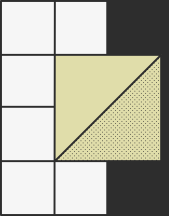




Digital Music Store Analysis

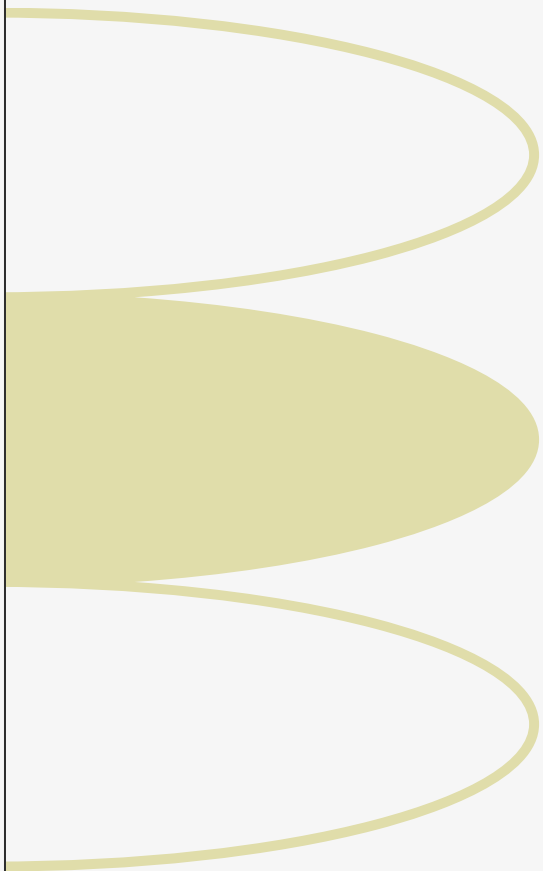
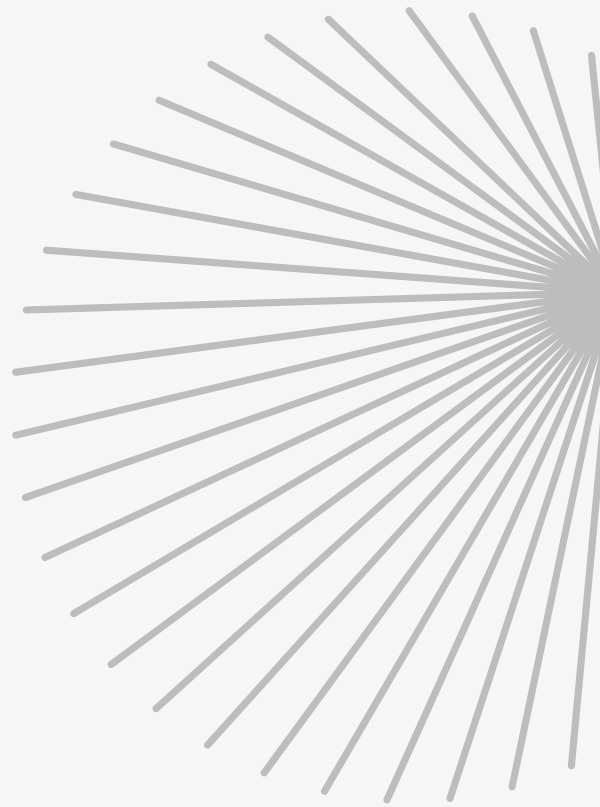
By
Chirag Modha

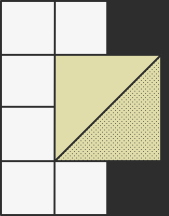




Contents

1. Introduction
2. Schema Overview
3. Questions Answered using SQL



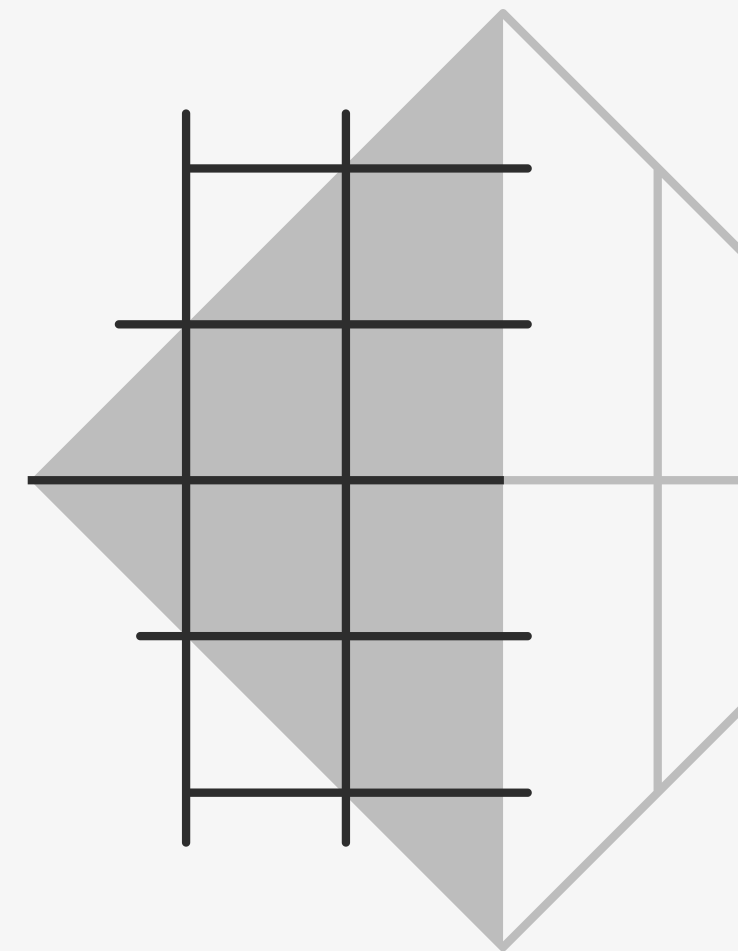


Hello, my name is Chirag Modha, an aspiring data analyst and in this project I've used SQL queries to solve questions related to pizza sales.

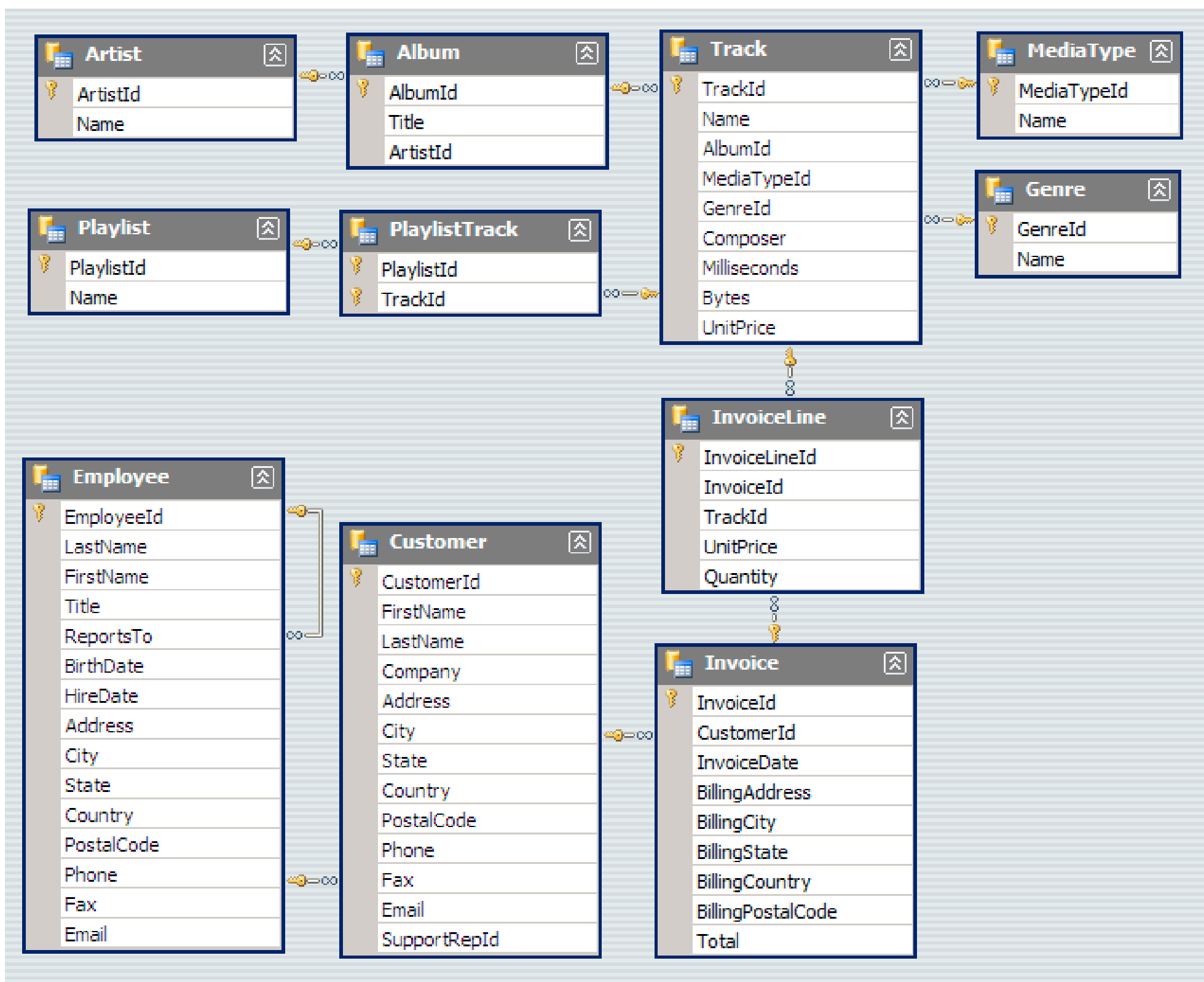
The dataset was provided by Mr. Rishabh Mishra.

