Graduate Enhancement Essay

Mike Pendleton

West Texas A@M University

Business intelligence in Decision support systems (BIDSS), is a method of organizational decision making results that can be one of the following methods according to (FRANCIS CHEMORION, 2020); Communicative driven, document driven, model driven and data driven. Data warehouse provides the means to store the collection of data. Post stored then the data can be used to visualize the data intended after it has been cleaned and ready for its purpose or intent.

Business intelligence contains the tools to be used for design support systems. When they are combined is when the BIDSS system is used together. For instance, Artificial intelligence is sometimes considered a part of business intelligence; however, AI is the decision support system leveraging the tools and data stored from business intelligence.

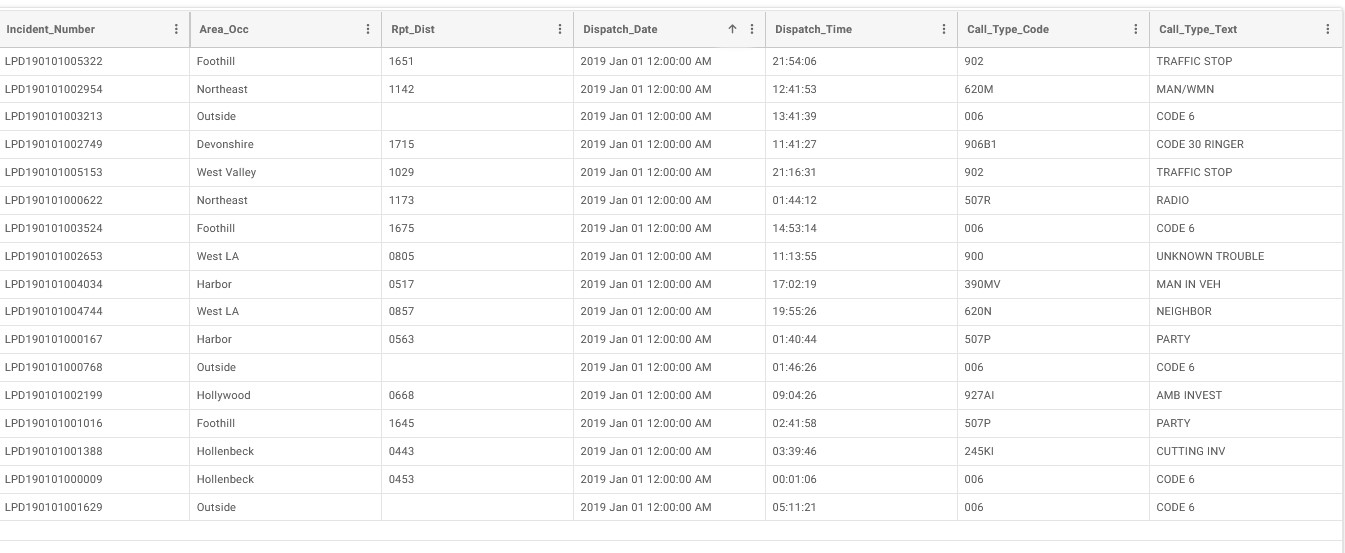
The project submitted was utilizing API data from LA County containing Dispatch data. BIDSS would be able to leverage a way to view this data where the data would make it more easily to determine what the data actually means for others that are not internal to LA County, but would need to understand how this data can help them.

For instance, one could make decisions on where they may want to move to depending on the amount of times each city has had an x amount of code 6 calls. Code 6 is when police are sent out to a residence or location. Individuals could then make a decision for the best city to live in as the lower amount of calls where police were dispatched, could be quantified to a lower crime rate. This coincidentally could also be something that either a new business owner, or franchises could see cities and/or districts that would pose a lower risk VS higher amount of dispatched cities.

Without knowledge of how LA County stores their data, then I would proceed with what is present where all data is stored into a single API json file. Both data sets used were identical regarding how the data was stored. This made visualizing data more of a task versus having a relational approach. Utilizing SQL or some other database structure would be able to normalize this data and create queries in order to get what the data’s use case is.

Based on my findings a complete restructure of the data would be needed. The major questions I had from the project, which I still have, as this data was not available. For instance, the man,woman, and other call type text fields, there is still an unknown to what exactly does this mean. Was the man the victim or who called? This would be the same questions that I have pertaining to woman and other. The best way to get the data needed in order to correct this would be meeting with the business team and discuss the findings, along with how to store the data in a relational structure.

Below shows a sample of what the data looks like:



An interesting result is that the json datatype for all fields were text fields, this included the date fields. When this data is then attempted to be imported into a database then if any of the data types were changed, then no data imported. What would need for this use case would be to import the data and then change the datatypes after the data exists in the new tables. I am expecting that the County only did this for the API JSON files and their actual internal data is normalized with the correct data types.

