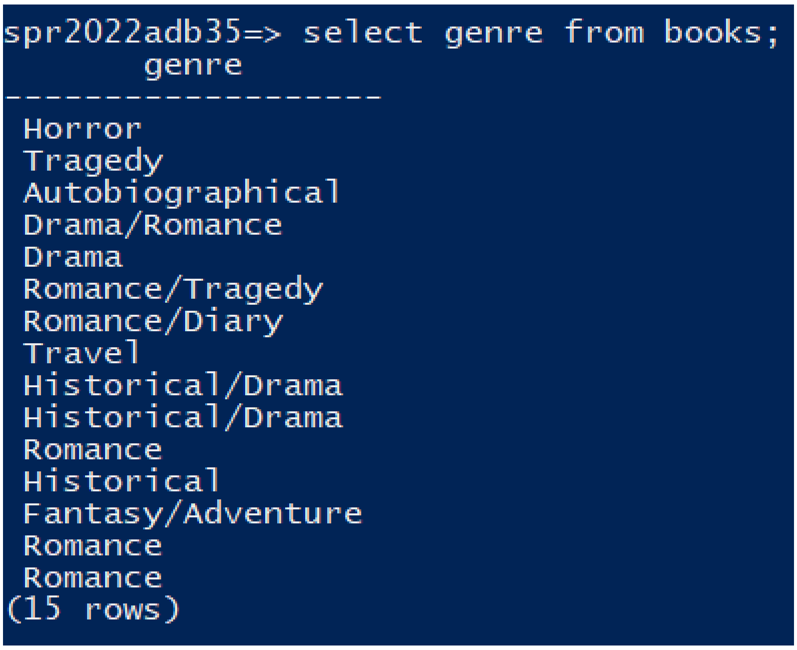
ACTIVITY 3.1: -

Question-1: - Find genres of books including duplicates

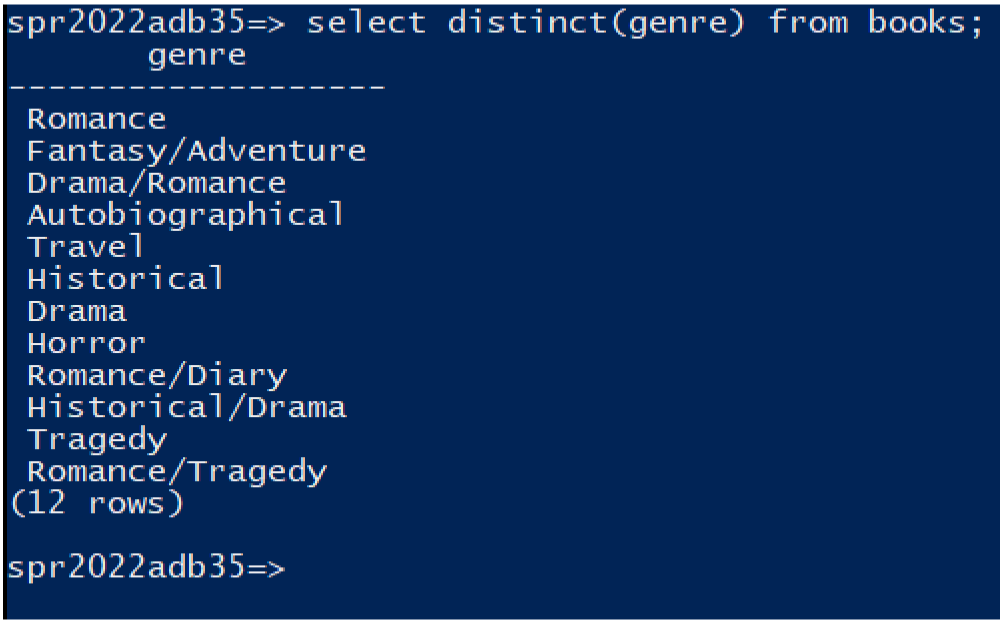
Answer: - select genre from books;



```
spr2022adb35=> select genre from books;
       genre
--------------------
 Horror
 Tragedy
 Autobiographical
 Drama/Romance
 Drama
 Romance/Tragedy
 Romance/Diary
 Travel
 Historical/Drama
 Historical/Drama
 Romance
 Historical
 Fantasy/Adventure
 Romance
 Romance
(15 rows)
```

Question-2: - Find genres of books excluding duplicates

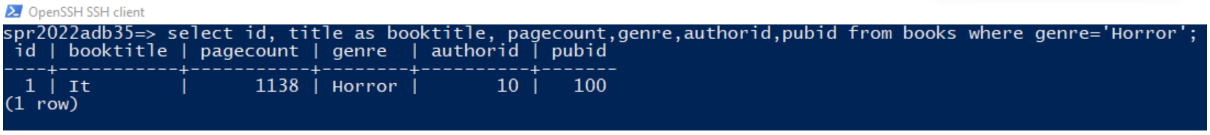
Answer: - select distinct(genre) from books;



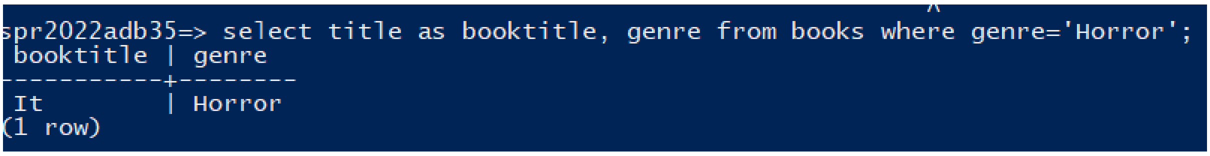
```
spr2022adb35=> select distinct(genre) from books;
       genre
--------------------
 Romance
 Fantasy/Adventure
 Drama/Romance
 Autobiographical
 Travel
 Historical
 Drama
 Horror
 Romance/Diary
 Historical/Drama
 Tragedy
 Romance/Tragedy
(12 rows)

spr2022adb35=>
```

Question-3: - Find all horror books with alias booktitle for title.

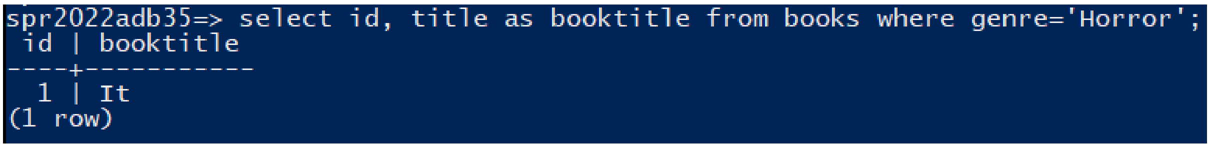
Answer: - select id, title as booktitle, pagecount,genre,authorid,pubid from books where genre='Horror';



```
spr2022adb35=> select id, title as booktitle, pagecount,genre,authorid,pubid from books where genre='Horror';
 id | booktitle | pagecount | genre  | authorid | pubid
----+-----------+-----------+--------+----------+-------
  1 | It        |      1138 | Horror |       10 |   100
(1 row)
```



