Grant Thornton
An instinct for growth™

Keys to extract value from the data analytics life cycle

There is a tremendous opportunity to obtain insights from all activities that make up the analytics life cycle. Value is not limited to the end results produced from data analyses.

The role of data analytics in business

Across industries, the applied use of data to inform business decisions has become a foundational, if not critical, element of business. In financial services, capital itself is effectively deployed to banking activities — from lending to product and service design, through the optimization provided by institutionalized data analytics. Mathematical models that forecast customer profitability, segmentation models that simulate credit losses driving regulatory capital, scorecards that assist in credit originations, models that simulate balance sheet impacts resulting from risk factor changes, and scenario -based models (2) that project potential impacts from broader operational or environmental risks (1) are now commonplace. More and more, predictive analytics that take forms such as key performance indicators (KPIs) and key risk indicators (KRIs) have become the norm.

Certain global institutions are pioneering applications of analytics to automate the cognitive processing of text-based audit reports while increasing consistency over human capacity (4). Others are exploring causal relationships between the risk event, audit and indicator data to reveal predictive insights. Analytics can be a formidable, competitive advantage in any function, whether used to reveal insights into revenue generation, day-to-day operations or risk management. Leading institutions are using data analytics at the enterprise level to increase effectiveness of decision-making, which can yield significant financial returns. It is worthwhile noting that a new economy exemplified by entire enterprises that rely on data, rather than tangible assets to drive revenue such as Uber, Google, Facebook, and Fintech, is on the rise. In financial services, regulatory mandates driving transparency and financial objectives requiring accurate understanding of customer needs have heightened the importance of data analytics to unprecedented levels

Performing data analyses

The purpose of any analysis is to simplify and derive understanding from some event that has been characterized through data. By deploying analytics, we aspire to translate observed phenomena into metrics we can understand and react to and then make informed decisions.

There is, however, an element of human bias that we inherently introduce. From this basic purpose, many different characterizations or objectives are typically pursued. These include but are not limited to (1) collapsing multiple dimensions into one — such as a scorecard or KRI, (2) introducing structure to enable understanding of a complex situation, (3) identifying causal drivers and types of outcomes, (4) seeking to identify control levers, (5) quantifying extreme values and limits, and (6) developing predictive and actionable insights.

Scorecards collapse multiple data inputs and reduce an analysis to one dimension. In credit scoring, for example, an applicant's credit attributes that could include income, debt load, past history of repayment and type of debt are combined to derive a single score from which one can easily set approval-escalating thresholds (5). Complex situations, such as the need to understand purchasing behavior, are simplified by the use of clustering — a technique used to define clusters or groupings of customers who exhibit like behaviors. Without this simplification, it would be very difficult or impossible to discern the behavior of one customer over the other based on tens if not hundreds of customer attributes.

One basic activity of data analysis is establishing cause-effect relationships between data. In fraud management, financial transaction patterns are used to determine potential for future fraud. In credit collections, data analytics is applied to determine contact frequency and defined procedures to optimize repayments. In effect, analytics unveils control points from which financial institutions can deploy resources or technology more efficiently. In the current regulatory environment, banks are employing scenario-based stress testing (6) to quantify capital requirements given "severely adverse" macroeconomic conditions (a type of extreme value analysis). Many financial institutions are now exploring "big data" as a way to reveal valuable and predictive insights about customers, operations, and products and services.

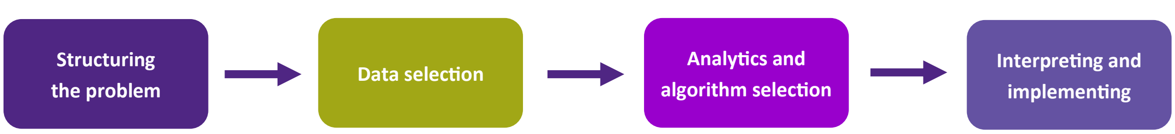
Example of Application of Analytics: Operational risk modeling in insurance claims processing

After one of its top executives' laptops had been stolen, a global insurance firm became hypersensitive to data privacy and technology risks. It engaged a consulting firm to develop a model that could quantify financial exposures and losses in its claims processing operation, hypothesizing that many data and technology risks were inherent in this part of the operation, which processed hundreds of millions of dollars in claims and premium transactions on a daily basis.

