Table of Content

Introduction

In this case study, I will perform many real-world tasks of a junior data analyst at a fictional Music Store Retailer, Beat Music Store. In order to answer the key business questions, I will follow the steps of the data process: Ask, Prepare, Process, Analyse, Share, and Act.

Background

Beat Music Store:

- Beat Music Store is a music retailer company. Beat Music Store sells its product globally.
- Beat Music store wants to understand their customers and grow more sales in 2023.

Scenario:

I am assuming to be a junior data analyst working in the marketing analyst team at Beat Music Store, a music retailer company. The director of marketing believes the company's future success depends on their customers and grow more sales in 2023. Therefore, my team wants to understand how different variables affected sales and growth of company in past.

From these insights, my team will design a new marketing strategy to understand customer base and increase sales in 2023. But first, Beat Music Store executives must approve our recommendations, so they must be backed up with compelling data insights.

Data Analysis Process

Ask

Business Task

Beat Music store wants to understand their customer base and grow more sales in 2023.

Analysis Questions and Queries.

1. who is the senior most employee base on job title?
2. which countries have the most invoices?
3. what are top 3 values of total invoice?
4. Which city has the best customers? We would like to throw a promotional Music Festival in the city we made the most money. Write a query that returns one city that has the highest sum of invoice totals. Return both the city name & sum of all invoice totals
5. Who is the best customer? The customer who has spent the most money will be declared the best customer. Write a query that returns the person who has spent the most money
6. Write query to return the email, first name, last name, & Genre of all Rock Music listeners. Return your list ordered alphabetically by email starting with A
7. Let's invite the artists who have written the most rock music in our dataset. Write a Query that returns the Artist name and total track count of the top 10 rock bands.
8. Return all the track names that have a song length longer than the average song length. Return the Name and Milliseconds for each track. Order by the song length with the longest songs listed first
9. Find how much amount spent by each customer on bestselling artist? Write a query to return customers name, artist name and total spent
10. We want to find out the most popular music Genre for each country. We determine the most popular genre as the genre with the highest amount of purchases. Write a query that returns each country along with the top Genre. For countries where the maximum number of purchases is shared return all Genres

11. Write a query that determines the customer that has spent the most on music for each country. Write a query that returns the country along with the top customer and how much they spent. For countries where the top amount spent is shared, provide all customers who spent this amount

Prepare

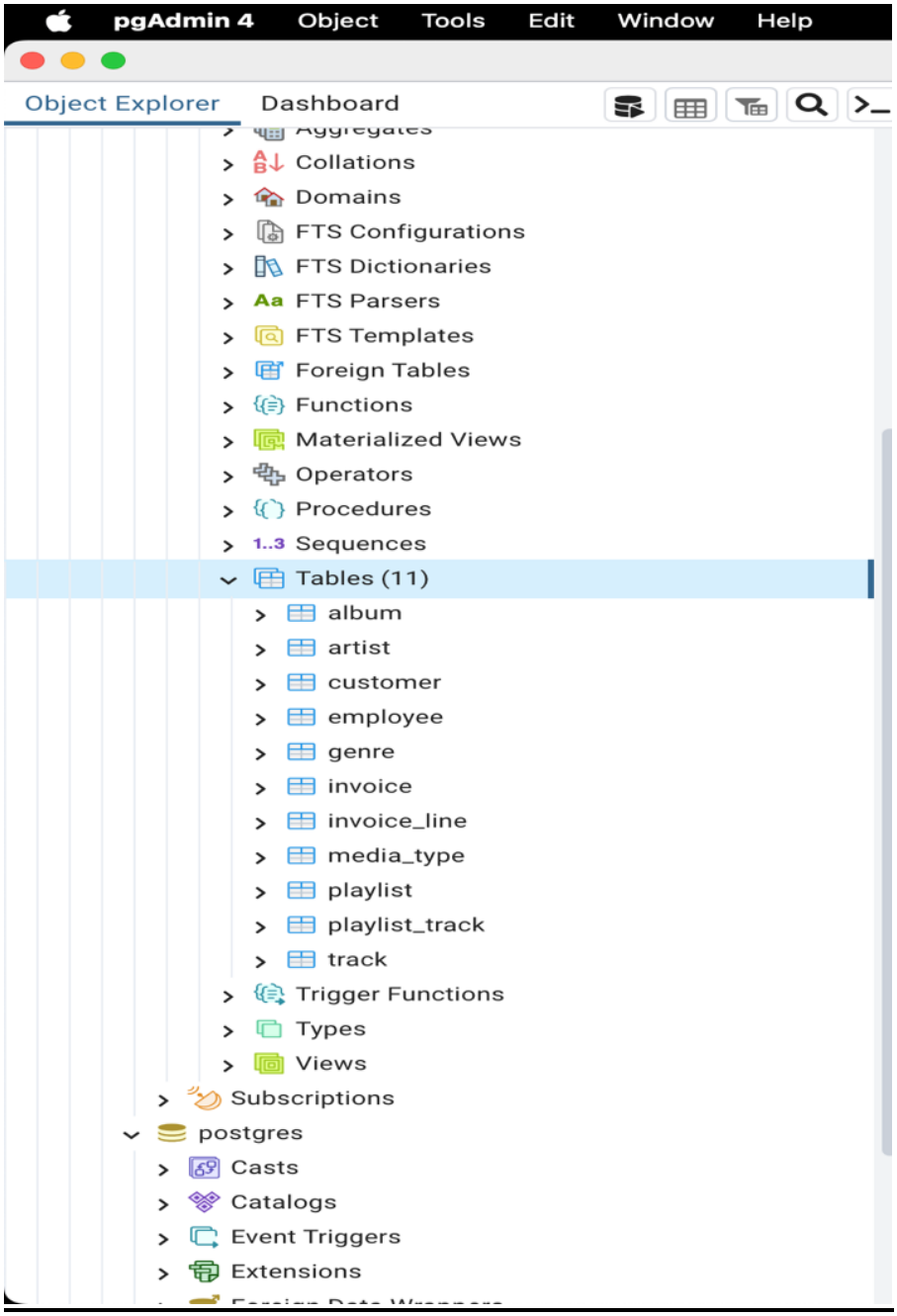
Data source

I will use Beat Music Store sales data to analyze and identify trends from Jan 2022 to Dec 2022 which can be downloaded from GitHub website.

