SQL Project Report

* Dataset

It is a public data set of Chicago public school data which I got accessed through the course which I did by IBM named “Databases and SQL for Data Science with python”. It includes a variety of data related to the schools, students, teachers, and administration within the CPS system. This data can encompass a wide range of information, such as:

1. School performance metrics
2. Enrolment statistics
3. Demographics of students and staff
4. Attendance records
5. Budget and spending details
6. Standardized test scores
7. Graduation rates
8. Disciplinary actions

* Learnings

I got to learn how to use IBM Db2, store data in it and apply queries on the table and download the output which we required. And also, how to use ibm\_db library in python to directly connect to the database and run query in the notebook.

Then I got to learn about SQLite database how to create your own serverless database and establish connection and use cursor to execute queries. And how to use magic SQL command in Jupiter notebook.

As the data needed some cleaning, So I cleaned it. And then I practiced as many queries as I can and try get maximum insights from the table.

The information I was able to extract from the dataset 1 are:

* Schools with the highest/lowest safety score
* Schools with the highest/lowest average attendance
* No. of student enrolled in the college as per community area
* Health certified school as per the area
* Schools with high teachers scores, environment score and much more