Big Data – Spring 2022

Prof. Arturo Castellanos

Brazilian E-Commerce Public Dataset by Olist

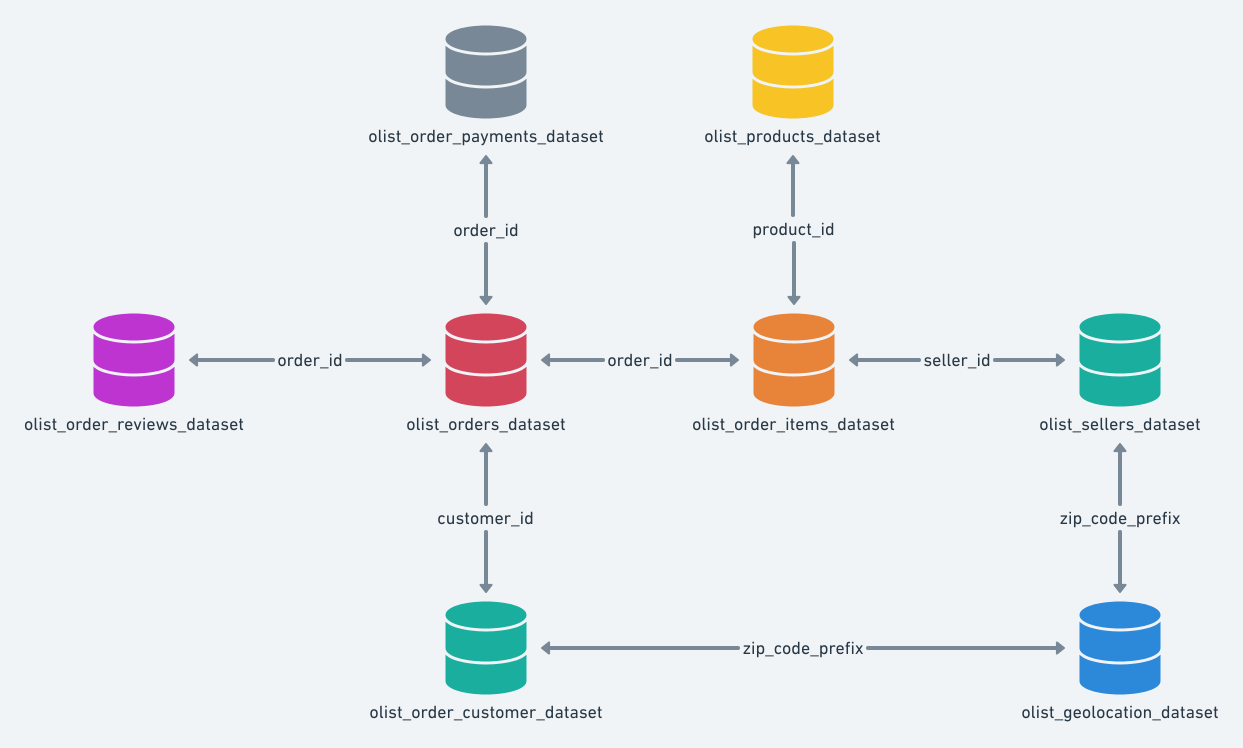
A Brazilian ecommerce company, [Olist Store](http://www.olist.com/), requires your expertise to find interesting insights regarding their sales and growth. The database provided contains information of about 100k orders from 2016 to 2018 made at multiple marketplaces in Brazil. The database allows you to see the order transactions from multiple dimensions: from order status, price, payment and freight performance to customer location, product attributes and finally reviews written by customers. They also provided geolocation information that relates Brazilian zip codes to latitude/longitude coordinates.

Olist ([www.olist.com](https://www.olist.com/)) is the largest department store in Brazilian marketplaces. Olist connects small businesses from all over Brazil directly to consumers. Merchants can sell their products through the Olist Store and ship them directly to the customers using Olist logistics partners.

