audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

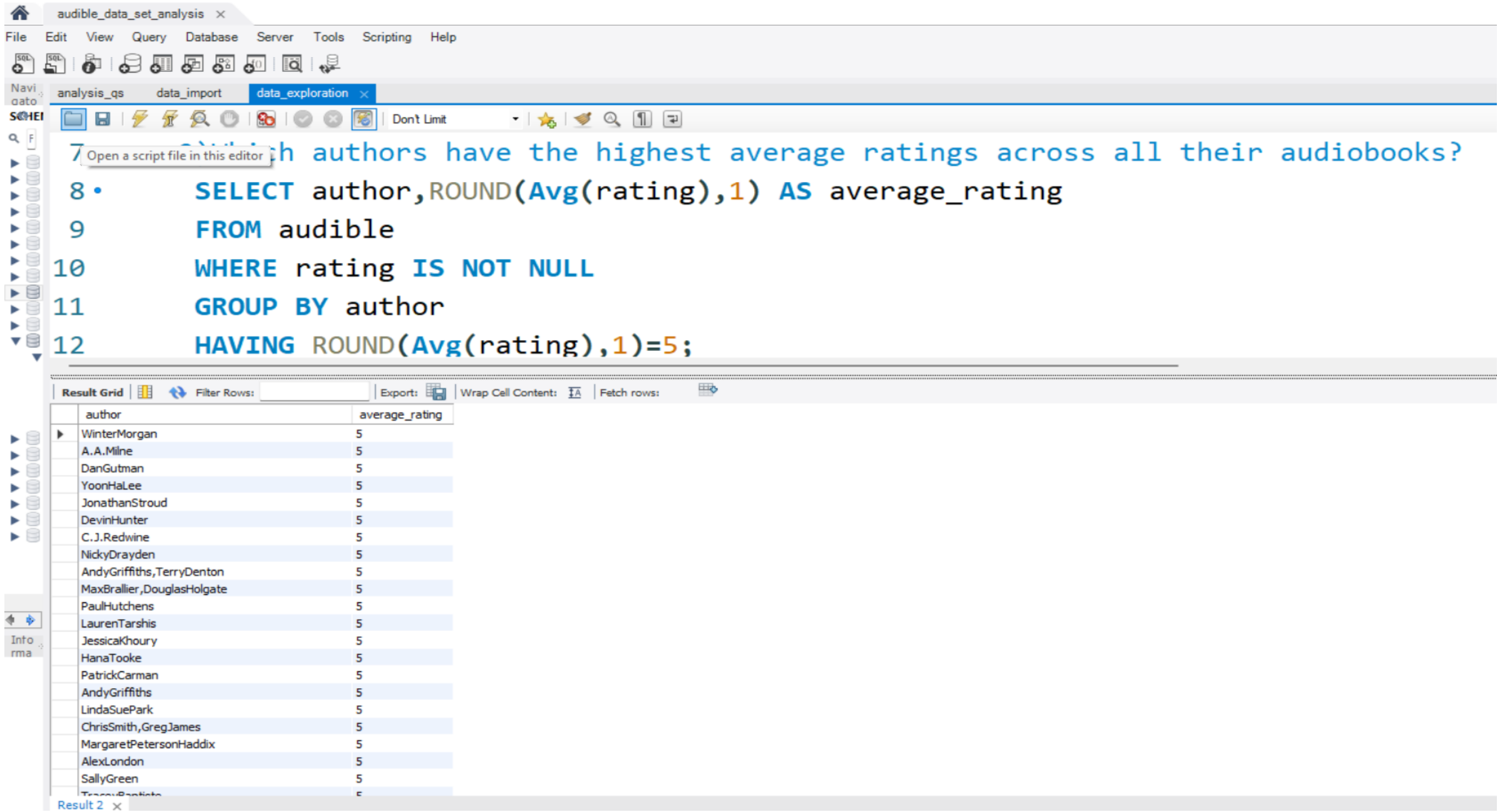
analysis_qs data_import data_exploration

