PROYECTO FINAL "OLIST: Brazilian E-commerce"

1. Fuente Base de Datos

La Base de Datos fue proporcionada por el profesor Andrés Ramírez.

2. Definición del problema

La directora del área de Mercadotecnia de OLIST está preocupada ante el aumento de reseñas negativas de los productos, indica que no desea perder ventas por esta situación, por el contrario, requiere estar informada para tomar medidas correctivas en bienestar de la empresa.

3. Definición Base de Datos

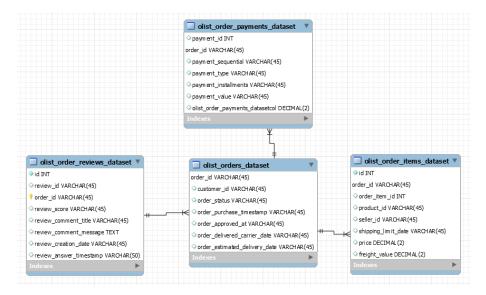
Los Datos provienen de cuatro tablas principales que son las siguientes:

olist_order_reviews_dataset: El archivo contiene la información referente a las puntuaciones/reseñas que los clientes han dado a la empresa Olist.

olist_orders_dataset: El archivo contiene la información de las ordenes de pedido, las fechas en las que se ha enviado, aprobado, cargado y su posible fecha de entrega.

olist order items dataset: El archivo contiene la información del precio y valor del flete de los productos.

olist_order_payments_dataset: El archivo contiene la información general del pago, forma de pago y valor.



4. Glosario

review_score: puntuación/reseña de calificación que los clientes indican al adquirir sus productos.

review comment title: de la puntuación/reseña es el título del comentario de los clientes.

review_comment_message: de la puntuación/reseña es el mensaje completo que dejan los clientes.

order_id: número de identificación de cada uno de los pedidos de los clientes.

order_status: estado del pedido en el que se encuentra un producto.

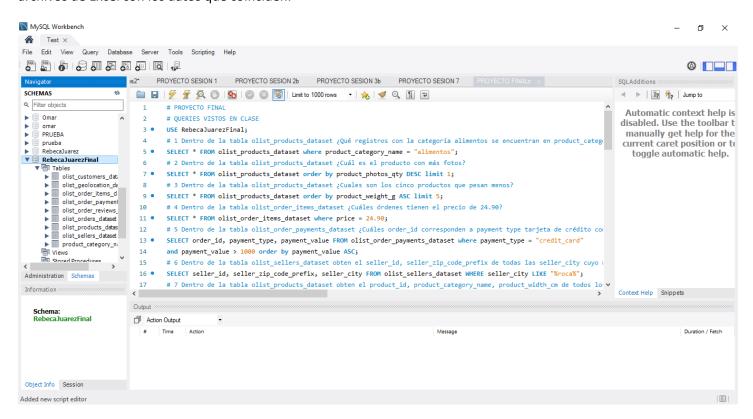
price: precio de cada orden de pedido.

payment_type: formas de pago de los clientes.

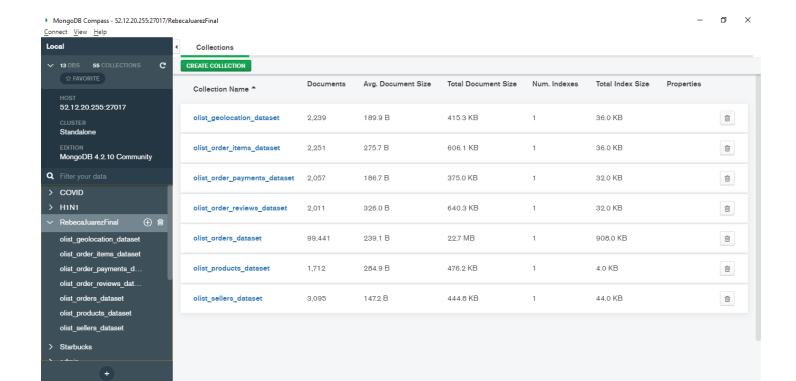
payment_value: valor del pago.

5. Creación Bases de Datos

Se realiza una depuración de tablas con 2,000 registros para que la carga en Workbench sea más rápida; limpiando los archivos de Excel con los datos que coinciden.



Se respetan las tablas que ya estaban listas para MySQL y se procede a subirlas; salvo la tabla olist_orders_dataset que cuenta con más documentos.



6. Consultas

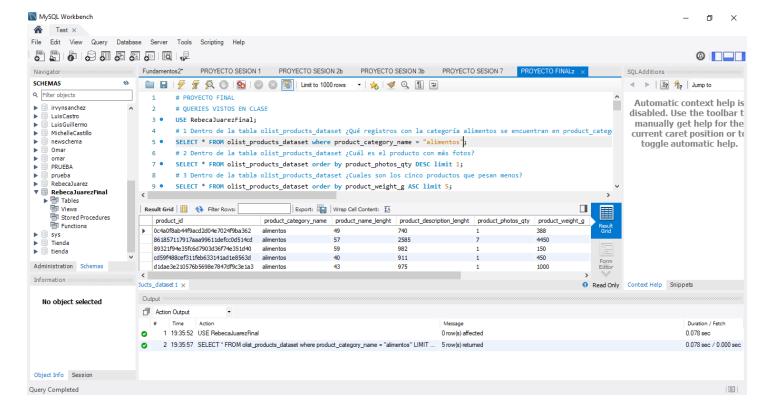
_MongoSH Beta

QUERIES VISTOS EN CLASE

USE RebecaJuarezFinal;

1 Dentro de la tabla olist_products_dataset ¿Qué registros con la categoría alimentos se encuentran en product_category_name?

