

# **Project A4 Report**

## **Crime Data Analysis**

**Marshall Jones: mjone032**

**SiraaJ Kuttarkar: skudt001**

**Muneeb Ghori: mghor004**

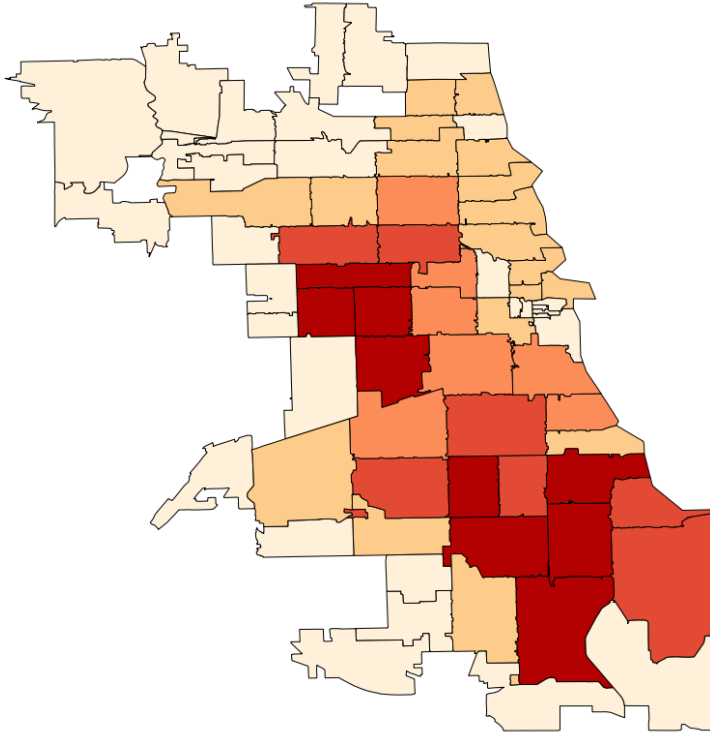
## Task 1

Parquet would be helpful for this project because it performs complex queries and analysis on large datasets by storing columns effectively. In tasks 2 and 3 respectively, spatial analysis and temporal analysis are performed, so Parquet would be a viable option to perform these analysis operations easily.

Dataset	CSV Size	Parquet Size
1,000	205 KB	72 KB
10,000	2 MB	546 KB
100,000	20.5 MB	4.6 MB

## Task 2

Visualization of the result for the 10k file from Task 1:



## Task 3

[Google Sheets File Source](#)

Visualization of 10k Dataset:

