
Academia Open



By Universitas Muhammadiyah Sidoarjo

Table Of Contents

Journal Cover	1
Author[s] Statement	3
Editorial Team	4
Article information	5
Check this article update (crossmark)	5
Check this article impact	5
Cite this article	5
Title page	6
Article Title	6
Author information	6
Abstract	6
Article content	7

Originality Statement

The author[s] declare that this article is their own work and to the best of their knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the published of any other published materials, except where due acknowledgement is made in the article. Any contribution made to the research by others, with whom author[s] have work, is explicitly acknowledged in the article.

Conflict of Interest Statement

The author[s] declare that this article was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright Statement

Copyright © Author(s). This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

Academia Open

Vol. 10 No. 1 (2025): June
DOI: 10.21070/acopen.10.2025.10928

EDITORIAL TEAM

Editor in Chief

Mochammad Tanzil Multazam, Universitas Muhammadiyah Sidoarjo, Indonesia

Managing Editor

Bobur Sobirov, Samarkand Institute of Economics and Service, Uzbekistan

Editors

Fika Megawati, Universitas Muhammadiyah Sidoarjo, Indonesia

Mahardika Darmawan Kusuma Wardana, Universitas Muhammadiyah Sidoarjo, Indonesia

Wiwit Wahyu Wijayanti, Universitas Muhammadiyah Sidoarjo, Indonesia

Farkhod Abdurakhmonov, Silk Road International Tourism University, Uzbekistan

Dr. Hindarto, Universitas Muhammadiyah Sidoarjo, Indonesia

Evi Rinata, Universitas Muhammadiyah Sidoarjo, Indonesia

M Faisal Amir, Universitas Muhammadiyah Sidoarjo, Indonesia

Dr. Hana Catur Wahyuni, Universitas Muhammadiyah Sidoarjo, Indonesia

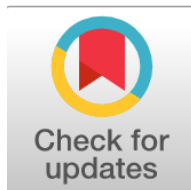
Complete list of editorial team ([link](#))

Complete list of indexing services for this journal ([link](#))

How to submit to this journal ([link](#))

Article information

Check this article update (crossmark)



Check this article impact ^(*)



Save this article to Mendeley



^(*) Time for indexing process is various, depends on indexing database platform

Creating a Website for Chicago City's Crimes Using MongoDB: Membuat Situs Web untuk Kejahatan di Kota Chicago Menggunakan MongoDB

Rusul J. Alsaedi, rusul@uoqasim.edu.iq (*)

Al-Qasim Green University, Babylon 51013, Iraq

(*) Corresponding author

Abstract

General Background: Crime data analysis plays a vital role in enhancing public safety, particularly in densely populated urban areas such as Chicago. **Specific Background:** The increasing complexity of socio-economic environments necessitates scalable tools for real-time data handling and visualization. MongoDB, a NoSQL database, offers advantages in managing large unstructured datasets for dynamic web applications. **Knowledge Gap:** Despite comparative studies between NoSQL and relational databases, there remains a lack of practical implementations integrating real-time visualization of crime data via web interfaces. **Aims:** This study aims to design and develop a prototype website utilizing MongoDB and PyMongo to manage and visualize Chicago crime data from 2001 to the present. **Results:** The system supports seven query operations, including insert, update, delete, and statistical queries by year and arrest status, optimized through indexing on a 6-million-record dataset. It enables CRUD operations and presents interactive visualizations such as bar and stacked charts. **Novelty:** Unlike previous works, this research integrates a full-stack solution combining efficient NoSQL querying with user-friendly visual analytics in a single platform. **Implications:** The prototype can be adapted for broader urban analytics applications, including demographic tracking and population census, offering a scalable framework for real-time data management and decision-making.