SELECT * FROM olist_products_dataset where product_category_name = "alimentos";

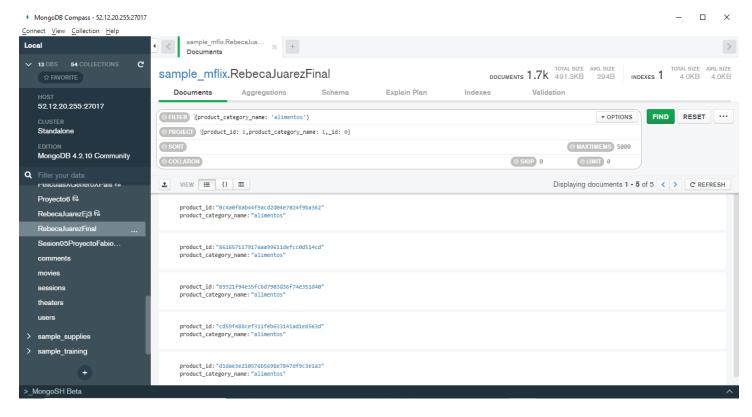


FILTER

```
{
product_category_name: 'alimentos'
}
```

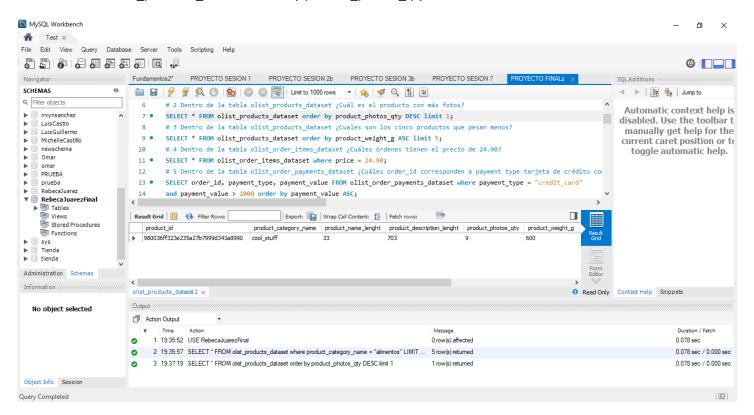
PROJECT

```
{
product_id: 1,
product_category_name: 1,
_id: 0
}
```



2 Dentro de la tabla olist_products_dataset ¿Cuál es el producto con más fotos?

SELECT * FROM olist_products_dataset order by product_photos_qty DESC limit 1;



PROJECT

```
{
product_id: 1,
```

```
product_photos_qty: 1,
_id: 0
}
SORT
product_photos_qty: -1
}
LIMIT

    MongoDB Compass - 52.12.20.255:27017/sample_mflix.RebecaJuarezFinal

 Connect View Collection Help
                                   sample_mflix.RebecaJua... +

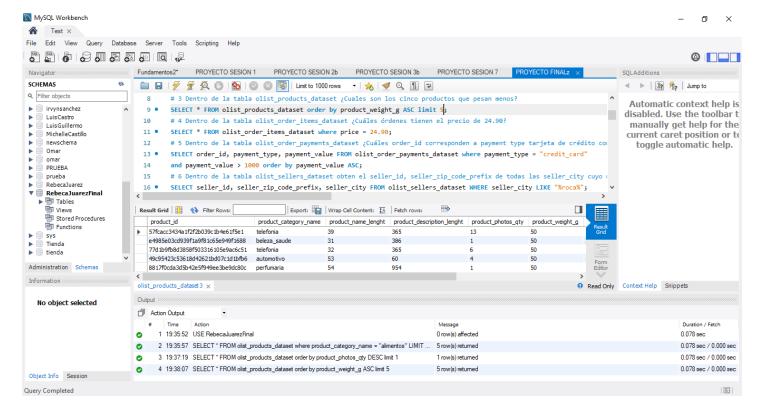
Documents
 Local
                                                                                                                            DOCUMENTS 1.7k TOTAL SIZE AVG. SIZE INDEXES 1 TOTAL SIZE AVG. SIZE 491.3KB 294B INDEXES 1 32.0KB 32.0KB
                                     sample_mflix.RebecaJuarezFinal
                                       Documents
                                                     Aggregations
                                                                                                Explain Plan
                                                                                                                     Indexes
                                                                                                                                       Validation
     52.12.20.255:27017
                                                                                                                                                                     FIND RESET ...
                                     6 FILTER
                                                                                                                                                         ▼ OPTIONS
                                     PROJECT { product_id: 1, product_photos_qty:1, _id: 0 }
                                     SORT |{ product_photos_qty: -1}
                                                                                                                                                (9 MAXTIMEMS) 5000
     MongoDB 4.2.10 Community
                                                                                                                                  SKIP 0

    COLLATION

                                                                                                                                                   G LIMIT 1
 Q Filter your data
                                     ± VIEW ∷ {} ⊞
                                                                                                                                             Displaying documents 1 - 1 of 1 ( ) C REFRESH
     FernandoHidalgoRe3 €
                                          product_id: "35537536ed2b4c561b4018bf3abf54e0"
product_photos_qty: "9"
     LuisCastroEi3
     PeliculasXGeneroXPais €
     Proyecto6 🗟
    RebecaJuarezEj3
     Sesion05ProyectoFabio.
  >_MongoSH Beta
```

3 Dentro de la tabla olist_products_dataset ¿Cuales son los cinco productos que pesan menos?

SELECT * FROM olist_products_dataset order by product_weight_g ASC limit 5;



PROJECT

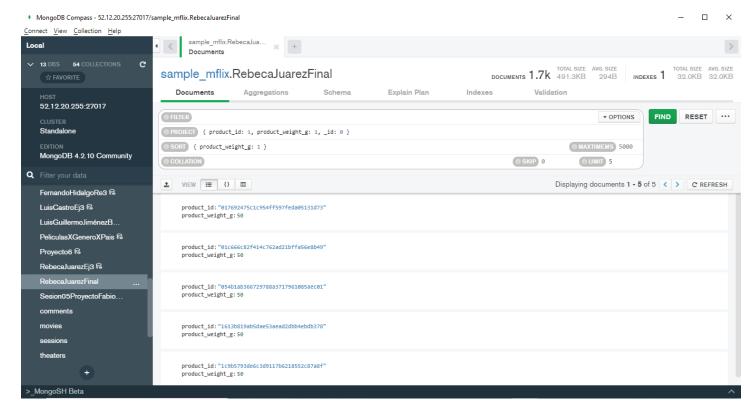
```
{
product_id: 1,
product_weight_g: 1,
_id: 0
}
```

