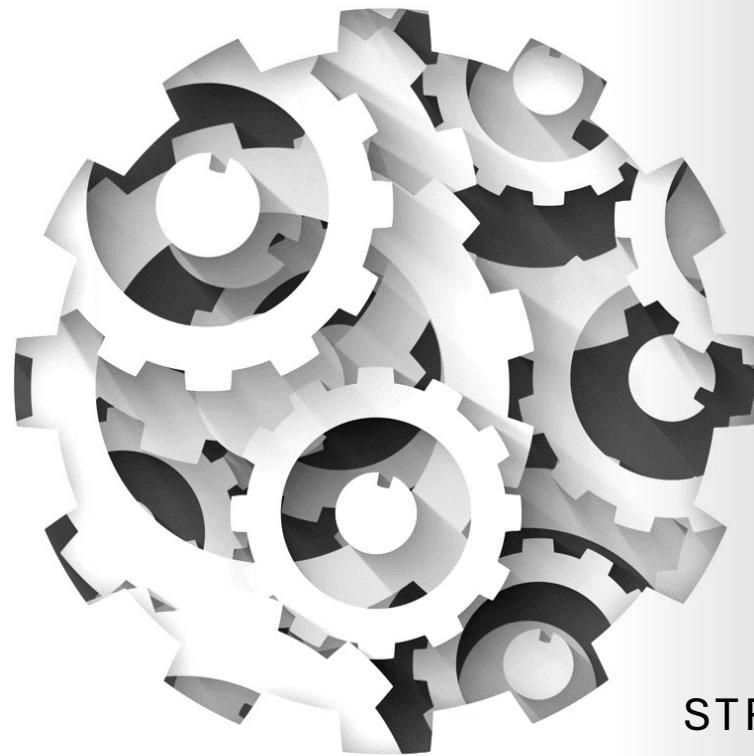


PART II



<SQL>
STRUCTURED QUERY
LANGUAGE



After this session you'll be able to:

Use SQLite statements to query data from your database

- Use SELECT, FROM and WHERE statements
- Use wildcards to retrieve data
- Use operators to modify data query
- Use sort functions
- Use aggregated functions
- Get insights from a relational database

Requirements from part I

| state_crime | |
|----------------------------|---------|
| ABC State | VARCHAR |
| 123 Crime_Year | INTEGER |
| 123 Population | INTEGER |
| 123 Rates_Property_Theft | REAL |
| 123 Rates_Violent_Robbery | REAL |
| 123 Totals_Property_Theft | INTEGER |
| 123 Totals_Violent_Robbery | INTEGER |

| state_workforce | |
|------------------------------------------|---------|
| ABC State | VARCHAR |
| 123 Population_Percent_Change | REAL |
| 123 Population_2014 | INTEGER |
| 123 Population_2010 | INTEGER |
| 123 Education_High_School_or_Higher | REAL |
| 123 Education_Bachelors_Degree_or_Higher | REAL |
| 123 Sales_Retail_Sales | INTEGER |
| 123 Mean_Travel_Time_to_Work | REAL |
| 123 Income_Median_Household_Income | INTEGER |
| 123 Income_Per_Capita_Income | INTEGER |
| 123 Income_Persons_Below_Poverty_Level | REAL |

| state_people | |
|--------------------------------|---------|
| ABC State | TEXT |
| 123 employment_firms_total | INT |
| 123 age_percent_under_18_years | DECIMAL |
| 123 age_percent_65_and_older | DECIMAL |

| state_computer_data | |
|------------------------------|---------|
| ABC State | TEXT |
| 123 Persons_per_household | DECIMAL |
| 123 Households_with_computer | DECIMAL |
| 123 Households_with_internet | DECIMAL |

Writing queries

How to retrieve and filter data



- Statements
- SELECT (attributes)**
- FROM (table)**
- WHERE (filter clause)**

Filtering Operators

language expressions

| | |
|---------|---------------------------|
| = | • Equal |
| >, >= | • Greater than |
| <, <= | • Less than |
| != | • Different than |
| NOT | • Negates a Boolean value |
| BETWEEN | • The value is in a range |
| LIKE | • Search of a pattern |

