LIQUIDITY RISK MANAGEMENT DASHBOARD FOR REGULATORY AND DATA QUALITY BREACHES

ALI HAIDER
CEU ID: 185957

Masters' in Business Analytics – Capstone Project
Project Summary*

*Details on all aspects of the project are described in detail in the technical documentation

Introduction

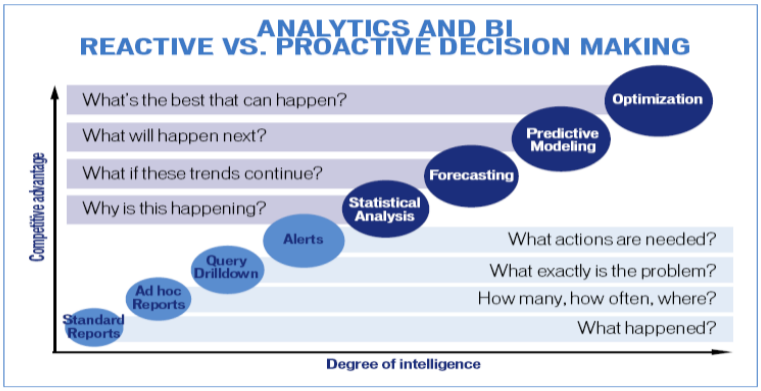
One of the major causes of the global financial crisis of 2008 was credit risk exposure by top investment banks from the Wall Street, which let subprime mortgages to individuals not being able to pay back their loans. The crisis which started off in the US banking and insurance sector, had far and wide ranging impacts on almost all developed and developing economies of the world – emphasizing on the fact of how closely knit the economies of today are. Some observers also relate the non-regulated derivatives markets at that time as fuel to the fire which had already erupted. Subsequently governments all around the world, had to step in and support the economies for them to survive. Henceforth it was inevitable that in the aftermath of the biggest financial crisis that the world has ever seen, more new regulations were enforced on the financial sector and the existing regulations were strengthened to avoid any similar collapse of banks in the future. Therefore, all investment banks had to refine and reinforce their risk management functions. This project deals with how some of the largest banks have taken the new risk management regulations in their stride. Big data can come to the rescue in this case whereby data monitoring can be made easy. We provide an example of an application which can be customized to automate monitoring of the most complex data sets thereby enabling large financial institutions to have critical figures at their fingertips constantly and thus contributing to a stable economic climate.

Study Background

One of the critical reasons behind the financial crisis of 2007-2008 was a buildup of massive uncollateralized risk in major American banks. In the years leading up to the crisis, mortgage lending sky rocketed whereby borrowers credit history was not scrutinized in the most risk averse manner before offering up loans. Banks assumed that the rise and fall of property markets across the nation would balance out thus keeping the loan repayments in checked. However, 2006 saw a property slump across the entire USA setting the stage for a seemingly endless series of defaults.

In response to the financial crisis, the federal reserve stepped in as it developed a new tranche of rules, regulation, checks and balances to establish greater control and stabilize the larger banks basically building a framework to prevent a similar crisis from happening again.

The above-mentioned events and their consequences were the backdrop against which my project was set. The main objective behind my research was to identify and explore a way in which big data can be used to help large, modern banks in meeting the regulations that have been set by the federal authorities not only in the US but also in EU and elsewhere, in the aftermath of the crisis and illustrate how big data modelling and user friendly data representation can be used to help a bank manage, protect, report and use the information to take astute business decisions and grow its business in previously unthinkable ways. In addition, I intended to enable the firm to move towards a more dynamic form of reporting that included predictive analytics and higher levels of user engagement.



While off the shelf solutions offered broader risk management programs, I worked with a major American investment bank for my project and I was lucky to already have been involved in Liquidity Risk Management so my project was also based around that specific risk stripe. Worth to mention here that liquidity risk is the risk that quantifies whether a business concern would be able to pay off its short-term liabilities or not. This is the basic definition of liquidity risk and clearly from the definition we can see it requires informed and quick decision making since it involves short term financial needs.