In the course of the engagement, multiple data sources were leveraged in the development of a dynamic model that connected all resources physically supporting this operation, including buildings, computer hardware and software, people and processes. The model successfully quantified potential exposures based on the probability of failed processes. However, the biggest value from this analytic exercise came from uncovering the fact that the two locations where customer data was stored were too close to each other geographically and both vulnerable to similar environmental hazards. In fact, in the discovery process, the firm did not have a robust process for business continuity in place, despite a history of impacts from weather events. The analytical exercise was successful in quantifying potential operational risk losses but the journey proved to deliver a much more valuable piece of insight.

The analytics life cycle

One way to ensure the success of any analytic exercise and explore the potential value of data analytics is by decomposing analytics into some type of lifecycle. The following presents a simple framework comprised of four components:



Structuring the problem

Proper structuring or framing of the business challenge is critical to achieving business-relevant and actionable outcomes. A key consideration in this is defining the unit of measure. In financial risk management analyses, there is often the need to translate risk exposures or impacts into a unit of measure that has some relevance and suggests economic value in terms of dollars.

The nature of the decision to be supported by analytics also needs to be considered. Take, for example, an analysis intended to predict the likelihood of an operational risk event occurring. A model could be structured to produce a "yes" or "no" answer — each with a certain probability. Alternatively, the problem could be structured to produce a single probability, which would be a more fitting structure. Oftentimes, in the absence of historical data or experience, it is difficult to place a dollar value on the potential impact of an event. The problem is then usually structured to produce a categorical output with a vague definition, such as a high, medium, and low impact, which would subsequently be difficult to develop implementable actions.

Sometimes, analytics requires inputs via scenarios — as is the case with regulatory stress testing. Organizations need to define business-relevant macroeconomic scenarios that will drive portfolio analyses. In structuring these scenarios, it is important that key stakeholders' buy-in is obtained upfront. Ultimately, any decision-making will depend on the trust stakeholders have in the scenarios driving the analytics performed.

Data selection

Data deserves special attention when it comes to deploying analytics. A critical analysis and examination of the data itself needs to be considered. All too often, data analyses make a number of presumptions about the data being used, which can introduce a number of biases and expensive mistakes. Selecting data that is used in an analysis, proper filtering of irrelevant data, ensuring high data integrity, and data element definition are core in any analytics process.

For example, regulatory capital required for operational risks mandate appropriate modeling of the tail in loss frequency distributions. Practitioners in the field recognize that to produce such accuracy would require institutions to gather about 100 years' worth of data, just to get statistically sufficient data points. However, one could then challenge the relevance of a data point that is 60, 40, even 20 years old to a current analysis (3). So lots of data or long histories don't necessarily translate into usable data for analytical purposes.

The characterization of data going into a model presents interesting analytical opportunities. Often missing values in data elements, which are typically seen as challenges, could be structured as an opportunity to "exercise" the model for extreme potential values. This technique could introduce a type of synthetic robustness in a model, which might otherwise be limited by data sufficiency. In statistics, this could be a bootstrapping-type technique to help smoothen holes in data sets, rather than drive to a homogenization of data.

Finally, intuition should be exercised in selecting and excluding certain data. In our experience, it is a healthy component equal to any set of assumptions that typically accompany any analysis. Wise analytics teams also find business value along the way in the modeling journey. Missing data becomes the opportunity to deploy human intuition through scenario approaches. Furthermore, sensitivity analyses may yield optimal conclusions as well.