Filtering Operators

language expressions

- AND** • AND operator
- OR** • OR Operator

Wildcards

- Wildcard for all elements
<used in select>
- Wildcard for 1 or more
characters <used in
LIKE>
- Wildcard 1 character
<Used in LIKE>

Aggregate functions

- COUNT()** • Counts the number of records
- MAX()** • Get maximum value
- MIN()** • Get minimum value
- SUM()** • Sums the field values
- AVG()** • Average of column values

Sort data

ORDER BY

ASC

- Sort ascending

DESC

- Sort Descending

QUERIES

HANDS-ON

Find out how many people take longer than 20 mins on average to get to work (hint: query the state_workforce table)

Sort the results to find out what states have the longest average time to get to work

Modify your query to show only the records of New York, New Jersey, New Hampshire and New Mexico

SQL JOINS

Orders

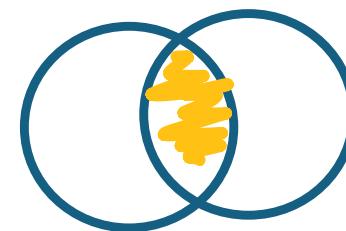
| order_no | Student_id | Product_sku |
|----------|------------|-------------|
| 1001 | 2 | 7ABC |
| 1002 | 3 | 58KJ6 |
| 1003 | 6 | SP500 |

Student

| StudentID | StudentName | Status |
|-----------|-------------|----------|
| 1 | James | Active |
| 2 | Jackie | Active |
| 3 | Robert | Inactive |

INNER JOIN

| order_no | Student_id | Product_sku | StudentID | StudentName | Status |
|----------|------------|-------------|-----------|-------------|----------|
| 1001 | 2 | 7ABC | 2 | Jackie | Active |
| 1002 | 3 | 58KJ6 | 3 | Robert | Inactive |



SQL JOINS

Orders

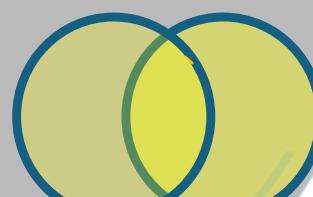
| order_no | Student_id | Product_sku |
|----------|------------|-------------|
| 1001 | 2 | 7ABC |
| 1002 | 3 | 58KJ6 |
| 1003 | 6 | SP500 |

Student

| StudentID | StudentName | Status |
|-----------|-------------|----------|
| 1 | James | Active |
| 2 | Jackie | Active |
| 3 | Robert | Inactive |

FULL JOIN

| order_no | Student_id | Product_sku | StudentID | StudentName | Status |
|----------|------------|-------------|-----------|-------------|----------|
| 1001 | 2 | 7ABC | 2 | Jackie | Active |
| 1002 | 3 | 58KJ6 | 3 | Robert | Inactive |
| 1003 | 6 | SP500 | NULL | NULL | NULL |
| NULL | NULL | NULL | 1 | James | Active |



SQL JOINS

Orders

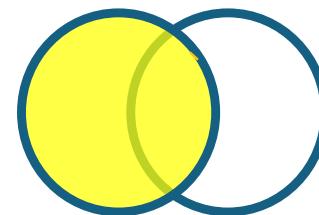
| order_no | Student_id | Product_sku |
|----------|------------|-------------|
| 1001 | 2 | 7ABC |
| 1002 | 3 | 58KJ6 |
| 1003 | 6 | SP500 |

Student

| StudentID | StudentName | Status |
|-----------|-------------|----------|
| 1 | James | Active |
| 2 | Jackie | Active |
| 3 | Robert | Inactive |

LEFT JOIN

| order_no | Student_id | Product_sku | StudentID | StudentName | Status |
|----------|------------|-------------|-----------|-------------|----------|
| 1001 | 2 | 7ABC | 2 | Jackie | Active |
| 1002 | 3 | 58KJ6 | 3 | Robert | Inactive |
| 1003 | 6 | SP500 | NULL | NULL | NULL |



SQL JOINS

Orders

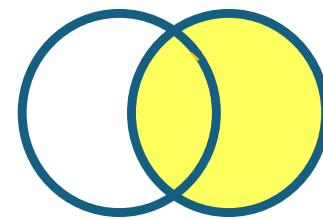
| order_no | Student_id | Product_sku |
|----------|------------|-------------|
| 1001 | 2 | 7ABC |
| 1002 | 3 | 58KJ6 |
| 1003 | 6 | SP500 |