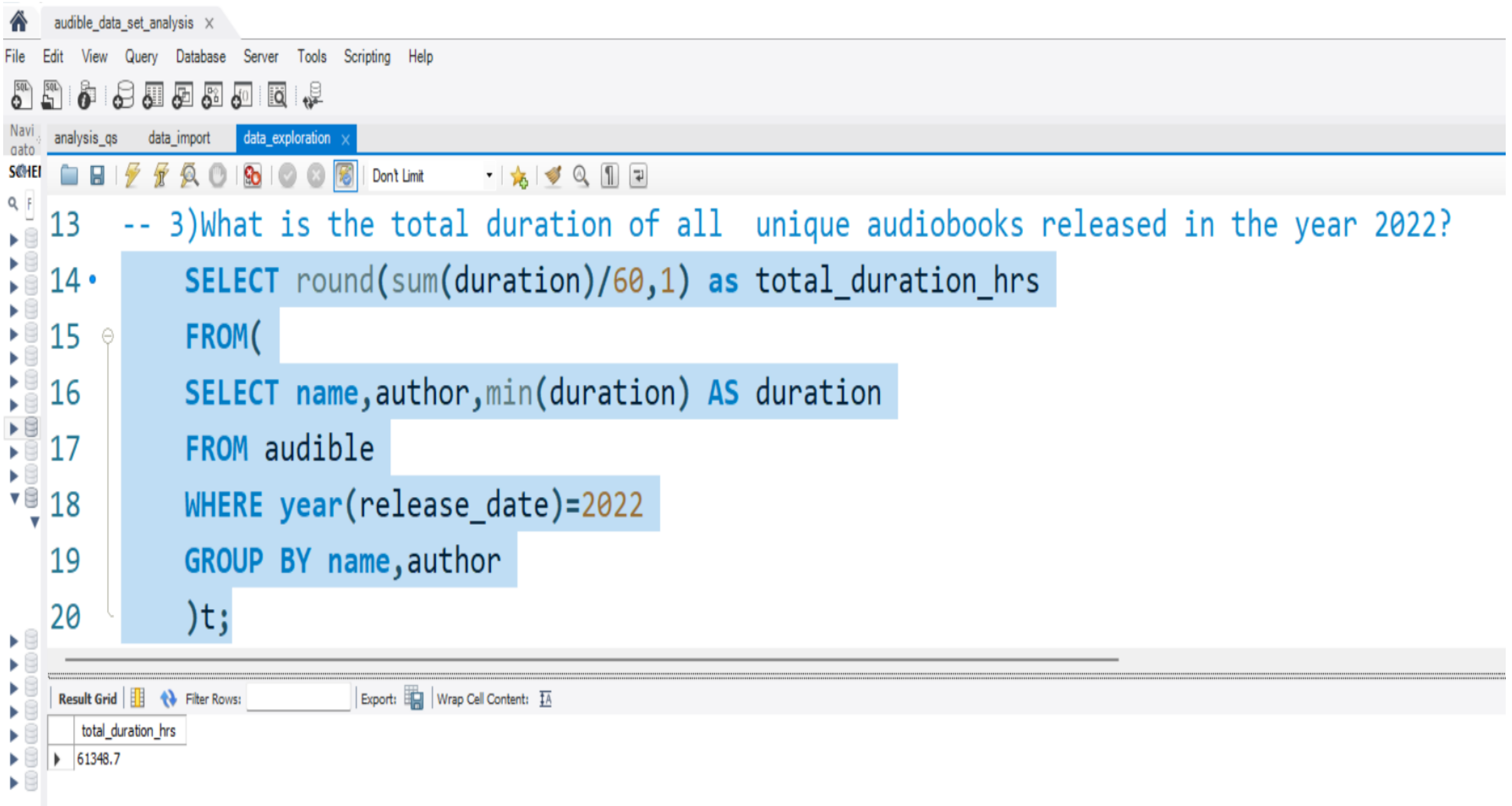
Don't Limit

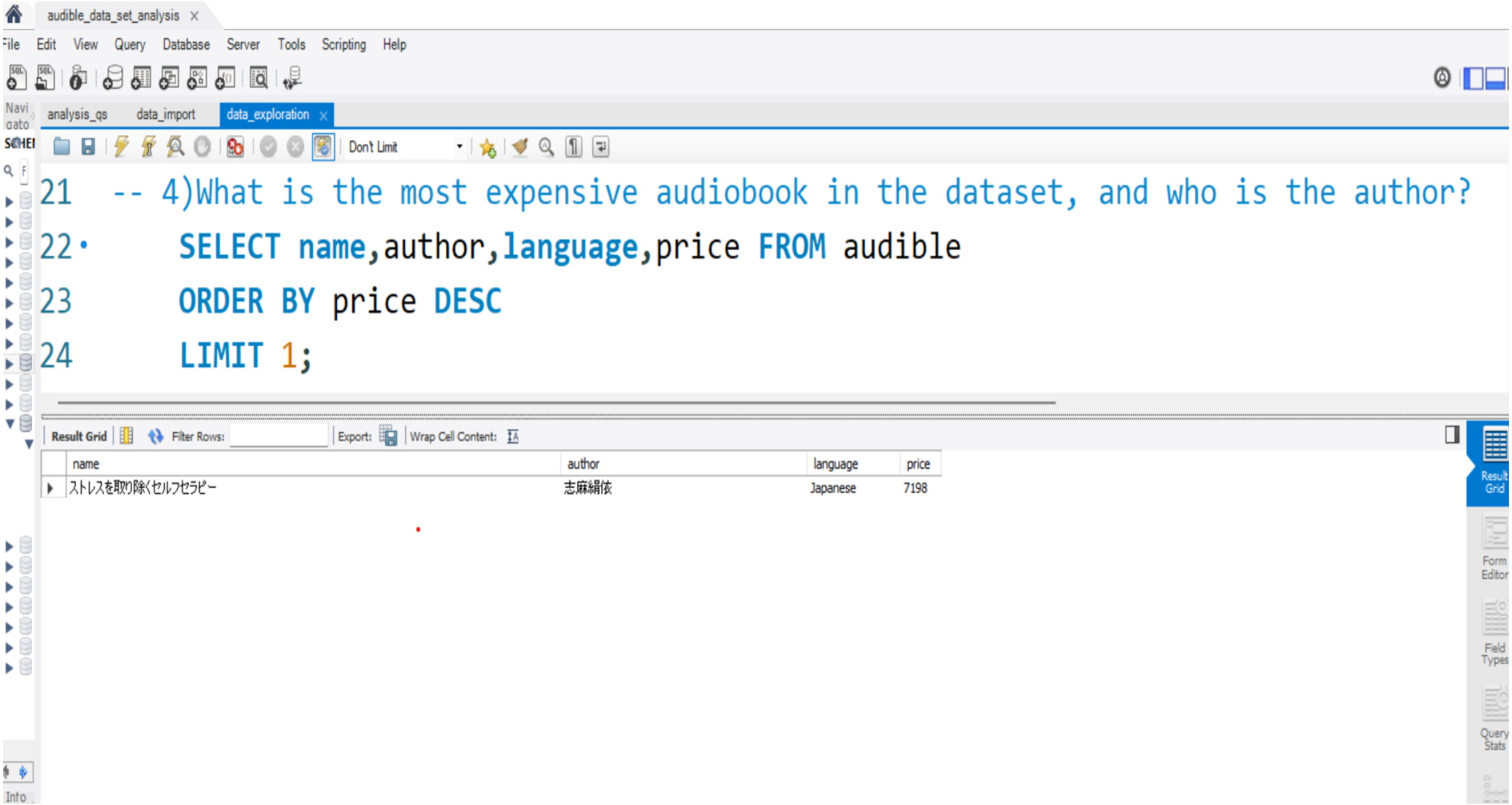
```sql
-- 1)Which 5 audiobooks have received the most ratings?
SELECT name,sum(no_of_ratings) as total_ratings FROM audible
WHERE no_of_ratings IS NOT NULL
GROUP BY name
ORDER BY 2 DESC
LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

name	total_ratings
Atomic Habits	25142
The Psychology of Money	11945
Ikigai	6981
Sapiens	5805
Rich Dad Poor Dad	5324


audible_data_set_analysis
File Edit View Query Database Server Tools Scripting Help
analysis_qs data_import data_exploration
Don't Limit
Open a script file in this editor
7 -- 2) Which authors have the highest average ratings across all their audiobooks?