Analytics and algorithm selection

The analytic framework, approach or algorithm is really the "lens through which the analyst chooses to view the problem." The analyst has a range of "funnels" through which data will be processed. It is, therefore, imperative that the analyst has an appreciation for a broad array of options and select the most appropriate one for the problem being analyzed. The analyst should also be cognizant of the benefits and limitations of each method and that problems are structured to produce understandable solutions and results. The following figure presents a selection of analytical approaches, as well as their potential applications, benefits and limitations.

Keys to extract value from the data analytics life cycle

Typical analytic methods applied in financial services

Analytic method	Potential applications	Benefits	Limitations
Scorecards	• Credit decisioning • Customer profitability • Collections management	• Simplifies decisioning into one dimension • Can incorporate feedback — "behavioral scorecards" • Can be based on simple rules, linear and nonlinear methods	• Needs to be refreshed to stay relevant and effective • Requires complete information to compute a score • Applications are usually very specific
Clustering and segmentation	• Marketing analyses • Customer behavior/profitability • Data mining	• Algorithms identify clusters with homogeneous properties • User-friendly tools • Alignment between intuition and typical outcomes	• Extra effort to document explanatory rationale • Visualization • Controlling for overfitting, prespecification of number of clusters
Logistic regression	• Probability modeling/KPIs/KRIs • Credit risk modeling • Predicting bankruptcy	• Standard approach in many financial services processes • Transparency of independent and dependent variables • Standard function in common statistical tools such as SAS	• Large amounts of data to achieve stability Multicollinearity • Limited outcome variables (e.g., not continuous)
Artificial intelligence/neural networks	• Predicting fraud in transactions • Voice recognition/IVR • Replication of human expertise	• Potentially superior to human capability to process data • Consistency • Cost savings	• Challenging to make transparent • Requires controls and monitoring • Perception of "black box" technology
Expert rules	• Setting credit limits • Exception handling • Application of human expertise	• Human knowledge directly applied • Simple to operationalize • Can be combined into other methods or by itself	• Rules may be limited to a specific application • Nonstatistical approach subject to error • Maintaining rule relevance in decisions supported
Systems dynamics	• Analysis of transaction processing • Cause and effect situations • Resource constraint situations	• Ideal for analyzing business processes • Fully contained model identifies all inputs/outputs • Inputs/outputs can be discrete/continuous	• Specialized software to process mathematical algorithms • Needs to exercise caution in intepreting results • Needs to account for assumptions
Scenario analysis	• Stress testing for regulatory capital • Business continuity planning • Wealth management and investments	• Enables explicit definition of inputs • Scenario inputs could be comprised of several variables • Allows analysis of possible ranges (e.g., min/max)	• Scenarios are subject to incorrect assumptions • Scenarios often defined by past experience • Scenarios may not be internally consistent
Extreme value theory	• Unexpected loss modeling • Predicting bank default • Operational and market risk/VAR	• Robust and rigorous mathematical theory • Straightforward to apply to data • Does not require strong assumptions prior to analysis	• Inefficient in processing data considered nonextreme • Assumptions about tail estimation
Text analytics	• Big data • Processing audit reports • Sentiment analysis	• Offers way to process text information in structured way • Enables processing of non-numerical information • Introduces behavioral, sentiment and other dimensions	• Requires codification and structuring of a model • Needs to filter for subjectivity and interpretation
Data visualization	• Communications tool • Network analysis • Business intelligence	• Easy to spot trends, relationships • Manipulation and interaction with data • Processes information in new ways	• Graphic introduces bias in how information is depicted • Simplification of information
Actuarial and econometrics	• Loss event prediction • Macroeconomic forecasting • Various aplications in insurance industry	• Data-driven analytical processes • Integrates economic theory with intuition • Leverages computational mathematics	• Analytics limited by data coverage • Challenging to incorporate effects of policy into analytics • Technical challenges such as model selection

Interpreting and implementing results

Finally, the interpretation and implementation of the results of an analysis or model are equally if not more important than the analysis itself. This implies that the results are properly translated into a business application, and any analytical or limitation of model outputs are recognized in the decision-making process. Any analytics soundness assessment requires validation techniques such as: (1) reviewing assumptions, (2) cross-checking outputs, (3) ensuring appropriate implementation, and (4) ensuring analytic is timely monitored and calibrated. This final stage in the life cycle of an analysis is about aligning the analytic to the business problem.

