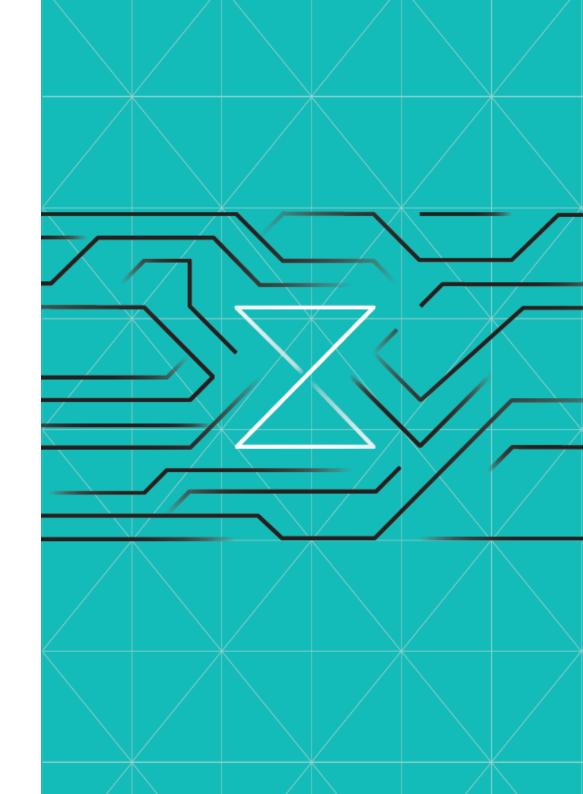
DB2

SQL and Relational Databases

- MANAGING DATA AS SETS RELATIONAL DATABASES
- 1 INSTALL DB2 VSCODE EXTENSION
- 2 INSTALL/CONFIGURE JAVA
- 3 SET UP A LITTLE MORE INFORMATION
- 4 MAKE A STATEMENT
- 5 RESULT!
- 6 GO EVEN DEEPER
- 7 WHERE YOU WANT TO BE
- 8 A SPECIAL SET OF RESULTS
- 9 MAKE IT COUNT (FOR REAL)
- 10 ALL ABOUT JOINS
- 11 JOINS HAPPEN
- 12 PURE FILTERED RESULTS
- 13 PUSH IT TO THE MAX
- 14 PUTTING IT ALL TOGETHER
- 15 LAST CHECK AND WRAP-UP



MANAGING DATA AS SETS - RELATIONAL DATABASES

The Challenge

Relational databases are an important part of just about all software today.

By organizing data into tables and then linking these tables based on common data, you can retrieve and organize data in order to better understand relationships in the data.

In this challenge, you will be using SQL to communicate with IBM Db2 relational databases, hosted on the IBM zOS system.

Before You Begin

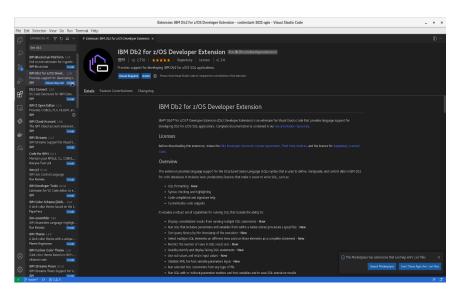
Not much, just make sure everything's right with VSCode and you should be good to go.

Investment

Steps	Duration	
15	60 minutes	

1 INSTALL DB2 VSCODE EXTENSION

In VSCode, on the "Extensions" launcher in the left-side menu, and search for IBM Db2.



Next, click on the "Install" button for "IBM Db2 for z/OS Developer Extension." This is an extension for VSCode that gives you the ability to work interactively with Db2 databases.

The install should only take a few seconds on most broadband internet connections.

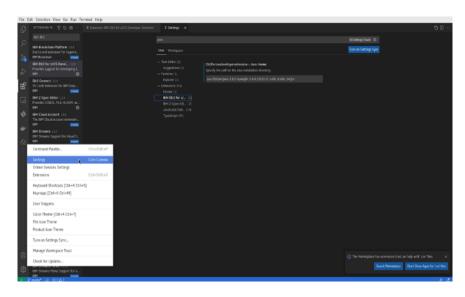
2 INSTALL/CONFIGURE JAVA

For the Db2 extension to be able to connect to the database server, you need to have a supported Java JDK installed.

Not just a Java Runtime (JRE), but the full Java Developer Kit (JDK).

You have options: you can use <a>Oracle Java JDK 8, or version 8 or 11 of the <a>OpenJDK.

