US Department of Agriculture and Rural Development Dataset in MySQL Database Security: Using Sisense Tool

Maytham N. Meqdad[1]
Al-Mustaqbal University College, Hillah, Babil, Iraq

Mustafa Habeeb Obaid Al Imari[2]
Al-Mustaqbal University College, Hillah, Babil, Iraq

Abstract:

In this paper we will analyze part of the information stored in the SQL database. The information is then converted into curves and diagrams using the Sisense Tool program. The data can be converted to curves and charts after a period of time. In the case of different shape curves and charts, this means that the data was broken.

Keywords: database security, Sisense Tool

Introduction

A Databases is a set of logical data elements linked to each other in a mathematical relationship, and the database consists of several tables. The table consists of several records and the record consists of a set of fields. For example, the record for a particular university student consists of several fields, such as student number, student name, grade level, grade of student, and student's science.

MySQL is a central component of the LAMP open-source web application software stack. LAMP is an acronym for "Linux, Apache, MySQL, Perl/PHP/Python". Applications that use the MySQL database include: TYPO3, MODx, Joomla, WordPress, phpBB, MyBB, and Drupal. MySQL is also used in many high-profile, large-scale websites, including Google (though not for searches), Facebook, Twitter, Flickr, and YouTube.[2][8]

The rate of data stored in databases is growing exponentially. Today, database technology has become a key element in many technological applications. It allows storage of large data, ease of data sharing electronically, ensuring the integrity of data and securing access to unauthorized persons. [9]

The Privacy Rights Clearing House (2010) reports confirmed that 345 million records were stolen for customers during 2005 since begin data theft incidents were recorded.[1]

The Ponemon Institute report also indicates that the average cost of the data mix per customer record is 202$.[5]

In 2009, a US criminal court issued criminal indictments against three data-hackers who stole more than 130 million credit cards by hacking a SQL database.[6]

In addition to the above, there are many electronic crimes and penetration of unregistered databases (for personal databases or companies) carried out by criminal groups for the purpose of money theft or defamation. Or to carry out terrorist operations that harm the security of states and people.

Database & Tools

A-MySQL Database

MySQL was created by a Swedish company, MySQL AB. Original development of MySQL by Widenius and Axmark began in 1994. The first version of MySQL appeared on 23 May 1995. MySQL can be built and installed manually from source code, but it is more commonly installed from a binary package unless special customizations are required. On most Linux distributions, the package management system can download and install MySQL with minimal effort, though further configuration is often required to adjust security and optimization settings.[2][7]

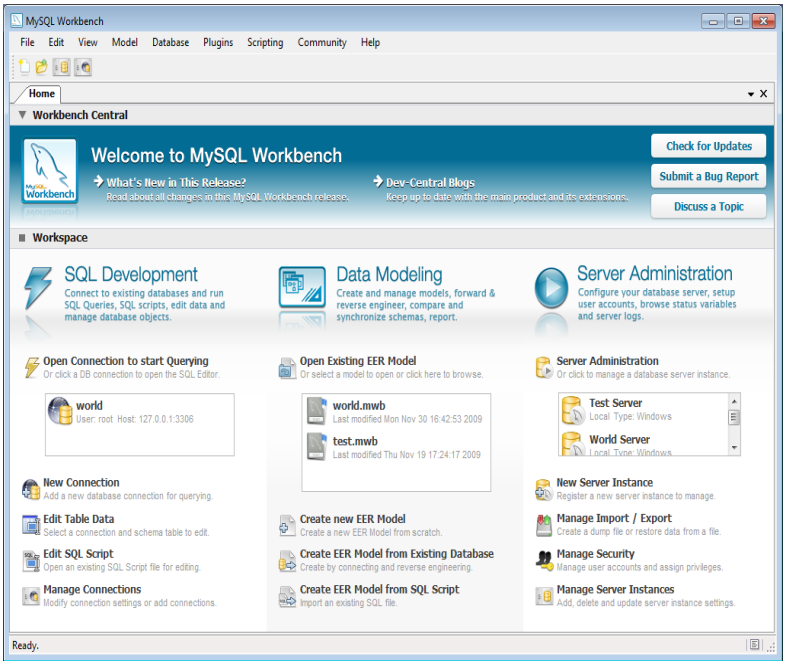

Figure 1. MySQL Database Interface

B-Sisense Prism Tool

Sisense Prism is a product suite for building business intelligence solutions. Prism is a single integrated environment providing an extensive toolset to accomplish the following:

-Analyzing data: - Formulate complex business queries using an intuitive visual interface, add custom calculations.