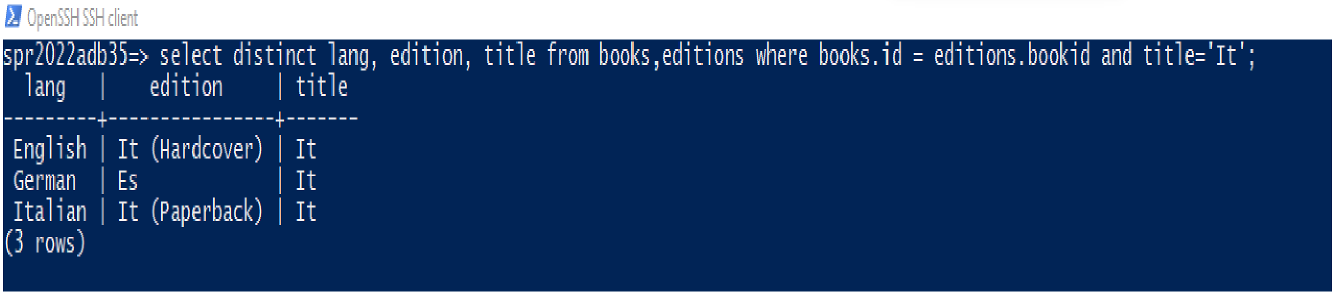
```
spr2022adb35=> select title as booktitle, genre from books where genre='Horror';
 booktitle | genre
-----------+--------
 It        | Horror
(1 row)
```



```
spr2022adb35=> select id, title as booktitle from books where genre='Horror';
 id | booktitle
----+-----------
  1 | It
(1 row)
```

Question-4: - Find the languages of all editions of “it”.

Answer: - select distinct lang, edition, title from books,editions where books.id = editions.bookid and title='It';

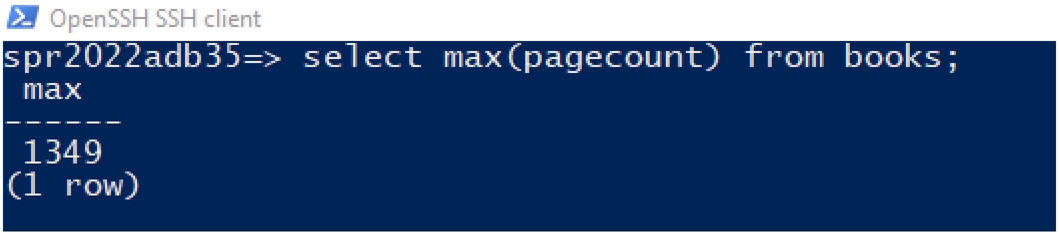


```
spr2022adb35=> select distinct lang, edition, title from books,editions where books.id = editions.bookid and title='It';
  lang   |    edition     | title
---------+----------------+-------
 English | It (Hardcover) | It
 German  | Es             | It
 Italian | It (Paperback) | It
(3 rows)
```

ACTIVITY-3.2(a) Aggregates: - Try a few queries

Question-1: - Try the select query

Answer: - select max(pagecount) from books;

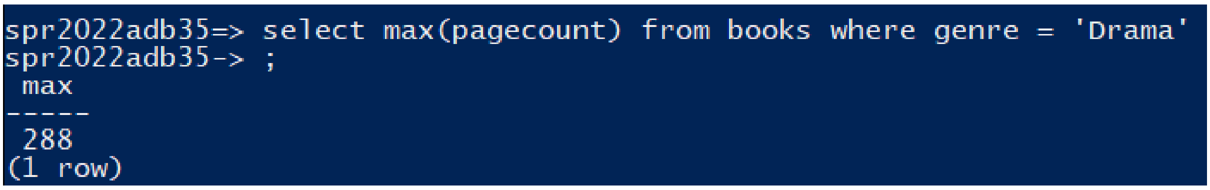


OpenSSH SSH client

```
spr2022adb35=> select max(pagecount) from books;
 max
------
 1349
(1 row)
```

Question-2: - Adding where clause to question-1.

Answer: - select max(pagecount) from books where genre = 'Drama'



```
spr2022adb35=> select max(pagecount) from books where genre = 'Drama'
spr2022adb35-> ;
 max
-----
 288
(1 row)
```

Question-3: - select query with error

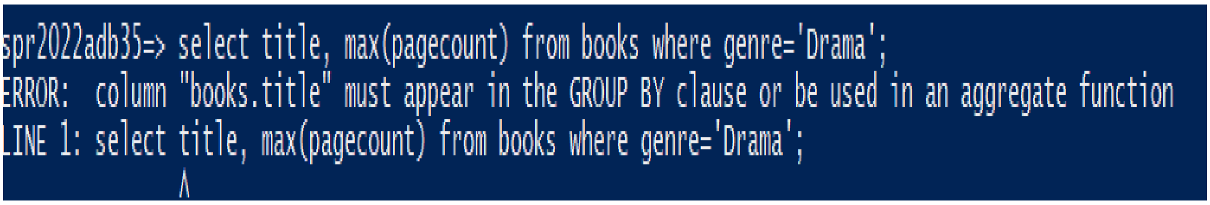
Answer: - select title, max(pagecount) from books where genre='Drama';

It gives an error message as given below: -

spr2022adb35=> select title, max(pagecount) from books where genre='Drama';

ERROR: column "books.title" must appear in the GROUP BY clause or be used in an aggregate function

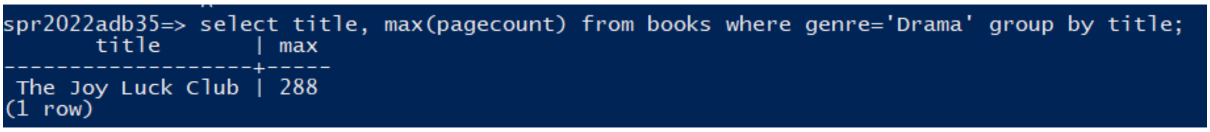
LINE 1: select title, max(pagecount) from books where genre='Drama';



```
spr2022adb35=> select title, max(pagecount) from books where genre='Drama';
ERROR:  column "books.title" must appear in the GROUP BY clause or be used in an aggregate function
LINE 1: select title, max(pagecount) from books where genre='Drama';
               ^
```

This error can be resolved by adding the group by clause as given below: -

select title, max(pagecount) from books where genre='Drama' group by title;



```
spr2022adb35=> select title, max(pagecount) from books where genre='Drama' group by title;
       title       | max
-------------------+-----
 The Joy Luck Club | 288
(1 row)
```