Database Link – <https://bit.ly/3wYyp88>

The project analyzes a Digital Music Store dataset using SQL to answer key business questions. The primary focus is on examining the data and help the store understand its business growth by answering the questions.



Schema



Who is the senior most employee based on job title?

```
select * from employee
order by levels desc
limit 1;
```

	employee_id [PK] character varying (50)	last_name character (50)	first_name character (50)	title character varying (50)
1	9	Madan	Mohan	Senior General Manager


Which countries have the most invoices?

```
select billing_country, count(billing_country) from invoice
group by billing_country
order by count(billing_country) desc;
```

	billing_country character varying (30) 🔒	count bigint 🔒
1	USA	131
2	Canada	76
3	Brazil	61
4	France	50
5	Germany	41
6	Czech Republic	30
7	Portugal	29
8	United Kingdom	28
9	India	21

What are top 3 values of total invoice?

```
select total from invoice  
order by total desc limit 3
```

	total double precision 
1	23.759999999999999998
2	19.8
3	19.8

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 total invoice. Return both the city name & sum of all invoice totals.

```
select billing_city, sum(total) as invoice_total from invoice
group by billing_city
order by invoice_total desc limit 1;
```

	billing_city character varying (30) 🔒	invoice_total double precision 🔒
1	Prague	273.240000000000007

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

```
select customer.customer_id, customer.first_name, customer.last_name, sum(invoice.total)from customer
join invoice on customer.customer_id = invoice.customer_id
group by customer.customer_id
order by sum(invoice.total) desc limit 1;
```

	customer_id [PK] integer	first_name character (50)	last_name character (50)	sum double precision
1	5	R	Madhav	144.540000000000002

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

```
select distinct email, first_name, last_name
from customer
join invoice on customer.customer_id = invoice.customer_id
join invoice_line on invoice.invoice_id = invoice_line.invoice_id
where track_id in(
    select track_id from track
    join genre on track.genre_id = genre.genre_id
    where genre.name like 'Rock'
)
order by email;
```

	email character varying (50)	first_name character (50)	last_name character (50)
1	aaronmitchell@yahoo.ca	Aaron	Mitchell
2	alero@uol.com.br	Alexandre	Rocha
3	astrid.gruber@apple.at	Astrid	Gruber
4	bjorn.hansen@yahoo.no	Bjørn	Hansen
5	camille.bernard@yahoo.fr	Camille	Bernard
6	daan.peeters@apple.be	Daan	Peeters
7	diego.gutierrez@yahoo.ar	Diego	Gutiérrez
8	dmiller@comcast.com	Dan	Miller
9	dominiquelefebvre@gmail.c...	Dominique	Lefebvre
10	edfrancis@yahoo.ca	Edward	Francis

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

```
select artist.artist_id, artist.name, count(artist.artist_id) as number_of_songs
from track
join album on album.album_id = track.album_id
join artist on artist.artist_id = album.artist_id
join genre on genre.genre_id = track.genre_id
where genre.name like 'Rock'
group by artist.artist_id
order by number_of_songs desc
limit 10;
```

	artist_id [PK] character varying (50)	name character varying (120)	number_of_songs bigint
1	22	Led Zeppelin	114
2	150	U2	112
3	58	Deep Purple	92
4	90	Iron Maiden	81
5	118	Pearl Jam	54
6	152	Van Halen	52
7	51	Queen	45
8	142	The Rolling Stones	41
9	76	Creedence Clearwater Revival	40
10	52	Kiss	35

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

```
select name, milliseconds
from track
where milliseconds > (
    select avg(milliseconds) as avg_track_length
    from track
)
order by milliseconds desc;
```

	name character varying (150)	milliseconds integer
1	Occupation / Precipice	5286953
2	Through a Looking Glass	5088838
3	Greetings from Earth, Pt. 1	2960293
4	The Man With Nine Lives	2956998
5	Battlestar Galactica, Pt. 2	2956081
6	Battlestar Galactica, Pt. 1	2952702
7	Murder On the Rising Star	2935894
8	Battlestar Galactica, Pt. 3	2927802
9	Take the Celestra	2927677
10	Fire In Space	2926593
11	The Long Patrol	2925008

