  
  
ALY6110-22824, Winter 2021  
DATA MANAGEMENT & BIG DATA  
MODA Case study

SUBMITTED BY: ANUPREETA MISHRA

NUID: 001050752

SUBMITTED TO: VALERIY SHEVCHENKO

DATE: 01/27/2021

CONTACT: mishra.anu@northeastern.edu

**Summary**

The case study talks about how New York City agencies are using Big Data to change the way they operate from being reactive to proactive. It talks about how this change came about and why it was necessary. What different city agencies like New York City Police Department (NYPD), the New York City Fire Department (FDNY), and Parks and Recreation used this system to hundreds of lives and improve the standard of living in New York. By 2018, 38 City agencies published 629 new datasets, including 38 new automation, in the Fiscal Year 2018[2]. It started in 1994 when William J. Bratton joined as a commissioner and introduced a tool called CompStat (“computerized comparison crime statistics”). This reduced the crime rate in New York by half in ten years. This was a great feat, the other agencies started to take notice and in 2007, FDNY joined in to add Data Analysis to prioritize building inspections and predict and prevent fires to optimize personnel and monetary resources. As more agencies started joining in, they realized there was a need to pool their data for the best output. This is when Mayor Bloomberg, elected in 2002, came into the picture. He set up a single point that would pool the resources of these different agencies and provide information as and when needed to different agencies. This was called the Mayor’s Office of Data Analytics (MODA). This agency has helped the chiefs make a data-driven decision that tends to be smarter and has made information easier to access, improving the efficiency of the government. However, this has also had a lot of privacy concerns being raised and tends to have a bias as it involves focus on historical data when things in the real world change daily, hence we might miss things without realizing and this may cost us.

**Analysis**

Data Analysis has become a huge part of the private and the public sectors. It has massive and multi-sourced data-driven capabilities that humans on their own just cannot match. It has great analytical and statistical capabilities which help give perspective and outcomes. We saw how William J. Bratton proved effective in his position as chief of New York’s Transit Police from 1990-92. He believed that if you allow lower-level crimes such as fare beating to occur without any consequences, it will lead to more serious crimes, and hence, he heavily enforced them. This action showed results in the first six months as subway crimes fell. I think it is a great idea as it shows everyday people that actions have consequences and instills in people a fear that worse things might happen if worse things are done. I think that people should be empathetic. What if someone is having a bad day and has to trespass for some reason like going to the washroom. But I am not someone to argue with the results.

NYPD had been one of the first US cities to start a 911 emergency call system. The department used data from these calls and other sources to create precinct-specific pictures of crime patterns. But the reports were compiled quarterly, so data was already four months old by the time it reached police commanders. The reports provided detailed “management information history” rather than a basis for decisions as real-world patterns change on weekly basis. This is exactly why the crime rate was not going down, as everything people were doing was reactive, not proactive.

FDNY had similar problems as NYPD as they also faced huge issues with the organization. Everybody was bold and self-sacrificing when they should have been more organized and focused on their goals. It is different in a field where you could hear the people screaming and smell the burning meat but just heading into action had already caused the deaths of hundreds of firemen. People died as the buildings were not safe in the first place.

Both NYPD and FDNY were trying to save lives but their target audience and approach were completely different. NYPD’s final goal was to reduce crimes while being organized, hence their KPI was completely different from FDNY who just wanted to lose fewer people in fire hazards.

Bloomberg was facing issues with sharing data between agencies and hence Michael Flowers was called in the smart data campaign. He identified technical, cultural, political and legal, issues in sharing data. His solution was to work within existing systems at each agency. For example, if the Buildings Department had an inspection system, Flowers’ team would not try to revamp it. Rather, it would use information from other city agencies, such as the Finance Department, to help Buildings improve the order, rather than the manner, in which it conducted inspections[1]. This led to 47 fire fatalities in 2013, a great decrease from 70 in 2012[1].

