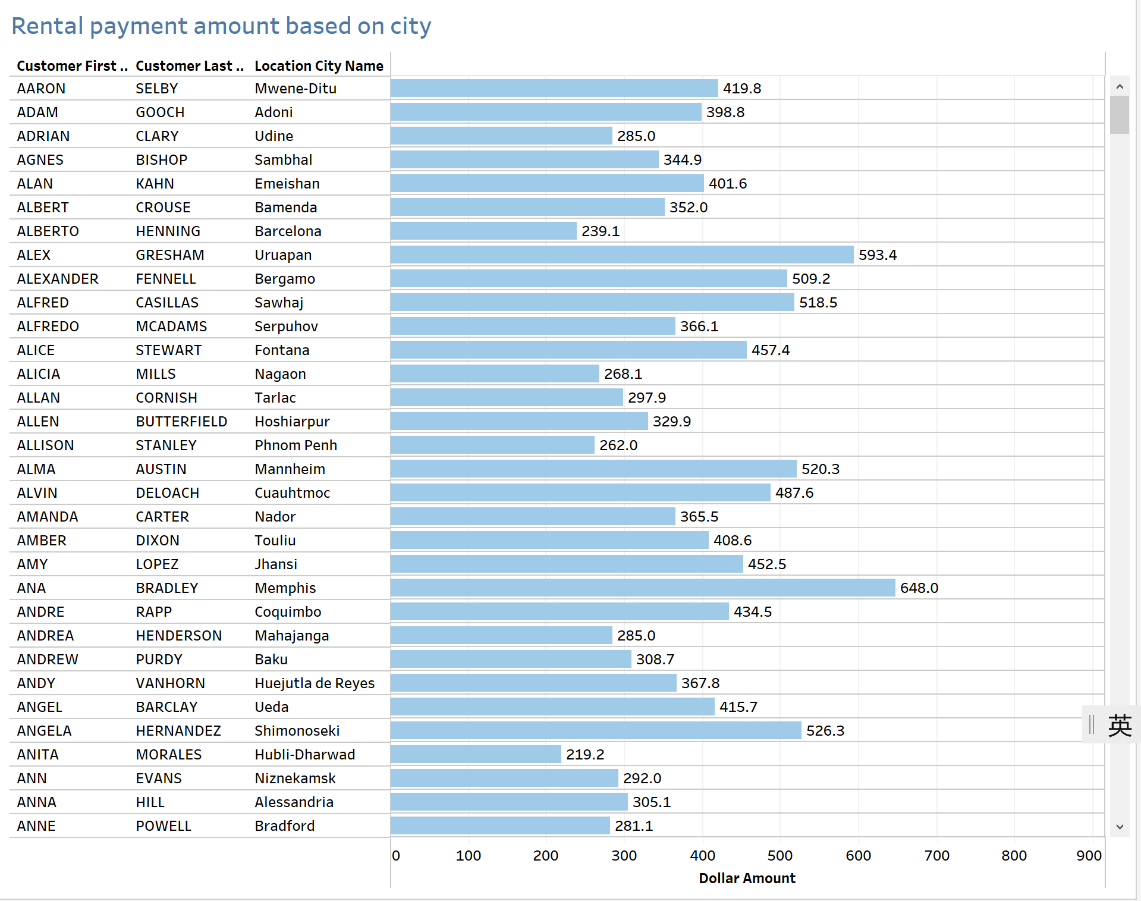
Solution is to delete the duplicated insertion query.