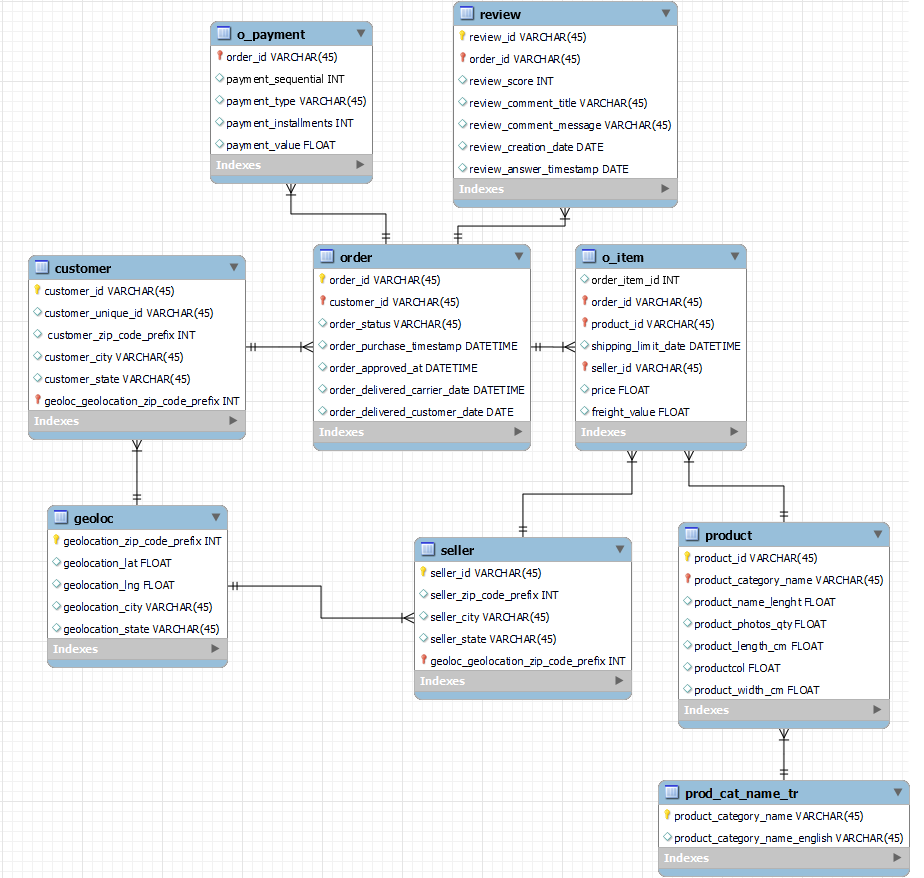
After a customer purchases the product from Olist Store a seller gets notified to fulfill that order. Once the customer receives the product, or the estimated delivery date is due, the customer gets a satisfaction survey by email where he can give a note for the purchase experience and write down some comments.

This is real commercial data, it has been anonymized, and references to the companies and partners in the review text have been replaced with the names of Game of Thrones great houses. The data was sourced from Kaggle and it is exclusively for learning purposes. To learn more on the terms of use go to [Kaggle](https://www.kaggle.com/olistbr/brazilian-ecommerce).