-Visualizing data: - use a wide range of data visualization widgets to visualize data.

-Preparing the data for analysis: - merge data from disparate sources, manipulate and cleanse data, full ETL capabilities.

-Support for large data volumes: - powered by ElastiCube technology, queries millions of rows of data in seconds.

-Data Exploration and Ad Hoc Analysis: - get insight into your data using a wide range of data exploration features.[4]

Figure 2. Sisense prism logo

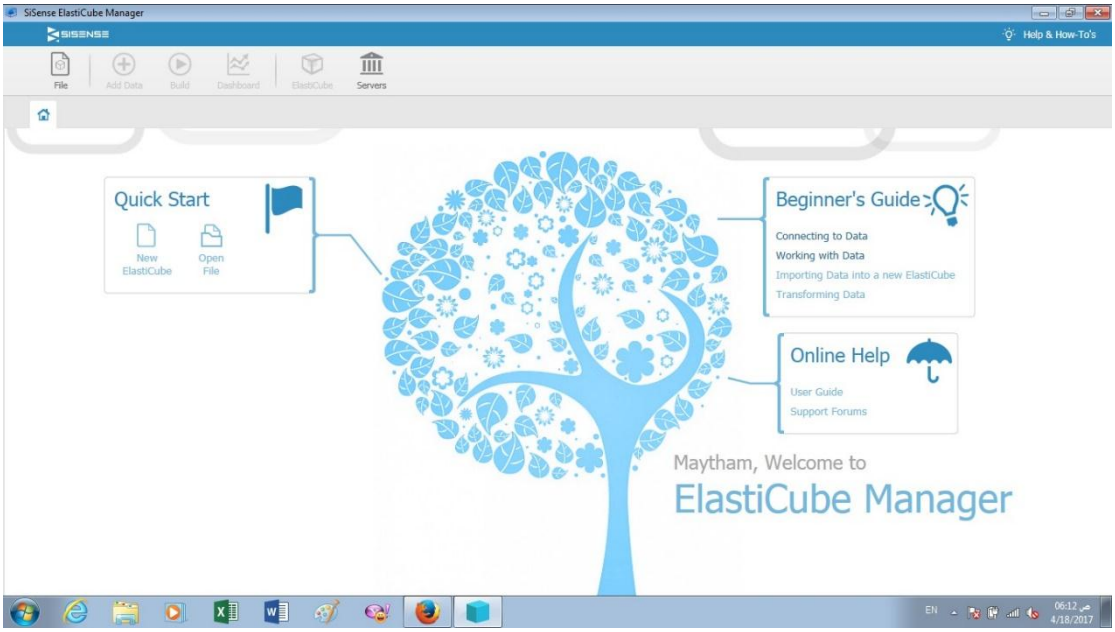

Figure 3. Sisense prism interface

In this article, we will use the feature visualizing data to indicate the characteristics of the data that we will analyze.

Data & Analysis

A-Data description

The data analyzed in this article are from the US Department of Agriculture and Rural Development. This data was last updated on 7 February 2017. These data show real estate sold by individuals or government. These properties may be homes owned by families or farms. These properties were sold by public auction or by other means depending on the type of property. The data is in a CSV file format. The file field consists of data, the name of the property, the date of sale, the address of the property, the seller's personal data, the seller's email, the specifications of the property, the price of the property and others. These data are controlled and stored in the MySQL database in the form of spreadsheets. [3]

B- Data analysis

By using the Sisense analysis tool, we have entered part of the data we have obtained from the US Department of Agriculture. Using some of the corals, the following figures appeared:

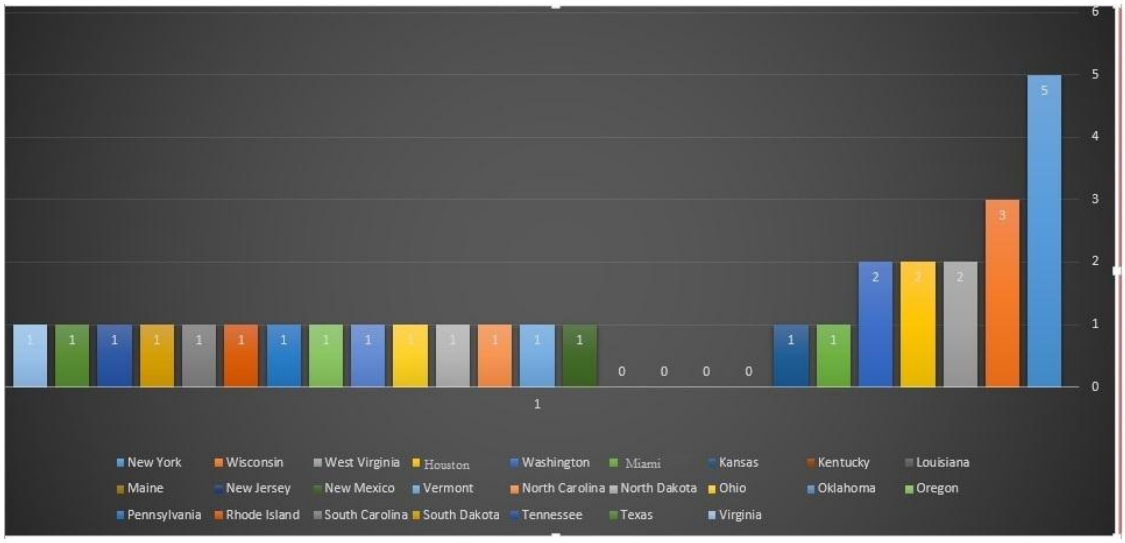

Figure 4. United States (Stacked Colum)

Figure 4 shows the Stacked Column of real estate sold in each state. As shown in the figure, New York State has the most selling properties.

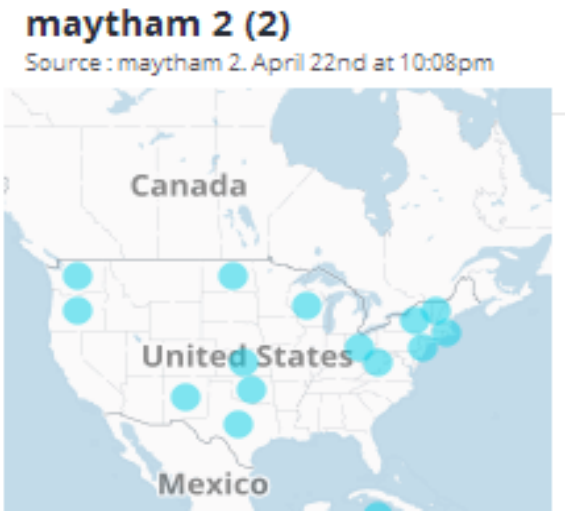

Figure 5. United States (Area Map)

Figure 5 shows the distribution of real estate on the US map.

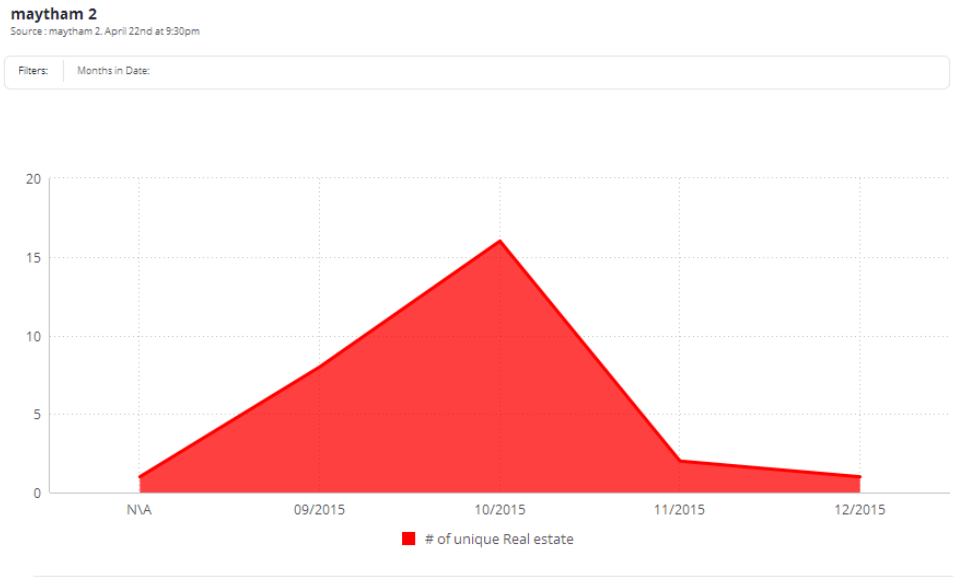

Figure 6. Date of sale of real estate (Area chart)

Figure 6 shows the real estate sales area chart during the period of time. As the figure shows, the most time period of sale is 10/2015.