This is public data that can be used to explore how different customer types buying different products of Beat Music Store from all over India.

Data organization

The Beat Music Store database contain 11 tables: album, artist, customer, employee, genre, invoice, invoice_line, media_type, playlist, playlist_track, track. These tables contains data of Beat Music Store.



Process

Data exploration

pgAdmin4 is used for data exploration and to get familiarize with the data

Data Cleaning

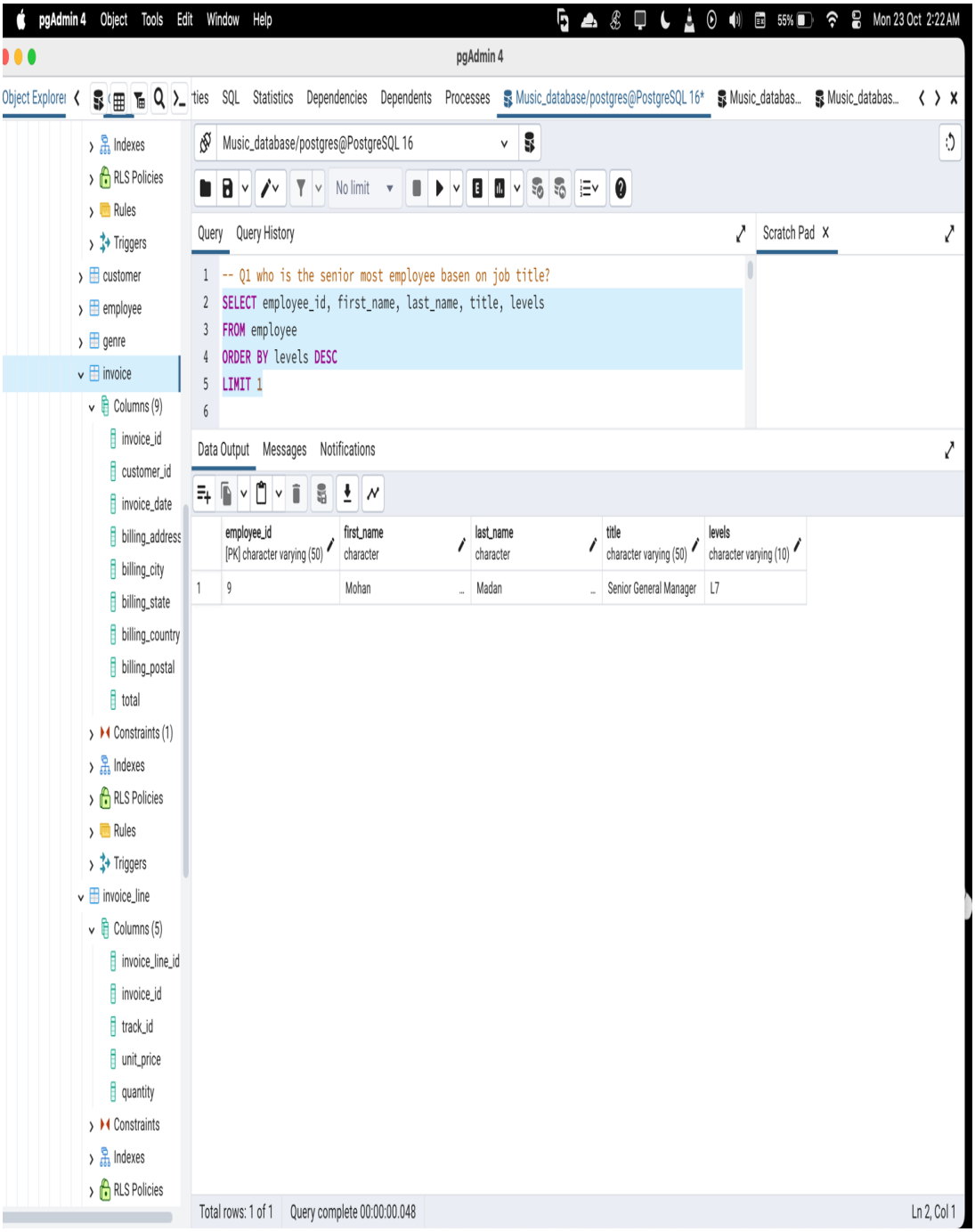
Data cleaning and modification is done with the help of SQL queries in pgAdmin4