SORT

```
{
product_weight_g: 1
}
```

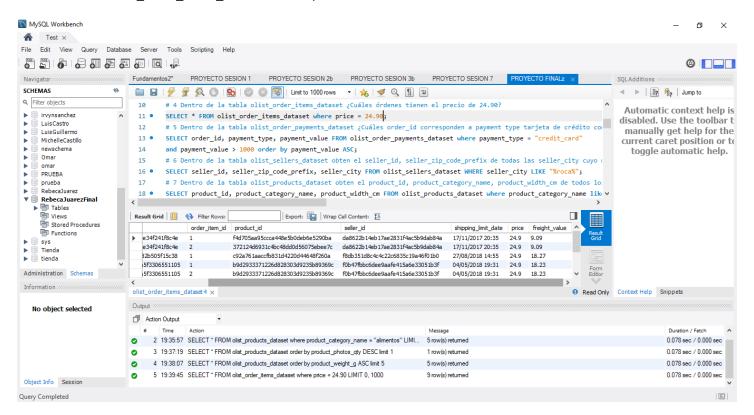
LIMIT

5



4 Dentro de la tabla olist_order_items_dataset ¿Cuáles órdenes tienen el precio de 24.90?

SELECT * FROM olist_order_items_dataset where price = 24.90;

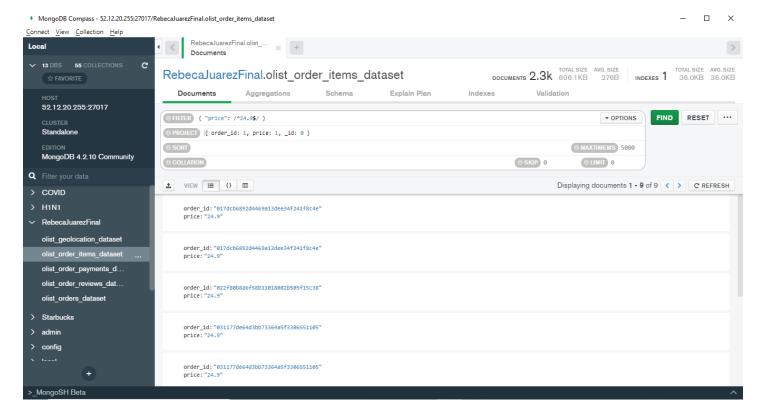


FILTER

```
{
price: RegExp('^24.9$')
```

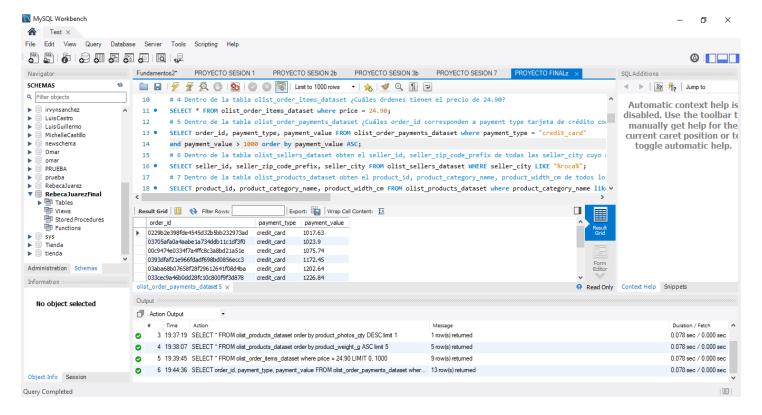
```
PROJECT

{
    order_id: 1,
    price: 1,
    _id: 0
}
```



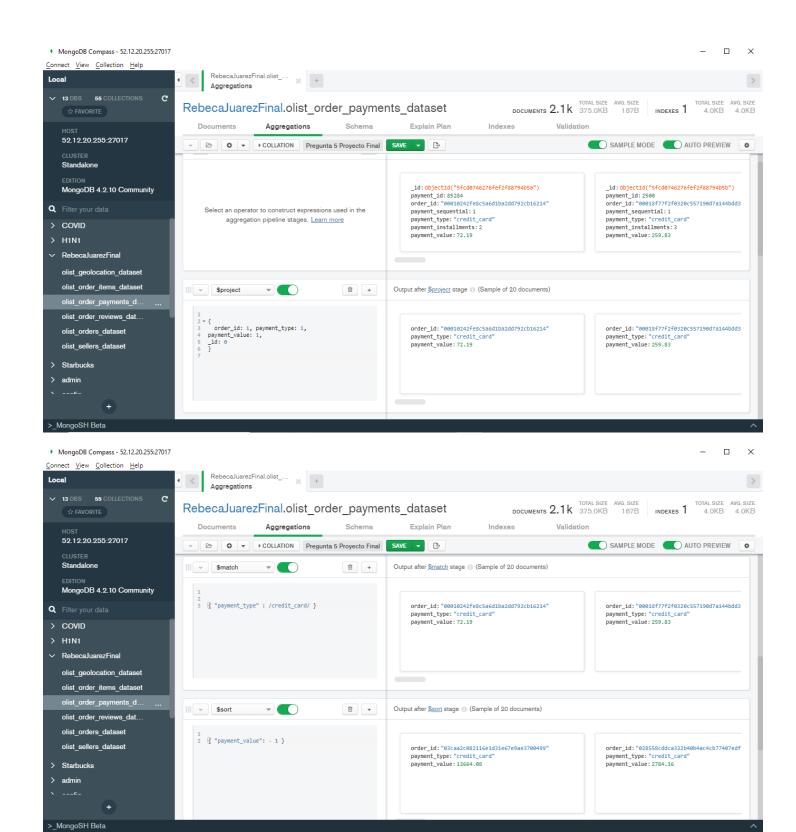
5 Dentro de la tabla olist_order_payments_dataset ¿Cuáles order_id corresponden a payment type tarjeta de crédito con valor de pago mayor que 1000?