Highlights:

Full-stack crime data system using MongoDB and PyMongo

Efficient queries with indexing on large datasets

Interactive visualizations for real-time urban insights

Keywords: Crime data visualization, MongoDB, NoSQL database, urban analytics, real-time web application

Published date: 2025-04-24

Introduction

Chicago is a famous, important and big city, so in this research, we used a dataset about the crimes of Chicago city to create our database. The database is a set of data that is saved, organized, and reached electronically. The database management system (DBMS) is the software that is used with the database, applications, and the end users to analyze and get data. The purpose of DBMS is to administrate the database and create, update, define, and query the database also [1]. The database has many types such as the relational database management system (RDBMS) that most of its systems use SQL (Structured Query Language) to query the database. The relational databases arrange the data into one or more tables or relations, and each table in the database consists of columns and rows. The columns are also called attributes, and the rows are also called records or tuples. The table has a unique key that identifies each row [2]. The other type of the database is the NoSQL database. The NoSQL database means non-relational or non SQL database . It is usually used in the real-time web applications and big data [3]. There are many examples of the NoSQL databases such as MongoDB, CouchDB, GimFire, Redis, Cassandra, memcached, Hazelcast, HBase, Mnesia, and Neo4j . In our research, we chose MongoDB to create our database. MongoDB is a NoSQL database that is an open source document [4]. It consists of collections that equalize to the tables in the RDBMS, and each collection is a set of MongoDB documents. The document equalizes to the row/ tuple in the RDBMS. The field in MongoDB equalizes to the column in the RDBMS [5]. Then, we used a website to show our results. The website is a set of related web pages; that is usually identified with a domain name, and it is published on one web server at least [6].

Therefore, in our research, we have created a website for Chicago city's crimes that allows the user to view all the documents in the database, insert new document to the database, delete one document from the database based on the case number field in the document, and update existing document in the database. Furthermore, our website allows the user to search documents by year. Our website also provides some visualization tools to see some statistical information about the number of crimes that were happened each year in Chicago city. The aim of this research is to implement a prototype that visualizes some statistical information about the crimes that were happened in Chicago city from 2001 to present. The other goal behind this research is to design a NoSQL database by using MongoDB and to use the indexing method in MongoDB to provide the efficient queries that retrieve information from the database. Then, we have implemented seven types of queries using pymongo.py. These queries are as follow: new, find, update, delete, find by year, count by year, and count by arrest. Our database was too big since it contains about 6000000 documents.

1.1 Paper Organization

We proceed as follows. We present the related work in section 2, and discuss the software description in section 3. Finally, we conclude in section 4 with a short discussion.

2. Literature Review

Chang Glasgow et al in section 2 showed how to create a website, and the basics that every simple website is needed to be generated. They showed in their work how to set the style of the website, and how to create the template of it. In addition to that, they showed how to edit the pages of the website [7]. In, they studied the 11 steps of the website creation to generate a successful website by starting from the low level study until reaching the advanced study of the website creation [8]. Mike Garcia showed how to create a simple webpage by using the basics and the principles of the HTML and CSS; also, he showed how to insert paragraphs and images into a webpage [9]. Sumitkumar Kanoje et al discussed the pros and the cons of MongoDB using in social networks since social networks are the biggest networks that are used recently, and they are needed very large storage of data. Therefore, they showed that MongoDB is good to be used in social networks in some points, and it is not good in other points [10]. Zhu Wei-ping et al tried to use the NoSQL database instead of the relational database, and compared between these two types of the databases. They showed that MongoDB (the NoSQL database) is more efficient than the relational database in terms of query big data [11]. Dipina Damodaran B et al explained the performance of the relational (MySQL) and non-relational (MongoDB) databases in the field of super market management system [12]. Comelia GYORODI et al and Sushil Soni et al tried to study the comparison between the relational and non-relational databases by choosing MySQL as an example of the relational database and MongoDB as an example of the non-relational database. Also, they showed in their study that MongoDB is more efficient than MySQL database [13][14].

Our paper is totally different from all the previous papers. We have created a website for Chicago city's crimes by using MongoDB. This website allows the user to view all the documents in the database, insert new document to the database, delete one document from the database based on the case number field in the document, and update existing document in the database. In addition to that, our website allows the user to search documents by year. Our website also provides some visualization tools to see some statistical information about the number of crimes that were happened each year in Chicago city. The aim of this research is to implement a prototype that visualizes some statistical information about the crimes that were happened in Chicago city from 2001 to present. The other goal behind this research is to design a NoSQL database by using MongoDB and to use the indexing method in MongoDB to provide the efficient queries that retrieve information from the database. Then, we have implemented seven types of queries using pymongo.py. These queries are as follow: new, find, update, delete, find by year, count by year, and count by arrest. The interesting point is our database was too big; it contains about 6000000 documents [15].

