动手实验室：计数、不同、限制

预计所需时间： 35 分钟

在本实验中，您将学习一些与 SELECT 语句一起使用的有用表达式。首先，您将学习 COUNT，它是一个聚合函数，用于检索与查询条件匹配的行数。接下来，您将学习 DISTINCT，它用于从指定结果集中删除重复值并仅返回唯一值。最后，您将学习 LIMIT，它用于限制从表中检索的行数。

本实验室使用的软件

在本实验中，您将使用Datasette，这是一个用于探索和发布数据的开源多功能工具。

本实验使用的数据库

本实验中使用的数据库来自以下数据集来源： PDDL下的旧金山电影地点：公共领域奉献和许可。

目标

完成本实验后，您将能够：

- 检索与查询条件匹配的行数
- 从结果集中删除重复值并返回唯一值
- 限制从表中检索的行数

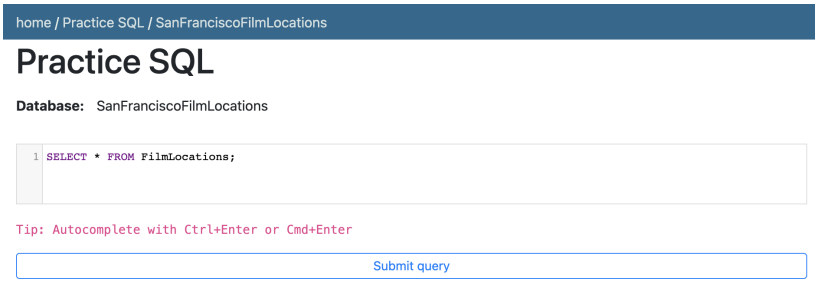
探索数据库

让我们首先使用**Datasette**工具探索**SanFranciscoFilmLocations**数据库：

1. 如果下面列出的第一个语句尚未出现在右侧的数据集文本框中，则通过单击下面代码块右下角的小复制按钮来复制下面的代码，然后使用以下任一方法将其粘贴到数据集工具的文本框中：**Ctrl+V**或右键单击文本框并选择**“粘贴”**。

```
1. 1
1. SELECT * FROM FilmLocations;
```

复制了！



2. 单击**提交查询**。

3. **现在，您可以向下滚动表并浏览FilmLocations**表的所有列和行，以了解该表的整体情况。

Title	ReleaseYear	Locations	FunFacts	ProductionCompany	Distributor	Director	Writer	Actor1	Actor2	Actor3
180	2011	Epic Roasthouse (399 Embarcadero)		SPI Cinemas		Jayendra	Umarji Anuradha, Jayendra, Aarthi Sriram, & Suba	Siddarth	Nithya Menon	Priya Anand
180	2011	Mason & California Streets (Nob Hill)		SPI Cinemas		Jayendra	Umarji Anuradha, Jayendra, Aarthi Sriram, & Suba	Siddarth	Nithya Menon	Priya Anand
180	2011	Justin Herman Plaza		SPI Cinemas		Jayendra	Umarji Anuradha, Jayendra, Aarthi Sriram, & Suba	Siddarth	Nithya Menon	Priya Anand
180	2011	200 block Market Street		SPI Cinemas		Jayendra	Umarji Anuradha, Jayendra, Aarthi Sriram, & Suba	Siddarth	Nithya Menon	Priya Anand
180	2011	City Hall		SPI Cinemas		Jayendra	Umarji Anuradha, Jayendra, Aarthi Sriram, & Suba	Siddarth	Nithya Menon	Priya Anand
180	2011	Polk & Larkin Streets		SPI Cinemas		Jayendra	Umarji Anuradha, Jayendra, Aarthi Sriram, & Suba	Siddarth	Nithya Menon	Priya Anand
180	2011	Randall Museum		SPI Cinemas		Jayendra	Umarji Anuradha, Jayendra, Aarthi Sriram, & Suba	Siddarth	Nithya Menon	Priya Anand
180	2011	555 Market St.		SPI Cinemas		Jayendra	Umarji	Siddarth	Nithya	Priya