Analyse , Share & Act

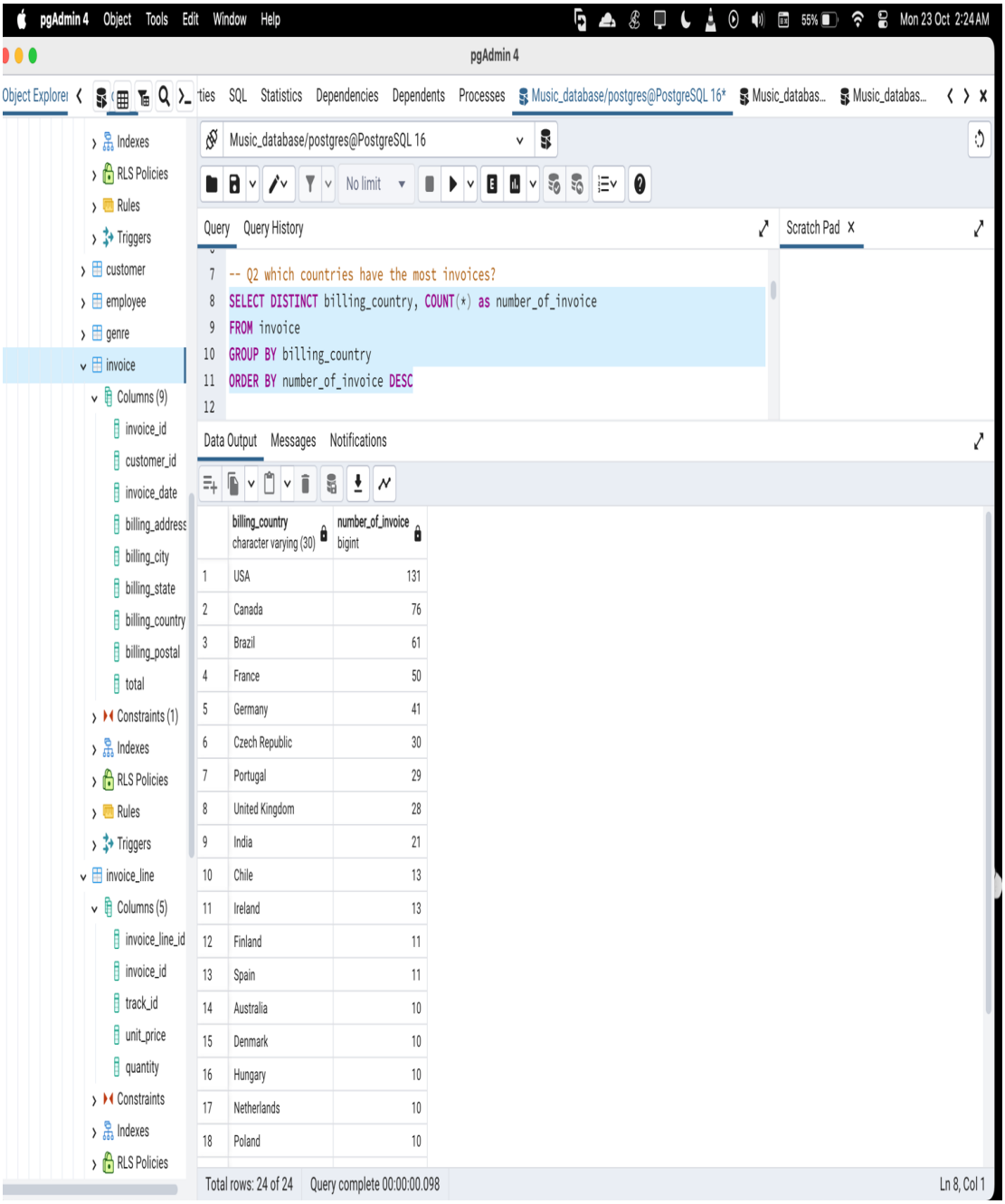
After cleaning the data, it is ready for analysis. pgAdmin4 is used for Data Analysis.

In data analysis we will find the answers to the Data Analysis Questions and Queries.

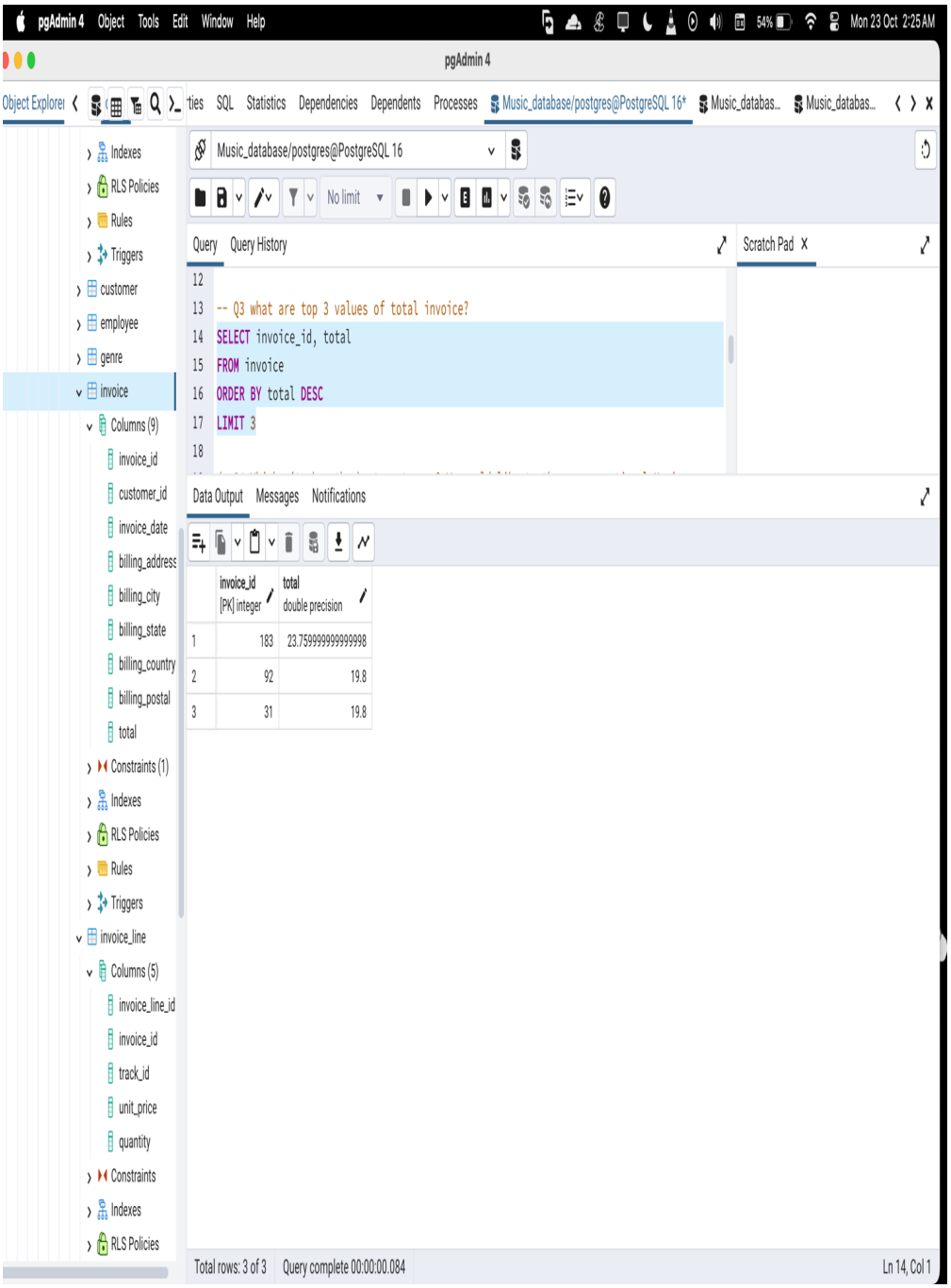
1. who is the senior most employee base on job title?