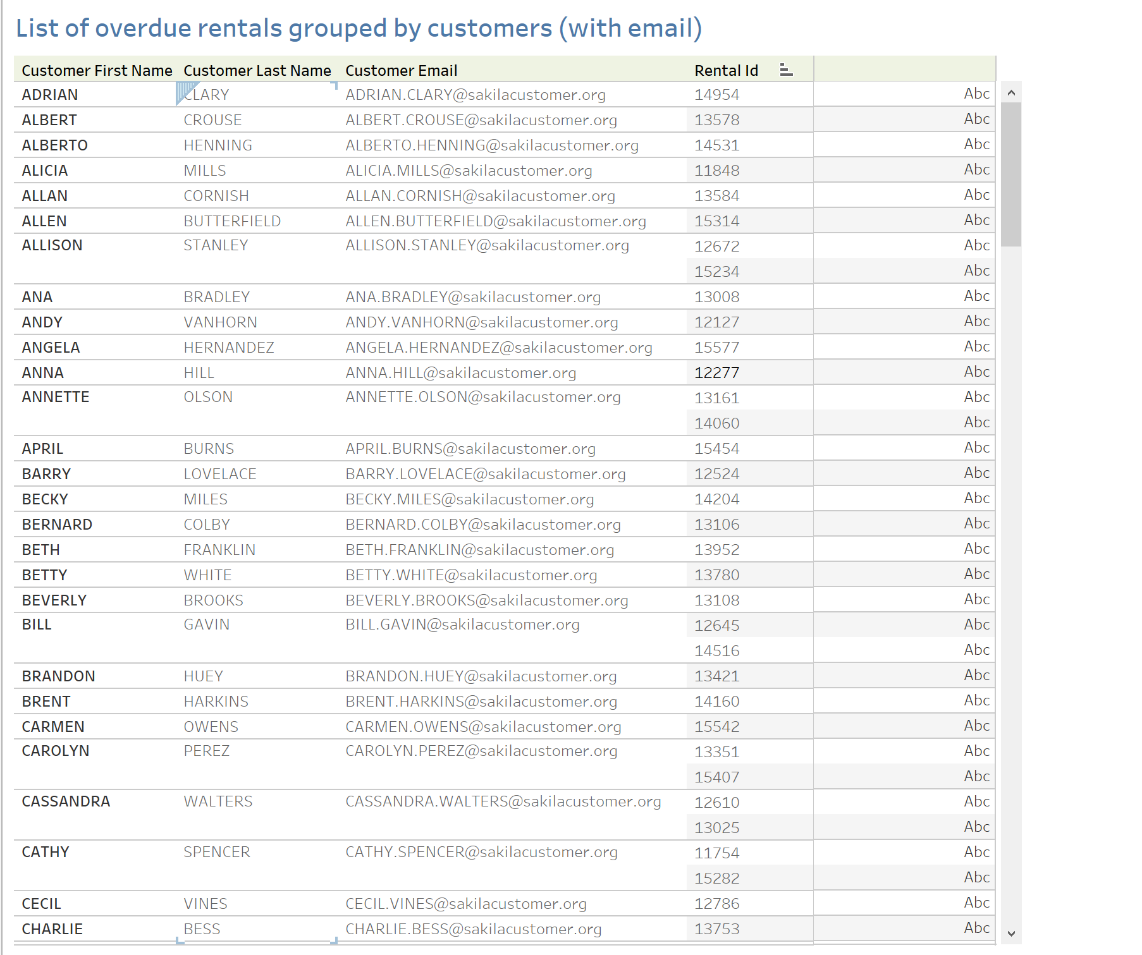
1. The current approach to indicate the film category in the dim\_film table is not very convenient.