Methodology

I had to study the liquidity metrics across all products of the financial institution and across all its subsidiaries, spread across the globe, to come up with a proposal that would be useful for updating the metrics identified at the right level of frequency and could be reported not only to the FED but also internally to the Board, hence providing transparency regarding the risk metrics. Important to mention here that during this study phase of the entire liquidity risk spectrum of the bank critical findings that would have a major impact on building a unified tool for the firm included:

(a) Multiple databases for different products, e.g. Unsecured funding, prime brokerage and secured funding stored in different databases with different data structures
(b) Timing and data availability that would affect the upload frequency, as different franchises are based at different locations
(c) Volume of data as there are millions of transactions happening every day and what level of granularity would be needed to be analyzed and reported through the tool
(d) Details of statistical analysis
(e) Data storage capacity and constraints
(f) Communication of breaches to relevant teams and risk managers
(g) Ease of use for the business

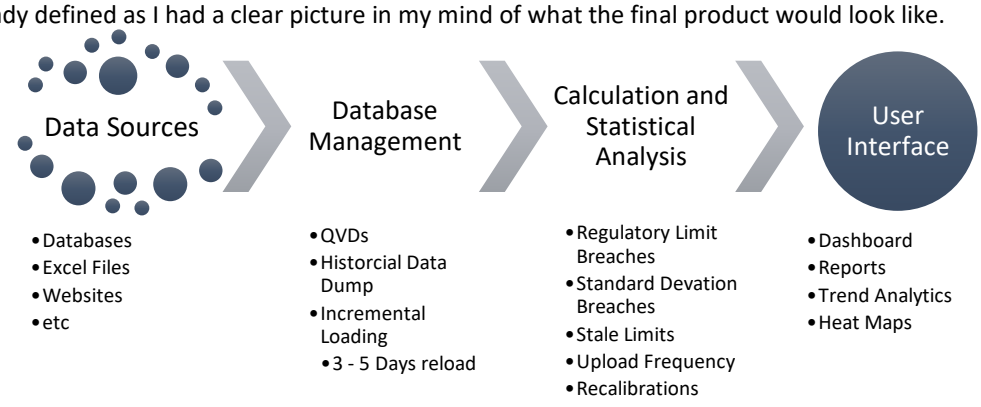
After looking at multiple options and doing the cost-benefit analysis between different tools and available products, I decided to use Qlikview as the desired tool through which the task at hand would be tackled with. Qlikview was chosen because it could process he volumes of data at a relatively high speed, being the pioneer in in-memory data processing, had its own database storage files called QVDs which occupy a fraction of disk storage when compared to the traditional files, thirdly and most importantly, it was the most compatible tool for joining different data sources and user had the option to join them in their own choice of schema.

Objectives & Implementation

The main objectives/goals for the project in conjunction with the stake holders were defined as follows:

(i) Replace existing static reporting with dynamic reporting
(ii) Create a historical database that would use least storage capacity
(iii) Updating of latest data using optimized techniques and with optimized querying
(iv) Calculation engine that can identify limit breaches set by the board and regulators
(v) Statistical engine that can identify data quality breaches
(vi) Trend analysis capabilities
(vii) Identify operational risk areas
(viii) Drill down functionality for informed decision making
(ix) User friendly interface for high user engagement

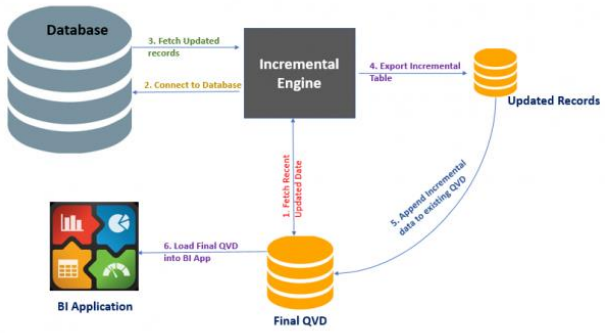
It was helpful during the coding and development phase that clear and concise business objectives and targets were already defined as I had a clear picture in my mind of what the final product would look like.