Activity 3.2(b) Aggregates: Average and Max

Question-1: - Example instance for different averages

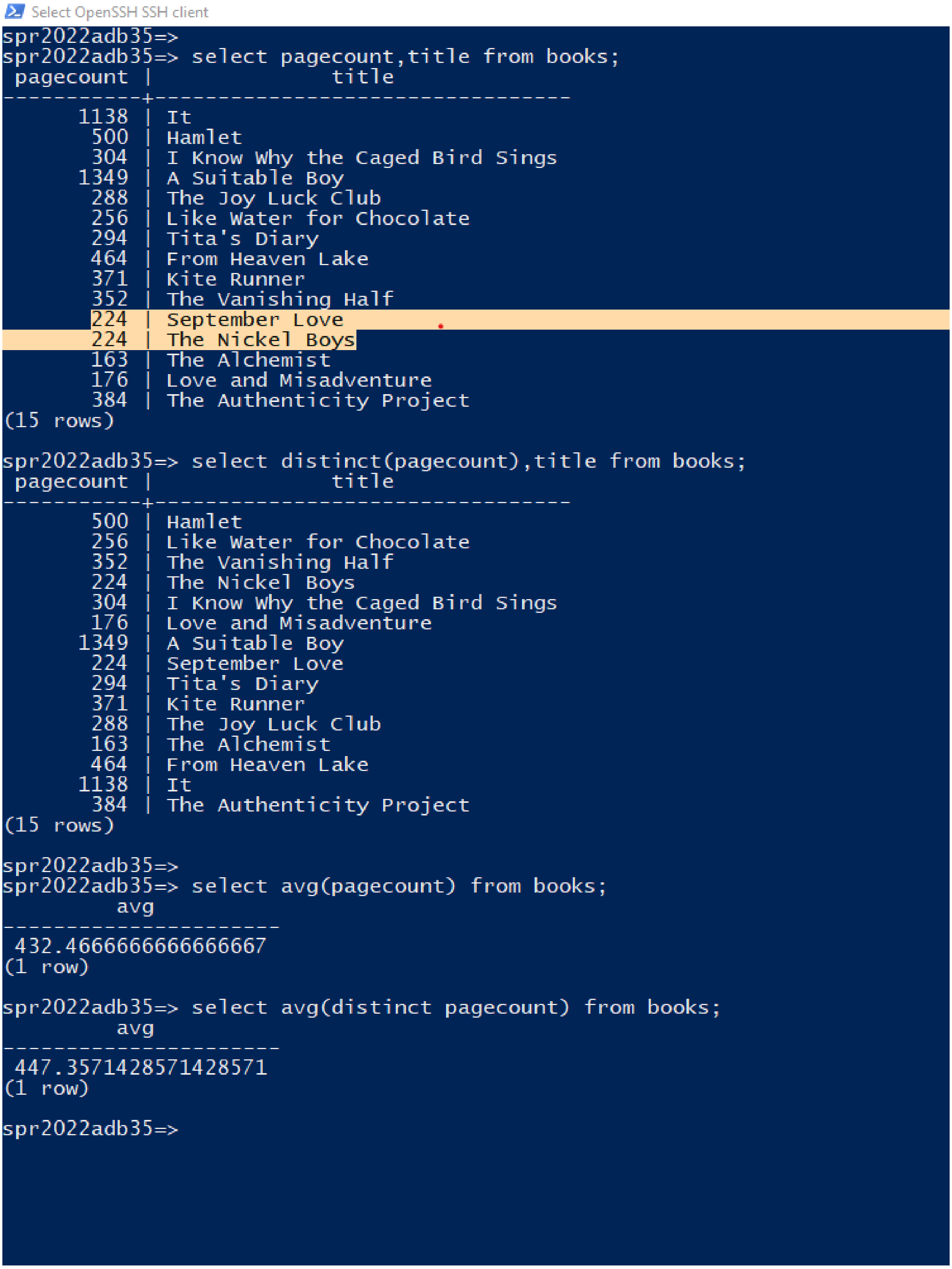
Answer: - Both of these queries will yield different results when there is a duplicate entry for pagecount.

In this case, there is a duplicate value as highlighted in the screenshot below: -

Pagecount-224 for book titles September love and The nickel boys.

This means when distinct pagecount was not used for calculating the average of pagecounts, then the result was 432.4667(approx.) whereas when distinct pagecount was used, then the result was 447.357 (approx.).

This is also present in the screenshot presented on the next page.



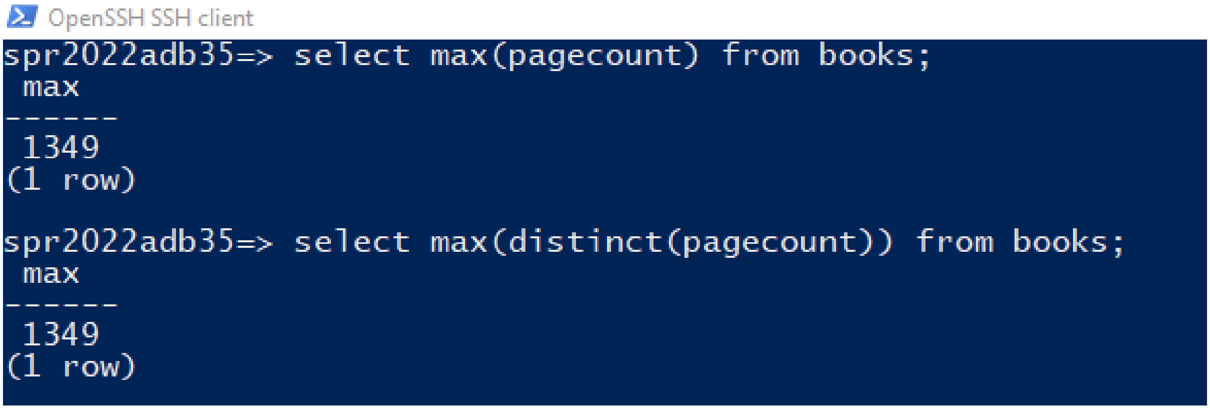
```
spr2022adb35=>
spr2022adb35=> select pagecount,title from books;
 pagecount |               title
-----------+---------------------------------
      1138 | It
       500 | Hamlet
       304 | I Know Why the Caged Bird Sings
      1349 | A Suitable Boy
       288 | The Joy Luck Club
       256 | Like Water for Chocolate
       294 | Tita's Diary
       464 | From Heaven Lake
       371 | Kite Runner
       352 | The Vanishing Half
       224 | September Love
       224 | The Nickel Boys
       163 | The Alchemist
       176 | Love and Misadventure
       384 | The Authenticity Project
(15 rows)

spr2022adb35=> select distinct(pagecount),title from books;
 pagecount |               title
-----------+---------------------------------
       500 | Hamlet
       256 | Like Water for Chocolate
       352 | The Vanishing Half
       224 | The Nickel Boys
       304 | I Know Why the Caged Bird Sings
       176 | Love and Misadventure
      1349 | A Suitable Boy
       224 | September Love
       294 | Tita's Diary
       371 | Kite Runner
       288 | The Joy Luck Club
       163 | The Alchemist
       464 | From Heaven Lake
      1138 | It
       384 | The Authenticity Project
(15 rows)

spr2022adb35=>
spr2022adb35=> select avg(pagecount) from books;
         avg
----------------------
 432.4666666666666667
(1 row)

spr2022adb35=> select avg(distinct pagecount) from books;
         avg
----------------------
 447.3571428571428571
(1 row)

spr2022adb35=>
```

Question-2: - can max(attribute) return different values or not?

Answer: - No, the value of these queries will always be the same and that is equal to the highest value of the attribute present in the table.

There is only one condition where the max(attribute) query can return different results (when distinct is used and when not) is only when a group by clause is used. This is because group by does a grouping according to the attribute and that can vary the result.