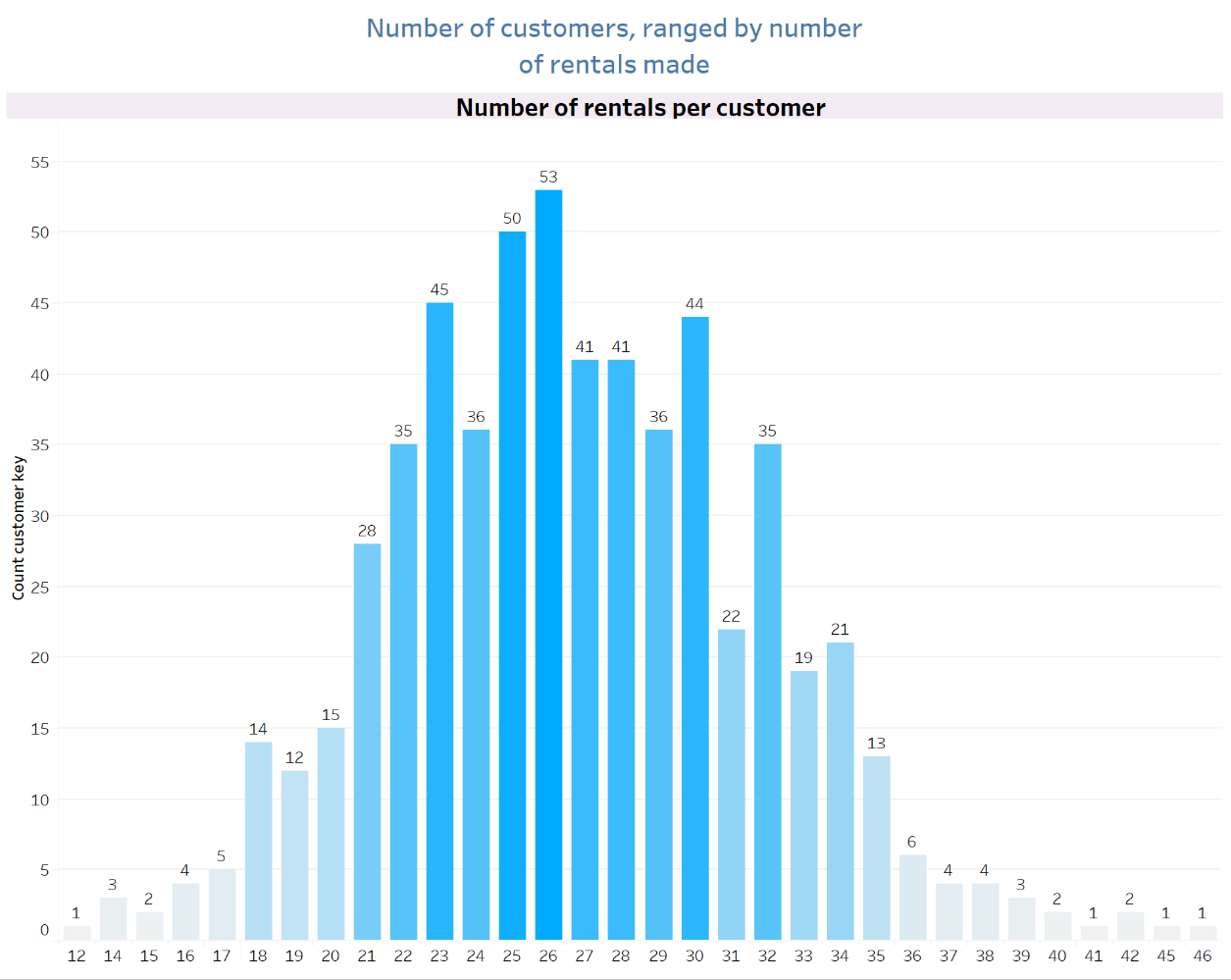
it would be better to use a new column category\_key (use different integers to represent different categories) instead of having these extra columns.

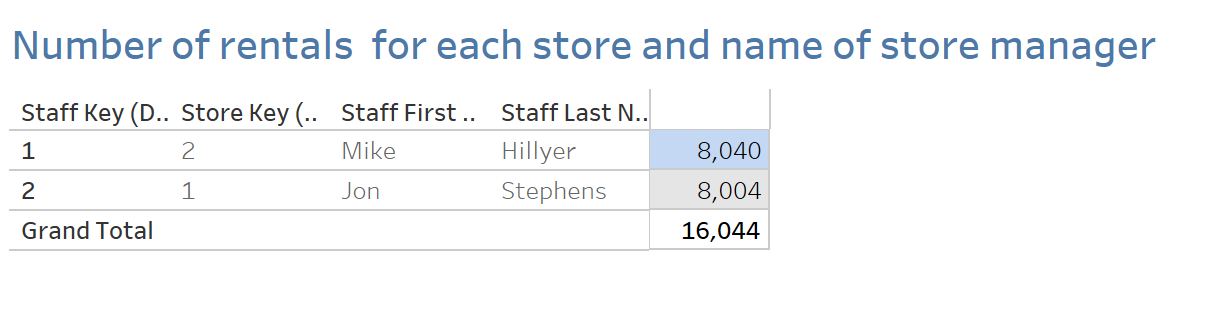
1. The current staff\_key and store\_key are not the same as their corresponding value in the old sakila table. (Mike Hillyer should have staff id of 1 and in the store with id 1, but it has store id 2 in the sakila\_snowflake).