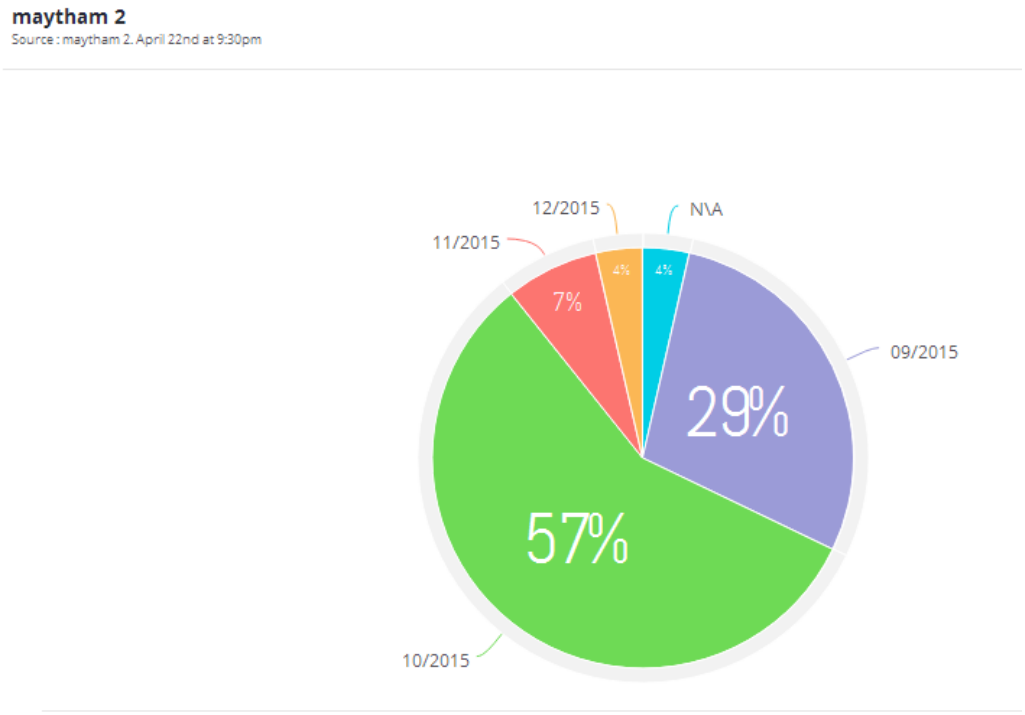

Figure 7. Date of sale of real estate (pie chart)

Figure 7 shows the percentage of real estate sold by time period. As the figure shows, the most time period of sale is 10/2015.

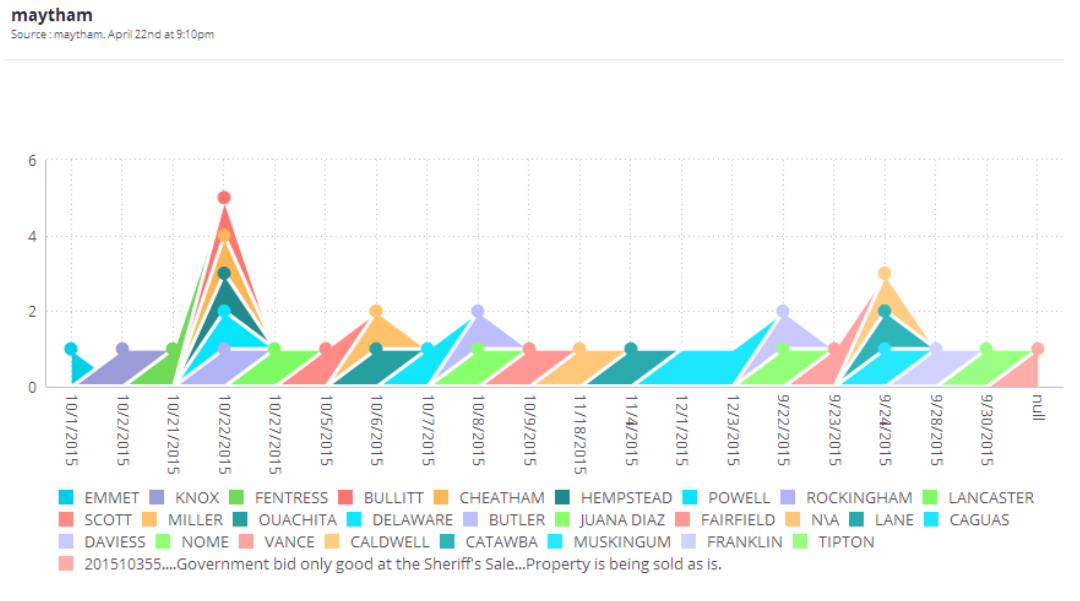

Figure 8. Date of sale of real estate (line chart)

Figure 8 shows the line chart of real estate sold by time period. As the figure shows, the most time period of sale is 22/10/2015.