Methodology

The methodology of this research was centered on the development of a dynamic, user-interactive website designed to

manage and visualize crime data in Chicago using MongoDB as the core database technology. The process began with the acquisition of a publicly available dataset containing detailed crime records in Chicago from 2001 to the present, comprising approximately six million entries. Each entry included 22 attributes, such as case number, date, location, type of crime, and arrest status. MongoDB was selected for its scalability and efficiency in handling large volumes of unstructured data, and a dedicated database named "Project" was created, housing a single collection for the crime records. To enhance query performance, two indexes were implemented: a single-field index on the case number and a compound index on the year and arrest fields. The research utilized Python and the PyMongo library to construct seven core queries new, find, update, delete, find by year, count by year, and count by arrest which were integrated into a website built using standard web development tools. The website enabled users to execute CRUD operations and retrieve statistical insights through interactive visualizations, including bar charts and grouped/stacked charts reflecting crime trends over time. Particular emphasis was placed on optimizing user experience and data accessibility through efficient query design and intuitive interface layout. The system was evaluated based on its ability to manage large-scale data with responsive performance, visual clarity, and ease of navigation. This methodology reflects a practical implementation of NoSQL database capabilities within a real-world urban analytics context.

Results and Discussion

We have got our idea from the fact that Chicago is implementing the idea of big data city. Also, since Chicago is an important city, it was a good idea to investigate the level of crimes in that city.

In this research, we have downloaded the crimes dataset. Our dataset is too big (it is about 6000000 records, and each record represents one crime with 22 fields). We used MongoDB to create a database called Project, which contains one collection with 5940000 documents (one document for each crime). To improve the performance of the queries, we have created two indexes in addition to the default index. The first one is based on the case number field, and the second one is compound index which is based on (Year and Arrest fields). The following figures show our project (Figure 1) (Figure 2).



Figure 1.

Figure 1: the home page of our website which contains several links (each one for one function).

The screenshot shows a web browser window with the address bar displaying 'localhost:8001/new'. The page has a blue header with the title 'Add New Document'. Below the header is a navigation bar with links: Home, View, Edit, New, Delete, Query By Year, and Statistic Visualization. The main content area contains a form with three columns of input fields. The first column contains fields for ID, Block, xfer, Description, Domestic, Ward, FBI Code, Year, and Longitude. The second column contains fields for Case Number, IUCR, Location Description, Beat, District, X Coordinate, Updated On, and Location. The third column contains fields for Date, Primary Type, Arrest, District, Community Area, Y Coordinate, Latitude, and a 'Submit' button. The form is styled with a light gray background and white text.

ID:	Case Number:	Date:
122334547	HY23456	11/12/2001 11:34:00 pm
Block:	IUCR:	Primary Type:
xfer	1410	Robbery
Description:	Location Description:	Arrest:
Without_Gun	street	?
Domestic:	Beat:	District:
?	0234	34
Ward:	District:	Community Area:
56	05	f
FBI Code:	X Coordinate:	Y Coordinate:
54	114356	234566
Year:	Updated On:	Latitude:
2001	2015	23.4566
Longitude:	Location:	Insert to the Database:
114.4564	Chicago	Submit

Figure 2.

Figure 2: the new page which allows the user to add one document to the database.

Figures (3 and 4) show the edit page that allows the user to edit specific document from the collection. This page asks the user to enter the case number; then, the system will automatically find and return that document from the collection; after that, it shows the values of all variables of that document. The user can modify the values and submit the changes to the system to update the document.

The screenshot shows a web browser window with the address bar displaying 'localhost:8001/edit'. The page has a blue header with the title 'Edit the Selected Document'. Below the header is a navigation bar with links: Home, View, Edit, New, Delete, Query By Year, and Statistic Visualization. The main content area contains a form with a 'Case Number' input field and a 'Submit' button. Below this is a table with three columns: ID, Case Number, and Date. The table contains the same data as the 'Add New Document' page. The form is styled with a light gray background and white text.