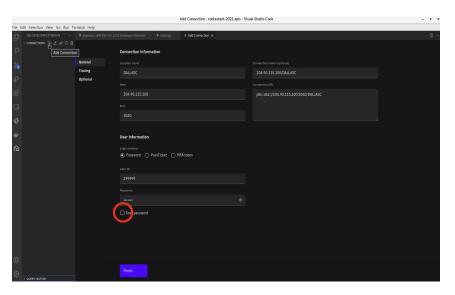
Once installed, navigate into the VSCode settings and search for "java" in the search bar. There should be an entry for Db2 under Extensions that you can click on and provide the path for your installed Java runtime libraries.



For example, on MacOS it is usually /Library/Java/JavaVirtualMachine/jdk1.8.0_181.jdk/Contents/Home
On Windows, it will probably be C:Files.0_181

3 SET UP A LITTLE MORE INFORMATION

In the VSCode left-side menu, navigate to "Db2 Developer Extension" (you may need to click on the 3 dots ". . ." to see it) and click the plus sign to add a new connection profile.

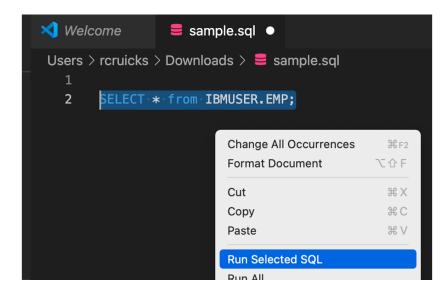


Enter the following values for the profile configuration: + DALLASC for location name + 204.90.115.200 for host + 5040 for port + and your z/OS login for User ID and password

Unless you want to type your password every time you start doing SQL queries, check the "Save password" box.

Then click Finish at the bottom and you should be setup and ready to use Db2!

4 MAKE A STATEMENT



Create a new folder on your laptop named "Db2 Challenges", open this folder in VSCode, and create a new file named sample.sql as shown above.

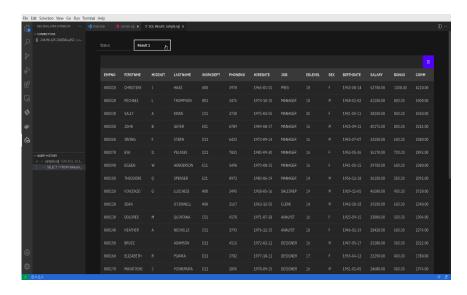
Now write a simple SQL statement. Type in the statement from the screenshot above:

SELECT * from IBMUSER.EMP;

Highlight the statement and right-click to see options to run.

Choose "Run Selected SQL".

5 RESULT!



The "SQL Results" tab will pop up after the statement is run and this is where you can view the set of records as a result of your SQL statement.

The "Status" section will display the statement you ran, and Return Code 0 indicating successful execution. The "Result 1" tab contains the resulting table.

As you can see, there are a lot of columns in the table. When you put the asterisk (*) after SELECT in your SQL statement, this tells the database to retrieve all the columns from the table ibmuser.emp, and include them in the results.

Note the "QUERY HISTORY" section in the left-side navigation area.

6 GO EVEN DEEPER

You may have guessed that if you don't want to work with all the columns, you can specify which columns you do want.

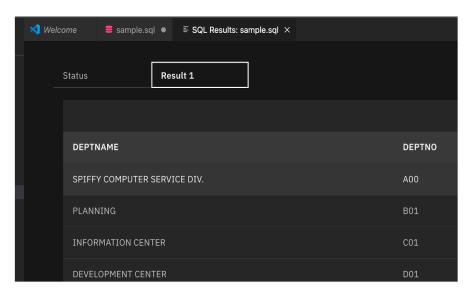
For example, you could display just the employees' job by typing select JOB from ibmuser.emp

You can also display multiple columns. Type in select LASTNAME, EMPNO, PHONENO from ibmuser.emp and see the result.

Note that the columns shows up in the exact order you specified.

Notice the table you've been using is called "ibmuser.emp" and contains some totally fictional employee data.

Well, this is another table you can use called "ibmuser.dept" that contains data about the different departments within this fictional company.





Look at the contents of the "ibmuser.dept" table and then write an SQL statement that will recreate the results shown in the screenshot above.

"So what exactly is Db2?"

IBM Db2 is what is known as a Relational Database Management System (RDBMS).

It uses the the concepts of relational databases to store many different types of data and then optimizes retrieval of that data.