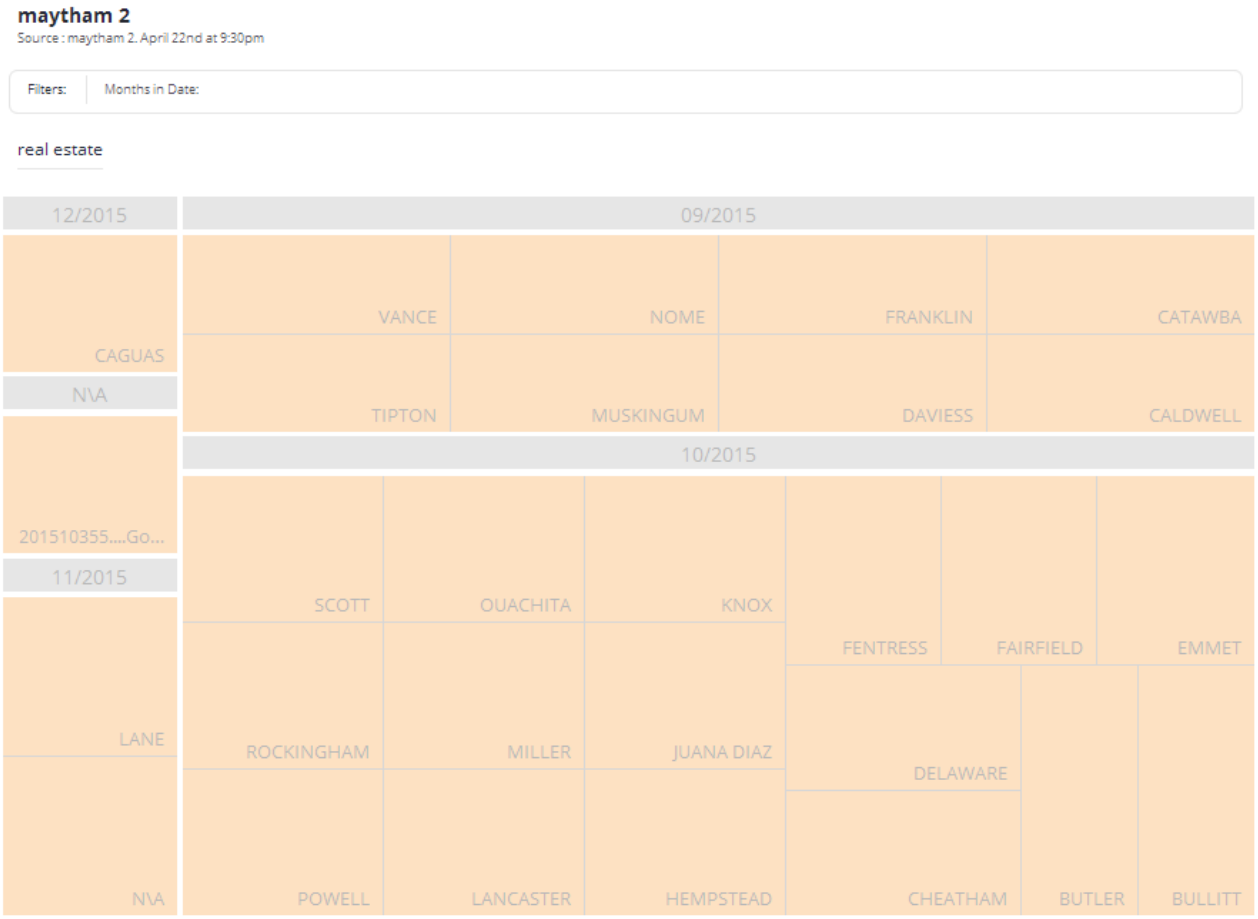

Figure 9. Date of sale of real estate (tree map)

Figure 8 shows the tree selling real estate during the time period. As the figure shows, the largest square for a period of time when the sale is 10 \ 2015.