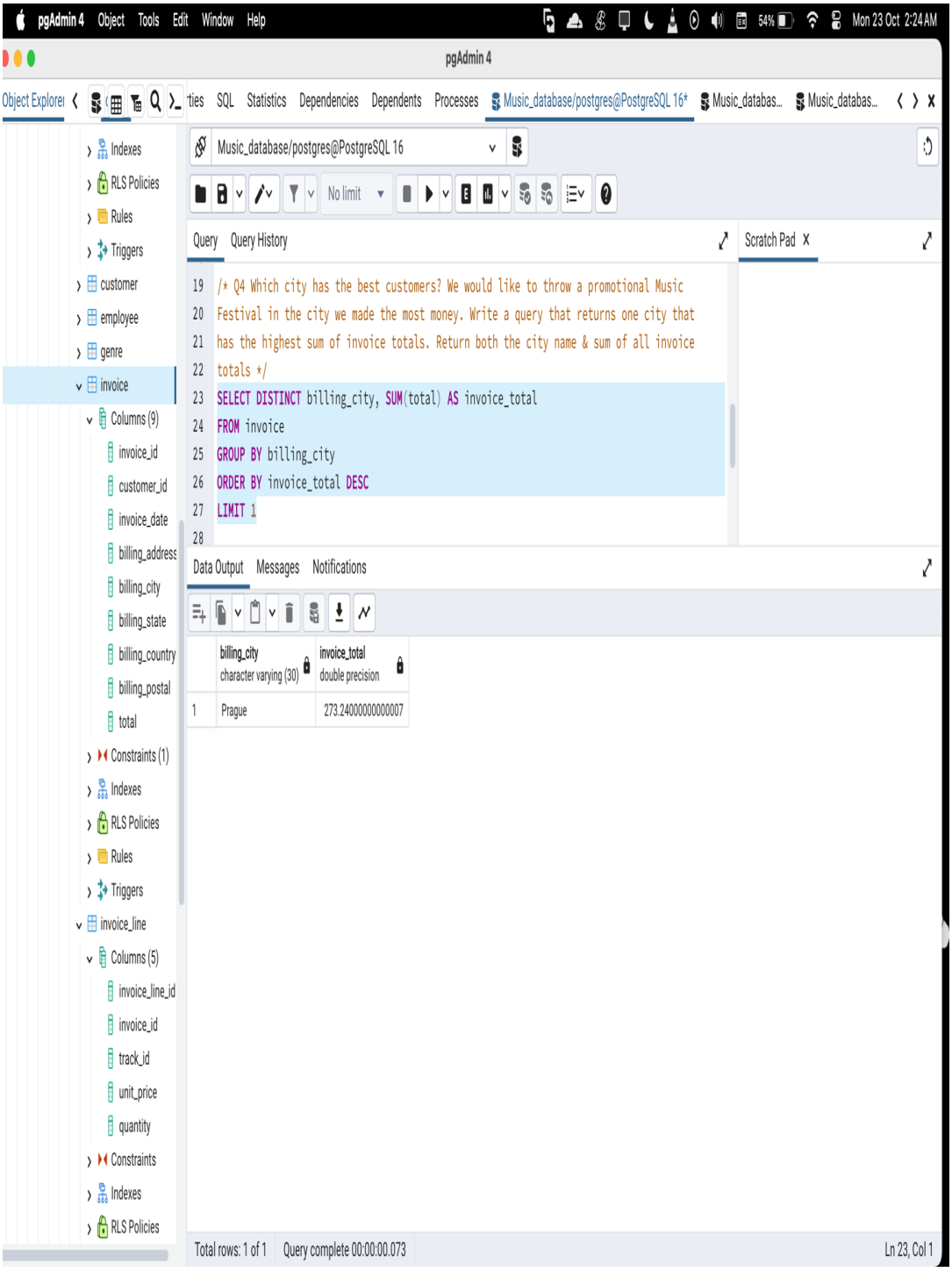
2. which countries have the most invoices?