Assumptions should be revisited to ensure that they don't disproportionately impact results in an unreasonable manner. Where possible, results should be cross-checked to ensure consistency. An analysis of severe scenarios should exhibit the same type of trend in the outputs. Implementation of the analytical solutions should also be carefully operationalized, and staff should be fully trained to execute decisions. Finally, analytics should include defined provisions to be refreshed so that its application remains relevant and under an acceptable operational risk appetite.

Interpreting and implementing results

Data analytics present a formidable tool for uncovering business insights and optimizing decision-making in financial services. There is an abundance of methods and tools that can be utilized to process and exploit data to provide insightful value in the decision-making process. Going through systematic processes involved in data analyses and modeling can provide valuable insights as resulting outcomes. Analytical expertise, along with an experienced-practitioner intuition, are very valuable, both when supporting business decisions derived from analytics outcomes and when managing the inherent data and methodological limitations. An effective internal control framework is essential in providing reliability to any analytical construct.

Special notes

The authors of this article present insights gained from extensive experience in the financial services sector over the past two decades, particularly as the availability of data, the emergence of analytics applied to financial services functions, and the capacity of computing facilities have created synergies in bringing these elements together. The authors have brought forth unique elements of analytics as applied across a broader spectrum of industries and functional decisioning applications that include, but are not limited to: (1) banking: retail credit decisioning; (2) public services: transportation planning and modeling; (3) business continuity: pandemic business impact analytics; (4) supply chain: systems dynamics modeling and value-at-risk; (5) artificial intelligence: neural networks; (6) marketing: cluster analysis and direct-mail solicitations; (7) financial services: loss forecasting and stress testing; and 8) aerospace: thermo-structural analysis of space vehicles.

Contacts

Jose Molina
Principal
Financial Services Advisory
Risk Center of Excellence
T 312.602.8330
E jose.molina@us.gt.com

Frank Saavedra-Lim
Director
Financial Services Advisory
Risk Center of Excellence
T 704.926.0365
E frank.saavedra-lim@us.gt.com

Nigel Smith
National Managing Principal
Financial Services Advisory
T 212.542.9920
E nigel.smith@us.gt.com

References

[1] Chen-Ritzo, Ching-Hua, An, Lianjun; Buckles, Stephen; et al. "Pandemic Business Impact Modeler," Proceedings of the 2007 INFORMS Simulation Society Research Workshop.

[2] Cheng, Feng; Jengte, Nitin; Min, Wanli; Ramachandran, Bala; Gamarnik, David. "Modeling Operational Risk in Business Processes," Journal of Operational Risk, IBM Thomas J. Watson Research Center and MIT Sloan School of Management, Aug. 14, 2007.

[3] Cope, Eric; Mignola, Giulio; Antonini, Gianluca; Ugoccioni, Roberto. "Challenges in Measuring Operational Risk from Loss Data," Global Association of Risk Professionals Conference, 2009.

[4] Finch, Glenn; Davidson, Steven; Kirschniak, Christian; et al. "Analytics: The Speed Advantage," Executive Report, IBM Global Business Services, 2014.

[5] Saavedra-Lim, Frank, and Breeden, Joseph. "Global Best Practices Course on Credit Scoring," Risk Management Association Course, 2005.

[6] Saavedra-Lim, Frank, and Molina, Jose. "CCAR/DFAST: Institutionalizing Stress Testing Into Everyday Banking Operations and Strategic Planning Processes," Grant Thornton LLP, 2016.

Grant Thornton

An instinct for growth™

Connect with us

grantthornton.com

@grantthorntonus

linkd.in/grantthorntonus