ID:	Case Number:	Date:
122334547	HY23456	11/12/2001 11:34:00 pm
Block:	IUCR:	Primary Type:
xfer	1410	Robbery
Description:	Location Description:	Arrest:
Without_Gun	street	?
Domestic:	Beat:	District:
?	0234	34
Ward:	District:	Community Area:
56	05	f
FBI Code:	X Coordinate:	Y Coordinate:
54	114356	234566
Year:	Updated On:	Latitude:
2001	2015	23.4566
Longitude:	Location:	Update the Document:
114.4564	Chicago	Submit

Figure 3.

Figure 3: edit page.

Edit the Selected Document

Home View Edit New Delete Query By Year Statistic Visualization

Case Number:

ID:	Case Number:	Date:
122334547	HY23456	11/12/2001 11:34:00 pm
Block:	IUCR:	Primary Type:
rfer	1410	Robbery
Description:	Location Description:	Arrest:
Without_Gun	street	<input "="" type="text" value="?"/>
Domestic:	Beat:	District:
<input "="" type="text" value="?"/>	0234	34
Ward:	District:	Community Area:
56	34	f
FBI Code:	X Coordinate:	Y Coordinate:
54	114356	234566
Year:	Updated On:	Latitude:
2001	2015	23.4566
Longitude:	Location:	Update the Document:
114.4564	Chicago	Submit

Figure 4.

Figure 4: edit page (with values).

Figures (5 and 6) show the query by year page. In this page, the user enters the specific year in the year text box and click the submit button. The system will show a table which contains all the crimes that were happened in the input year.

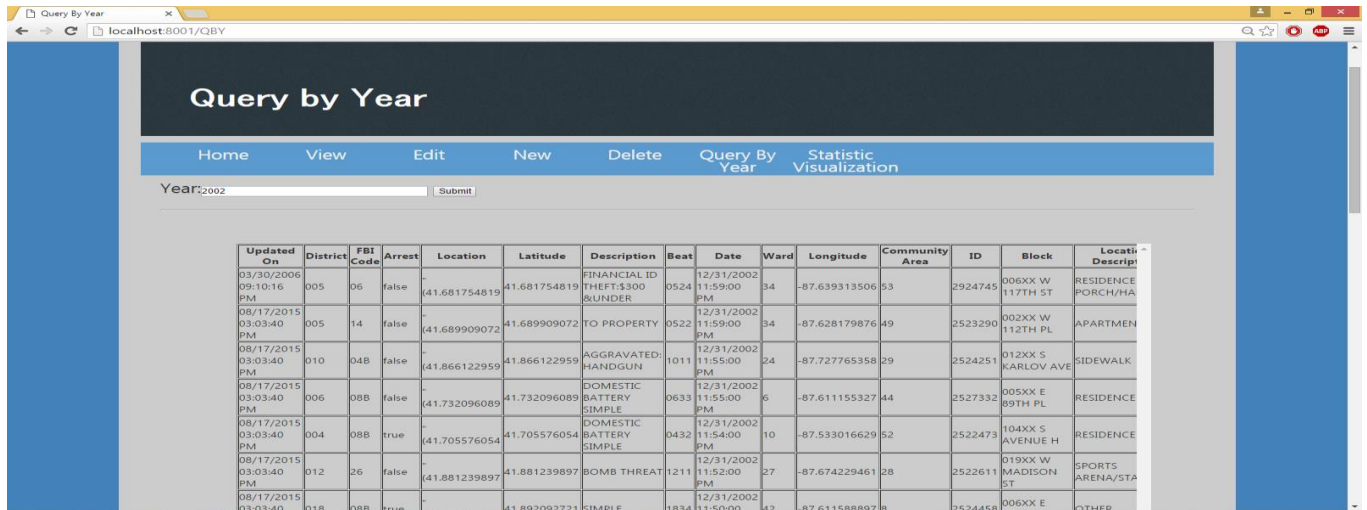
Query by Year

Home View Edit New Delete Query By Year Statistic Visualization

Year:

Figure 5.