8 • SELECT author,ROUND(Avg(rating),1) AS average_rating
9 FROM audible
10 WHERE rating IS NOT NULL
11 GROUP BY author
12 HAVING ROUND(Avg(rating),1)=5;
Result Grid Filter Rows: Export: Wrap Cell Content: Fetch rows:
author average_rating
WinterMorgan 5
A.A.Milne 5
DanGutman 5
YoonHaLee 5
JonathanStroud 5
DevinHunter 5
C.J.Redwine 5
NickyDrayden 5
AndyGriffiths,TerryDenton 5
MaxBrallier,DouglasHolgate 5
PaulHutchens 5
LaurenTarshis 5
JessicaKhoury 5
HanaTooke 5
PatrickCarman 5
AndyGriffiths 5
LindaSuePark 5
ChrisSmith,GregJames 5
MargaretPetersonHaddix 5
AlexLondon 5
SallyGreen 5
Result 2


audible_data_set_analysis
File Edit View Query Database Server Tools Scripting Help
analysis_qs
data_import
data_exploration
Don't Limit
-- 3)What is the total duration of all unique audiobooks released in the year 2022?
SELECT round(sum(duration)/60,1) as total_duration_hrs
FROM(
SELECT name,author,min(duration) AS duration
FROM audible
WHERE year(release_date)=2022
GROUP BY name,author
)t;
Result Grid
Filter Rows:
Export:
Wrap Cell Content:
total_duration_hrs
61348.7


audible_data_set_analysis
File Edit View Query Database Server Tools Scripting Help
analysis_qs
data_import
data_exploration
Don't Limit
21 -- 4)What is the most expensive audiobook in the dataset, and who is the author?
22 SELECT name,author,language,price FROM audible
23 ORDER BY price DESC
24 LIMIT 1;
Result Grid
Filter Rows:
Export:
Wrap Cell Content:
name
author
language
price
ストレスを取り除くセルフセラピー
志麻絹依
Japanese
7198
Result Grid
Form Editor
Field Types
Query Stats

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

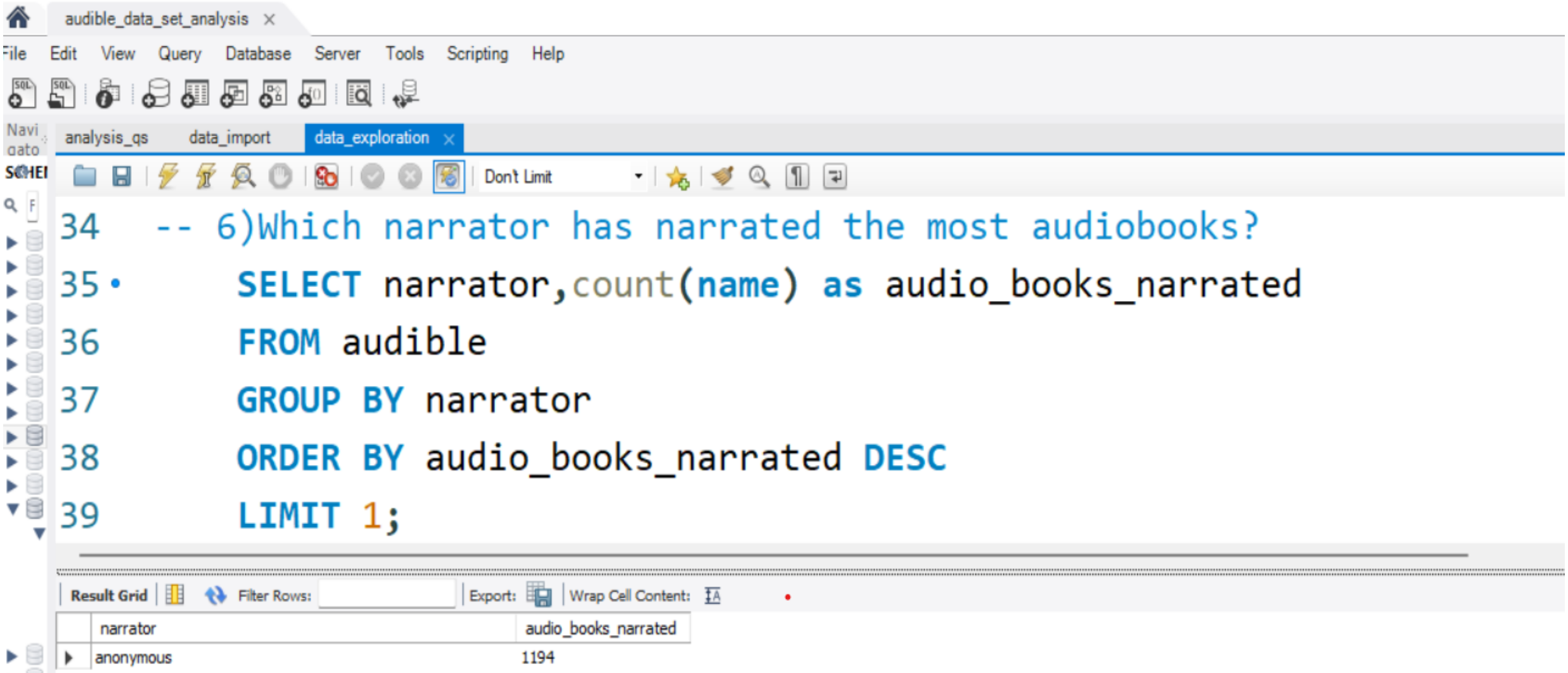
analysis_qs data_import data_exploration

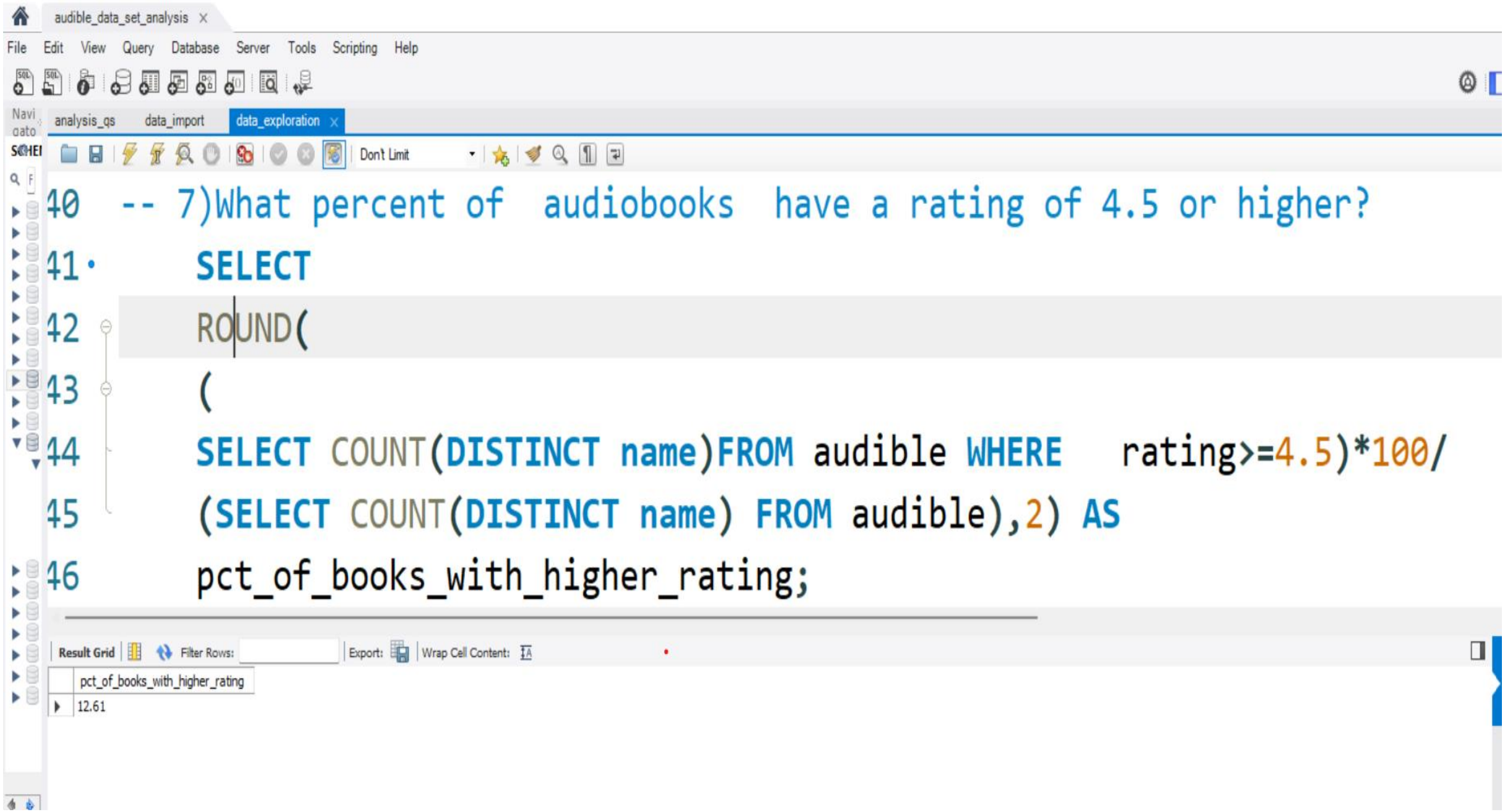
Don't Limit

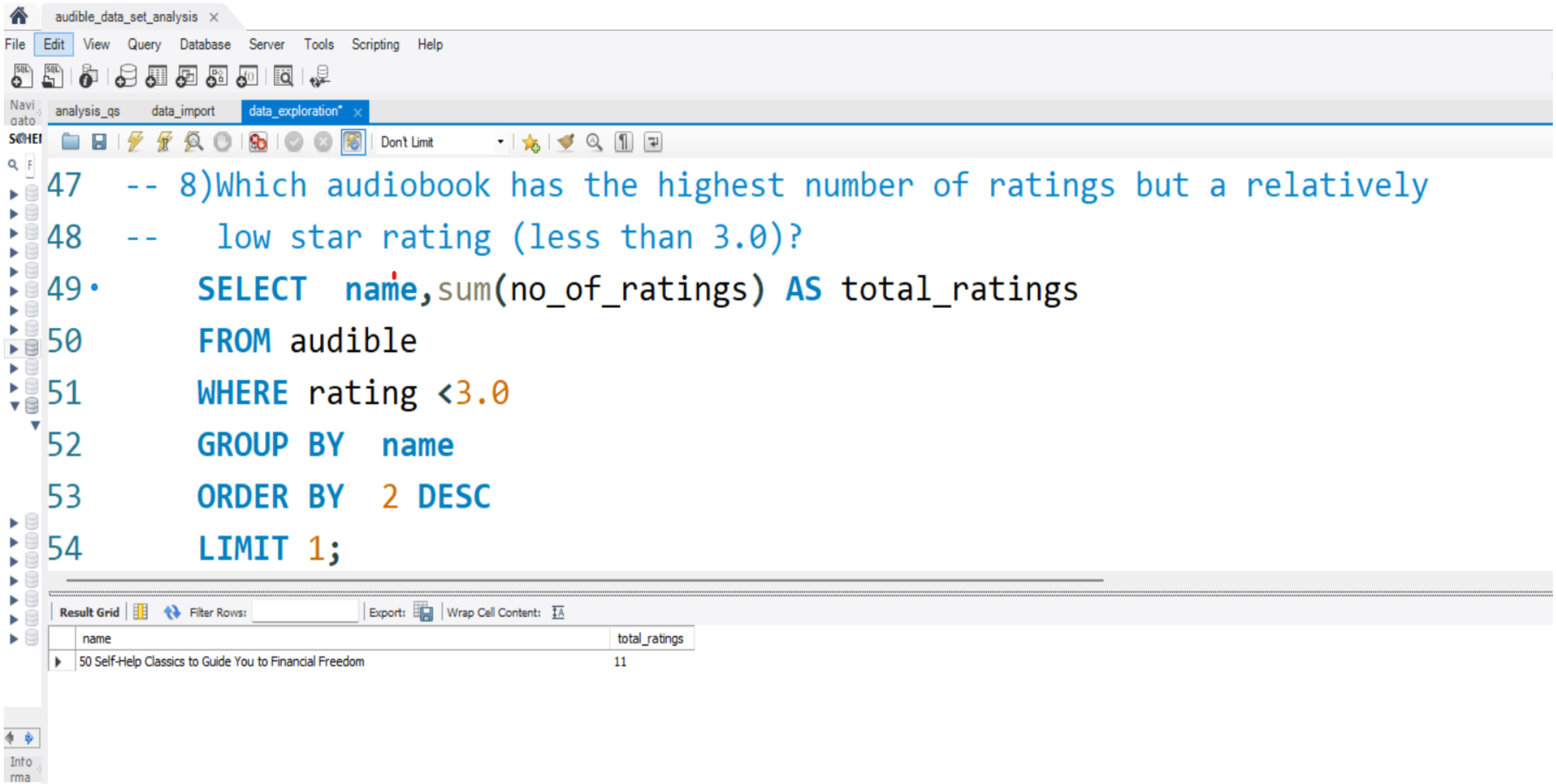
```sql
-- 5)How many unique audiobooks were released each year, and what is the average price for each year?
SELECT year,count(author) as no_of_audio_books,round(avg(price),1) as average_price_INR
FROM
(
SELECT DISTINCT author ,year(release_date) AS year,price
FROM audible
) T
GROUP BY year
ORDER BY year;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