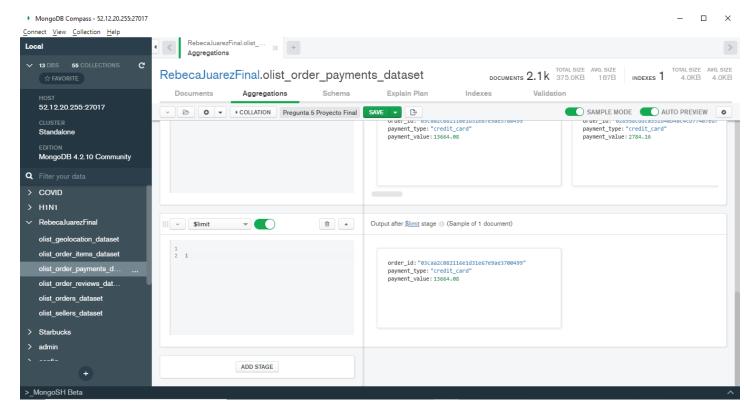
SELECT order_id, payment_type, payment_value FROM olist_order_payments_dataset where payment_type = "credit_card" and payment_value > 1000 order by payment_value ASC;



6 Dentro de la tabla olist_order_payments_dataset ¿Cuáles order_id corresponden a payment type tarjeta de crédito y payment_value mayor?

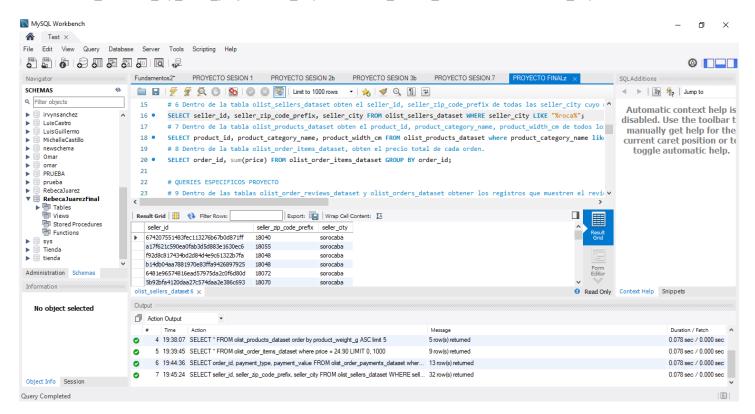
```
[{$project:
{
    order_id: 1, payment_type: 1,
    payment_value: 1,
    _id: 0
}
}, {$match:
{ "payment_type" : /credit_card/ }}, {$sort:
{ "payment_value": - 1 }}, {$limit:
1}]
```





7 Dentro de la tabla olist_sellers_dataset obten el seller_id, seller_zip_code_prefix de todas las seller_city cuyo nombre incluye la cadena roca

SELECT seller_id, seller_zip_code_prefix, seller_city FROM olist_sellers_dataset WHERE seller_city LIKE "%roca%";



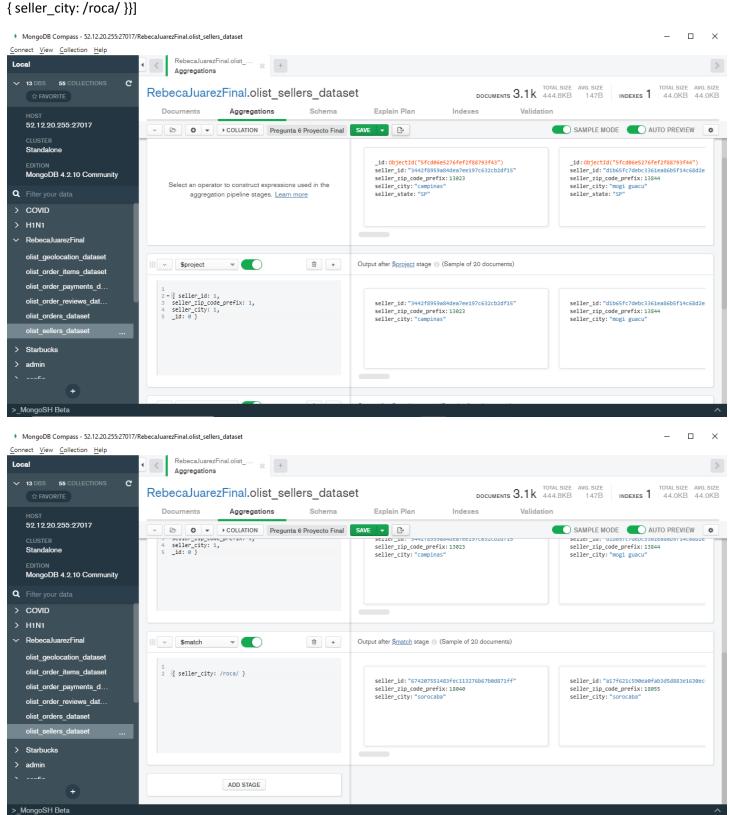
[{\$project:

{ seller_id: 1,

seller_zip_code_prefix: 1,

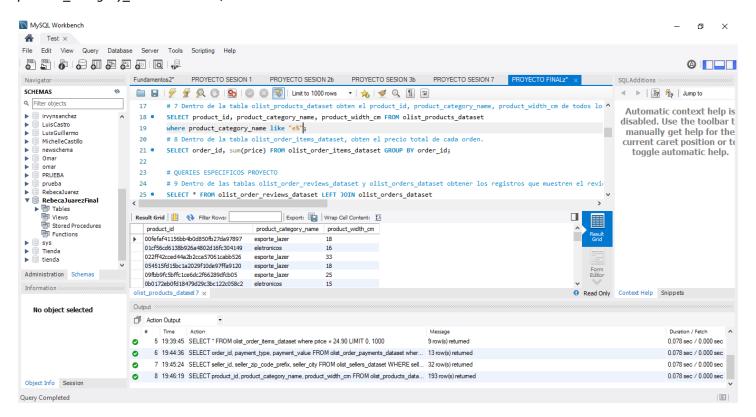
seller_city: 1,

_id: 0 }}, {\$match:



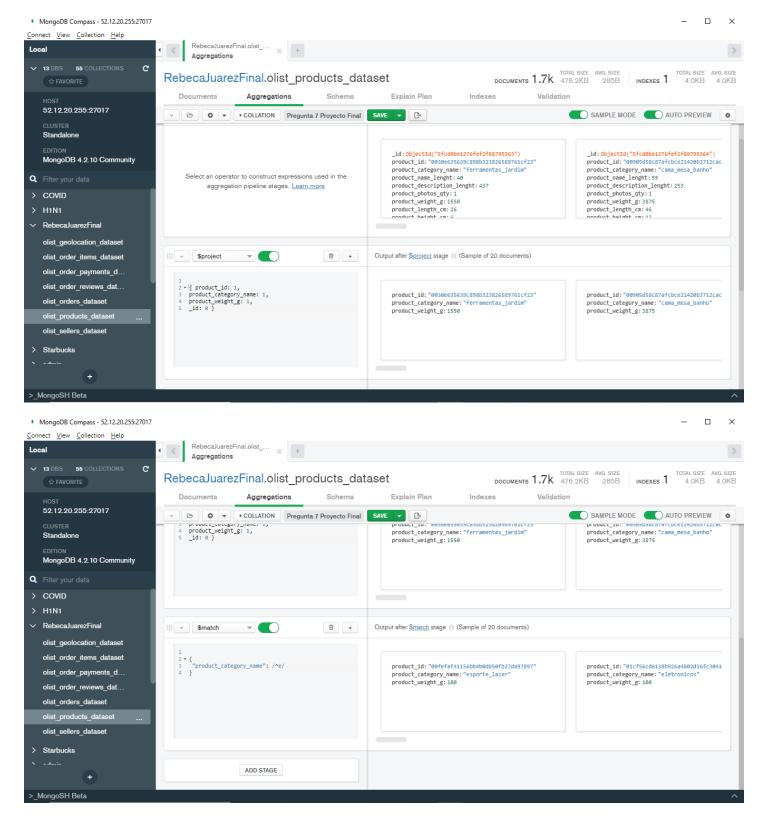
8 Dentro de la tabla olist_products_dataset obten el product_id, product_category_name, product_width_cm de todos los product_category_name cuyo nombre inicia con e

SELECT product_id, product_category_name, product_width_cm FROM olist_products_dataset where product_category_name like "e%";

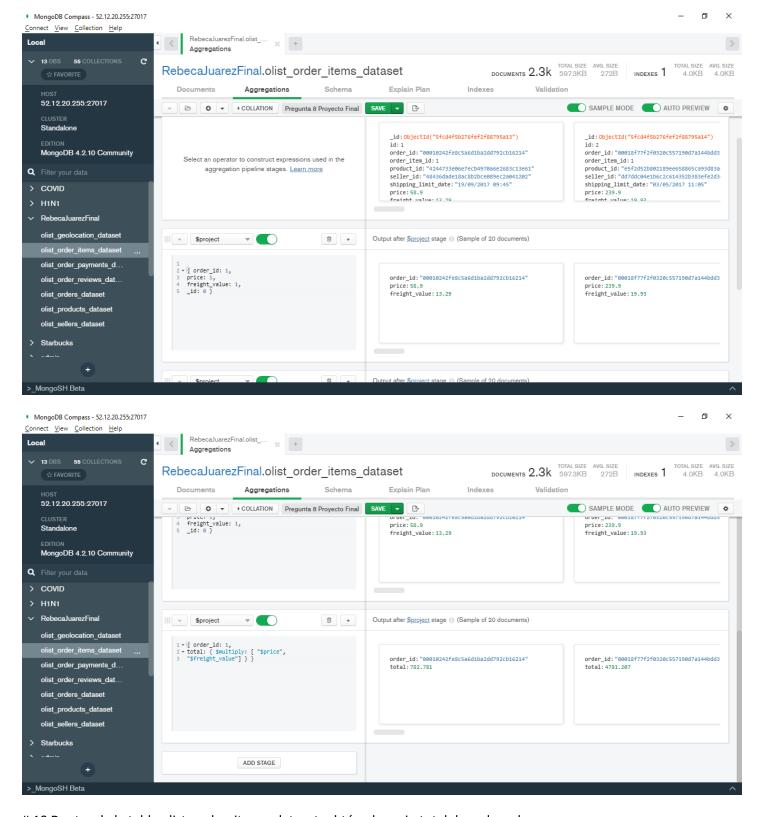


```
[{$project:
```

```
{ product_id: 1, product_category_name: 1, product_weight_g: 1, _id: 0 }}, {$match: { "product_category_name": /^e/ }}]
```

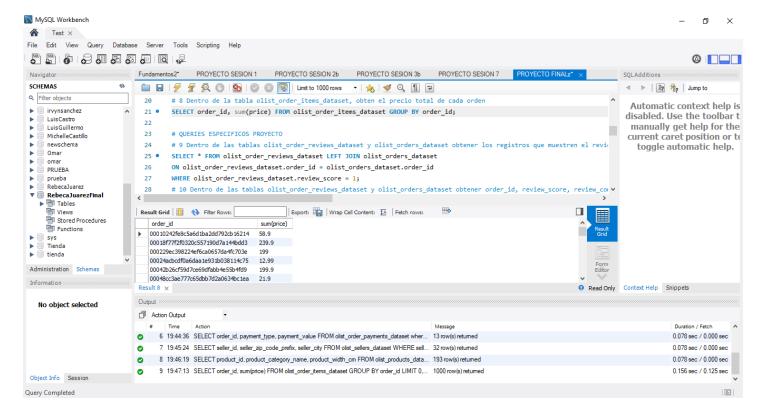


9 Dentro de la table olist_order_items_dataset, obtén por order_id el precio total con flete



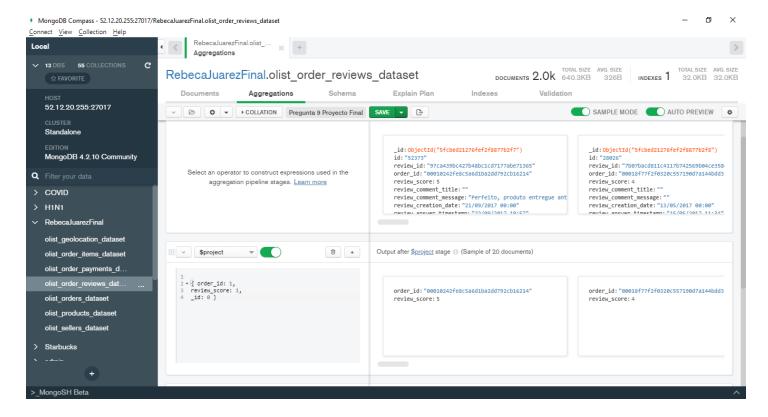
10 Dentro de la tabla olist_order_items_dataset, obtén el precio total de cada orden