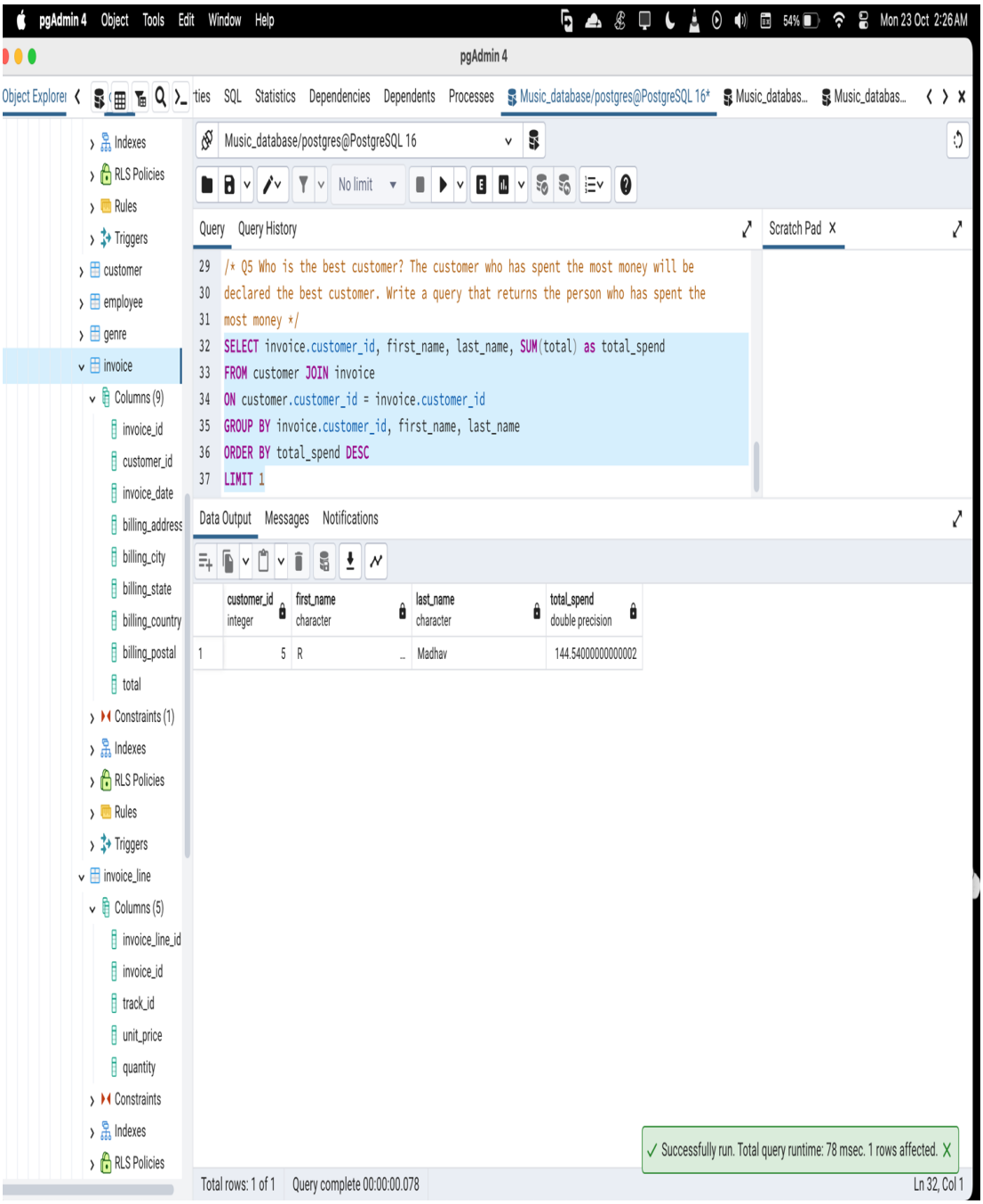
3. what are top 3 values of total invoice?



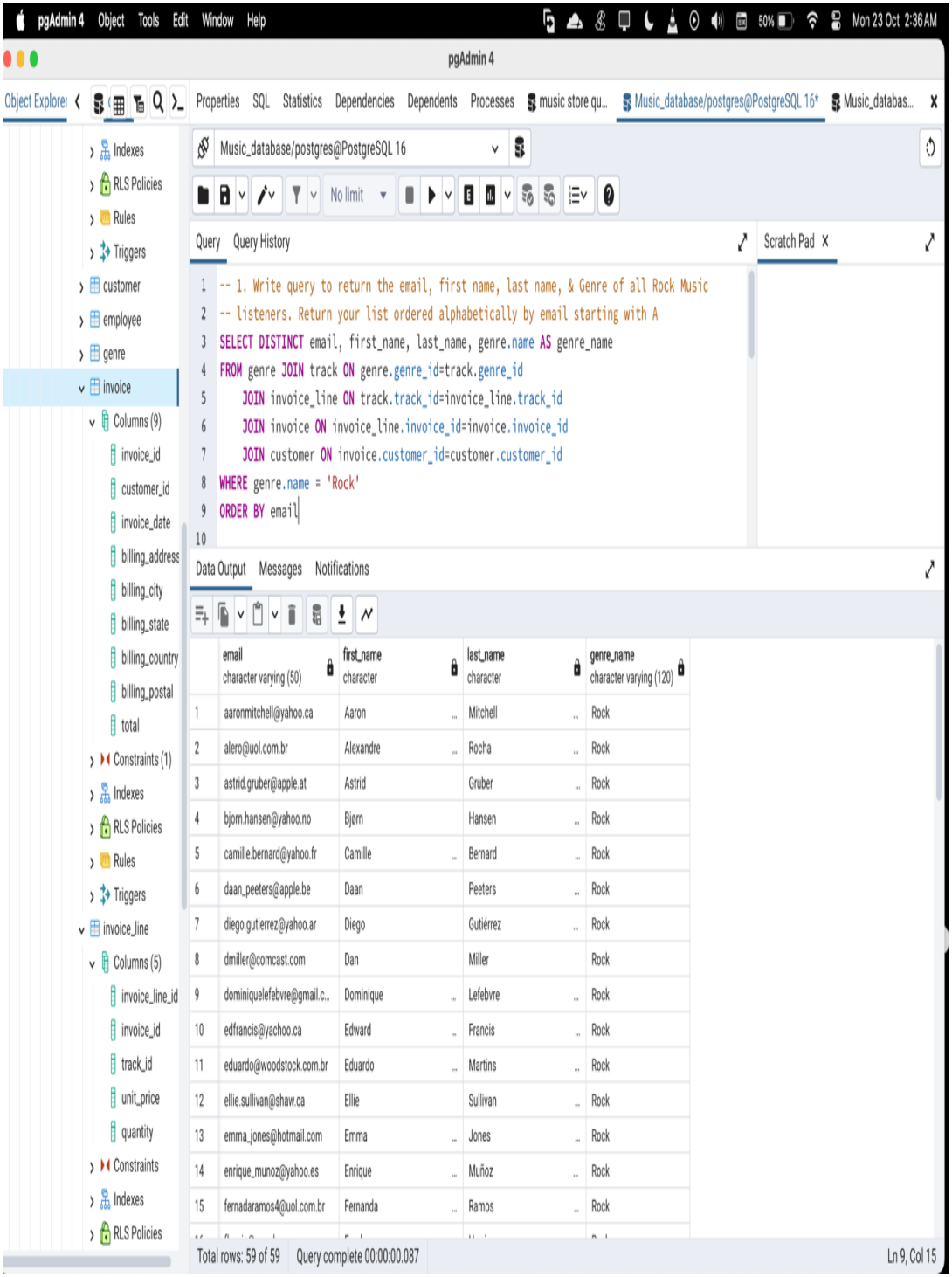
4. Which city has the best customers? We would like to throw a promotional Music Festival in the city we made the most money. Write a query that returns one city that has the highest sum of invoice totals. Return both the city name & sum of all invoice totals