Student

| StudentID | StudentName | Status |
|-----------|-------------|----------|
| 1 | James | Active |
| 2 | Jackie | Active |
| 3 | Robert | Inactive |

RIGHT JOIN

| order_no | Student_id | Product_sku | StudentID | StudentName | Status |
|----------|------------|-------------|-----------|-------------|----------|
| NULL | NULL | NULL | 1 | James | Active |
| 1001 | 2 | 7ABC | 2 | Jackie | Active |
| 1002 | 3 | 58KJ6 | 3 | Robert | Inactive |

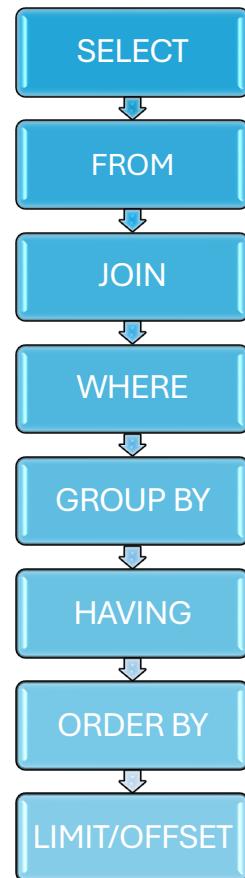


SUBQUERIES

HANDS-ON

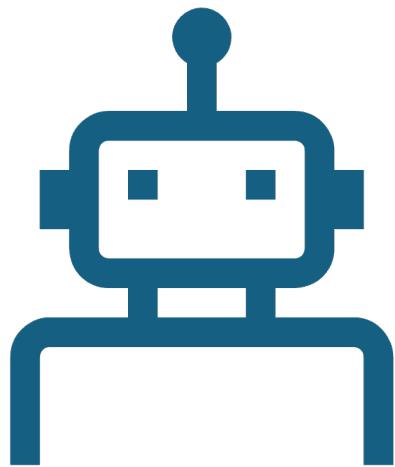
Find the maximum percentage of people with education of high school or higher from states where the rate of property theft is above 2,500