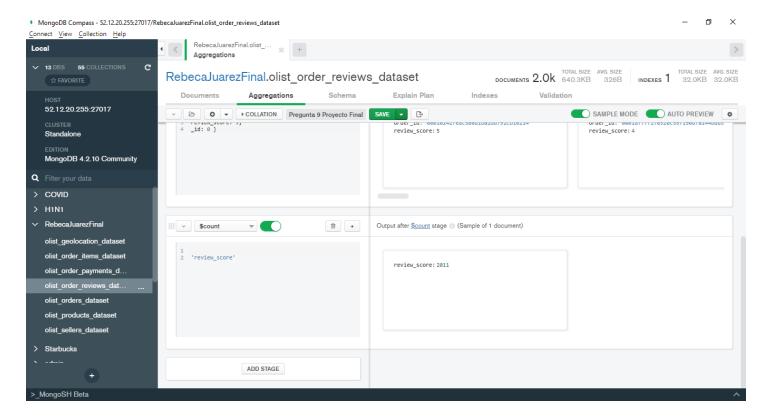
SELECT order_id, sum(price) FROM olist_order_items_dataset GROUP BY order_id;



QUERIES ESPECIFICOS PROYECTO

11 De la tabla de olist_order_reviews_dataset obtener con MongoDB el número total de reseñas

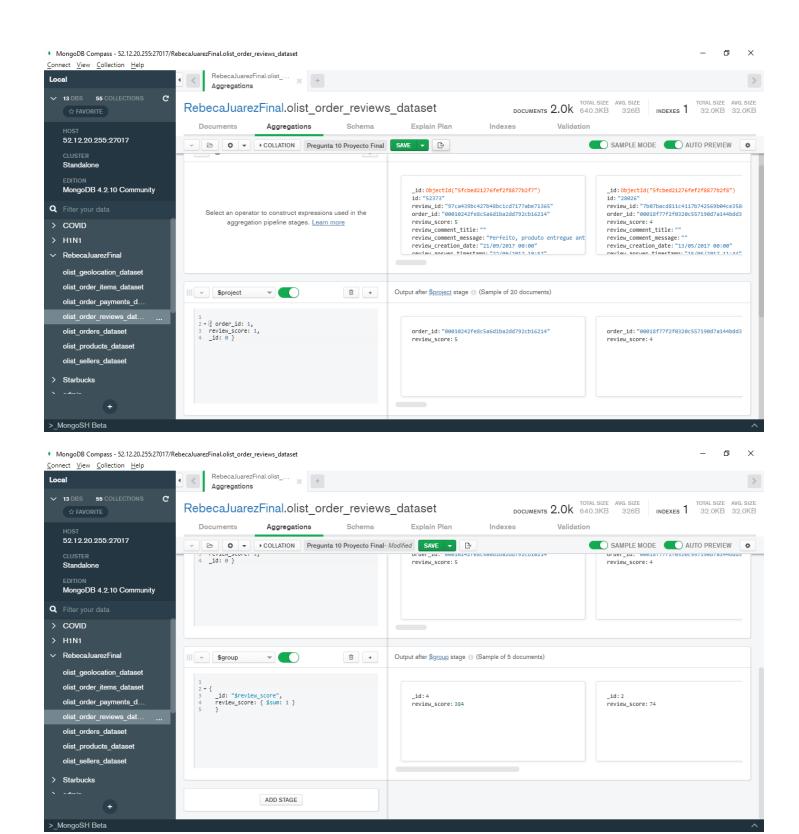


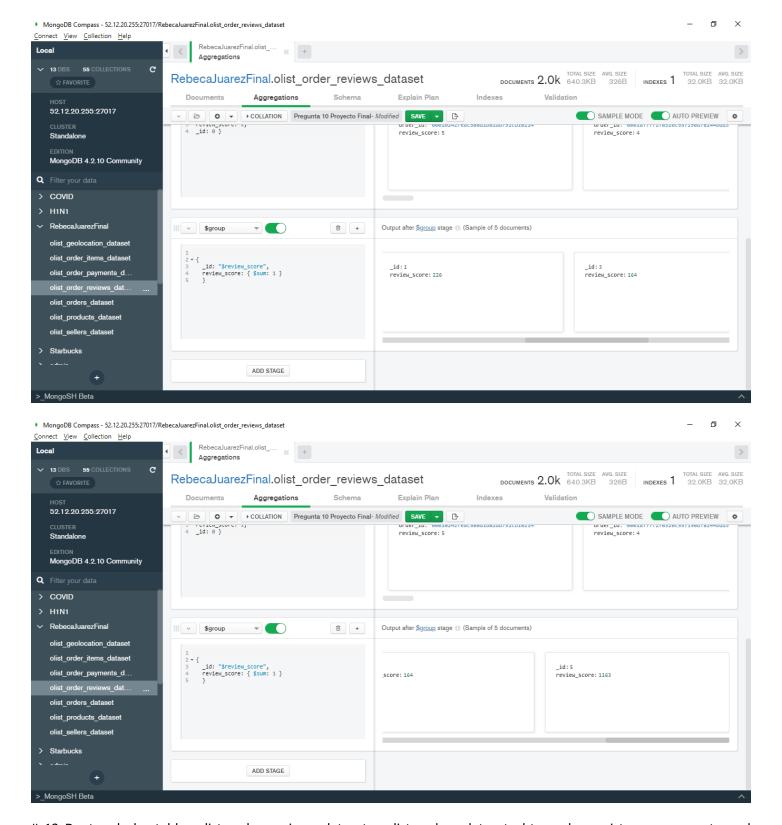


12 De la tabla de olist_order_reviews_dataset obtener con MongoDB el número de review_score por calificación

```
{ order_id: 1,
review_score: 1,
_id: 0 }}, {$group:
{
_id: "$review_score",
review_score: { $sum: 1 }
}}]
```