This can also be proved by using the screenshot below: -



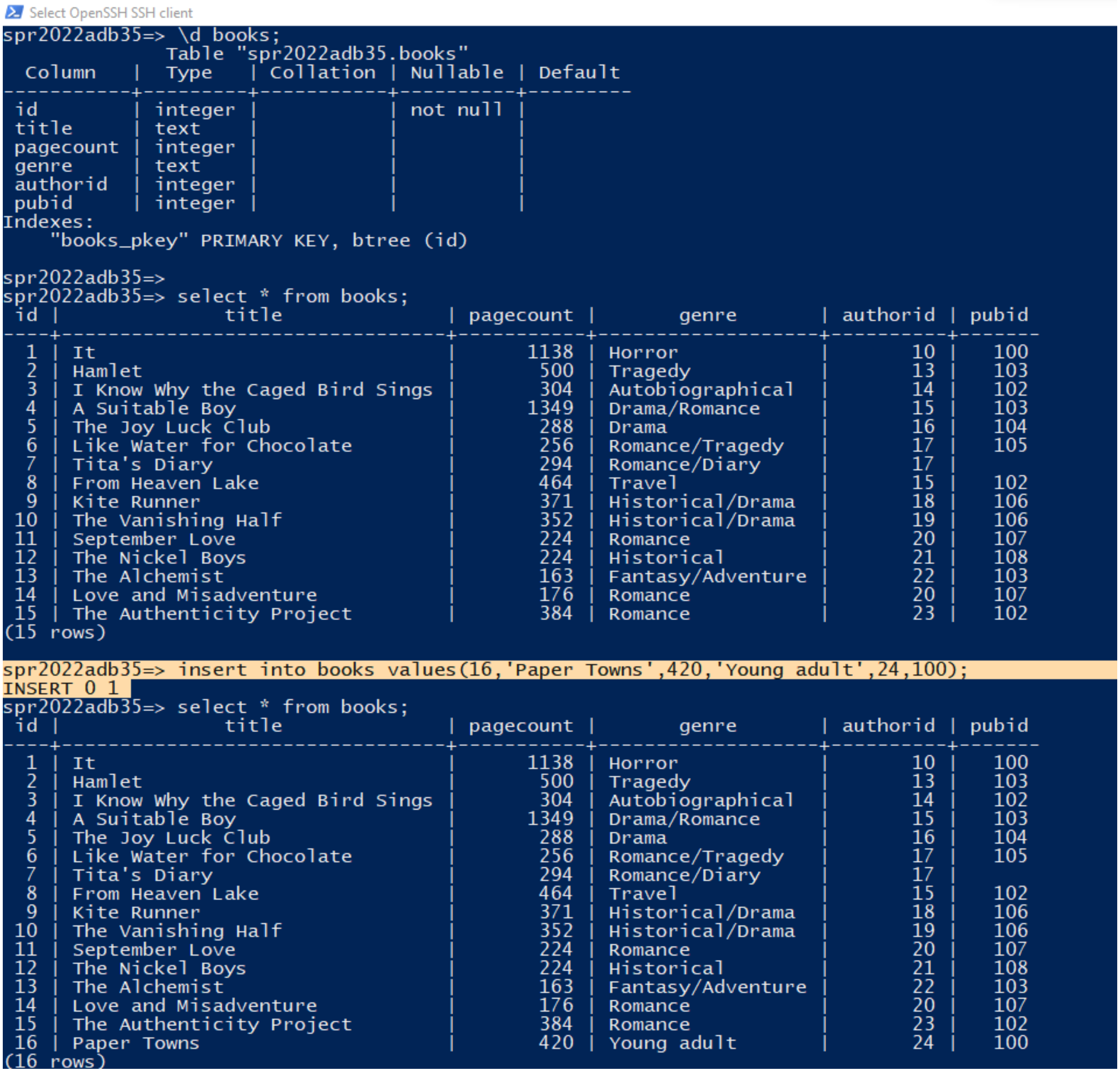
```
spr2022adb35=> select max(pagecount) from books;
 max
------
 1349
(1 row)

spr2022adb35=> select max(distinct(pagecount)) from books;
 max
------
 1349
(1 row)
```

ACTIVITY 3.3(A): - INSERTS

Question-1: - Insert query for books

Answer: - insert into books values(16,'Paper Towns',420,'Young adult',24,100);

This is also highlighted in the screenshot below along with the select * to confirm the insertion.

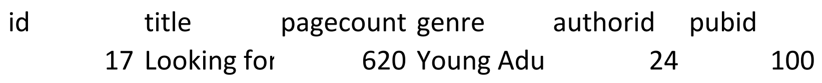


```
spr2022adb35=> \d books;
              Table "spr2022adb35.books"
  Column   |  Type   | Collation | Nullable | Default
-----------+---------+-----------+----------+---------
 id        | integer |           | not null |
 title     | text    |           |          |
 pagecount | integer |           |          |
 genre     | text    |           |          |
 authorid  | integer |           |          |
 pubid     | integer |           |          |
Indexes:
    "books_pkey" PRIMARY KEY, btree (id)

spr2022adb35=>
spr2022adb35=> select * from books;
 id |              title               | pagecount |       genre        | authorid | pubid
----+----------------------------------+-----------+--------------------+----------+-------
  1 | It                               |      1138 | Horror             |       10 |   100
  2 | Hamlet                           |       500 | Tragedy            |       13 |   103
  3 | I Know Why the Caged Bird Sings  |       304 | Autobiographical   |       14 |   102
  4 | A Suitable Boy                   |      1349 | Drama/Romance      |       15 |   103
  5 | The Joy Luck Club                |       288 | Drama              |       16 |   104
  6 | Like Water for Chocolate         |       256 | Romance/Tragedy    |       17 |   105
  7 | Tita's Diary                     |       294 | Romance/Diary      |       17 |
  8 | From Heaven Lake                 |       464 | Travel             |       15 |   102
  9 | Kite Runner                      |       371 | Historical/Drama   |       18 |   106
 10 | The Vanishing Half               |       352 | Historical/Drama   |       19 |   106
 11 | September Love                   |       224 | Romance            |       20 |   107
 12 | The Nickel Boys                  |       224 | Historical         |       21 |   108
 13 | The Alchemist                    |       163 | Fantasy/Adventure  |       22 |   103
 14 | Love and Misadventure            |       176 | Romance            |       20 |   107
 15 | The Authenticity Project         |       384 | Romance            |       23 |   102
(15 rows)

spr2022adb35=> insert into books values(16,'Paper Towns',420,'Young adult',24,100);
INSERT 0 1
spr2022adb35=> select * from books;
 id |              title               | pagecount |       genre        | authorid | pubid
----+----------------------------------+-----------+--------------------+----------+-------
  1 | It                               |      1138 | Horror             |       10 |   100
  2 | Hamlet                           |       500 | Tragedy            |       13 |   103
  3 | I Know Why the Caged Bird Sings  |       304 | Autobiographical   |       14 |   102
  4 | A Suitable Boy                   |      1349 | Drama/Romance      |       15 |   103
  5 | The Joy Luck Club                |       288 | Drama              |       16 |   104
  6 | Like Water for Chocolate         |       256 | Romance/Tragedy    |       17 |   105
  7 | Tita's Diary                     |       294 | Romance/Diary      |       17 |
  8 | From Heaven Lake                 |       464 | Travel             |       15 |   102
  9 | Kite Runner                      |       371 | Historical/Drama   |       18 |   106
 10 | The Vanishing Half               |       352 | Historical/Drama   |       19 |   106
 11 | September Love                   |       224 | Romance            |       20 |   107
 12 | The Nickel Boys                  |       224 | Historical         |       21 |   108
 13 | The Alchemist                    |       163 | Fantasy/Adventure  |       22 |   103
 14 | Love and Misadventure            |       176 | Romance            |       20 |   107
 15 | The Authenticity Project         |       384 | Romance            |       23 |   102
 16 | Paper Towns                      |       420 | Young adult        |       24 |   100
(16 rows)
```

Question-2: - Insertion using the /COPY command

Answer: - The first step is to create the csv file on the local PC.