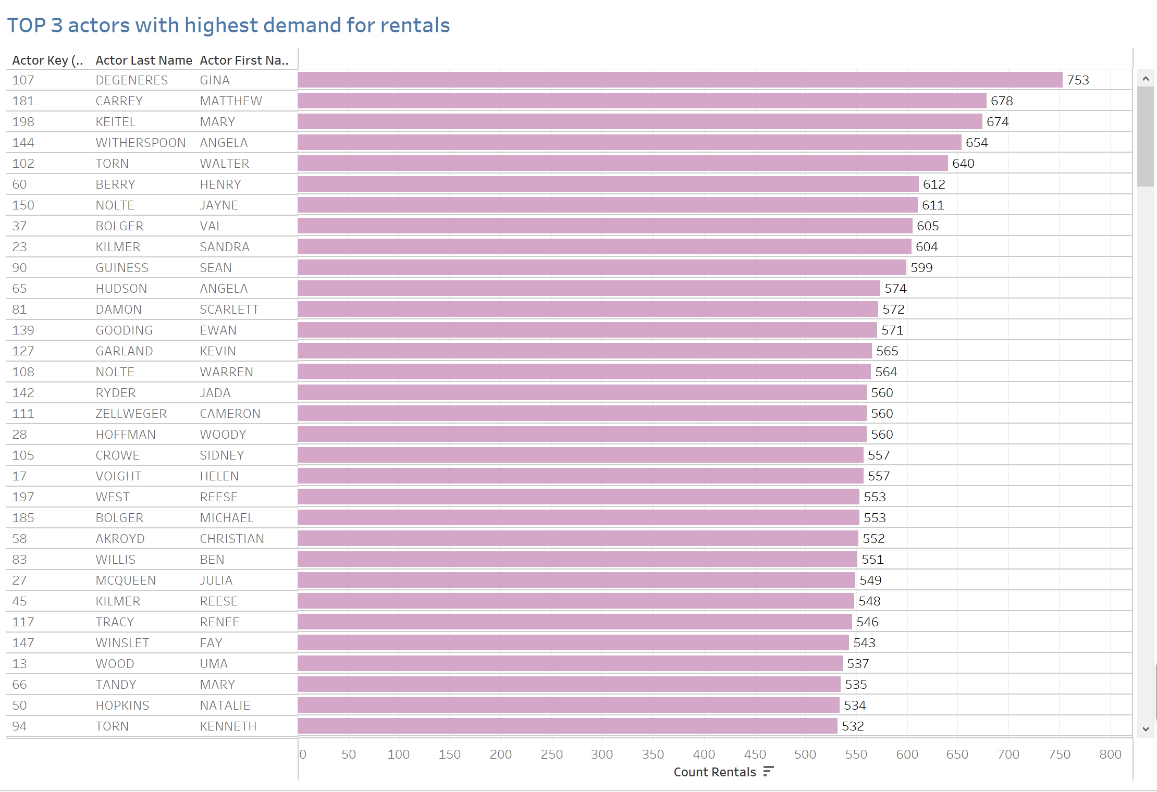
In the dim\_store, it should use `staff\_key` in dim\_staff as the identifier for different staffs instead of the old manager\_staff\_id from sakila.staff. But this doesn’t have very important effect on the data.

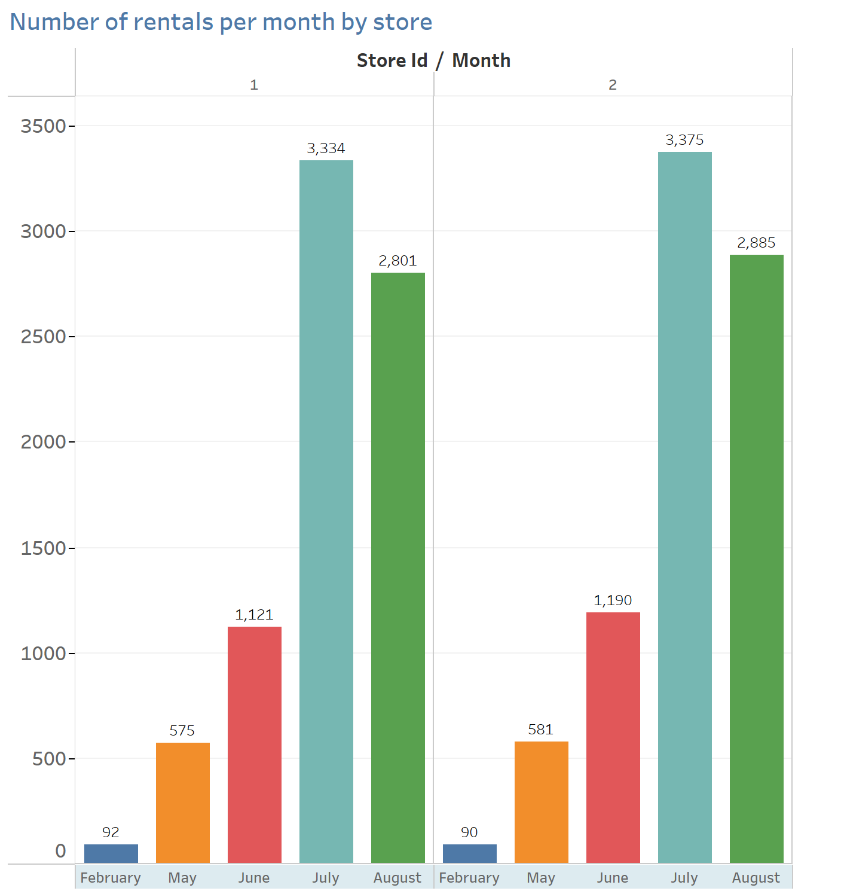
* **Screenshots of all Tableau reports and dashboard(s)**
* **Reports**

