

# Hands on Lab: Joins

Estimated time needed: **30** minutes

## Objectives

After completing this lab, you will be able to:

1. Determine the correct type of join to use for a given problem.
2. Write and execute joins to query data from multiple tables.

## Scenario

In this hands-on lab, you will work with three datasets that are available on the City of Chicago's Data Portal:

- Socioeconomic indicators in Chicago
- Chicago public schools
- Chicago crime data

You must download each dataset, create a table for each one, and load the appropriate dataset through the Db2 console. You should not reuse similar tables with other names from other exercises or labs, as they may not create the correct results.

### Important note:

If you have **not** yet downloaded the three datasets from the City of Chicago's Data Portal, created the required tables, and loaded the data, please follow the instructions in this section.

## City of Chicago Datasets

### 1. Socioeconomic indicators in Chicago

This dataset contains a selection of six socioeconomic indicators of public health significance and a “hardship index,” for each Chicago community area, for the years 2008 – 2012. A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:

<https://data.cityofchicago.org/Health-Human-Services/Census-Data-Selected-socioeconomic-indicators-in-C/kn9c-c2s2>

## 2. Chicago public schools

This dataset shows all school level performance data used to create CPS School Report Cards for the 2011-2012 school year. A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:

<https://data.cityofchicago.org/Education/Chicago-Public-Schools-Progress-Report-Cards-2011-/9xs2-f89t>

## 3. Chicago crime data

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days. A detailed description of this dataset and the original dataset can be obtained from the Chicago Data Portal at:

<https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-present/ijzp-q8t2>

## Store the datasets in database tables





The lab requires you to have these three tables populated with a subset of the whole datasets. Download the ‘ChicagoCensusData.csv’, ‘ChicagoPublicSchools.csv’, and ‘ChicagoCrimeData.csv’ datasets below and load the data into your Db2 On Cloud database.

[Chicago Census Data](#)


[Chicago Public Schools](#)


[Chicago Crime Data](#)


You need to create a new table for each dataset. As you load each dataset, click on “(+) New Table”, specify the name of the table you want to create, and then click “Next”.


 IBM Db2 on Cloud  Storage: 29% [Cookie Preferences](#) [Discover](#)  

LOAD DATA

 Source


 Target


 Define

 Finalize

You are loading the file ChicagoCensusData.csv

Select a load target


Schema  New Schema


Find a schema 

AUDIT

DB2INST1

ERRORSCHEMA *Sample*

Table **1**  New Table

Find a table in MXC01472 

CHICAGO\_CRIME\_DATA


CHICAGO\_PUBLIC\_SCHOOLS

Create a new Table

CENSUS\_DATA **2**

Create **3**

Back

**4**  Next

Name the new tables as follows:

1. CENSUS\_DATA
2. CHICAGO\_PUBLIC\_SCHOOLS
3. CHICAGO\_CRIME\_DATA

After you have created the tables, review the data in each table by using the View Data feature in the Db2 On Cloud console.

## Problems

When you have confirmed that the three datasets are loaded into the CENSUS\_DATA, CHICAGO\_PUBLIC\_SCHOOLS, and CHICAGO\_CRIME\_DATA tables, you can proceed. Write and execute SQL queries to solve the problems below.

## Problem 1

List the case number, type of crime and community area for all crimes in community area number 18.

### ▼ Hint 1

Use tables CHICAGO\_CRIME\_DATA and CENSUS\_DATA.

### ▼ Hint 2

Use an inner join.

### ▼ Hint 3

The column PRIMARY\_TYPE contains the crime type.

## Problem 2

List all crimes that took place at a school. Include case number, crime type and community name.

### ▼ Hint 1

Use tables CHICAGO\_CRIME\_DATA and CENSUS\_DATA.

### ▼ Hint 2

Use a left join or a right join.

### ▼ Hint 3

The column LOCATION\_DESCRIPTION will help you find the crime location.

## Problem 3

For the communities of Oakland, Armour Square, Edgewater and CHICAGO list the associated community\_area\_numbers and the case\_numbers.

### ▼ Hint 1

Use tables CHICAGO\_CRIME\_DATA and CENSUS\_DATA.

### ▼ Hint 2

Use a full outer join.

### ▼ Hint 3

Use COMMUNITY\_AREA\_NUMBER from CENSUS\_DATA.

Copyright © 2020 IBM Corporation. All rights reserved.

## Author(s)

**Ramesh Sannareddy**

## **Contributor(s)**

**Rav Ahuja**

## **Change log**

| <b>Date</b> | <b>Version</b> | <b>Changed by</b> | <b>Change Description</b>  |
|-------------|----------------|-------------------|--|
| 2023-05-04  | 1.3            | Benny Li          | Fixed copyright date   |
| 2021-01-29  | 1.2            | Rav Ahuja         | Fixed license statement  |
| 2021-01-28  | 1.1            | Rose Malcolm      | Converted to markdown. Added instructions to skip table setup if done in previous lab. |
| 2020-11-25  | 1.0            | Ramesh Sannareddy | New lab created  |

**© IBM Corporation 2023. All rights reserved.**