The file’s contents are attached below along with the file: -



id	title	pagecount	genre	authorid	pubid
17	Looking for	620	Young Adu	24	100

books-copy.csv

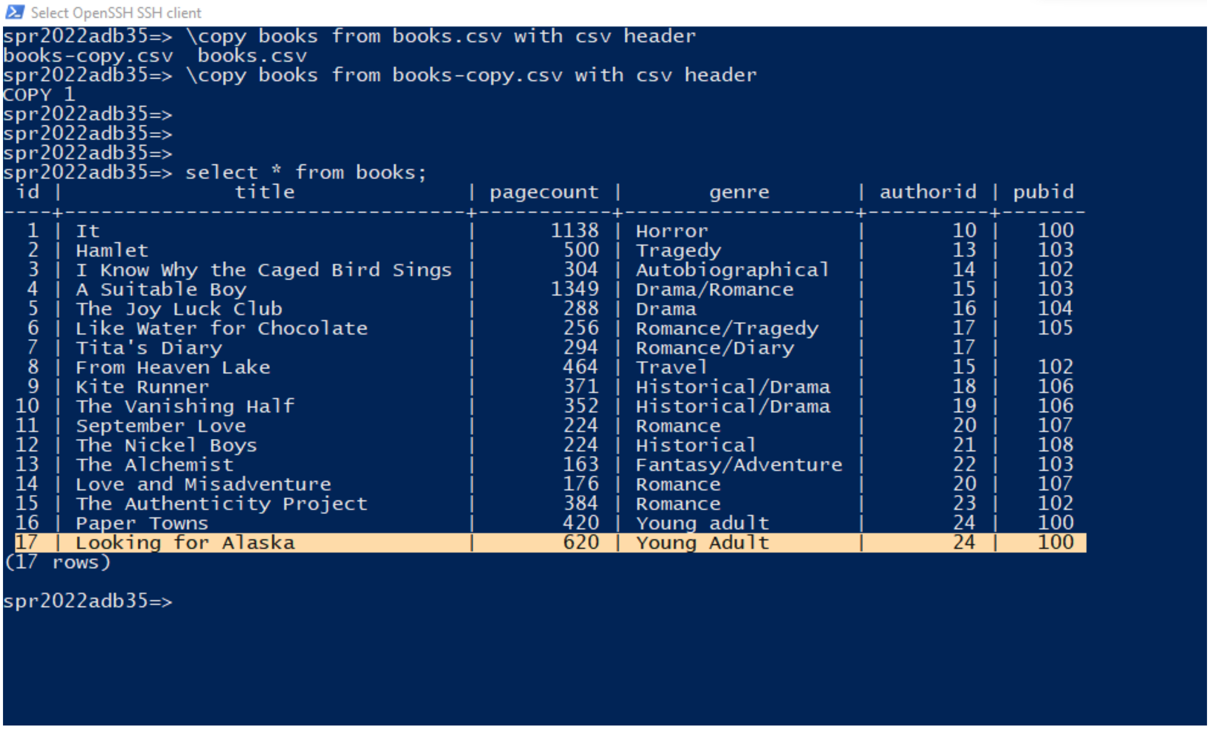
After this, we must copy this file from local machine to the remote machine (linux.pdx.cs.edu). This will be done by the following command: -

```
scp books-copy.csv parth2@131.252.208.103:/u/parth2/DBMS/bookschema
```

After this, we will use the following command to copy the contents of the file into the database: -

```
\copy books from books-copy.csv with csv header
```

The output along with the select command highlighting the inserted field is given below: -



```
spr2022adb35=> \copy books from books.csv with csv header
books-copy.csv  books.csv
spr2022adb35=> \copy books from books-copy.csv with csv header
COPY 1
spr2022adb35=>
spr2022adb35=>
spr2022adb35=>
spr2022adb35=> select * from books;
 id |               title               | pagecount |       genre        | authorid | pubid
----+-----------------------------------+-----------+--------------------+----------+-------
  1 | It                                |      1138 | Horror             |       10 |   100
  2 | Hamlet                            |       500 | Tragedy            |       13 |   103
  3 | I Know Why the Caged Bird Sings   |       304 | Autobiographical   |       14 |   102
  4 | A Suitable Boy                    |      1349 | Drama/Romance      |       15 |   103
  5 | The Joy Luck Club                 |       288 | Drama              |       16 |   104
  6 | Like Water for Chocolate          |       256 | Romance/Tragedy    |       17 |   105
  7 | Tita's Diary                      |       294 | Romance/Diary      |       17 |
  8 | From Heaven Lake                  |       464 | Travel             |       15 |   102
  9 | Kite Runner                       |       371 | Historical/Drama   |       18 |   106
 10 | The Vanishing Half                |       352 | Historical/Drama   |       19 |   106
 11 | September Love                    |       224 | Romance            |       20 |   107
 12 | The Nickel Boys                   |       224 | Historical         |       21 |   108
 13 | The Alchemist                     |       163 | Fantasy/Adventure  |       22 |   103
 14 | Love and Misadventure             |       176 | Romance            |       20 |   107
 15 | The Authenticity Project          |       384 | Romance            |       23 |   102
 16 | Paper Towns                       |       420 | Young adult        |       24 |   100
 17 | Looking for Alaska                |       620 | Young Adult        |       24 |   100
(17 rows)

spr2022adb35=>
```

Activity-3.3(B): -Default and null values

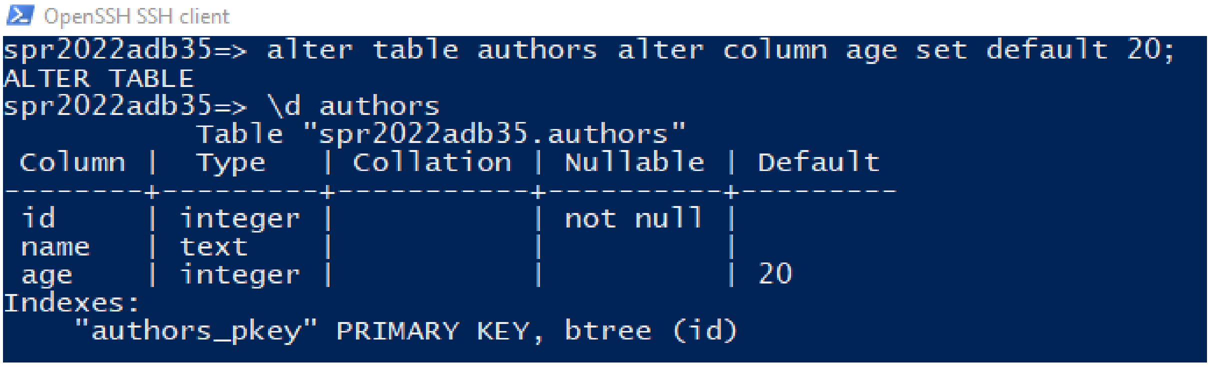
Question: - add default value to age in author, insert a author without an age and query to return author added above

Answer: - The ***first task*** to be done is to add a default value to the author's age in author table.

This will be done by the following query: -

alter table authors alter column age set default 20;

This change can be seen in the screenshot below: -

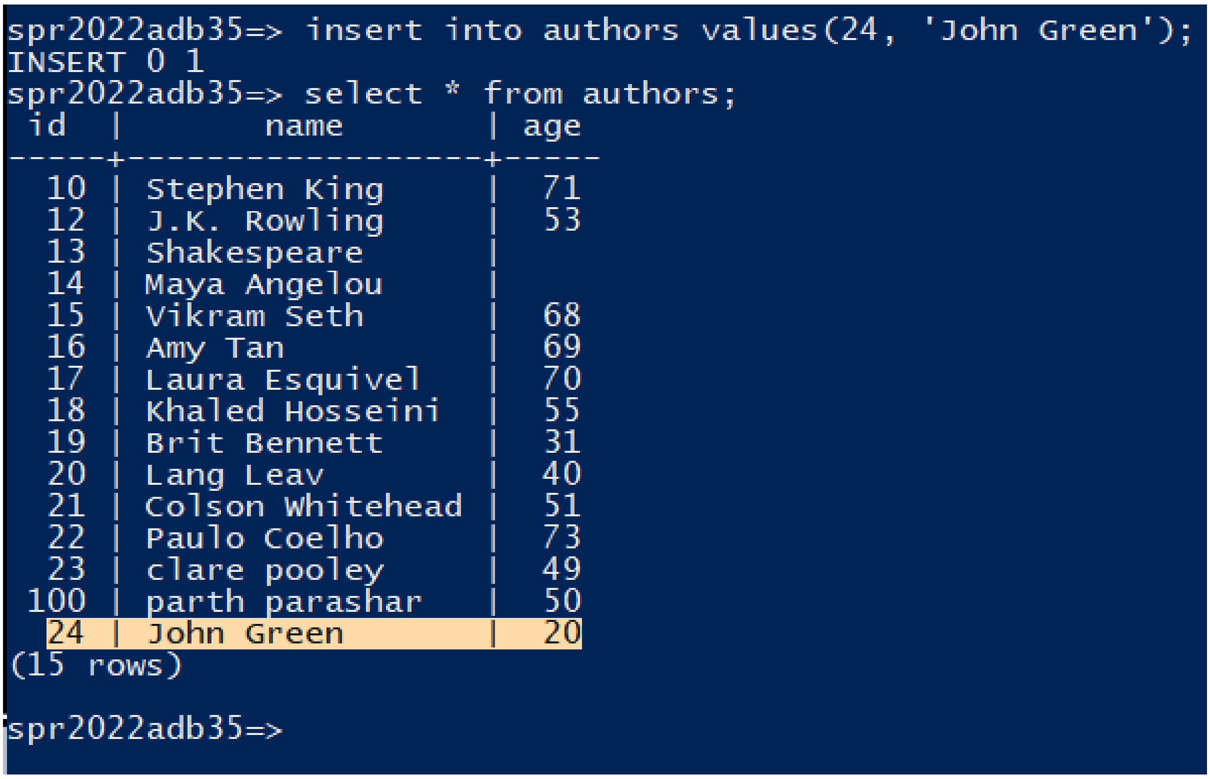


```
spr2022adb35=> alter table authors alter column age set default 20;
ALTER TABLE
spr2022adb35=> \d authors
           Table "spr2022adb35.authors"
 Column |  Type   | Collation | Nullable | Default
--------+---------+-----------+----------+---------
 id     | integer |           | not null |
 name   | text    |           |          |
 age    | integer |           |          | 20
Indexes:
    "authors_pkey" PRIMARY KEY, btree (id)
```

The ***second task*** to do is to insert an author without an age so that it sets to the default age. This will be done by the following query: -

insert into authors values(24, 'John Green');

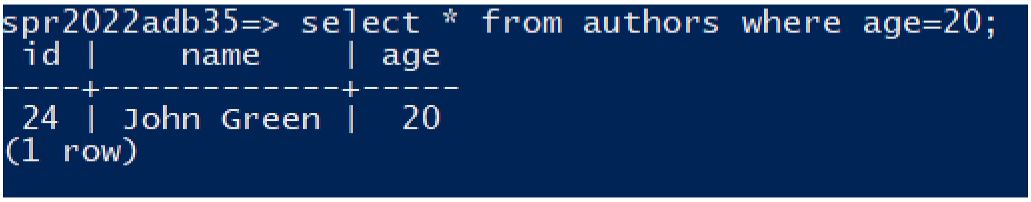
To verify the insertion, we have the screenshot containing the highlighted insertion.



```
spr2022adb35=> insert into authors values(24, 'John Green');
INSERT 0 1
spr2022adb35=> select * from authors;
 id  |       name        | age
-----+-------------------+-----
  10 | Stephen King      |  71
  12 | J.K. Rowling      |  53
  13 | Shakespeare       |
  14 | Maya Angelou      |
  15 | Vikram Seth       |  68
  16 | Amy Tan           |  69
  17 | Laura Esquivel    |  70
  18 | Khaled Hosseini   |  55
  19 | Brit Bennett      |  31
  20 | Lang Leav         |  40
  21 | Colson Whitehead  |  51
  22 | Paulo Coelho      |  73
  23 | clare pooley      |  49
 100 | parth parashar    |  50
  24 | John Green        |  20
(15 rows)

spr2022adb35=>
```

The ***third task*** to complete is to write a select query to return the author added in the second task: -

select * from authors where age=20;



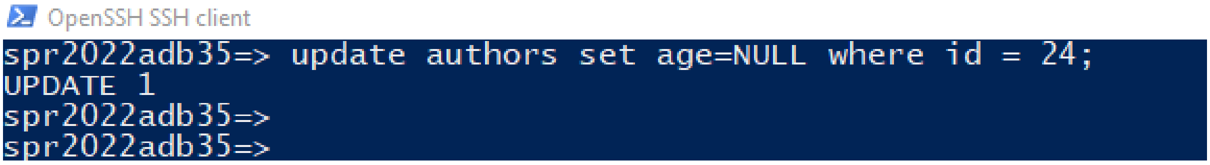
```
spr2022adb35=> select * from authors where age=20;
 id |    name    | age
----+------------+-----
 24 | John Green |  20
(1 row)
```

ACTIVITY-3.4: - Updates and Null Values

Question-1: - Update the tuple inserted in 3.4 to set their age to null?

Answer: - update authors set age=NULL where id = 24;

The screenshot for this query is given below: -

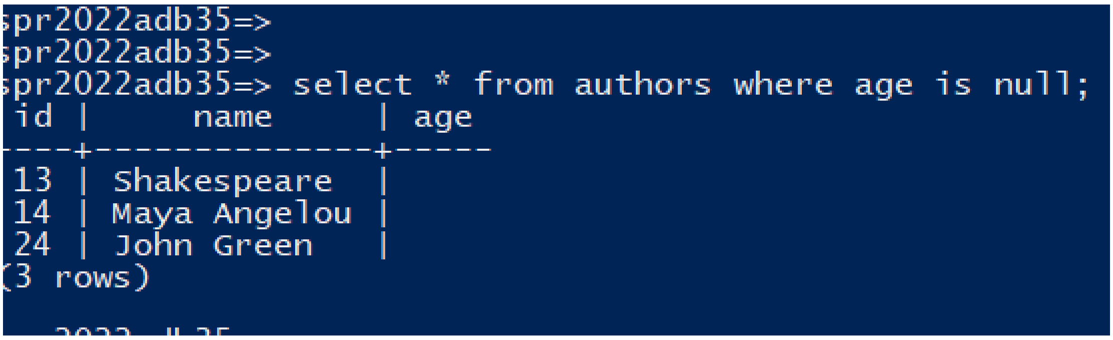


```
spr2022adb35=> update authors set age=NULL where id = 24;
UPDATE 1
spr2022adb35=>
spr2022adb35=>
```

Question-2: - Write a query to find authors where age is null?