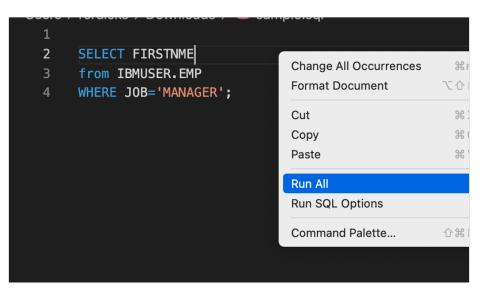
There are many different types of RDBMS, Db2 is just IBM's version - it has been around since the beginning of relational databases. Actually the concept of a relational database was developed at IBM by computer scientist Edgar Frank "Ted" Codd.

SQL (Structured Query Language) is the language we use to talk to Db2, and is what pretty much all modern RDBMSs use.

7 WHERE YOU WANT TO BE

Not only can you specify the columns you want to select, but you can specify the rows as well.

This is done by specifying the value of one (or more) of the columns so that only rows that meet the column values are displayed.



Can you guess what the above query will return?

Run the multi-line statement in the screenshot above using the WHERE clause and find out.

```
SELECT FIRSTNME
from IBMUSER.EMP
WHERE JOB='MANAGER';
```

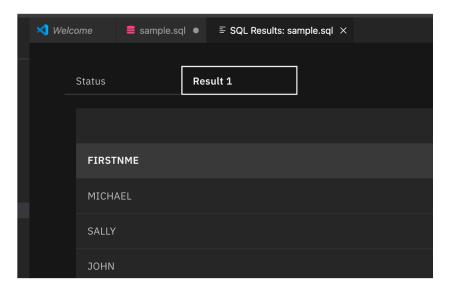
What do you think would happen for JOB='manager'?



82|230222-15

8 A SPECIAL SET OF RESULTS

Look at the result from the query, and see what you have.



Notice that only those employees that are managers show up. Is that what you expected to happen?

The WHERE clause is a filter that extracts only records that fulfill the specified condition.

Not only can you use strings, but you can use numbers to specify and other operators like ">" to filter by a greater-than range or "<=" to filter by a less-than-or-equal-to range.

9 MAKE IT COUNT (FOR REAL)

There are also some functions that allow you to count, average, or sum the rows - these are called "aggregate" functions.

```
sample.sql
1 SELECT COUNT(*) from ibmuser.emp
```

The syntax is pretty simple; you just write COUNT(column_name), AVG(column_name), or SUM(column_name) after SELECT. There's an example in the screenshot above.

Each of these functions takes the result of the query and performs the function on each row of the resulting column.

So if you add a WHERE to the query, it will only aggregate the amount of rows that meet the condition.

Now you are getting closer to turning normal questions someone might ask "How many managers are there?" into actual SQL queries.

"Can I make the display name of aggregate functions nicer?"

Yup! All you have to do is add whatever you want displayed as the name after the aggregate function.

Like this:

SELECT COUNT(*) as NUM_EMPLOYEES from IBMUSER.EMP

Also something to note - these aggregate functions work on numerical and non-numerical columns, and they count all the rows in that column that have an entry.

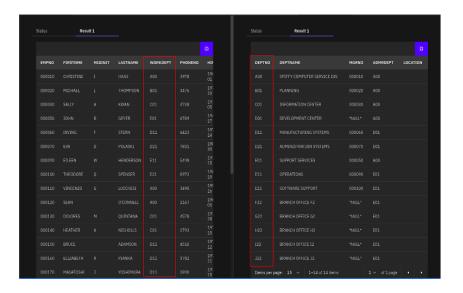
When the asterisk is used in any of the functions, it counts the total number of rows.

10 ALL ABOUT JOINS

A "join" is the process of combining data from two or more tables based on some common domain of information.

Rows from one table are paired with rows from another table when information in the corresponding rows match based on the joining criteria.

In order to do this, a WHERE clause comes in handy again.



You simply have to relate two common columns from two different tables by setting them equal to each other in the WHERE clause. The example above is from the DEPT and EMP tables; these tables can be joined on their common columns - WORKDEPT and DEPTNO.

Note that it is the data values in the columns that are common - not necessarily the column names.

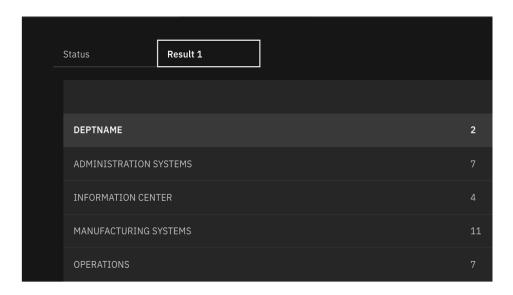
11 JOINS HAPPEN

Run the example above.