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

```
WITH best_selling_artist as (  
    select artist.artist_id as artist_id, artist.name as artist_name,  
    sum(invoice_line.unit_price*invoice_line.quantity) as total_sales  
    from invoice_line  
    join track on track.track_id = invoice_line.track_id  
    join album on album.album_id = track.album_id  
    join artist on artist.artist_id = album.artist_id  
    group by 1  
    order by 3 desc  
    limit 1  
)  
select c.customer_id, c.first_name, c.last_name, bsa.artist_name,  
sum(il.unit_price*il.quantity) as amount_spent  
from invoice i  
join customer c on c.customer_id = i.customer_id  
join invoice_line il on il.invoice_id = i.invoice_id  
join track t on t.track_id = il.track_id  
join album alb on alb.album_id = t.album_id  
join best_selling_artist bsa on bsa.artist_id = alb.artist_id  
group by 1,2,3,4  
order by 5 desc;
```

	customer_id integer	first_name character (50)	last_name character (50)	artist_name character varying (120)	amount_spent double precision
1	46	Hugh	O'Reilly	Queen	27.7199999999999985
2	38	Niklas	Schröder	Queen	18.81
3	3	François	Tremblay	Queen	17.82
4	34	João	Fernandes	Queen	16.8300000000000002
5	53	Phil	Hughes	Queen	11.88
6	41	Marc	Dubois	Queen	11.88
7	47	Lucas	Mancini	Queen	10.89
8	33	Ellie	Sullivan	Queen	10.89
9	20	Dan	Miller	Queen	3.96
10	5	R	Madhav	Queen	3.96
11	23	John	Gordon	Queen	2.9699999999999998
12	54	Steve	Murray	Queen	2.9699999999999998
13	31	Martha	Silk	Queen	2.9699999999999998
14	16	Frank	Harris	Queen	1.98

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 its top genre. For countries where the maximum number of purchases is shared return all Genres.

```
with popular_genre as (  
    select count(invoice_line.quantity) as purchases, customer.country, genre.name, genre.genre_id,  
    row_number() over(partition by customer.country order by count(invoice_line.quantity) desc) as RowNo  
    from invoice_line  
    join invoice on invoice.invoice_id = invoice_line.invoice_id  
    join customer on customer.customer_id = invoice.customer_id  
    join track on track.track_id = invoice_line.track_id  
    join genre on genre.genre_id = track.genre_id  
    group by 2,3,4  
    order by 2 asc, 1 desc  
)  
select * from popular_genre where RowNo <= 1
```

	purchases bigint	country character varying (50)	name character varying (120)	genre_id character varying (50)	rowno bigint
1	17	Argentina	Alternative & Punk	4	1
2	34	Australia	Rock	1	1
3	40	Austria	Rock	1	1
4	26	Belgium	Rock	1	1
5	205	Brazil	Rock	1	1
6	333	Canada	Rock	1	1
7	61	Chile	Rock	1	1
8	143	Czech Republic	Rock	1	1
9	24	Denmark	Rock	1	1

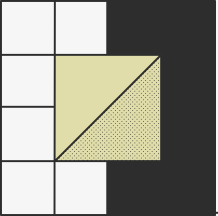
Write a query that determines the customer that has spent the most on music for each country. Write a query that returned the country along with the top customer and how much they spent. For countries where the top amount is shared, provide all customers who spent the amount.

```
with recursive
customer_with_country as (
    select customer.customer_id,first_name,last_name,billing_country,sum(total) as total_spending
    from invoice
    join customer on customer.customer_id = invoice.customer_id
    group by 1,2,3,4
    order by 1,5 desc
),

country_max_spending as (
    select billing_country,max(total_spending) as max_spending
    from customer_with_country
    group by billing_country
)

select cc.billing_country, cc.total_spending, cc.first_name, cc.last_name, cc.billing_country, cc.customer_id
from customer_with_country cc
join country_max_spending ms
on cc.billing_country = ms.billing_country
where cc.total_spending = ms.max_spending
order by 1;
```

	billing_country character varying (30)	total_spending double precision	first_name character (50)	last_name character (50)	billing_country character varying (30)	customer_id integer
1	Argentina	39.6	Diego	Gutiérrez	Argentina	56
2	Australia	81.18	Mark	Taylor	Australia	55
3	Austria	69.3	Astrid	Gruber	Austria	7
4	Belgium	60.389999999999999	Daan	Peeters	Belgium	8
5	Brazil	108.899999999999998	Luís	Gonçalves	Brazil	1
6	Canada	99.99	François	Tremblay	Canada	3
7	Chile	97.020000000000001	Luis	Rojas	Chile	57
8	Czech Republic	144.540000000000002	R	Madhav	Czech Republic	5
9	Denmark	37.619999999999999	Kara	Nielsen	Denmark	9



Thank You