Support

4. **这些是FilmLocations表中的列属性描述：**

```
1. 1
2. 2
3. 3
4. 4
5. 5
6. 6
7. 7
8. 8
9. 9
10. 10
11. 11
12. 12
13. 13
1. FilmLocations(
2.     Title:              titles of the films,
3.     ReleaseYear:        time of public release of the films,
4.     Locations:          locations of San Francisco where the films were shot,
5.     FunFacts:           funny facts about the filming locations,
6.     ProductionCompany:  companies who produced the films,
7.     Distributor:        companies who distributed the films,
8.     Director:           people who directed the films,
9.     Writer:             people who wrote the films,
10.    Actor1:             person 1 who acted in the films,
11.    Actor2:             person 2 who acted in the films,
12.    Actor3:             person 3 who acted in the films
13. )
```

复制了！

练习 1：计数

在本练习中，您将首先了解在查询中使用 COUNT 的一些示例，然后使用它解决一些练习问题。

任务A

COUNT 的示例练习

让我们看一下 COUNT 相关查询的一些示例：

1. 在此示例中，假设我们要计算“FilmLocations”表的记录数或行数。
    1. 问题：

        从“FilmLocations”表中检索行数。

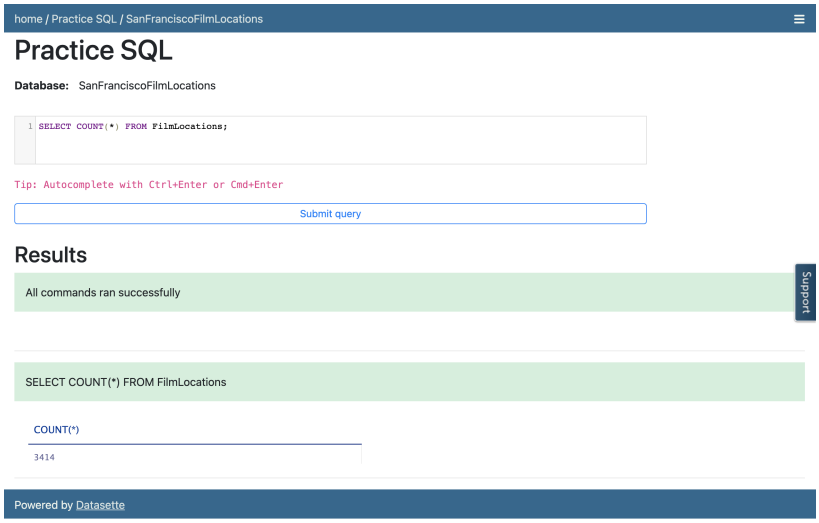
2. 解决方案：

```
1. 1
1. SELECT COUNT(*) FROM FilmLocations;
```

复制了！

3. 单击下面代码块右下角的小复制按钮复制上面的解决方案代码，并将其粘贴到数据集工具的文本框中。然后单击**提交查询**。

4. 您的输出结果集应如下图所示：



2. 在此示例中，现在我们要计算电影的拍摄地点数量。但我们还想限制输出结果集，以便我们只检索某个作家所写的电影的地点数量。

1. 问题：

   检索詹姆斯·卡梅隆编剧的电影的地点数量。

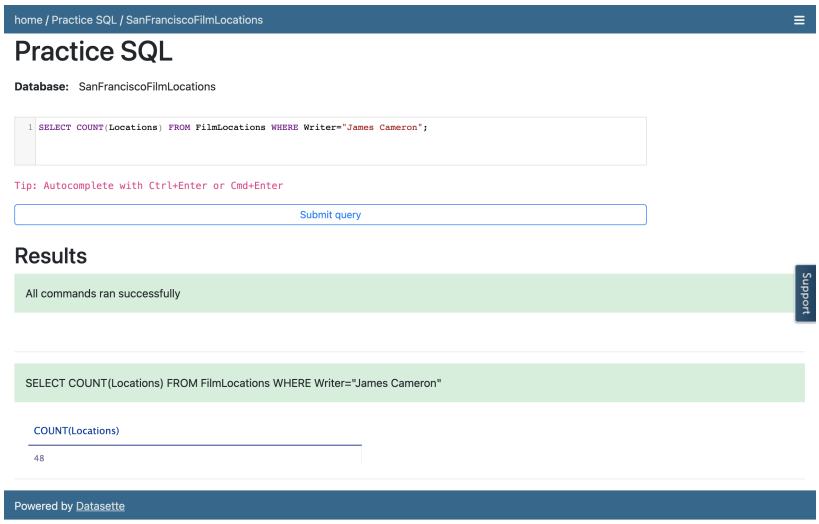
2. 解决方案：

```
1. 1
1. SELECT COUNT(Locations) FROM FilmLocations WHERE Writer="James Cameron";
```