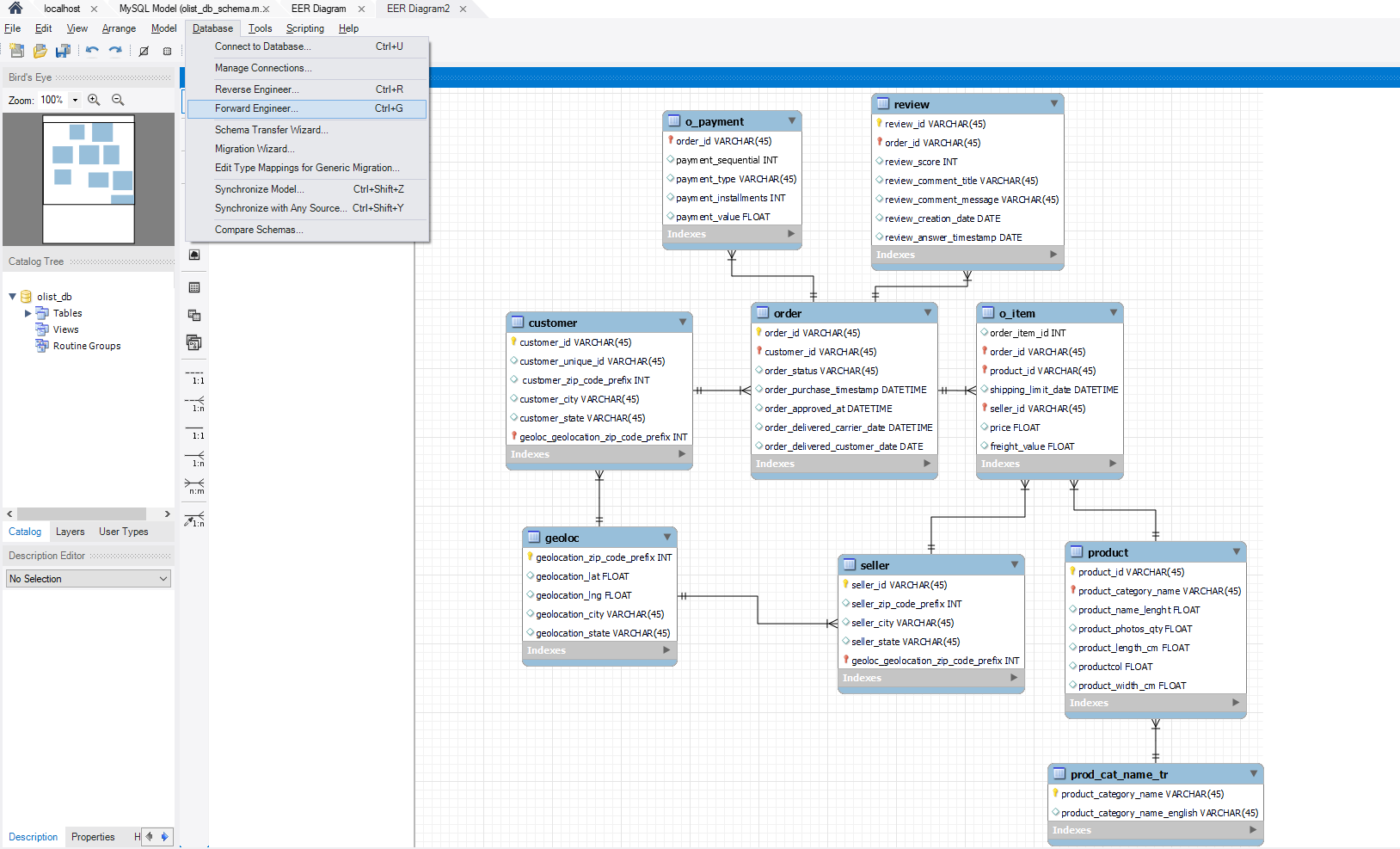
**Data Schema:**

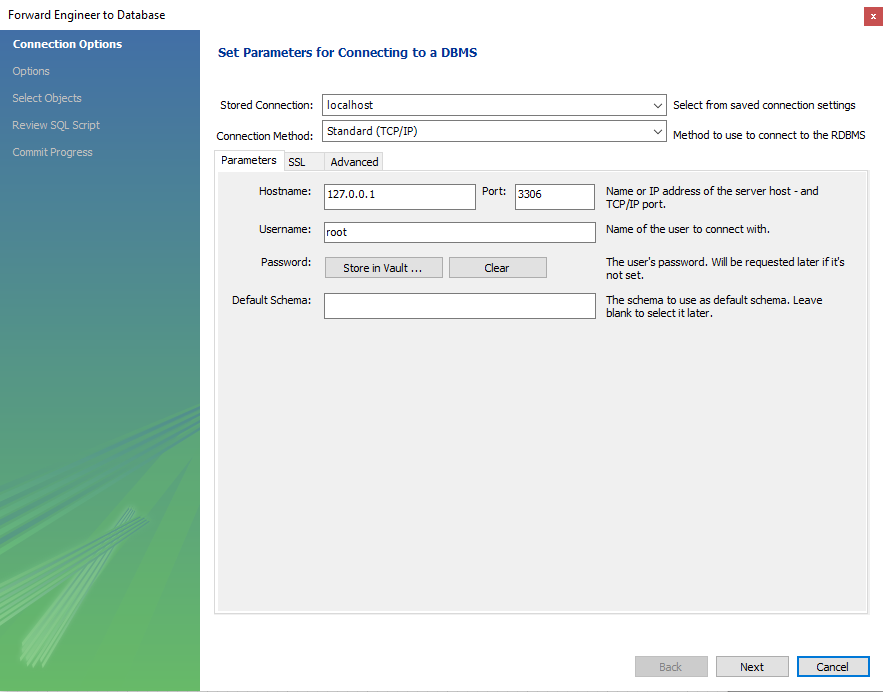