[{\$project:



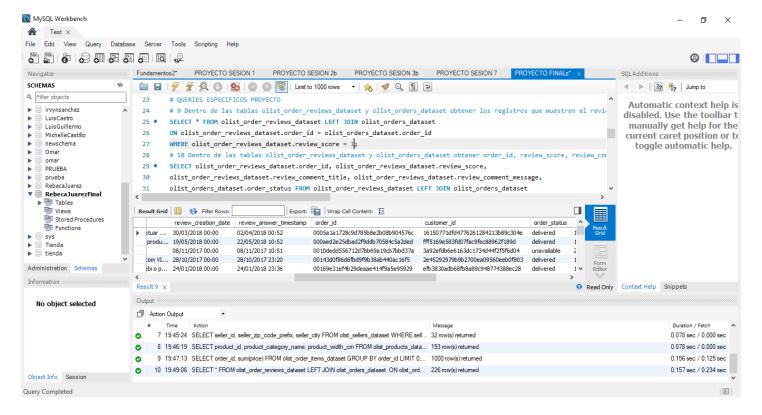


13 Dentro de las tablas olist_order_reviews_dataset y olist_orders_dataset obtener los registros que muestren el review_score más bajo con su respectivo order_status

SELECT * FROM olist_order_reviews_dataset LEFT JOIN olist_orders_dataset

ON olist_order_reviews_dataset.order_id = olist_orders_dataset.order_id

WHERE olist_order_reviews_dataset.review_score = 1;



14 Dentro de las tablas olist_order_reviews_dataset y olist_orders_dataset obtener order_id, review_score, review_comment_title, review_comment_message con su respectivo order_status

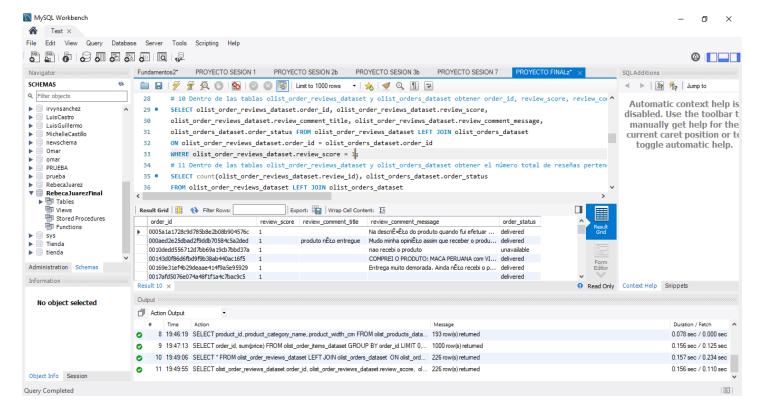
SELECT olist_order_reviews_dataset.order_id, olist_order_reviews_dataset.review_score,

olist_order_reviews_dataset.review_comment_title, olist_order_reviews_dataset.review_comment_message,

olist_orders_dataset.order_status FROM olist_order_reviews_dataset LEFT JOIN olist_orders_dataset

ON olist_order_reviews_dataset.order_id = olist_orders_dataset.order_id

WHERE olist order reviews dataset.review score = 1;



15 Dentro de las tablas olist_order_reviews_dataset y olist_orders_dataset obtener el número total de reseñas pertenecen a cada uno de los order_status

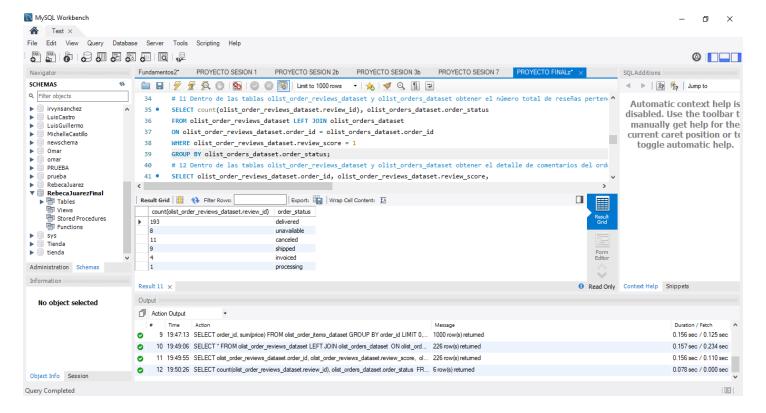
SELECT count(olist_order_reviews_dataset.review_id), olist_orders_dataset.order_status

FROM olist_order_reviews_dataset LEFT JOIN olist_orders_dataset

ON olist_order_reviews_dataset.order_id = olist_orders_dataset.order_id

WHERE olist_order_reviews_dataset.review_score = 1

GROUP BY olist orders dataset.order status;



16 Dentro de las tablas olist_order_reviews_dataset y olist_orders_dataset obtener el detalle de comentarios del order_status delivered para entregar detalle a la Directora

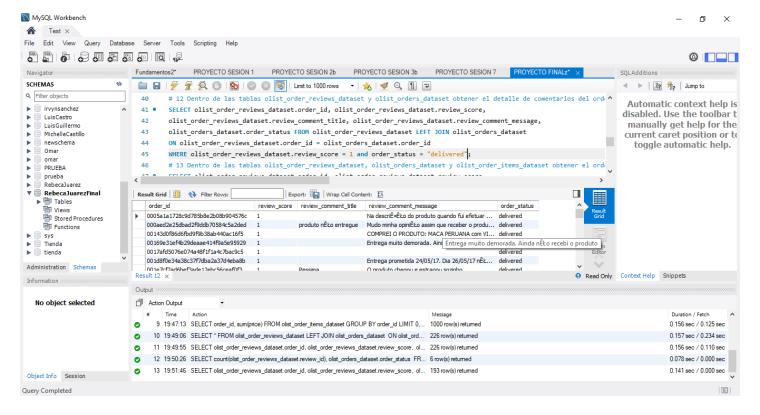
SELECT olist_order_reviews_dataset.order_id, olist_order_reviews_dataset.review_score,

olist_order_reviews_dataset.review_comment_title, olist_order_reviews_dataset.review_comment_message,

olist_orders_dataset.order_status FROM olist_order_reviews_dataset LEFT JOIN olist_orders_dataset