复制了！

3. 单击下面代码块右下角的小复制按钮复制上面的解决方案代码，并将其粘贴到数据集工具的文本框中。然后单击**提交查询**。

4. 您的输出结果集应如下图所示：



任务B

在 COUNT 上进行练习

现在，让我们练习创建和运行一些 COUNT 相关查询。

1. 问题：

   检索伍迪·艾伦执导的电影的拍摄地点数量。

   - ▶ 暗示
   - ▶ 解决方案
   - ▶ 输出

2. 问题：

   检索在俄罗斯山拍摄的电影数量。

   - ▶ 暗示
   - ▶ 解决方案
   - ▶ 输出

3. 问题：

   从“FilmLocations”表中检索发行年份早于 1950 年的行数。

   - ▶ 暗示
   - ▶ 解决方案
   - ▶ 输出

练习 2：独特

在本练习中，您将首先了解在查询中使用 DISTINCT 的一些示例，然后使用它解决一些练习问题。

任务A

DISTINCT 练习示例

让我们看一下 DISTINCT 相关查询的一些示例：

1. In this example, we want to retrieve the title of all films in the table in such a way that duplicates will be discarded in the output resultset.

    1. Problem:

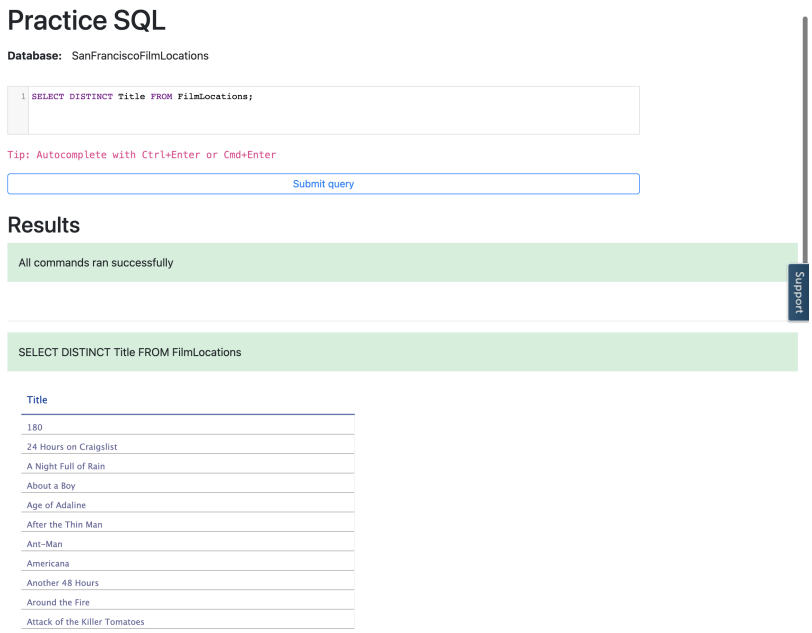
        Retrieve the name of all films without any repeated titles.

    2. Solution:

    ```
    1. 1
    1. SELECT DISTINCT Title FROM FilmLocations;
    ```
    Copied!

    3. Copy the solution code above by clicking on the little copy button on the bottom right of the codeblock below and paste it to the textbox of the Datasette tool. Then click **Submit query**.

    4. Your output resultset should look like the image below:



2. In this example, we want to retrieve the count of release years of the films produced by a specific company in such a way that duplicate release years of those films will be discarded in the count.