Figure 5: query by year.



Updated On	District	FBI Code	Arrest	Location	Latitude	Description	Beat	Date	Ward	Longitude	Community Area	ID	Block	Location Description
03/30/2006 09:10:16 PM	005	06	false	(41.681754819	41.681754819	FINANCIAL ID THEFT \$300 & UNDER	0524	12/31/2002 11:59:00 PM	34	-87.639313506	53	2924745	006XX W 117TH ST	RESIDENCE PORCH/HA
08/17/2015 03:03:40 PM	005	14	false	(41.689909072	41.689909072	TO PROPERTY	0522	12/31/2002 11:59:00 PM	34	-87.628179876	49	2523290	002XX W 112TH PL	APARTMEN
08/17/2015 03:03:40 PM	010	04B	false	(41.866122959	41.866122959	AGGRAVATED HANDGUN	1011	12/31/2002 11:55:00 PM	24	-87.727765358	29	2524251	012XX S KARLOV AVE	SIDEWALK
08/17/2015 03:03:40 PM	006	08B	false	(41.732096089	41.732096089	DOMESTIC BATTERY SIMPLE	0633	11:55:00 PM	6	-87.611155327	44	2527332	005XX E 89TH PL	RESIDENCE
08/17/2015 03:03:40 PM	004	08B	true	(41.705576054	41.705576054	DOMESTIC BATTERY SIMPLE	0432	12/31/2002 11:54:00 PM	10	-87.533016629	52	2522473	104XX S AVENUE H	RESIDENCE
08/17/2015 03:03:40 PM	012	26	false	(41.881239897	41.881239897	BOMB THREAT	1211	12/31/2002 11:52:00 PM	27	-87.674229461	28	2522611	019XX W MADISON ST	SPORTS ARENA/STA
08/17/2015 03:03:40 PM	018	08B	true	(41.892092721	41.892092721	SIMPLE	1834	12/31/2002 11:50:00 PM	42	-87.611588897	8	2524458	006XX E	OTHER

Figure 6.

Figure 6: the result of the query by year.

Figure 7 shows statistical visualization page that shows an interactive bar chart for the number of crimes that were happened each year in Chicago city from 2001 to present. In the bar chart, the x-axis represents the year, and the y-axis represents the number of crimes.



Figure 7.

Figure 7: number of crimes each year.

Figures (8 and 9) show the grouped and stacked bar chart which can be used to do the comparison between the number of crimes and number of arrested criminals per year in Chicago city. These two visualization tools are interactive (the name and the value of each bar can be shown by moving the mouse over the bars).

