INTRODUCTION

Five local healthy restaurant chains were selected in Austin, Texas. Each restaurant's top 7 menu items were compared against the others based on their nutritional values.

EXTRACT

Restaurant nutrition data was extracted from the Nutritionix API (https://www.nutritionix.com/business/api) to determine restaurant menu content.

Recommended dietary values were extracted from USDA's Food Data Central website (https://fdc.nal.usda.gov/download-datasets.html#bkmk-1) to determine "healthy" parameters.

TRANSFORM

Food Data Central pdfs were converted to CSVs for data processing purposes.

Nutritionix's API data was transformed from JSON to text and then exported to CSV for data processing, using Jupyter Notebook (see NutrientsConverter.ipnyb).

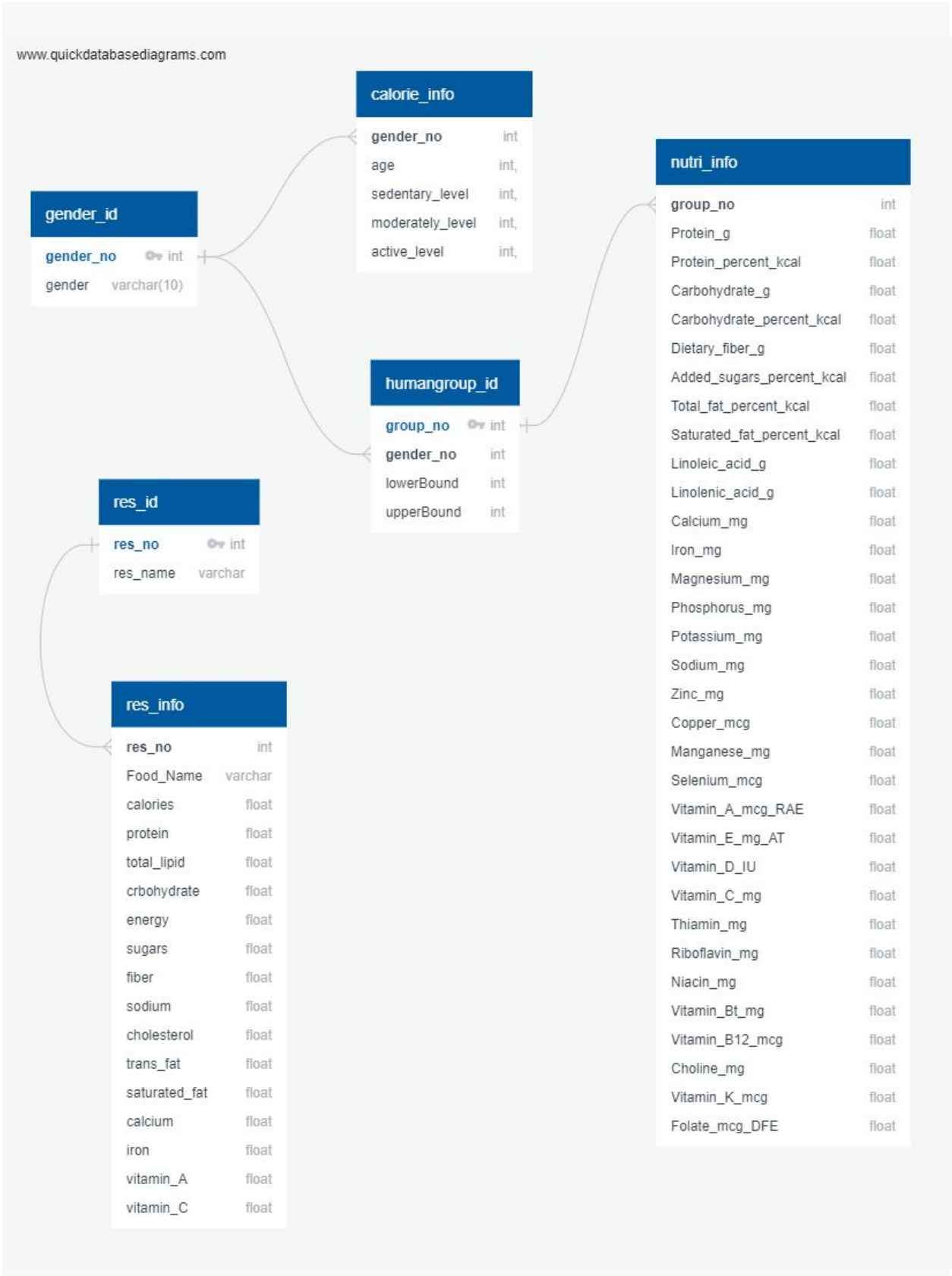
Nutritionix data was further filtered to display only the restaurants needed for this project (see NutrientsConverter.ipnyb).

LOAD

We loaded our data into a postgres database via pgAdmin4. We decided to go with this method because we had a limited dataset of structured data. MongoDB would not have been appropriate for this project since we didn't obtain large amounts of unstructured data. This relational database provided easy querying abilities that allowed us to have a smooth load process.

Before we could load our data, we had to transpose the nutri_info table in Excel. Initially, this table did not match our existing SQL schema due to the rows and columns being inverted. Using the transpose function quickly solved this issue.

Below is an entity relationship diagram that demonstrates how our tables are structured.


www.quickdatabasediagrams.com
calorie_info
gender_no int
age int,
sedentary_level int,
moderately_level int,
active_level int,
gender_id
gender_no int
gender varchar(10)
nutri_info
group_no int
Protein_g float
Protein_percent_kcal float
Carbohydrate_g float
Carbohydrate_percent_kcal float
Dietary_fiber_g float
Added_sugars_percent_kcal float
Total_fat_percent_kcal float
Saturated_fat_percent_kcal float
Linoleic_acid_g float
Linolenic_acid_g float
Calcium_mg float
Iron_mg float
Magnesium_mg float
Phosphorus_mg float
Potassium_mg float
Sodium_mg float
Zinc_mg float
Copper_mcg float
Manganese_mg float
Selenium_mcg float
Vitamin_A_mcg_RAE float
Vitamin_E_mg_AT float
Vitamin_D_IU float
Vitamin_C_mg float
Thiamin_mg float
Riboflavin_mg float
Niacin_mg float
Vitamin_Bt_mg float
Vitamin_B12_mcg float
Choline_mg float
Vitamin_K_mcg float
Folate_mcg_DFE float
humangroup_id
group_no int
gender_no int
lowerBound int
upperBound int
res_id
res_no int
res_name varchar
res_info
res_no int
Food_Name varchar
calories float
protein float
total_lipid float
crbohydrate float
energy float
sugars float
fiber float
sodium float
cholesterol float
trans_fat float
saturated_fat float
calcium float
iron float
vitamin_A float
vitamin_C float