ON olist_order_reviews_dataset.order_id = olist_orders_dataset.order_id

WHERE olist order reviews dataset.review score = 1 and order status = "delivered";



17 Dentro de las tablas olist_order_reviews_dataset, olist_orders_dataset y olist_order_items_dataset obtener el order_id, review_score y price del review_score más bajo

SELECT olist_order_reviews_dataset.order_id, olist_order_reviews_dataset.review_score,

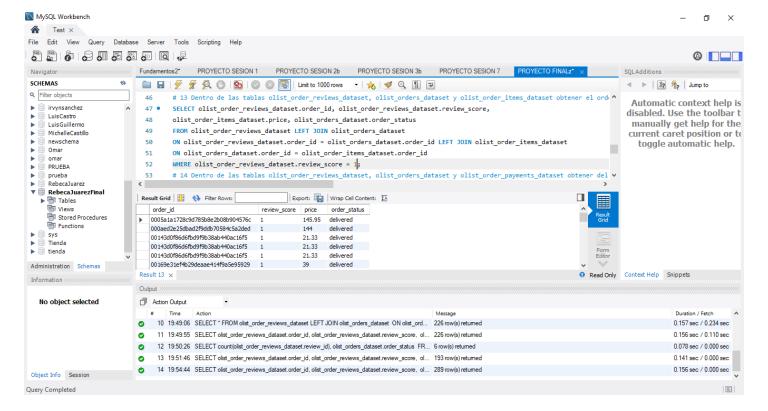
olist_order_items_dataset.price, olist_orders_dataset.order_status

FROM olist_order_reviews_dataset LEFT JOIN olist_orders_dataset

ON olist_order_reviews_dataset.order_id = olist_orders_dataset.order_id LEFT JOIN olist_order_items_dataset

ON olist_orders_dataset.order_id = olist_order_items_dataset.order_id

WHERE olist_order_reviews_dataset.review_score = 1;



18 Dentro de las tablas olist_order_reviews_dataset, olist_orders_dataset y olist_order_payments_dataset obtener del order_status el tipo de pago y el valor del pago por order_id

SELECT olist_order_reviews_dataset.order_id, olist_order_reviews_dataset.review_score,

olist_orders_dataset.order_status, olist_order_payments_dataset.payment_type,

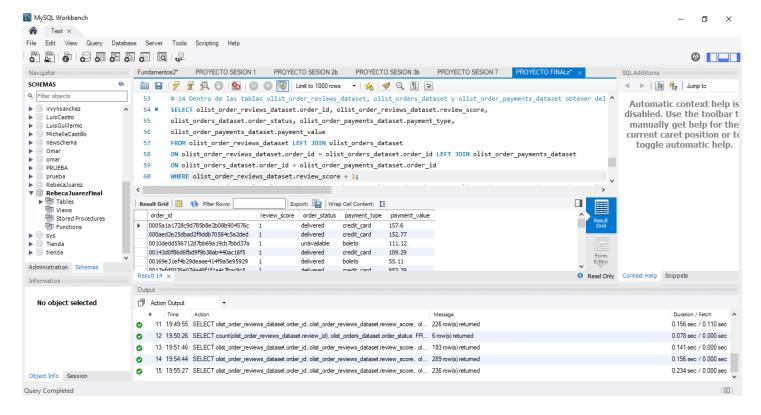
olist_order_payments_dataset.payment_value

FROM olist_order_reviews_dataset LEFT JOIN olist_orders_dataset

ON olist order reviews dataset.order id = olist orders dataset.order id LEFT JOIN olist order payments dataset

ON olist_orders_dataset.order_id = olist_order_payments_dataset.order_id

WHERE olist_order_reviews_dataset.review_score = 1;



19 Dentro de las tablas olist_order_reviews_dataset, olist_orders_dataset y olist_order_payments_dataset obtener los 10 payment_value más altos

SELECT olist_order_reviews_dataset.order_id, olist_order_reviews_dataset.review_score,

olist_orders_dataset.order_status, olist_order_payments_dataset.payment_type,

olist_order_payments_dataset.payment_value

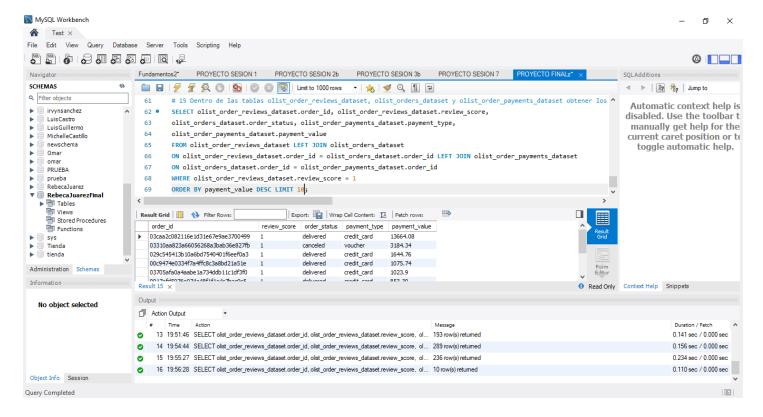
FROM olist_order_reviews_dataset LEFT JOIN olist_orders_dataset

ON olist order reviews dataset.order id = olist orders dataset.order id LEFT JOIN olist order payments dataset

ON olist_orders_dataset.order_id = olist_order_payments_dataset.order_id

WHERE olist_order_reviews_dataset.review_score = 1

ORDER BY payment_value DESC LIMIT 10;



20 Utilizando la posición actual de la filial Olist en Sao Paulo coordenadas -23.519327, -46.305635, encuentra las localizaciones de los clientes más cercanas.

```
[{$project:
{id: 1,
geolocation_lat: 1,
geolocation_lng: 1,
geolocation_city: 1,
_id: 0 }}, {$match:
{
    $and:
[
    {geolocation_lat: { $gte: "-23.5" }},
    {geolocation_lng: { $gte: "-46.6" }},
    {geolocation_lat: { $lte: "-23.51" }},
}
{ geolocation_lng: { $lte: "-46.63" }}
]
```