The first step was building the data infrastructure on which the whole dashboard would be developed. Important tasks at this stage were to determine how to connect different data sources that had varying data models and fields into a single unified table. For example, in prime brokerage it was important for the business to know what client exposure does each legal entity within the firm has, called client internalization, but the same does not hold true for unsecured funding due to varying business nature. The findings from the study of each product

type and subsidiary and source of data for each section, I decided to use the star schema for the data model of the dashboard. The central master table would hold each metric / KPI and tagged by the product it belongs to, along with other master data information details as a tag for subsidiary, responsible risk manager, limit level and so on, which then would be connected to the transactional tables. The advantage of using a scheme like this would-be disk storage and processing optimization as I would not need to store complete KPI and metric names in each table for each subsidiary. Having more than 600 metrics to report on, with twenty-five plus legal entities, this decision led not only to huge space saving but also reduction in time for fetching the data from the servers and databases every day to update the metrics.

After defining the data sources, incremental loading technique was used to feed the database, optimizing data storage and processing of the tool. As shown in the illustration, incremental loading meant that the tool was not querying the database for historical data but only for the rows that have been modified or are new transactions. The modified rows would then be updated on the existing database where as the new records would be appended into the whole data set. This was part of the incremental engine developed in which key fields that could identify new or old records were defined for each data source. The database created through these steps would then be used to feed into the calculation and statistical analysis engine that would perform calculations on:



(1) Limits breaches as defined by regulators and the board

The most important functionality of the data quality and analytics dashboard for liquidity risk was to enable the risk managers to identify and breaches on the 600+ metrics defined. The limits or allowed exposures of these metrics were already available to us and they were just connected to the actual exposures calculated for each day by the tool itself. An example of a limit breach would be for example, the firm should only be exposed to 20% of unsecured funding as a proportion of its total sources of funding. If, however on 10th of December, the unsecured funding exposure as a percentage of total funding sources rises to 20.5%, the dashboard should be able to identify this breach. In this scenario, the 20% would be the limit and the 20.5% is the limit exposure or limit actual for that date. The limits themselves are calibrated by the risk managers every year based on regulations and historical exposures and the limit actuals or the actual exposures need to be calculated by the tool through the normalized data QVDs. This required extensive coding and back-end Qlikview scripting.

(2) Standard deviation and data quality breaches

Another functionality included calculating the standard deviation breaches for each risk metric separately for the latest reporting day. Since we I had fetched and stored in QVDs the complete historical data, I was able to calculate the actual risk exposure for each metric for each day in the past. This allowed me to visualize the trend of the risk exposures and with varying degrees of certainty, different standard deviations, and predict how much the exposure would be on a day in the future if the metric exposures are normally distributed.

(3) Trend Analysis

The last step in development of the tool was to use the calculation engines results and display them through an interface that is user friendly. The following dashboard was eventually created that gives would give a birds' eye view to a risk manager on the performance of the firm on liquidity risk related metrics on a given day, with the ability to drill down into any aspect they feel needs to be investigated:

Conclusion

The learnings and outcomes for me from undertaking this project and since it was capstone to the Master's program in Business Analytics at CEU can be broadly categorized into technical or professional, educational and personal development related.

It goes without saying that the program and the coursework enabled me to accomplish this gigantic task with much ease as it equipped me with the knowledge of tools and techniques that were needed to handle big data. We learnt about database infrastructures from Professor Zoltan Toth enabling me to apply the principles to implement techniques such as incremental loading at the financial organization that had not been practiced before. The data analysis courses spread across the whole year of the master's program were key to develop the calculations engine for this project about which I have talked in detail in the fifth chapter of this report. Statistical techniques, trend analysis and standard deviation models were all methods which I studied about in the program and applied at the firm. Finally, the data visualization course, in which we learnt about Tableau, opened the window for me to be able to visualize data in ways that allows user driven analysis, leading to more informed decision making and high user engagement. Just from a technical stand point, I feel I have grown immensely over the past two years and now having practically implemented the learnings from this time through this project I feel confident about my technical expertise.

The educational and personal development learnings from the whole experience were immense. Engaging with different stake holders, making sure they are all as engaged in the development of the project as yourself and crossing the finish line within the given deadlines not only taught me inter-personal skills but also time management and personal organizational skills. Being involved and learning first-hand about the corporate bureaucracy would also something that will stay with me, as I learnt the delivery on any project is directly linked to these aspects.

One important learning specifically from the project itself would be that spending more time at the initial data gathering phase instead of just starting the project and gathering requirements as the development goes on, would save a lot of time and resource. Each project phase must be well defined and time bound, and as much time should be spent at understanding the problem and business case as is spent of development of the tool/project for the resolution of the problem.