Answer: - select * from authors where age is null;

The screenshot for the same is given below: -

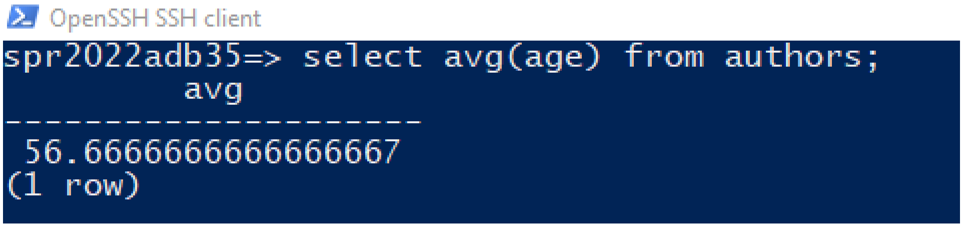


```
spr2022adb35=>
spr2022adb35=>
spr2022adb35=> select * from authors where age is null;
 id |      name      | age
----+----------------+-----
 13 | Shakespeare    |
 14 | Maya Angelou   |
 24 | John Green     |
(3 rows)
```

Question-3: - Write a query to find average age of authors? How do null values affect this average?

Answer: - select avg(age) from authors;

The result for this query is given in the screenshot below: -

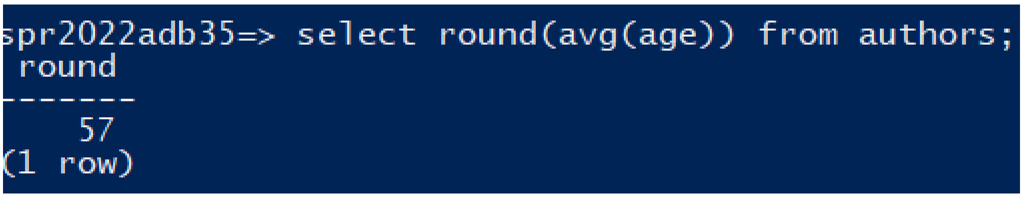


```
spr2022adb35=> select avg(age) from authors;
         avg
---------------------
 56.6666666666666667
(1 row)
```

We can also find the round value of the average age using the following query:

select round(avg(age)) from authors;

The screenshot for the same is given below: -



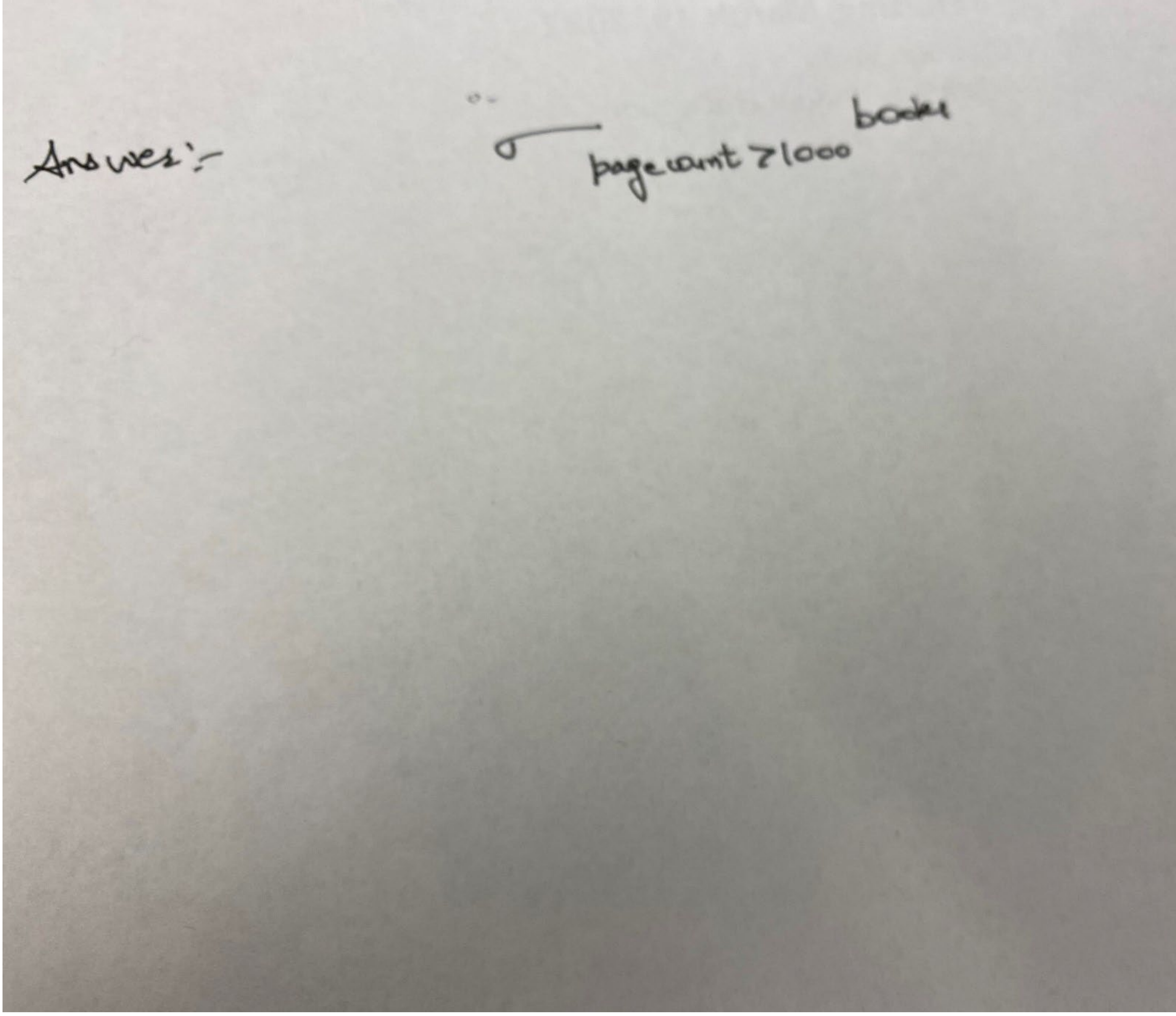
```
spr2022adb35=> select round(avg(age)) from authors;
 round
-------
    57
(1 row)
```

Null values have no effect on the average age of authors. This is because NULL values are not included by the AVG aggregate function when calculating the average.