How to best design the data would be by discussing the initial questions with the Business Unit. Due to the more generic descriptions of some in which they have their own codes, then it would be better to further break it down. This may end up with a dozen or more codes, but the view used by the dispatchers would be easier is there would be more of a drop down effect and not writing the text directly.

See the below diagram which has a Dispatch DB and all this data would be housed in three tables.



There are a couple other questions that would arise here, but the scope of LA County is to provide this information to the public. Data containing names, age, address and any other personal identifiable information would not be included. However, by adding a gender column in the Dispatch\_Call table the man and women codes would no longer be needed. If there is a specific need for what the man,woman or other existent data entails, then progressing with how to properly address these types of data in the new design.

A complete revamp of the design could be simplified as well. Since LA Counties data is stored via API, then all their data sets can be imported into a new database. Once done, then there would be two more tables created and then move the according data to the appropriate table and correct the datatypes for future use as stated above.

As part of the design, the City table has city and district. The purpose on why I would implement this is you can have a many to one relationship, as there are multiple districts per city. For instance if Fullerton id of 5, then the data if you were to select \* from city where city id =5, then what would be returned would look like the following:



According to (City of Fullerton, n.d.), this particular city has five Districts. Though this could be viewed as duplicate data, an other option would be to have a District table where it would reference a city id. Then a primary key would like from city id to district id. Since California is extremely unlikely to create new districts or new cities, I would proceed the direction of the static data in one table. Any further defining or redefining of data (adding columns, tables) would be followed through the consideration of the business units to discuss future changes.

With the design on how to break the data out, then discussing with the business unit further to determine how to break out the codes as the dispatcher would only need to select a drop down box and select what the appropriate field would be. The candidates to select in order to implement a design refresh would be the dispatchers, dispatch management, sample size of police, and organizational management. I would include top level organizational management is this data may be used by other business units and not solely just for dispatch employees. Any change to business process would need to be discussed at several senior levels to ensure data integrity is taken into account. If this data is changed and another department utilizes the data, then their processes would be impacted.

This is also known as business process management. For the scope of the paper, the integration would be with integration-centric BPM. According to (What is Business Process Management? An Introduction to BPM and its benefits, n.d.), “**Integration-centric BPM** focuses on processes require little human involvement. Instead, these processes rely on APIs and automated tools that integrate data across systems. Often, these types of BPM integrate enterprise tools such as Enterprise Resource Management (ERP), Human Resource Management (HRM) or Customer Relationship Management (CRM) suites.” While researching this concept, my ideas on who would need to contact is also referenced in the article. This business article states to have governance of the process roles defined.

One of the visualizations at the time of completing my project that would benefit significantly from this design is the time series. The times and cities was difficult to determine a big picture on what the data is actually providing. Adding the district would help with viewing this data, as one of the questions that an individual may consider is the times that police are dispatched. Since the majority of crimes take place late at night, then the data can be filtered further down after the design is integrated and used.

Leveraging Jupyter Notebook, with pandas, matplotlib, and API’s then if this data was stored in three different API urls, then you could parse the data through pandas like you would be able to with sql. This provides a way to publicly provide the intended data and not providing access to the backend database.

As an additional option if the County of LA wanted to have the public read access to the database, then pushing a Virtual Machine in the DMZ and replicate the data from the internal to external database or they could leverage a VPC virtual machine in order to accommodate this. Synchronizing data would control what data is stored in the database housed in the DMZ. This would also help address risks, vulnerabilities, since data availability is controlled, and housed in the DMZ where the organization would only allow internal to external, and not bi-directionally. Knowing how large the county is, I am assuming that their infrastructure is already established for secure internal and external communication containing Virtual Machines.

# References

City of Fullerton. (n.d.). *City of Fullerton*. Retrieved from find-my-district: https://www.cityoffullerton.com/government/departments/city-clerk/elections/find-my-district?locale=en

FRANCIS CHEMORION. (2020). *BUSINESS INTELLIGENCE IN DECISION SUPPORT SYSTEMS*. Retrieved from /www.pathwaysinternational.com: https://www.pathwaysinternational.com/2017/05/12/business-intelligence-decision-support-systems/

*What is Business Process Management? An Introduction to BPM and its benefits*. (n.d.). Retrieved from www.softwareag.com: https://www.softwareag.com/en\_corporate/resources/process-management/guide/business-process-management.html?utm\_source=google&utm\_medium=cpc&utm\_campaign=bt\_aris&utm\_region=hq&utm\_subcampaign=stg-1&utm\_content=guide\_bpm-intro&gclid=Cj0KCQjwwvilBhCFARIsADv