C-Purpose of data analysis

The data shows the movement of real estate sales from 9/2015 to 12/2015 with some unknown sales. The data shows the movement of real estate sales to some areas of the United States of America.

After using the Sisense analysis tool on the data, the following results appeared:

Figure 4 shows the Stacked Column of real estate sold in each state. As shown in the figure, New York State has the most selling properties. Figure 5 shows the distribution of real estate on the US map. Figure 6 shows the real estate sales area chart during the period of time. As the figure shows, the most time period of sale is 10/2015. Figure 7 shows the percentage of real estate sold by time period. As the figure shows, the most time period of sale is 10/2015 shows the line chart of real estate sold by time period. As the figure shows, the most time period of sale is 22/10/2015. Figure 8 shows the line chart of real estate sold by time period. As the figure shows, the most time period of sale is 22/10/2015. Figure 8 shows the tree selling real estate during the time period. As the figure shows, the largest square for a period of time when the sale is 10 \ 2015.

Certainly these results will be fixed and non-changeable because they are sales transactions on fixed and previous dates. Any change to these data is a change of ownership illegally. This process is therefore a theft of real estate and public and private funds.

The purpose of the US Department of Agriculture and Rural Development data analysis is to know the sales ratios during the time period where we can keep this ratio and curves for the purpose of applying the analysis periodically (daily, weekly or monthly). To ensure that the data has not been hacked and counterfeited so we can maintain the electronic records of real estate and thus the preservation of property and public and private property.

If we assume that we have re-applied the algorithms Sisense analysis tool on the same data in another period and we have shown different forms, this confirms that the data has been subjected to forgery and penetration

We can compare these data accurately with the use of image processing programs to check data drawings so that they are similar. Any difference in ratio or curves means that data may be subject to forgery and piracy.

In the future we can program an application that performs the process of self-analysis as well as continuous comparison of results to detect cases of penetration and falsification of data.

Conclusion & Future work

In this paper we analyzed the data stored in the SQL database. The purpose of the analysis is the easy and quick way to find out if the data is broken, deleted or added. By converting the data into curves and charts, the database manager can easily detect the differences if the shape of the charts and curves changes after the duration.

The future work is to program the image processing program. The purpose of this program is to detect any change in the curves and charts. Hence the knowledge of the violations that are exposed to the information within the database.

References

[1] Jason D. Collins Vincenzo A. Sainato David N. Khey Loyola University New Orleans, USA "Organizational Data Breaches 2005-2010: Applying SCP to the Healthcare and Education Sectors" Collins, et.al - Organizational Data Breaches 2005-2010: Applying SCP to the Healthcare and Education Sectors Copyright © 2011 International Journal of Cyber Criminology (IJCC) ISSN: 0974 – 2891 Jan – July 2011, Vol 5 (1): 794–81

[2]"Available Today: Preview Release of the SQL Server ODBC Driver for Linux". *SQL Server Team Blog. 2011-11-28*. Retrieved 2013-06-15.

[3] US Rural Development. Retrieved, October 30, 2018, from https://catalog.data.gov/dataset/usda-rural-development-resale-properties-foreclosure

[4] Sisense Prism V4.7 Documentation. Retrieved, October 30, 2018, from https://documentation.sisense.com/

[5] Ponemon Institute LLC Attn: Research Department 2308 US 31 North Traverse City, Michigan 49686 USA 1.800.887.3118 "2017 Cost of Data Breach Study" Ponemon Institute Research Report

[6] David Jones, Jim Finkle, REUTERS, October 30, 2018, from https://www.reuters.com/article/us-usa-hackers-creditcards/u-s-indicts-hackers-in-biggest-cyber-fraud-case-in-history-idUSBRE96O0RI20130725

[7] Krishna Kulkarni, Jan-Eike Michels "Temporal features in SQL: 2011" SIGMOD Record, September 2012 (Vol. 41, No. 3)

[8] Joseph D'Antoni (Denny Cherry & Associates Consulting), Stacia Varga (Data Inspirations) "Security and Azure SQL Database" Published: October 2015, Applies to: Microsoft Azure SQL Database

[9] Kaushik Sekaran, Maytham N. Meqdad, Pardeep Kumar, Soundar Rajan, Seifedine Kadry "Smart agriculture management system using internet of things" Published: June2020, (Vol.18, No.3, June2020, pp. 1276~1285), TELKOMNIKA Telecommunication, Computing, Electronics and Control