ACTIVITY-4.1: - RA – Select Activity

Question: - Find all books with pagecount>1000

Answer: -



Answer:-

$\sigma_{pagecount>1000}\ books$

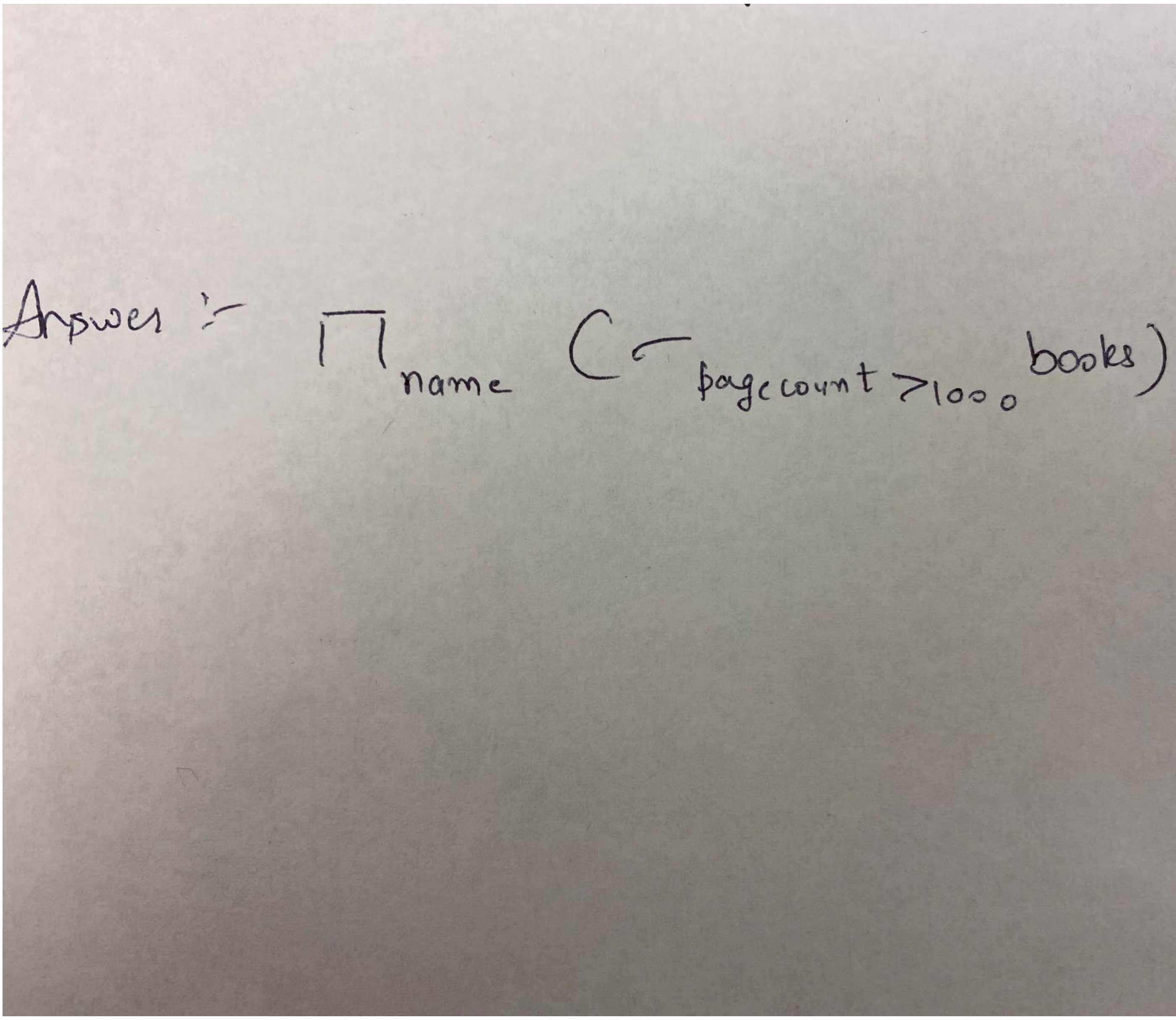
The equivalent SQL command is: -

Select * from books where pagecount>1000;

ACTIVITY-4.2 RA-Project Activity

Question-1: - Find all names of books with pagecount>1000.

Answer: -



Answer :- $\Pi_{name} (\sigma_{pagecount>1000}\ books)$

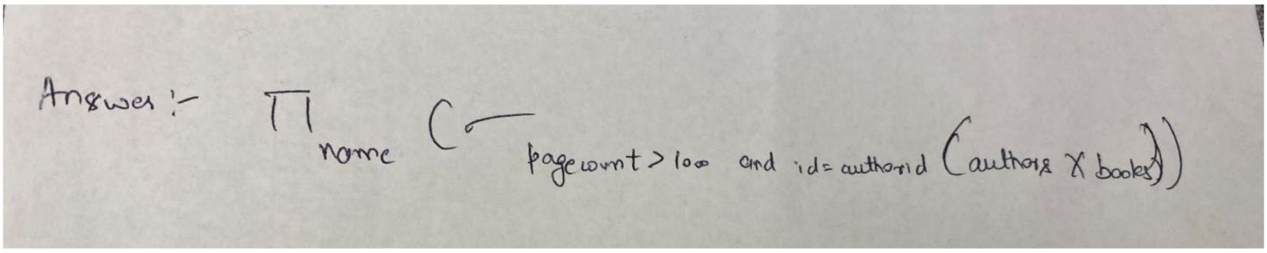
The equivalent SQL statement is: -

Select name from books where pagecount>1000;

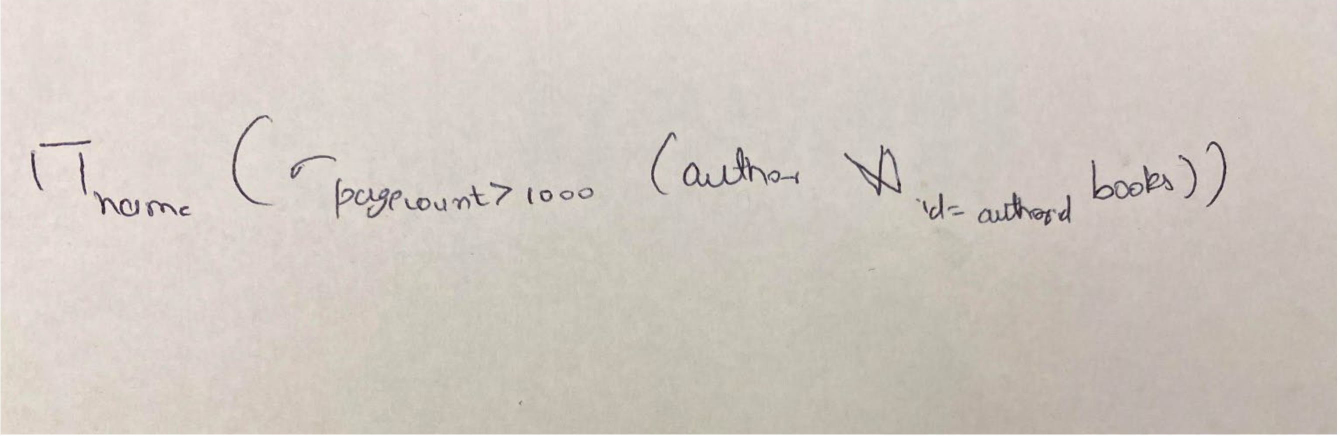
ACTIVITY-4.3-RA-Join Activity

Question: - Find the names of the author with books of pagecount>1000.

Answer: -



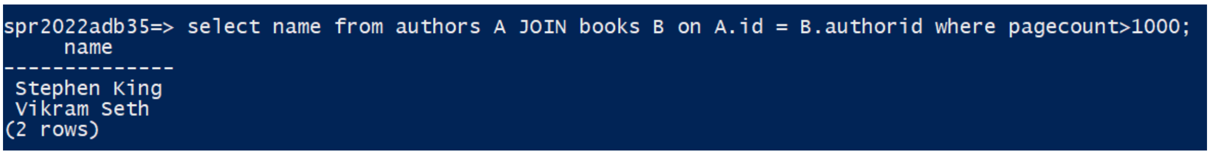
Answer :- $\Pi_{name}(\sigma_{pagecount>1000 \text{ and } id=authorid}(authors \times books))$



$\Pi_{name}(\sigma_{pagecount>1000}(author \bowtie_{id=authorid} books))$

The equivalent SQL query is: -

select name from authors A JOIN books B on A.id = B.authorid where pagecount>1000;



```
spr2022adb35=> select name from authors A JOIN books B on A.id = B.authorid where pagecount>1000;
     name
--------------
 Stephen King
 Vikram Seth
(2 rows)
```