Notice how this query joins IBMUSER.DEPT and IBMUSER.EMP on the columns DEPTNO and WORKDEPT. Also notice how there is a variable provided in the FROM clause that represents each table (**D** represents ibmuser.dept and **E** is for ibmuser.emp) - these are "correlation identifiers".

There are a couple other syntax things to look at in detail after you run the query.

12 PURE FILTERED RESULTS



When you look at the result, it successfully counts the number of employees from each department.

In order to do this, the query had to use the GROUP BY keyword.

This clause groups rows that have the same values into summary rows. This is commonly used for aggregate functions (COUNT, SUM, AVG, MAX, MIN), as in the earlier example.

GROUP BY can be a little confusing, so experiment with the query to see what works, what does not, and what different combinations produce.

The last piece of the puzzle is the **ORDER BY** keyword; this simply displays the information in alphabetical order by the column specified.

82|230222-1

"Can I join more than two tables?"

You can! This is called a "multi-table join".

It is very similar in structure to the two-table join but you add an AND clause after WHERE that specifies a join to another table.

It looks like this:

```
SELECT company, stock_num, menu_code
FROM customer c, orders o, items i
WHERE c.customer_num = o.customer_num
AND o.order num = i.order num;
```

13 PUSH IT TO THE MAX

One last, quick addition to your growing knowledge of SQL will be the MIN and MAX functions.

Just as they sound, these functions return the smallest value and the greatest value respectively from the specified columns. They are used in the same way COUNT, SUM, and AVG are used.

```
sample.sql

SELECT MAX(SALARY) SAL

from ibmuser.emp

3
```

You can also name the output of these functions anything you want by adding a name immediately after the function clause as shown in the screenshot above.

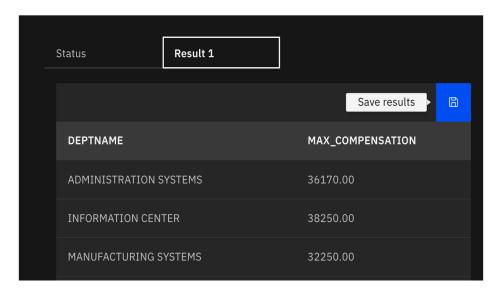
Again, play around with the functions to get some more familiarity before you move onto the challenge!

Experimenting on your own to observe the results of different combinations of syntax is one of the best ways you can learn!

14 PUTTING IT ALL TOGETHER

Your challenge is to query the maximum compensation (**all the ways** of paying someone) in each department sort by department name.

Display the results as two columns - **DEPTNAME** and **MAX_COMPENSATION**.



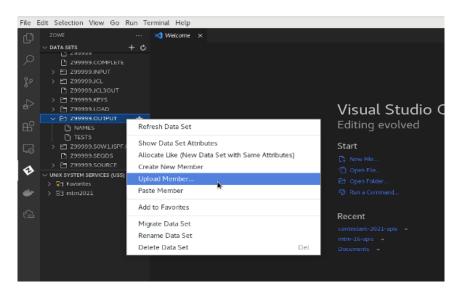
Make sure you're thinking about all of the SQL syntax you have learned - you will need it!

Make sure you know what all of the columns in each table are; perhaps write a list of the columns in each table. When you have the output you want, download it as a .csv file by clicking the File icon in the results tab. This will download it to your laptop/workstation.

Hint: there is more than one column that counts as "compensation"!

15 LAST CHECK AND WRAP-UP

It's now time to validate the challenge, so after you've downloaded the .csv file, right-click on your Zxxxxx.OUTPUT data set and click on "Upload Member".



Then upload your .csv file. After uploading, rename the file as **DB2OUT** so the checker job can find it, and you should be good to go!

Now it's time to run the check to see if your query was correct. Submit the **CHKSQL** member from **ZXP.PUBLIC.JCL** and look for a completion code of 0000!

Nice job - let's recap	Next up
You just learned a ton of new information about relational databases and SQL syntax. Keep in mind, Db2 is just one relational database system - there are many more, but SQL stays the same! So you now know how to query all kinds of information from tables, as well as how to join tables together so you can relate information across different tables. Pretty powerful stuff!	Check out the other available Advanced challenges