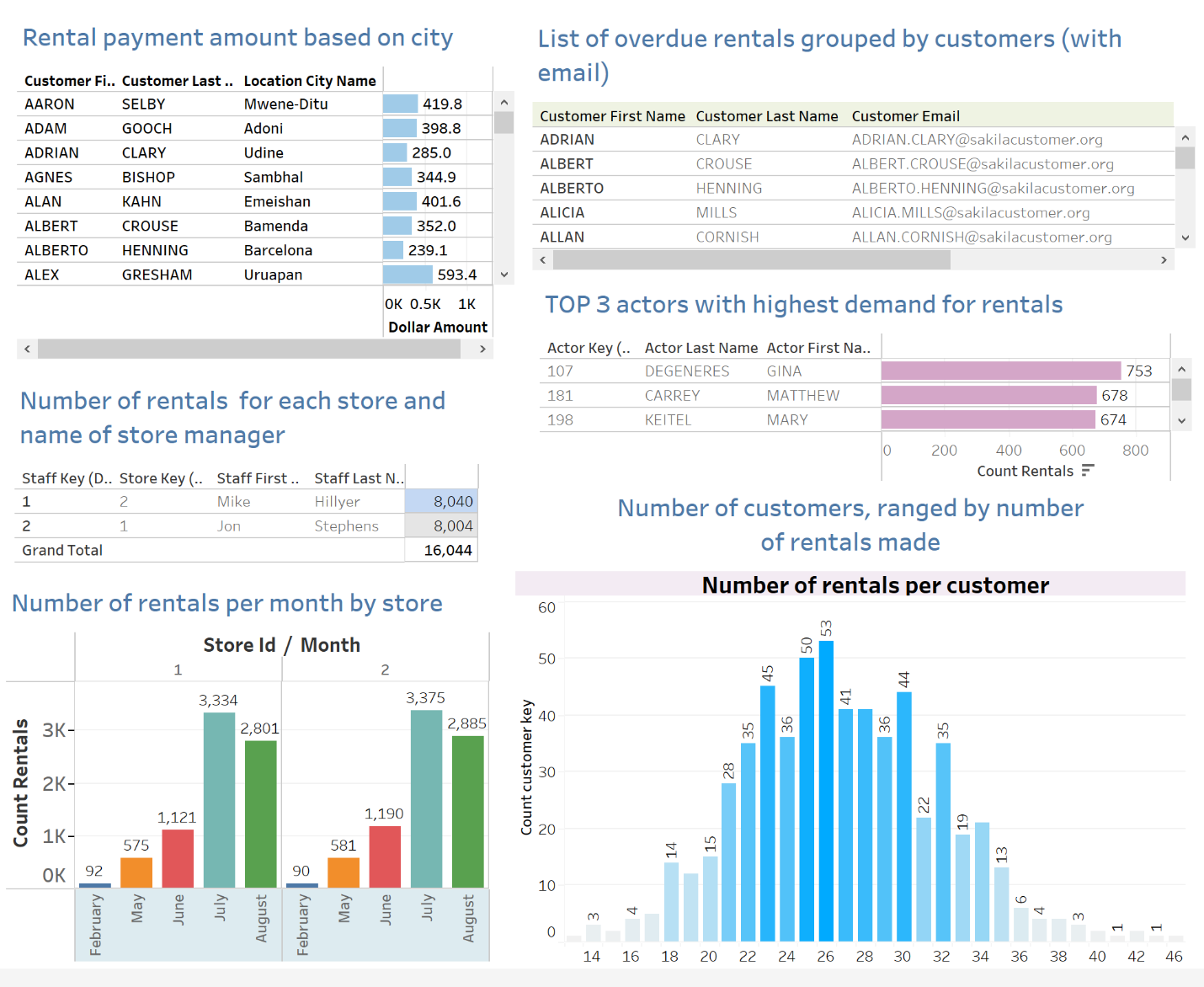
****

****

****

****

****

* **Dashboard:**