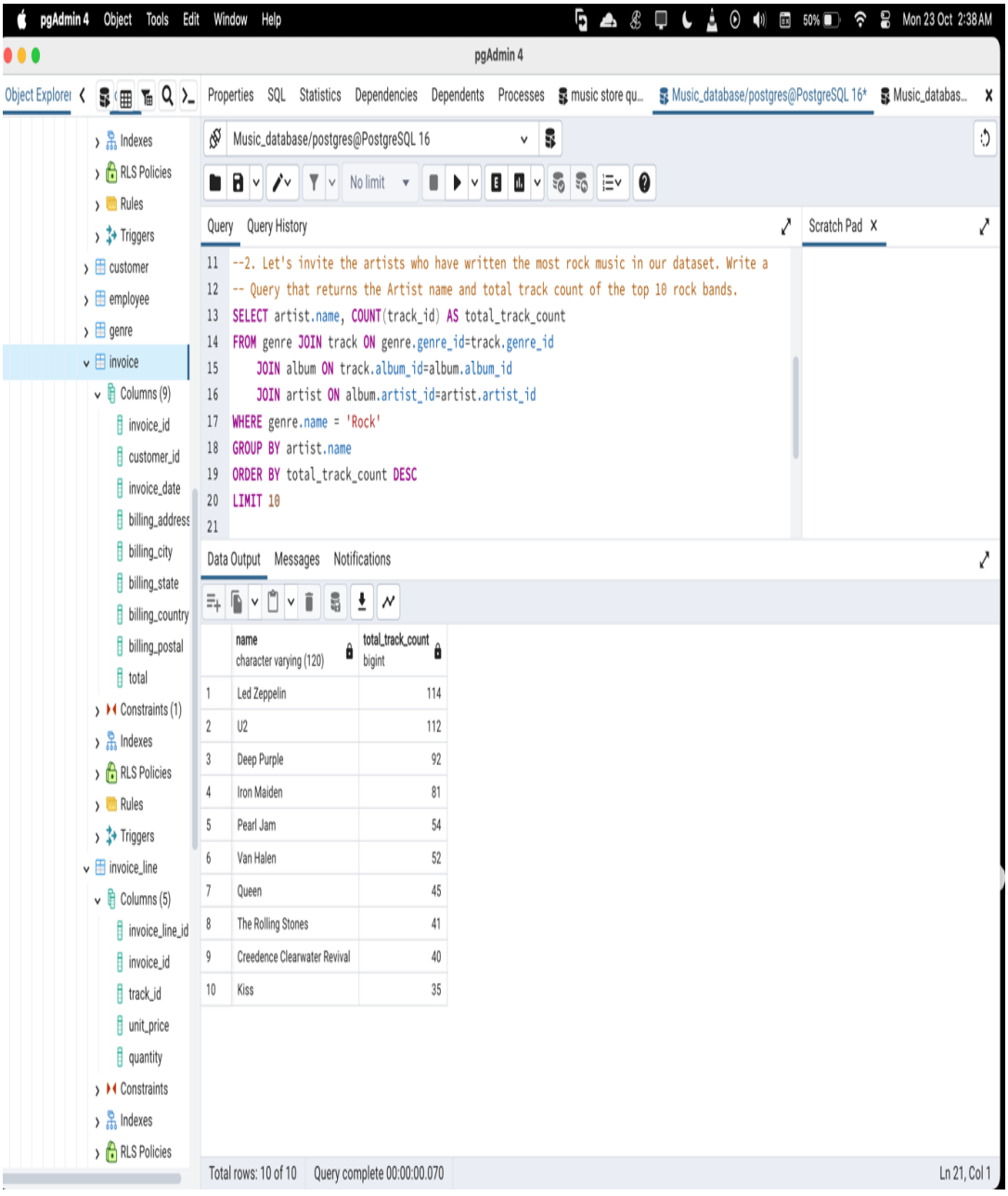
5. Who is the best customer? The customer who has spent the most money will be declared the best customer. Write a query that returns the person who has spent the most money