year	no_of_audio_books	average_price_INR
2004	268	679.4
2005	314	617.8
2006	566	620.3
2007	531	578.5
2008	890	596.5
2009	1237	586.6
2010	1158	614.4
2011	1306	639
2012	1595	615.5
2013	2711	657.1
2014	2277	671.2
2015	2511	636.8
2016	2653	648
2017	3743	644.9
2018	5000	626.5
2019	7076	613.7
2020	8712	612.7
2021	18225	590.1
2022	7876	520.7


audible_data_set_analysis
File Edit View Query Database Server Tools Scripting Help
analysis_qs
data_import
data_exploration
Don't Limit
34 -- 6)Which narrator has narrated the most audiobooks?
35 • SELECT narrator,count(name) as audio_books_narrated
36 FROM audible
37 GROUP BY narrator
38 ORDER BY audio_books_narrated DESC
39 LIMIT 1;
Result Grid
Filter Rows:
Export:
Wrap Cell Content:
narrator
audio_books_narrated
anonymous
1194


audible_data_set_analysis
File Edit View Query Database Server Tools Scripting Help
analysis_qs
data_import
data_exploration
Don't Limit
-- 7)What percent of audiobooks have a rating of 4.5 or higher?
SELECT
ROUND(
(
SELECT COUNT(DISTINCT name)FROM audible WHERE rating>=4.5)*100/
(SELECT COUNT(DISTINCT name) FROM audible),2) AS
pct_of_books_with_higher_rating;
Result Grid
Filter Rows:
Export:
Wrap Cell Content:
pct_of_books_with_higher_rating
12.61


audible_data_set_analysis
File Edit View Query Database Server Tools Scripting Help
analysis_qs data_import data_exploration*
Don't Limit
-- 8)Which audiobook has the highest number of ratings but a relatively
-- low star rating (less than 3.0)?
SELECT name,sum(no_of_ratings) AS total_ratings
FROM audible
WHERE rating <3.0
GROUP BY name
ORDER BY 2 DESC
LIMIT 1;
Result Grid Filter Rows: Export: Wrap Cell Content:
name total_ratings
50 Self-Help Classics to Guide You to Financial Freedom 11

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

analysis_qs data_import data_exploration*

Don't Limit

```sql
-- 9)What are the top 3 most expensive audiobooks that are longer than 8 hours?
    SELECT name FROM
    (
    SELECT  name,min(duration),min(price)
    FROM audible
    WHERE duration>480
    group by name
    ORDER BY 3 DESC
    ) t
    LIMIT 3;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

name
せかい伝記図書館 全集
せかい童話図書館 全集
シャーロック・ホームズ傑作選

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

analysis_qs data_import data_exploration*

Don't Limit