To solve the issue of rogue pharmacies in Statin Island, Flowers’ did a basic analysis of the redemptions for those specific high concentration oxy, found that one percent—about 20 of the pharmacies—were responsible for about 80 percent, 90 percent of the [oxycontin and oxycodone] distribution, at least for Medicaid redemptions. Then, further tested that by having HRA train their audit capacity on those pharmacies, and about 19 out of the 20 turned out to be up to no good.

MODA made government activities more visible to the public through user-friendly data and made its services more accessible. It is a great way to show performance and give people a sense of security. Though it is a centralized concept that has its pros and cons. I recommend that although we have MODA, each agency should have its Data Analysis team. It will inspire a desire to build an analytical mindset and make people’s thinking change from being reactive to proactive.

**Answers and Recommendations**

* Before CompStat was created, the police officers used to rate their performance by how long it took them to answer a 911 call, which was not as important as hitting where it hurts the crime rate. Data Analysis makes you question these metrics. It asks how a policeman should be evaluated, what is the most important thing that tells us the performance quality of these agencies. I think, having metrics helps as you have a clear goal and if you have a clear goal, it is easier to think of steps towards accomplishing it. I also think that these goals need to keep being reevaluated as the only constant in the real world is change. There is a big risk of your goals and hence your metrics becoming obsolete. It also narrows down the things you need to do, which is both, a good and a bad thing. The good thing is it gives purpose, and it is easier to achieve. It also means that things may go unnoticed.
* The case study makes some great points on how MODA has solved a lot of problems. For example [1], different agencies had different ways to identify buildings. The Post Office used addresses. The Department of Buildings used unique building identification numbers. The Finance Department used the lot number for the land a given building sat on. Finally, emergency response agencies such as the police and fire departments used latitude and longitude. I think it is unavoidable. One point for all this sensitive data needs to be kept very secure. It also raises a lot of privacy issues that need to be addressed. I think it is worth it to lose a bit of privacy and not get murdered on the street, but others might have a different opinion.
* MODA was a great solution to the issues at hand. It has given proven results and dramatically improved the standard of living of the citizens of New York. In my opinion, this was the best approach for the time. It has opened doors for increased and efficient communication and knowledge sharing among different agencies. However, it needs to be kept safe.
* I recommend that although we have MODA, each agency should have it’s own Data Analysis team. It will inspire a desire to build an analytical mindset and make people’s thinking change from being reactive to proactive. Certain things are also missed if there is a single point source that would get caught if it were being looked at by a person from a certain team. For example, a policeman would notice something about an area that a MODA employee might not in the same graph. I think they should follow a hybrid approach and that MODA is still a great idea.

**Conclusion:**

I think that this was a great approach that solved issues that were almost to solve without technology. I also think New York was the perfect place to start this as it already had a lot of the needed architecture already set up. It has had great achievement and should be incorporated into these agencies. However, any Data Analysis is only as good as the data provided and we should be aware of any bias generated in this data and how relevant it is in the current world. It needs to be tested against reality constantly and should not be relied on solely. One point for all this sensitive data needs to be kept very secure. It also raises a lot of privacy issues that need to be addressed. I think it is worth it to lose a bit of privacy and not get murdered on the street, but others might have a different opinion. In conclusion, I think it is a great initiative and should be supported as long as there is no data leak and we do not rely solely on it.

**Reference**:

1. From Compstat to Gov 2.0: Big Data in New York City Management. (n.d.). Retrieved January 27, 2021, from <https://northeastern.instructure.com/courses/66489/files/7107232?wrap=1>
2. Office, N. (2019, February 04). NYC Mayor's Office of Data Analytics: 2018 Year in Review and Look Ahead. Retrieved January 27, 2021, from https://medium.com/@NYCMayorsOffice/nyc-mayors-office-of-data-analytics-2018-year-in-review-and-look-ahead-1382903ec202