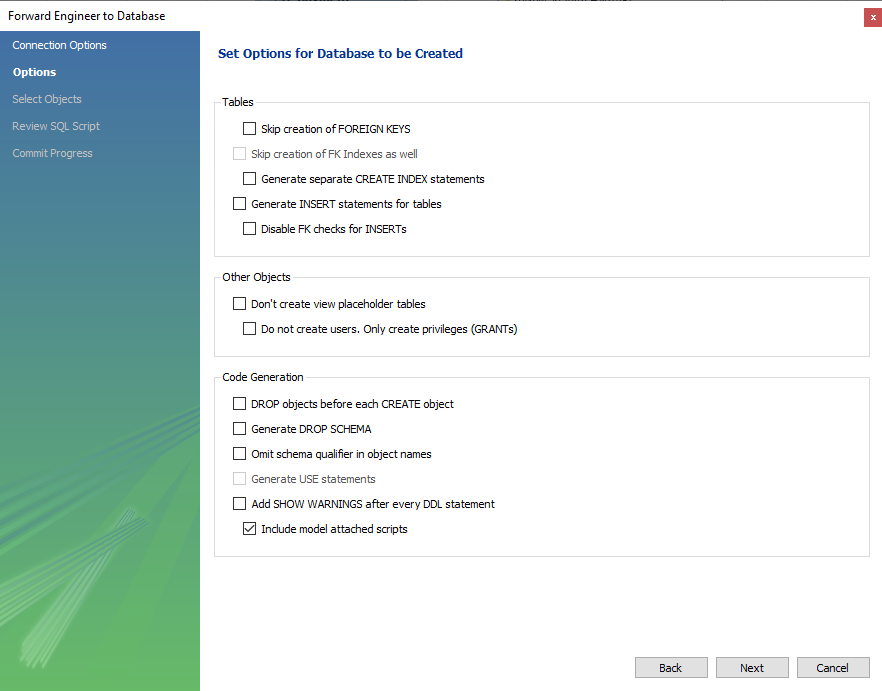
MySQL Diagram:

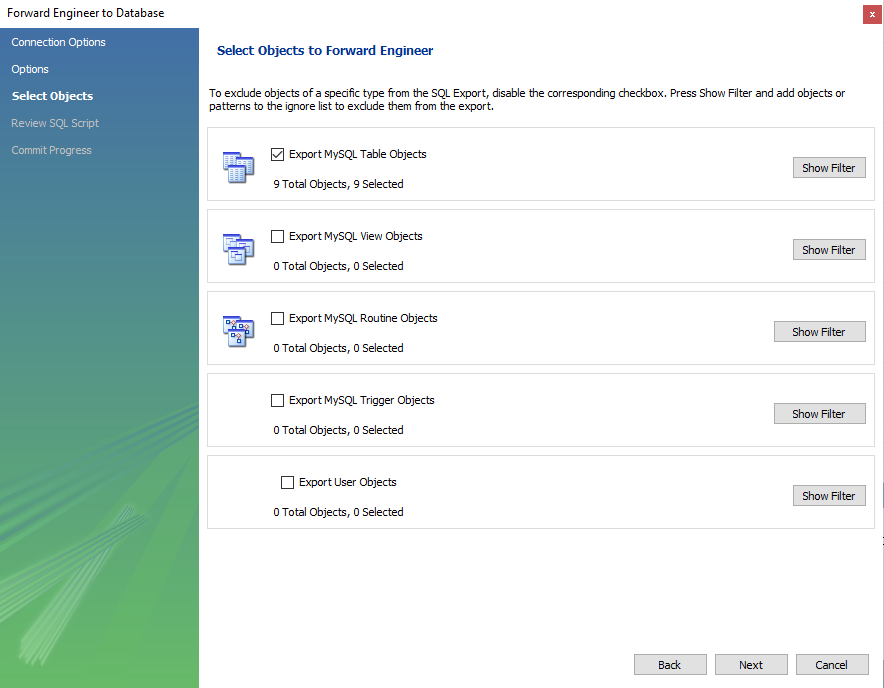


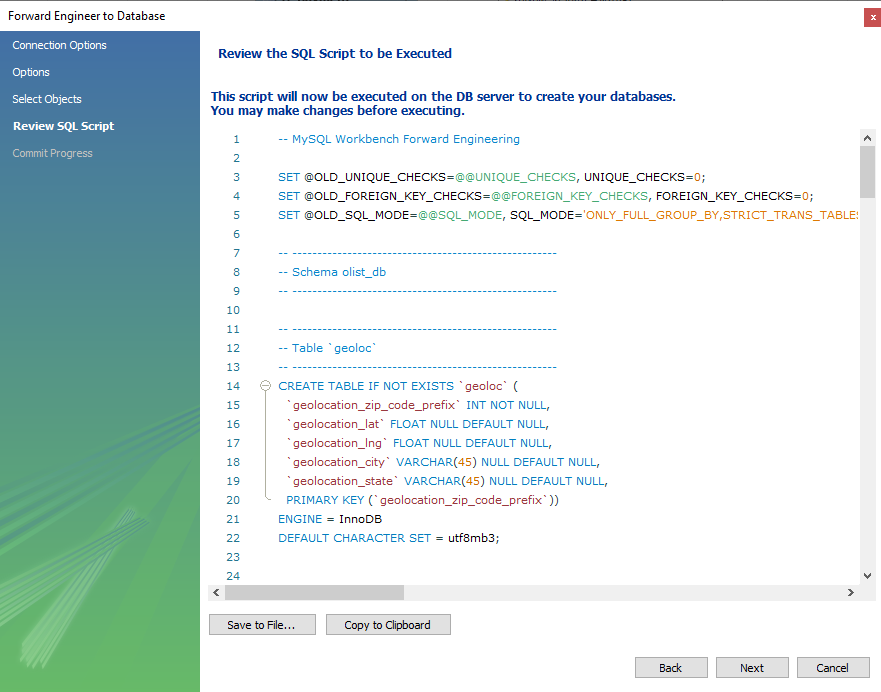
You can *forward engineer* the model above and create the database in MySQL Workbench.











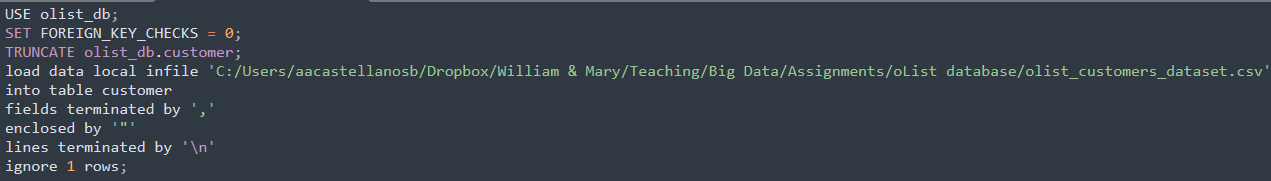
2 Methods:

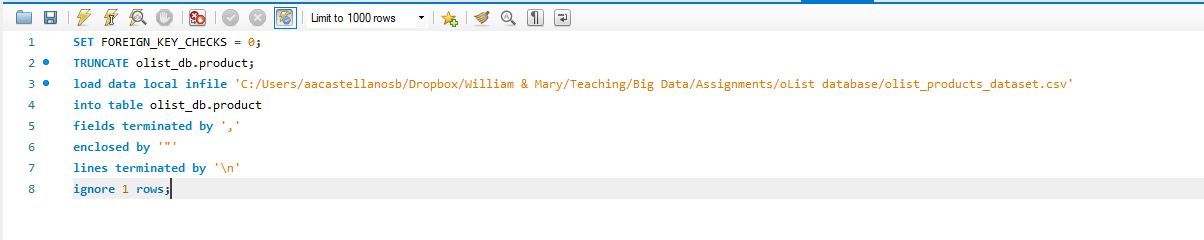
1. Using command line to insert data into database:

Open powershell or terminal:



Enter: mysql -u *user* -p –local-infile *db\_name*



1. Using the MySQL workbench 

Repeat the process for each table, note that not all were provided, you need to write the commands for **two** of the tables. Make sure to change the path and extension of the files, were applicable.