    1. Problem:

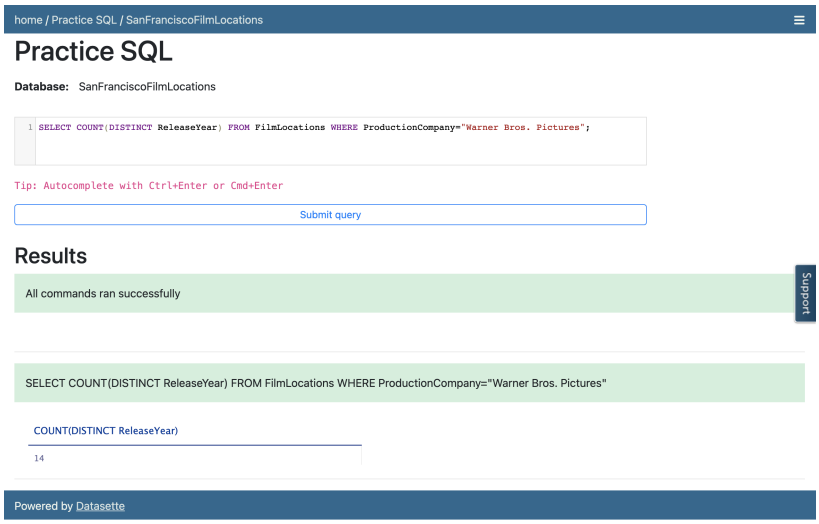
        Retrieve the number of release years of the films distinctly, produced by Warner Bros. Pictures.

    2. Solution:

    ```
    1. 1
    1. SELECT COUNT(DISTINCT ReleaseYear) FROM FilmLocations WHERE ProductionCompany="Warner Bros. Pictures";
    ```
    Copied!

    3. Copy the solution code above by clicking on the little copy button on the bottom right of the codeblock below and paste it to the textbox of the Datasette tool. Then click **Submit query**.

    4. Your output resultset should look like the image below:



Task B

Practice exercises on DISTINCT

Now, let us practice creating and running some DISTINCT related queries.

1. Problem:

   Retrieve the name of all unique films released in the 21st century and onwards, along with their release years.

   ▶ Hint
   ▶ Solution
   ▶ Output

2. Problem:

   Retrieve the names of all the directors and their distinct films shot at City Hall.

   ▶ Hint
   ▶ Solution
   ▶ Output

3. Problem:

   Retrieve the number of distributors distinctly who distributed films acted by Clint Eastwood as 1st actor.

   ▶ Hint
   ▶ Solution
   ▶ Output

Exercise 3: LIMIT

In this exercise, you will first go through some examples of using LIMIT in queries and then solve some exercise by using it.

Task A: Example exercises of LIMIT

Let us go through some examples of LIMIT related queries:

1. In this example, let us retrieve a specific number of rows from the top of the table in such a way that rows other than those are not in the output resultset.

    1. Problem:

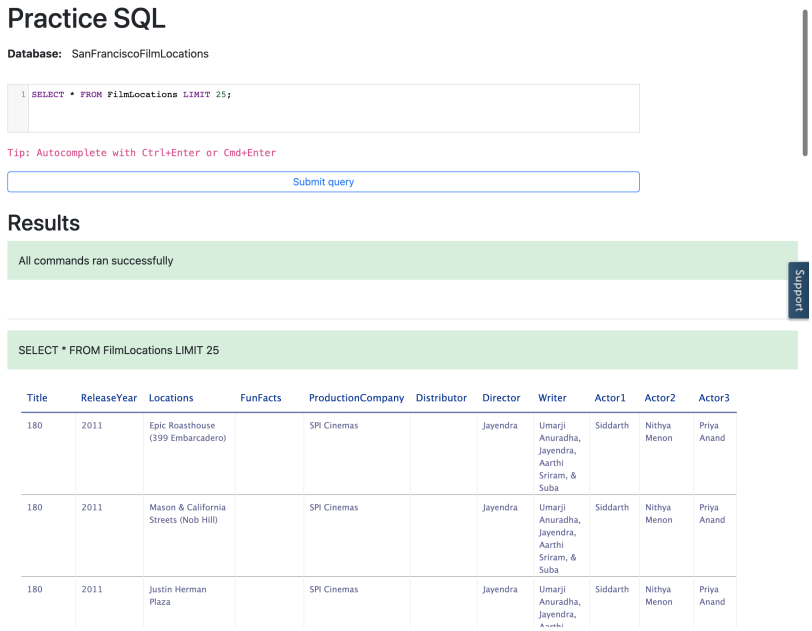
        Retrieve the first 25 rows from the "FilmLocations" table.