```
65  10)Which authors have produced more than 10 audiobooks, and
66     what is their average rating?
67  SELECT author,count(name) AS no_of_audibooks_produced,
68  round(avg(rating),1) as average_rating
69  FROM
70  (
71  SELECT DISTINCT author,name,rating
72  FROM audible
73  WHERE rating  IS NOT NULL
74  ) t
75  GROUP BY author
76  HAVING count(name)>10
77  ORDER BY 2 DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

author	no_of_audibooks_produced	average_rating
InnovativeLanguageLearning	55	4.1
P.G.Wodehouse	52	4.6
HarvardBusinessReview	44	4.5
RuskinBond	39	4.5
NoraRoberts	39	4.5
GertrudeChandlerWarner	38	4.8
BrianTracy	36	4.7
Osho	33	4.8

Result 11

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

analysis_qs data_import data_exploration*

Don't Limit

```
89  12)Find the  top 5 narrators who narrated the most number of audiobooks.
90  SELECT narrator,count(name) AS no_of_audiobooks_narrated
91  FROM audible
92  GROUP BY  1
93  ORDER BY  2 DESC
94  LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

narrator	no_of_audiobooks_narrated
anonymous	1194
矢島雅弘,石橋遊	874
Intuitive	465
uncredited	326
中西貴之,BJ	311

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

analysis_qs data_import data_exploration*

Don't Limit

```
89  12)Find the  top 5 narrators who narrated the most number of audiobooks.
90  SELECT narrator,count(name) AS no_of_audiobooks_narrated
91  FROM audible
92  GROUP BY  1
93  ORDER BY  2 DESC
94  LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

narrator	no_of_audiobooks_narrated
anonymous	1194
矢島雅弘,石橋遊	874
Intuitive	465
uncredited	326
中西貴之,BJ	311

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

analysis_qs data_import data_exploration*

Don't Limit

```
95  · 13)What percentage of audiobooks have a rating of at least 4 stars?
96  •  SELECT
97     ROUND(
98     (SELECT COUNT(DISTINCT name) FROM audible WHERE rating>=4)*100/
99     (SELECT COUNT(DISTINCT name) FROM audible),2) AS pct_of_audiobooks_atleast_4_stars;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

pct_of_audiobooks_atleast_4_stars
15.74

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

analysis_qs data_import data_exploration*

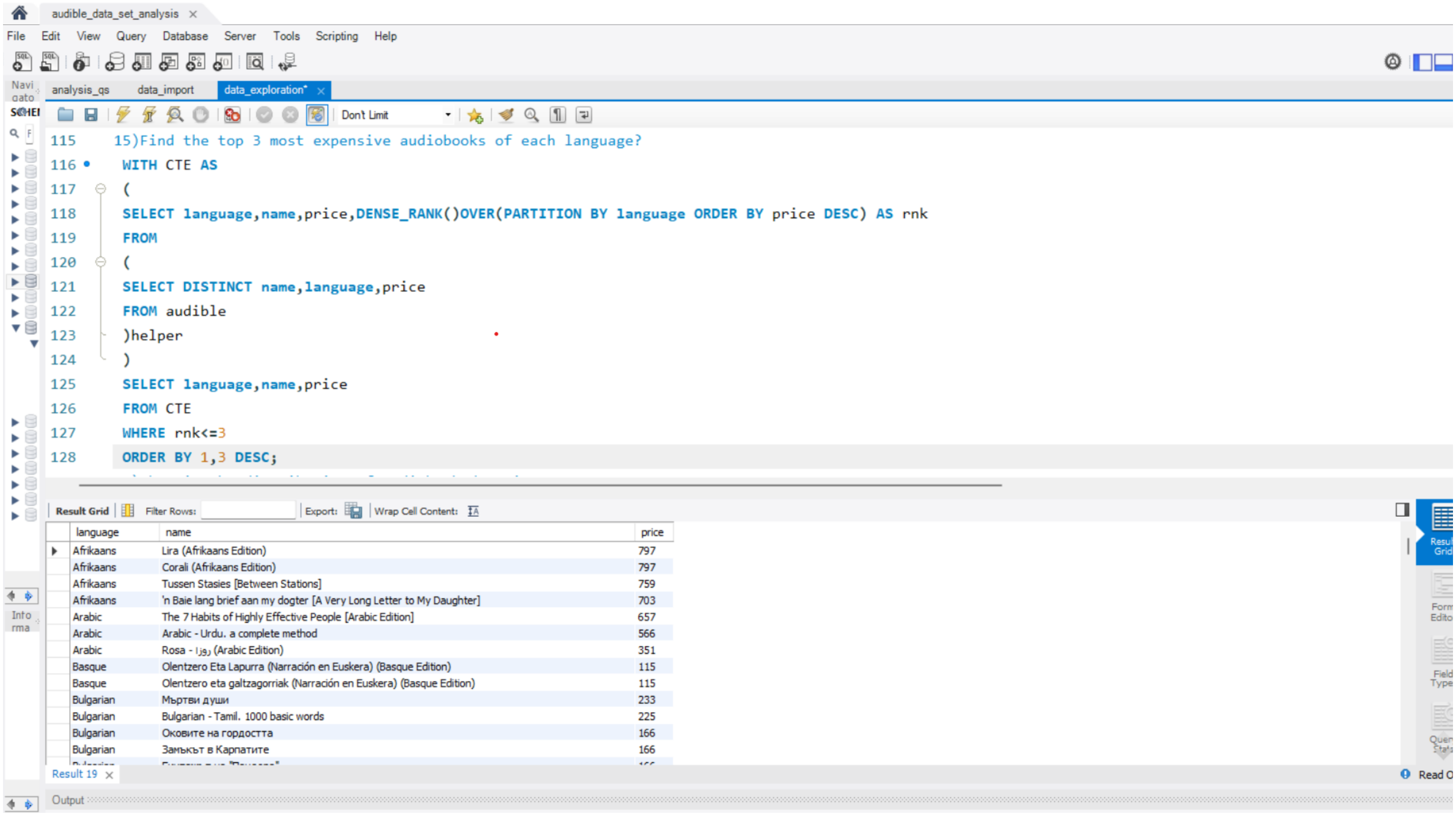
Don't Limit