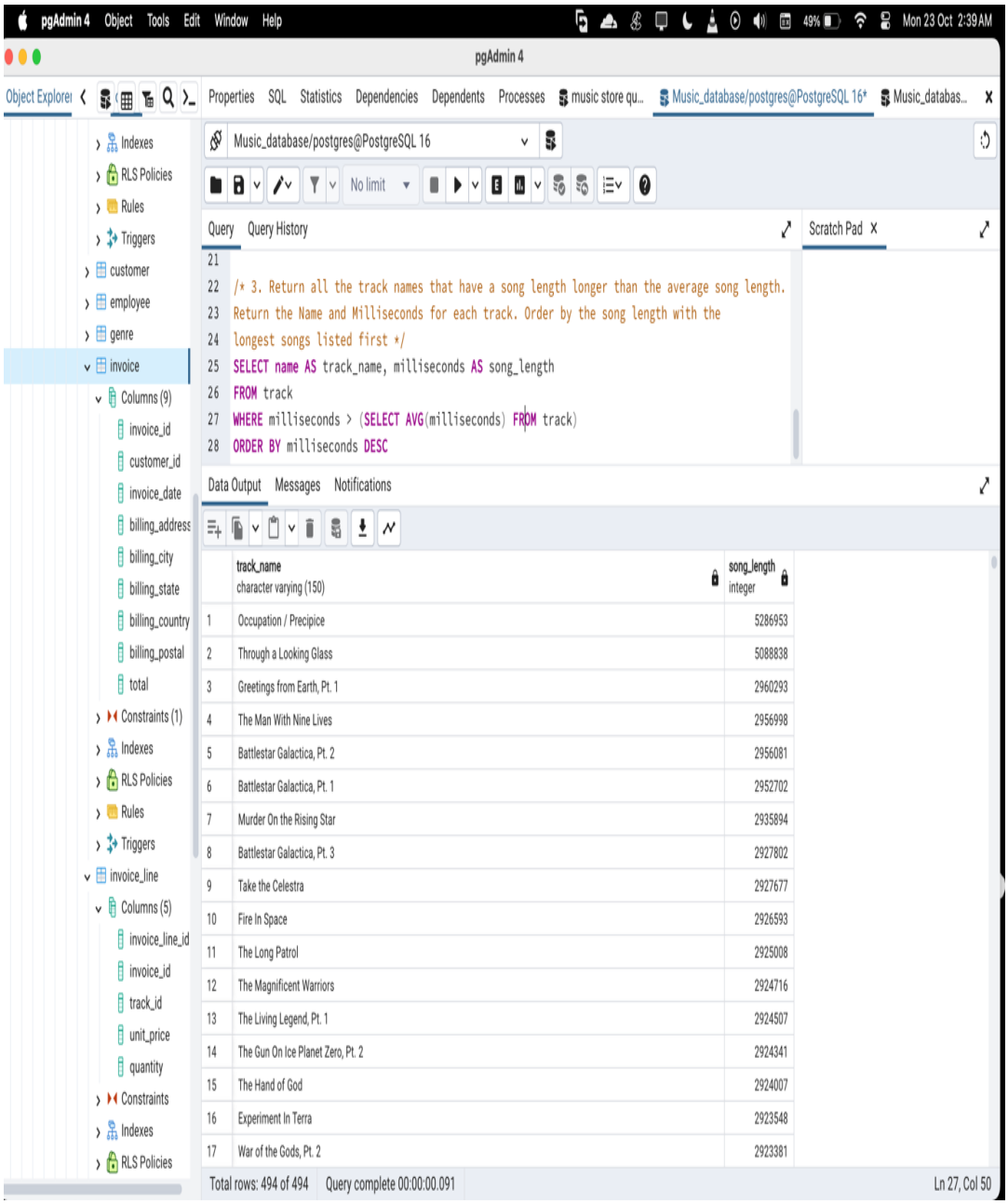
6. Write query to return the email, first name, last name, & Genre of all Rock Music listeners. Return your list ordered alphabetically by email starting with A



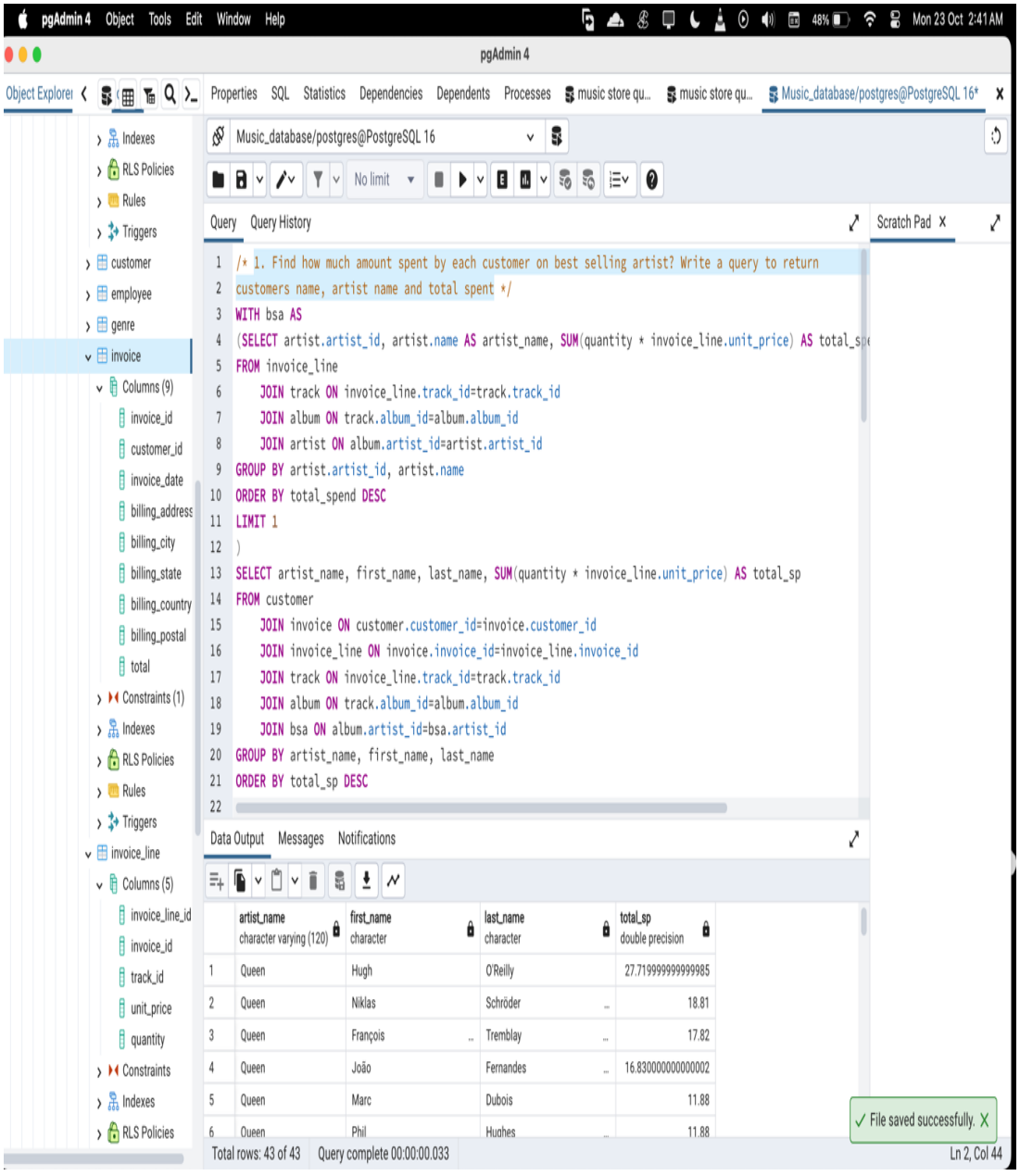
7. Let's invite the artists who have written the most rock music in our dataset. Write a Query that returns the Artist name and total track count of the top 10 rock bands.