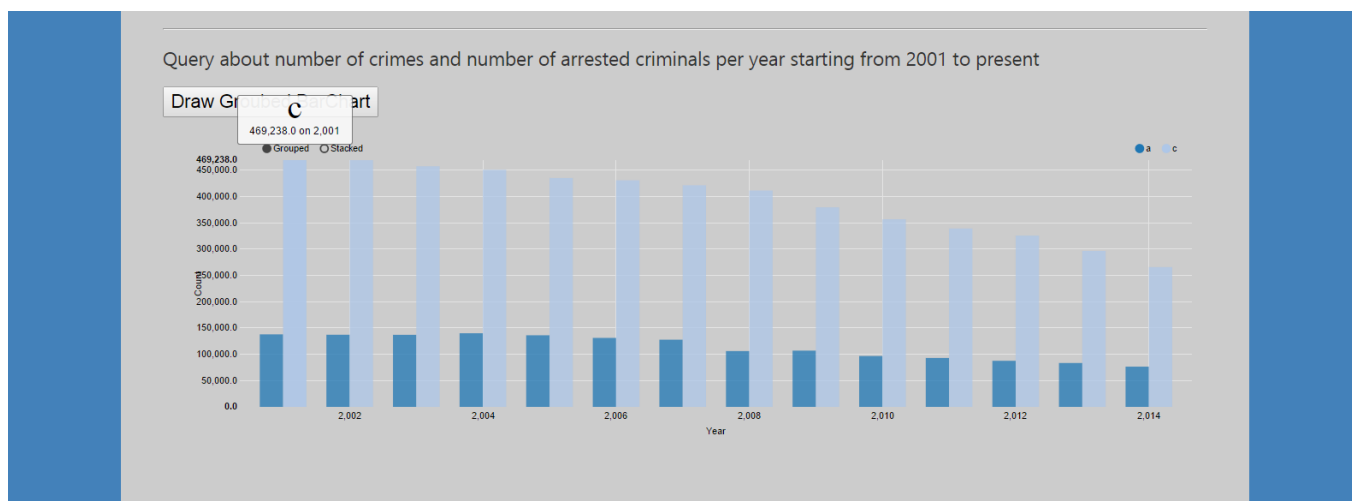


Figure 8.

Figure 8: grouped bar chart.

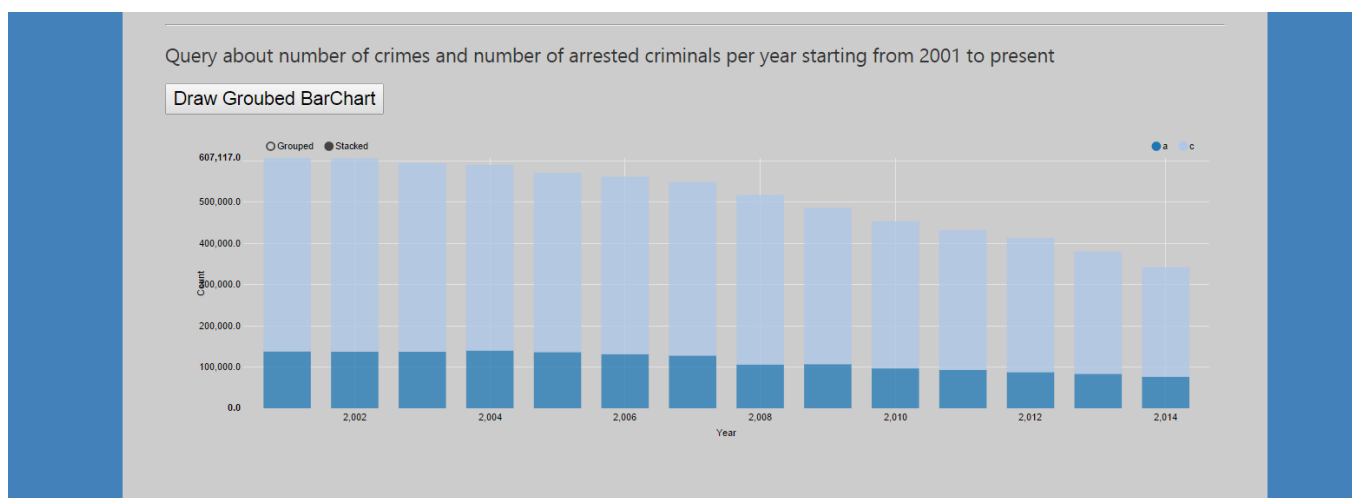


Figure 9.

Figure 9: stacked bar chart.

The interesting point is our database was too big since it contains about 6000000 documents; this makes the work interesting by trying to improve the performance of the queries by doing multiple types of indexing for the database. Also, the results of the queries were very interesting; for example, in one query, we have found that the number of crimes each year is too big compared with the number of arrested criminals.

Conclusion

In this research, we worked on the dataset of the crimes of Chicago city since Chicago is an important and big city. As a result, we have created a website that allows the user to view all the documents in the database, insert new document to the database, delete one document from the database based on the case number field in the document, and update existing document in the database. Furthermore, our website allows the user to search documents by year. Moreover, our website

provides some visualization tools to see some statistical information about the number of crimes that were happened each year in Chicago city. The first aim of this research is to design a NoSQL database by using MongoDB and to use the indexing method in MongoDB to provide the efficient queries that retrieve information from the database. Then, we have implemented seven types of queries using pymongo.py. These queries are as follow: new, find, update, delete, find by year, count by year, and count by arrest. The other goal behind this research is to implement a prototype that visualizes some statistical information about the crimes that were happened in Chicago city from 2001 to present.