Syntax Order



Operation Order

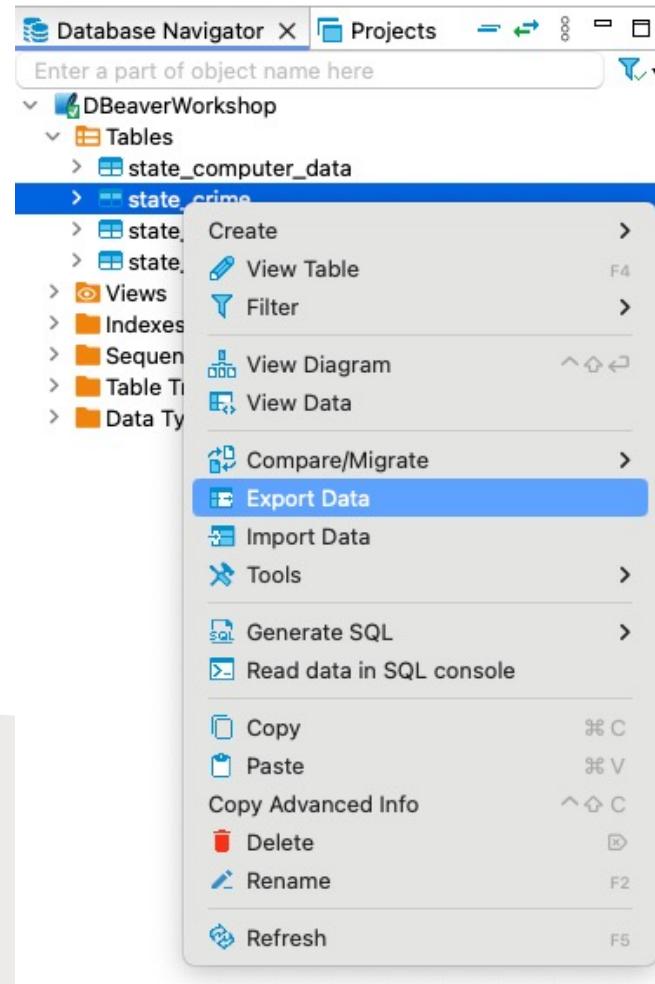




HANDS-ON

1. Create an inner join using aliases with tables state_crime and state_people. Make sure to observe attributes from:
 - State
 - Rates of Property theft
 - Crime Year
 - Employment_Firms_Total
2. Add a new row to state people for state 'US' with totals of employment firms and averages of age percentages
3. Add a new 'division' column as a division of 100 * rates of property theft by employment firms total
4. Create other types of joins and see the differences
5. Examine what happens with aggregated functions and null values

EXPORT YOUR DATA FILE



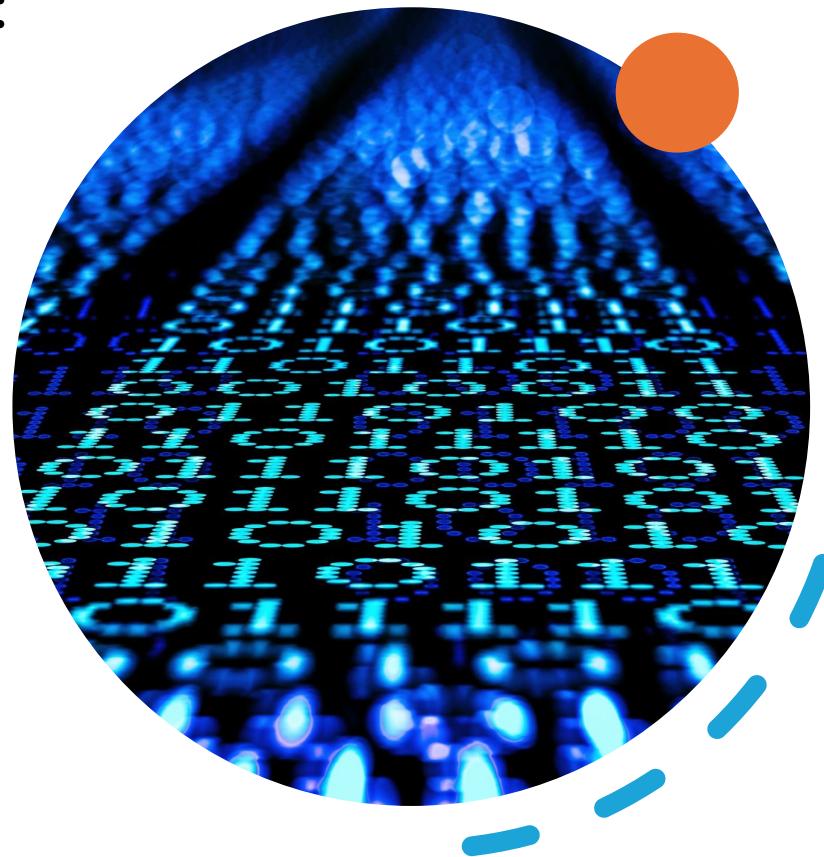
Extra challenge

Select top 10 states with highest rates of property theft and display sorted by population 2014:
median household income,
age percent under 18,
average persons per household

FURTHER EXPLORATION: SUBQUERY

Executes multiple queries, perform innermost query first

```
SELECT sc.State , sc.Population, sc.Totals_Violent_Robbery  
FROM state_crime sc  
WHERE sc.State IN (  
    SELECT scd.State  
    FROM state_computer_data scd  
    WHERE scd.Households_with_computer >= 93  
);
```



THANK YOU

Resources

- Dbeaver Wiki - <https://github.com/dbeaver/dbeaver/wiki>
- SQLite library - <https://www.sqlite.org/>
- W3schools - <https://www.w3schools.com/sql/>
- Codecademy - <https://www.codecademy.com/learn/learn-sql>
- Geeks for Geeks - <https://www.geeksforgeeks.org/>

Fun way to learn

- SQL Murder Mystery - <https://mystery.knightlab.com/>

Daniel Perez
QCL Graduate Fellow
danperez@cmc.edu