```
14)Find the percentage breakdown of audiobooks by each language
WITH CTE AS
(
SELECT language,count(name) as num,max(total) AS total
FROM
(
SELECT distinct name,language,
(SELECT count(distinct name) FROM audible) AS total
FROM audible
)T
GROUP BY 1
)
SELECT language,ROUND(num*100/total,2) as  pct_of_audiobooks
FROM cte
ORDER BY 2 DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

language	pct_of_audiobooks
English	69.83
German	9.79
Spanish	4.14
Japanese	3.79
Italian	3.22
French	2.85
Russian	2.17
Danish	1.13
Portuguese	0.64
Swedish	0.62
Hindi	0.52

Result 16


audible_data_set_analysis
File Edit View Query Database Server Tools Scripting Help
analysis_qs data_import data_exploration*
Don't Limit
15)Find the top 3 most expensive audiobooks of each language?
WITH CTE AS
(
SELECT language,name,price,DENSE_RANK()OVER(PARTITION BY language ORDER BY price DESC) AS rnk
FROM
(
SELECT DISTINCT name,language,price
FROM audible
)helper
)
SELECT language,name,price
FROM CTE
WHERE rnk<=3
ORDER BY 1,3 DESC;
Result Grid Filter Rows: Export: Wrap Cell Content:
language name price
Afrikaans Lira (Afrikaans Edition) 797
Afrikaans Corali (Afrikaans Edition) 797
Afrikaans Tussen Stasies [Between Stations] 759
Afrikaans 'n Baie lang brief aan my dogter [A Very Long Letter to My Daughter] 703
Arabic The 7 Habits of Highly Effective People [Arabic Edition] 657
Arabic Arabic - Urdu. a complete method 566
Arabic Rosa - ﻭﺭﺩ (Arabic Edition) 351
Basque Olentzero Eta Lapurra (Narración en Euskera) (Basque Edition) 115
Basque Olentzero eta galtzagorriak (Narración en Euskera) (Basque Edition) 115
Bulgarian Мъртви души 233
Bulgarian Bulgarian - Tamil. 1000 basic words 225
Bulgarian Оковите на гордостта 166
Bulgarian Замъкът в Карпатите 166
Result 19
Output

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

analysis_qs | data_import | data_exploration*

Don't Limit

```
129   16)What is the distribution of audiobook durations
130   (e.g., how many fall into categories like less than 0-2 hours, 2-4 hours, 4-6 hours,6+ hours)?
131 • WITH CTE3 AS
132   (
133   WITH CTE2 AS
134   (
135   WITH CTE1 AS
136   (
137   SELECT name,round(duration/60,2) as length_hrs,count(*)over() as total
138   FROM
139   (
140   SELECT name,min(duration) AS duration FROM audible
141   GROUP BY name,duration
142   )t
143   )
```

audible_data_set_analysis

File Edit View Query Database Server Tools Scripting Help

analysis_qs | data_import | data_exploration*

Don't Limit

```
SELECT
CASE WHEN length_hrs <2.00 THEN "0-2 hours"
WHEN length_hrs>=2.00 AND length_hrs<4.00 THEN "2-4 hours"
WHEN length_hrs>=4.00 AND length_hrs<6.00 THEN "4-6 hours"
ELSE "6+ hours"
END AS bins,total
FROM CTE1
)
SELECT bins,count(bins) AS num,max(total) AS total
FROM CTE2
GROUP BY 1
ORDER BY 1
)
SELECT bins,ROUND(num*100/total,2) AS  pct_distribution
FROM CTE3;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

bins	pct_distribution
0-2 hours	22.97
2-4 hours	11.02
4-6 hours	12.59
6+ hours	53.42