

База данных, которую я использовала в работе, создана на основе дата сетов <https://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+data#>

В этих дата сетах содержится информация о ресторанах, их рейтинге и посетителях. Эти наборы данных использовались для разработки рекомендательной системы на основании предпочтений пользователя. Для своей работы я немного упростила эти наборы данных, убрала часть переменных.

Две основных таблицы unic_placeid и unic_userid содержат списки уникальных id ресторанов и пользователей.

Таблица rest2 содержит подробную информацию о ресторанах:

placeID – id ресторана,
name – название ресторана,
alcohol – наличие или отсутствие алкоголя в продаже,
smoking_area – наличие или отсутствие возможности курить,
dress_code – дресс-код,
accessibility - доступность,
price – уровень цен,
Rambience - обстановка,
Franchise – сетевой ресторан или нет.

Таблица chefmozaccepts содержит id ресторана и типы платежных систем.

Таблица chefmozcuisine содержит id ресторана и информацию о типе кухни ресторана.

Таблица user2 содержит подробную информацию о посетителях:

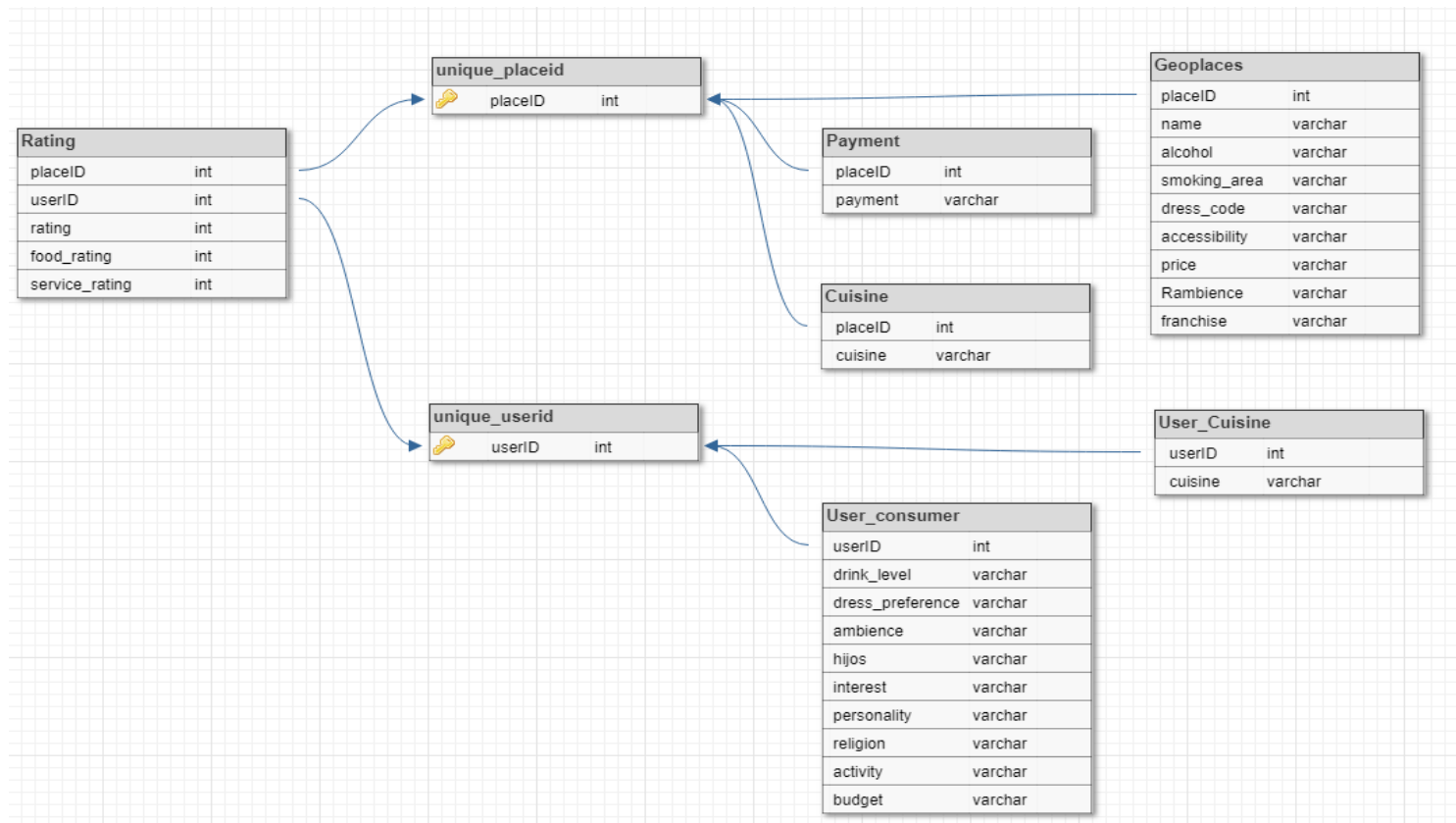
userID – id посетителя,
drink_level – употребляет алкоголь или нет,
dress_preference – предпочтения в одежде,
ambience – в каком окружении предпочитает посещать заведения,
hijos – с детьми или без,
interest - интересы,
personality – тип личности,
religion – религиозная принадлежность,
activity - занятость,
budget – уровень бюджета.

Таблица usercuisine содержит информацию о предпочтениях посетителей в еде.

И таблица rating_final содержит информацию об итоговом рейтинге, который поставили посетители ресторанам. В рейтинге 3 оценки: за еду, за обслуживание и за общее впечатление. По

каждому пункту выставлена оценка от 0 до 2 баллов. То есть всего ресторан может получить от одного пользователя максимум 6 баллов.

Схема базы данных:



Запросы на создание таблиц содержатся в файле Создание и заполнение таблиц. Запросы на обработку и анализ содержатся в файле Запросы.