    2. Solution:

```
1. 1
1. SELECT * FROM FilmLocations LIMIT 25;
```
Copied!

    3. Copy the solution code above by clicking on the little copy button on the bottom right of the codeblock below and paste it to the textbox of the Datasette tool. Then click **Submit query**.

    4. Your output resultset should look like the image below:



2. In this example, let us take the first example to a more advanced level. Now we want to retrieve a specific number of rows from the table, but thid time, not from the top of the table. This time we want to retrieve a specific number of rows starting from a specific row in the table.

    1. Problem:

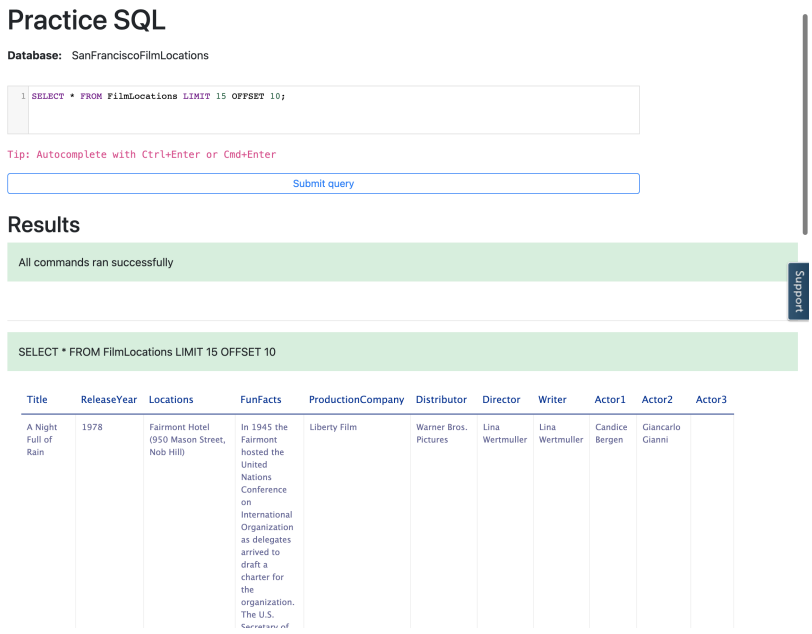
        Retrieve the first 15 rows from the "FilmLocations" table starting from row 11.

    2. Solution:

```
1. 1
1. SELECT * FROM FilmLocations LIMIT 15 OFFSET 10;
```
Copied!

    3. Copy the solution code above by clicking on the little copy button on the bottom right of the codeblock below and paste it to the textbox of the Datasette tool. Then click **Submit query**.

    4. Your output resultset should look like the image below:



Task B: Practice exercises on LIMIT

Now, let us practice creating and running some LIMIT related queries.

1. Problem:

   Retrieve the name of first 50 films distinctly.

   - ▶ Hint
   - ▶ Solution
   - ▶ Output

2. Problem:

   Retrieve first 10 film names distinctly released in 2015.

   - ▶ Hint
   - ▶ Solution
   - ▶ Output

3. Problem:

   Retrieve the next 3 film names distinctly after first 5 films released in 2015.

   - ▶ Hint
   - ▶ Solution
   - ▶ Output

Congratulations! You have completed this Lab.

Author(s)

- Sandip Saha Joy

Other Contributor(s)

Changelog

Date	Version	Changed by	Change Description
2023-05-11	1.6	Eric Hao & Vladislav Boyko	Updated Page Frames
2023-05-10	1.5	Eric Hao & Vladislav Boyko	Updated Page Frames
2023-05-10	1.4	Eric Hao & Vladislav Boyko	Updated Page Frames
2023-05-05	1.3	Benny Li	Reformatted and republished
2022-07-27	1.2	Lakshmi Holla	Updated html tag
2020-12-23	1.1	Steve Ryan	ID Review
2020-11-24	1.0	Sandip Saha Joy	Initial version created