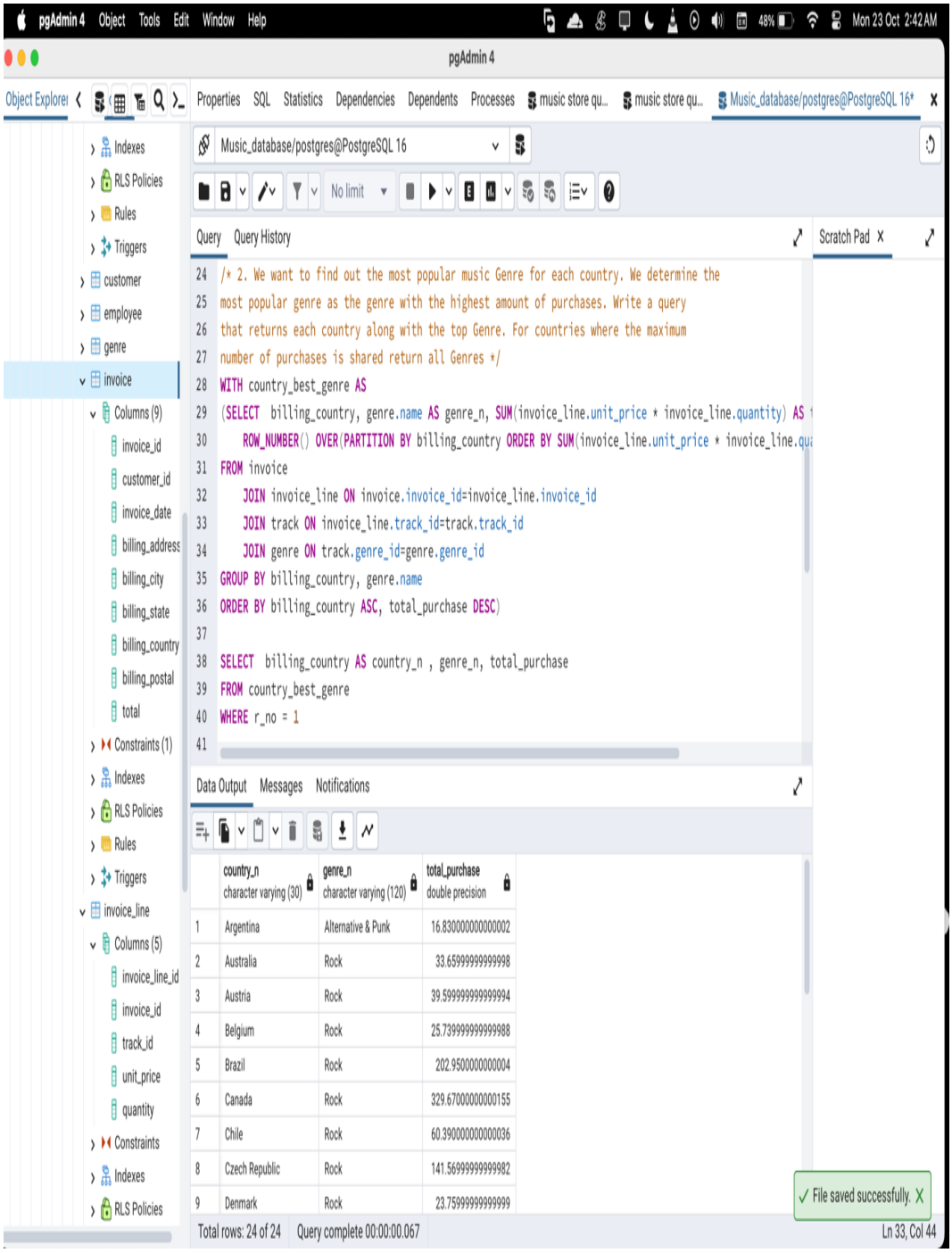
8. Return all the track names that have a song length longer than the average song length. Return the Name and Milliseconds for each track. Order by the song length with the longest songs listed first



9. Find how much amount spent by each customer on bestselling artist? Write a query to return customers name, artist name and total spent



10. We want to find out the most popular music Genre for each country. We determine the most popular genre as the genre with the highest amount of purchases. Write a query that returns each country along with the top Genre. For countries where the maximum number of purchases is shared return all Genres



11. Write a query that determines the customer that has spent the most on music for each country. Write a query that returns the country along with the top customer and how much they spent. For countries where the top amount spent is shared, provide all customers who spent this amount