For future work, we are going to add more queries on the crimes database. Also, we are going to use other visualization methods to visualize the results of the queries. For example, we are going to use the heat-map to show the crimes that were happened in each location in Chicago city based on the latitude and longitude fields that we have in our database. Also, this research can be applied in any city or country in the world and about any topic related to this type of a NoSQL database that contains huge amount of information. One of the examples is to create a website that shows the information of the population of Iraq in any city, and this can add the new information by adding the new bourns through the new query in the database design, and any information can be updated by the edit query. This will benefit in population census in Iraq. However, this example requires the online dataset or Excel sheet of the population of Iraq to be applied.

References

1. K. Santhiya and V. Bhuvaneswari, "An Automated MapReduce Framework for Crime Classification of News Articles Using MongoDB," *Int. J. Appl. Eng. Res.*, vol. 13, no. 1, pp. 131-136, 2018.
2. V. Jain, A. K. Dubey, A. Jain, M. Malhotra, and S. Rastogi, "Crime Pattern Recognition in Chicago City Using Hadoop Multinode Cluster," *J. Inf. Optim. Sci.*, vol. 40, no. 2, pp. 587-601, 2019.
3. N. Baviskar, *Smart City Development Using Data Analytics*, Doctoral dissertation, California State Univ., Sacramento, 2017. [Online]. Available: <https://spring.io/understanding/NoSQL>
4. K. Banker, D. Garrett, P. Bakum, and S. Verch, *MongoDB in Action: Covers MongoDB Version 3.0*. New York, NY, USA: Simon and Schuster, 2016.
5. N. E. Stone, *Social Media Canvassing Using Twitter and Web GIS to Aid in Solving Crime*, Master's thesis, Univ. of Southern California, 2017.
6. G. Chang, G. Juhasz, and L. Stephan, "Creating Your Personal Website," Mac Edition, 2006. [Online]. Available: <http://andrew.cmu.edu/70-271-htmlman>
7. Webware Staff, "11 Steps to Create a Successful Website," *StartupNation/CNET*, 2007.
8. M. Garcia, "Creating a Webpage Using HTML & CSS," *ULN Internship Program, PCL Media Lab*, 2015.
9. S. Kanoje, V. Powar, and D. Mukhopadhyay, "Using MongoDB for Social Networking Website," in *Proc. IEEE 2nd Int. Conf. Innovations Inf. Embedded Commun. Syst. (ICIIECS)*, 2015.
10. W.-P. Zhu, M.-X. Li, and H. Chen, "Using MongoDB to Implement Textbook Management System Instead of MySQL," in *Proc. 2011 IEEE 3rd Int. Conf. Commun. Softw. Netw.*, pp. 497-501, 2011.
11. D. D. B. Dipina, S. Salim, and S. M. Vargese, "MongoDB vs. MySQL: A Comparative Study of Performance in Super Market Management System," *Int. J. Comput. Sci. Inf. Technol. (IJCSITY)*, vol. 4, no. 2, pp. 1-10, 2016.
12. C. Gyorodi, R. Gyorodi, G. Pecherle, and A. Olah, "A Comparative Study: MongoDB vs. MySQL," in *Proc. 2015 13th Int. Conf. Eng. Mod. Electr. Syst. (EMES)*, pp. 89-94, 2015.
13. S. Soni, M. Ambavane, S. Ambre, and S. Maitra, "A Comparative Study: MongoDB vs. MySQL," *Int. J. Sci. Eng. Res.*, vol. 8, no. 5, pp. 1701-1705, 2017.
14. S. M. Kiio, *Apache Spark Based Big Data Analytics for Social Network Cybercrime Forensics*, Doctoral dissertation, Univ. of Nairobi, 2017.
15. A. Shaker, N. Mandela, and A. K. Agrawal, "Review on Analyzing and Detecting Crimes," in *Proc. Int. Conf. Commun. Netw. Comput.*, Cham, Switzerland: Springer Nature, pp. 116-127, Dec. 2022.